



Quo vadis morphology?

MMM10 On-line Proceedings

Edited by:

Jenny Audring
Francesca Masini
Wendy Sandler



Quo vadis morphology?

MMM10 On-line Proceedings

Edited by:

Jenny Audring

Francesca Masini

Wendy Sandler

On-line Proceedings of the
Tenth Mediterranean Morphology Meeting (MMM10)
Haifa, Israel, 7-10 September 2015

The MMM Permanent Committee

Jenny Audring (Leiden University)

Geert Booij (Leiden University)

Nikos Koutsoukos (Université catholique de Louvain)

Francesca Masini (University of Bologna)

Angela Ralli (University of Patras)

Sergio Scalise (University of Bologna)

ON-LINE PROCEEDINGS OF THE
MEDITERRANEAN MORPHOLOGY MEETINGS
ISSN: 1826-7491

Table of contents

Foreword	v
Abstracts	vi
Suspended affixation with derivational suffixes and lexical integrity	
Faruk Akkuş	1
The Semitic templates from the perspective of reciprocal predicates	
Elitzur A. Bar-Asher Siegal	16
Motivated phonological templates in Sign Language	
Gal Belsitzman, Wendy Sandler	31
The morphosyntax of definiteness agreement in Neo-Aramaic and Central Semitic	
Edit Doron, Geoffrey Khan.....	45
Telicity makes or breaks verb serialization	
Kazuhiko Fukushima	55
The role of stem frequency in morphological processing	
Hélène Giraud, Serena Dal Maso, Sabrina Piccinin.....	64
“Romanes eunt domus”: where you can go with Latin morphology	
Variation in motion expression between system and usage	
Claudio Iacobini, Luisa Corona	73
Morphology: the base processor	
Aysun Kunduracı, Aslı Göksel	88
Lexical blends and lexical patterns in English and in American Sign Language	
Ryan Lepic	98
Grammaticalization is not the full story: a non-grammaticalization account of the emergence of sign language agreement morphemes	
Irit Meir	112
A paradigmatic analysis of the Italian verbal derivation	
Fabio Montermini, Matteo Pascoli	125

A comparison of roots as units of analysis in Modern Hebrew and Spanish: exploring a remnant approach to defining roots	
Ignacio L. Montoya	136
Inflected and periphrastic features: issues of comparison and modelling	
Gergana Popova	146
The lexicalization of complex constructions: an analysis of adjective-noun combinations	
Marcel Schlechtweg, Holden Härtl.....	159
Spatial reduplication in Sicilian: lexicon or grammar?	
Giuseppina Todaro, Fabio Montermini	169

Foreword

Since 1997, the Mediterranean Morphology Meetings (MMM) have been jointly organized by Geert Booij (Leiden University), Angela Ralli (University of Patras), and Sergio Scalise (University of Bologna). As of 2013, organization is in the hands of Jenny Audring (University of Leiden), Nikos Koutsoukos (Université catholique de Louvain / University of Patras) and Francesca Masini (University of Bologna). MMM10 was made possible thanks to the excellent local organizing committee chaired by Wendy Sandler (University of Haifa).

The aim of MMM is to provide an informal setting for morphologists to present and discuss their work. The single-session setup guarantees maximal interaction between researchers, and gives young linguists the chance to present their work at a conference of moderate size, where fruitful contacts with senior linguists can be established. The first ten meetings – in 1997 (Mytilene, Greece), 1999 (Lija, Malta), 2001 (Barcelona, Spain), 2003 (Catania, Sicily), 2005 (Fréjus, France), 2007 (Ithaca, Greece), 2009 (Nicosia, Cyprus), 2011 (Cagliari, Sardinia), 2013 (Dubrovnik, Croatia) and 2015 (Haifa, Israel) – have proven the success of this formula.

In good tradition, this volume continues the MMM Online Proceedings series with a selection of papers presented at MMM10, which took place on September 7-10, 2015 in Haifa (Israel). Many good abstracts were submitted, the attendance was high, and a number of leading researchers participated, including invited speakers Stephen R. Anderson (Yale University), Mark Aronoff (Stony Brook University) and Ray Jackendoff (Tufts University). The editors of this volume wish to thank the many authors who submitted their papers to the MMM Online Proceedings and Nikos Koutsoukos for his constant support throughout the process.

Abstracts

Faruk Akkuş: *Suspended affixation with derivational suffixes and lexical integrity*

In this paper, I investigate suspended affixation, particularly suspended affixation formed with derivational suffixes with a focus on Turkish. Showing that this phenomenon is attested in both the nominal and verbal domain, I defend the argument that such constructions cannot be reduced to an account of natural coordination or be explained as being part of the lexical word formation. Accordingly, I discuss the relevance of the suspended affixation for the morphology-syntax interface. I conclude that it might have significant implications for the Lexical Integrity Hypothesis (LIH).

Keywords: suspended affixation, Lexical Integrity Hypothesis, derivational suffixes, coordination.

Elitzur A. Bar-Asher Siegal: *The Semitic templates from the perspective of reciprocal predicates*

One of the main questions that theories about Semitic verbal morphology aim to answer concerns the relationship between verbs with different but related meanings that share the same phonological root but appear in different templates. The goal of this paper is to shed some light on this broader question by considering the so-called “reciprocal verbs” (*rec-predicates*), i.e., verbs with certain morphology that allegedly encodes reciprocal relations. Such verbs often appear in the T-templates, across the Semitic languages, thus this paper would like to examine their relation to other verbs with the same root.

All previous analyses of verbal reciprocals assume that *rec-predicates* are at some level of analysis derivatives of more basic-predicates. Furthermore, most of the reciprocals in Hebrew are in the T-template, and the assumption in various theories about verbs in this template is that they are derivative of functions, that have as their input either the root or verbs in other templates.

This paper argues that there is no grammatical relation between *rec-predicates* and other transitive verbs in the same root, by pointing out problems with previous derivational analyses, by analyzing the argument structure of these predicates and by providing a semantic account for the various readings the *rec-constructions* have. The differences between the current proposal and the previous ones stem from the fact that the current analysis does not consider the *rec-predicate* to be an encoding of reciprocal/symmetric relations and offer, therefore, an alternative portrayal of the relationship between them and the symmetric events they denote.

Since previous studies on the morphology of the templates pay only little attention to the verbal expressions of reciprocity, the goal of this paper is to see what theories concerning the morphology of the templates in Modern Hebrew would have to account for with respect to these verbs.

Keywords: reciprocals, reflexives, templates, Hebrew, collective and distributive, root.

Gal Belsitzman, Wendy Sandler: *Motivated phonological templates in Sign Language*

A basic design feature of language is duality of patterning, the existence of a meaningless level of elements that combine to create meaningful morphemes and words (Hockett 1960). Although the signs of sign languages have iconic origins, Stokoe (1960) showed that sign languages do have a meaningless level, akin to phonology, setting the stage for much subsequent linguistic research on sign languages at all levels (Sandler and Lillo-Martin 2006). Here we show, following Lepic et al. (2016), that part of the phonological structure across sign languages is often motivated by meaning. Specifically, two-handedness is motivated, as are details of structure in two-handed signs previously believed to be strictly phonological and hence meaningless, such as symmetry, dominance, type of movement, and patterns of contact between the two hands. We further develop a templatic model of sign structure (Sandler 1986, 1989) to reveal which aspects of the phonological form in two-handed signs are often so motivated. Noting ways in which meaning penetrates phonology in spoken languages too, we conclude that the line between the dual levels of patterning in language is not as sharp as is commonly believed.

Keywords: sign language, two-handed signs, motivation in phonology, duality of patterning.

Edit Doron, Geoffrey Khan: *The morphosyntax of definiteness agreement in Neo-Aramaic and Central Semitic*

The article describes the progression of the Neo-Aramaic dialects along the first stage of Greenberg's cycle, where demonstrative pronouns turn into definite articles. We suggest that the same progression might have originally taken place in Central Semitic, and that it is this process which accounts for the multiple marking of definiteness in the Central Semitic noun phrase. The article describes the two main factors of the change in Neo-Aramaic: First, the syntactic status of the definite article changes from a phrasal demonstrative to a lexical determiner head. Second, the attachment of the definite article to the attributive adjective originally marks the adjective as contrastive. These two factors put in motion a process whereby an original demonstrative phrase evolves into a marker of definiteness agreement. When the article is still a phrasal constituent, it attaches within the noun phrase either to the adjective or the noun, never to both. In subsequent stages, where the demonstrative pronoun has turned into a determiner, it may attach to both noun and adjective within the same noun phrase. Eventually, the latter option may grammaticalize into a marker of definiteness agreement, as it did in such Central Semitic languages as Arabic, Aramaic, and Hebrew.

Keywords: article, determiner, demonstrative, demonstrative cycle, definite, adjective, contrastive, deictic, anaphoric, definiteness agreement, Central Semitic, Neo Aramaic.

Kazuhiko Fukushima: *Telicity makes or breaks verb serialization*

The major focus in the research of lexical V_1 - V_2 compounds in Japanese has been on the conditions for argument matching between two verbs. This paper sheds light on the role of telicity as an additional condition regarding verb serialization with Japanese as a test case. First, it is shown that V_1 cannot be telic unless V_2 is also telic. Any other patterns are possible. Second, (im)possible aspectual combinations are shown to be a consequence of aspectual composition based on Dowty's (1986) aspectual classification of predicates.

Keywords: lexical V-V compounds, telicity, verbal aspects, verb serialization, temporal iconicity, Japanese.

Hélène Giraud, Serena Dal Maso, Sabrina Piccinin: *The role of stem frequency in morphological processing*

The aim of the present research is to investigate frequency effects in morphological processing and to provide insights into the role of the stem in lexical access. To this aim we conducted a masked priming experiment associated with a lexical decision task (Forster and Davis 1984) focused on Italian suffixed words, in which we manipulated the frequency of the stem of the target words with respect to the derivative primes. More precisely, we opposed high frequency stem targets (e.g., *trasferimento* ‘transfer’- *trasferire* ‘to transfer’) to low frequency stem targets (e.g., *motivazione* ‘motivation’ - *motivare* ‘to motivate’); the frequency of the primes, on the other hand, was comparable. Results show full morphological priming effects for both types of targets, irrespective of stem frequency. This suggests that suffixed words are accessed holistically and not through the stem and therefore via a decomposition process, as suggested by previous studies based on simple lexical decision tasks. We argue that, while the lexical decision task is not suited to explore the very early stages of word recognition, given that the derived word is perceived consciously, masked priming focuses on fast automatic non-conscious mechanisms of lexical access, as the activation of the masked prime is constrained by strictly determined time limits. The masked priming paradigm therefore does not examine the entire process of word recognition as the lexical decision task does, but represents a window on the transfer of activation between a prime and a target.

Keywords: morphological processing, stem frequency, Italian, masked priming.

Claudio Iacobini, Luisa Corona: *“Romanes eunt domus”*: where you can go with Latin morphology. *Variation in motion expression between system and usage*

In this paper, based on a corpus analysis of Classical Latin texts, we show that, although Latin displays at the system level a wide array of linguistic resources characterizing Satellite-Framed languages, the actual usage of the strategies employed in motion encoding significantly differs from what is expected in a typical Satellite-Framed language.

Our results claim in favour of a neat distinction between system and usage in the typological classification of motion events, since a rich set of linguistic means available for motion encoding at the system level is not a sufficient condition for assessing the actual strategies being used in a language.

The findings about the preferred strategies of motion encoding in Latin can contribute to a better understanding of typological change in motion expression. We believe that the limited use of manner verbs, the lack of complexity of Path, together with the semantic congruence of Path (expressed in the prefixed verb) and prepositional phrase, may constitute conditions paving the way for the typical expression of dislocation motion in Verb-Framed languages, in which the function of indicating the Path is carried out by the verb, whereas prepositional phrases tend to express neutral meanings with respect to the static / dynamic distinction.

Keywords: motion events, typological change, Latin linguistics, system, usage.

Aysun Kunduracı, Aslı Göksel: *Morphology: the base processor*

One of the motivations behind the Lexical Integrity Hypothesis (LIH) was to draw a distinction between word structure and phrase structure as evidence for a separate morphological component. In this study, we claim that drawing such a line between morphology and syntax is not contingent on whether word structure can be built on purely morphological items; rather, the crucial issue seems to be whether an operation is manipulated by a dedicated morphological mechanism, and on the properties of the output(s). We start from the premise that morphological well-formedness cannot be derived from syntax: morphology (the word-formation component) must be autonomous with its own principles (as in Anderson 1992, Aronoff 1994, Beard 1995, Pounder 2000 among others). We then turn our attention to the function of morphology. We show that a morphological component responsible for creating and inflecting lexemes and performing activities upon syntactic and prosodic phrases is necessary. When performing its tasks, morphology uses base operations and provides the bases on which morphological processes (i.e. word-formation and inflection processes) operate. Moreover, we show that these bases can be stems, word-forms, syntactic constituents, and prosodic constituents. Thus one aspect of the LIH which does not allow postsyntactic morphology (No-Phrase Constraint, Botha 1983) must be eliminated so that we can account for both simple and complex outputs of morphology. We conclude that the fact that complex morphological expressions are possible does not necessarily show that there is no (need for a) morphological component. Rather, morphology is the only component in which base operations take place to generate word formations.

Keywords: lexical integrity, autonomous morphology, base types.

Ryan Lepic: *Lexical blends and lexical patterns in English and in American Sign Language*

Lexical blending has long been recognized as a creative and productive strategy for coining new words. English *brunch*, for example, is an established lexical blend of *breakfast* and *lunch*, and *cronut* is a more recent blend of *croissant* and *donut*. Lexical blending has also traditionally been viewed as a largely unpredictable process. However, recent studies have demonstrated that, though blending is probabilistic rather than categorical, blend structure is indeed constrained by phonological and semantic considerations. This paper examines some consequences of lexical blending for morphology and for morphological theory, particularly issues stemming from the fact that parts of existing words seem to develop new or specialized meanings as a result of the blending process. Here I examine smaller, less-established blending patterns in two languages, English and American Sign Language. I argue that, though many individual blends are unlikely to become established, conventional lexical items, the general mechanism that allows speakers to produce and interpret blends is clearly part of morphology: Lexical blending forges and reinforces connections between whole words. Accordingly, blending is most appropriately viewed as a step between compounding and derivation, an analogical process that typically creates and exploits paradigmatic lexical relationships. These findings are situated in the framework of Construction Morphology, which allows for the straightforward description of morphological patterns as structural alignments of form and meaning and as templates for the creation of new words.

Keywords: American Sign Language, analogy, construction morphology, English, lexical blends, word-creation, word-formation.

Irit Meir: *Grammaticalization is not the full story: a non-grammaticalization account of the emergence of sign language agreement morphemes*

Many studies show that grammatical morphemes and categories emerge by means of grammaticalization. Here I argue that grammaticalization cannot account for all cases of the emergence of grammatical elements. Drawing on the development of agreement verbs in Israeli Sign Language, I suggest another morphological mechanism for morpheme emergence that supplements grammaticalization: carving morphemes from monomorphemic words. Sign language agreement verbs constitute a class of verbs with a shared semantic component – verbs denoting transfer, and a shared morphological structure - marking agreement with their subject and object arguments by copying the arguments' R-loci onto their initial and final locations. A diachronic study of verbs of transfer in ISL (Meir 2012) reveals that agreement verbs were initially monomorphemic, and eventually developed morphological complexity, by means of reanalyzing the initial and final locations of the signs as morphemic. I suggest that the key property underlying this process is that verbs of transfer in the manual modality share both *a meaning component* and *a physical formational component*. By acknowledging meaning-form resemblance across words and reanalyzing the shared phonological segments as morphemic, a morpheme is extracted from a formerly monomorphemic word. While processes along these lines have been attested in spoken languages as well, sign language agreement verbs show that such a process can create not only new morphemes but also a new grammatical category.

Keywords: grammaticalization, sign language agreement verbs, grammatical categories.

Fabio Montermini, Matteo Pascoli: *A paradigmatic analysis of the Italian verbal derivation*

Word and Paradigm theories of inflection can be classified as inferential-realizational theories, according to the classification of Stump (2001), in that the associations between morphosyntactic properties and exponents of morphology are not listed in the lexicon, but are identified by rules which relate the inflected form with the root, and are selected by their morphosyntactic properties (inferential theories); and morphosyntactic properties are not added to the word by an exponent, but these properties select the exponents that realize them (realizational theories).

It has been observed (cf. *inter alia* Maiden 1992; Pirrelli and Battista 2000) how alternations, on verbal stems that present allomorphy, meet a surprisingly regular distribution, which is not dictated by the phonological context. This regularity reflects the organization of the verbal paradigm, or the set of all inflected forms for each lexeme, into *morphomes*, purely formal entities independent from morphosyntactic features (Aronoff 1994).

In the last twenty years, there has been much interest in studying the *paradigmatic distribution of allomorphy*, or the way in which variation between forms (the traditional “irregularity”) of a paradigm rests on regular patterns.

Practically, these studies aim at analysing the paradigmatic structure of inflection, i.e. to decompose the paradigm into zones where forms are realized on possibly distinct basic stems, and to examine the formal relations (on the phonological level) between these basic stems, looking for predictability chains allowing to handle both regular and irregular lexemes.

In the present work, we examine the formal relations between some verbal derivatives in Italian and the basic stems of the related verbs with the goal of extending the study of paradigmatic distribution to derivation.

In Latin, verbal derivatives in *-tionem* (event/ result), *-torem* (agent/ instrument), *-tura* (event/ result), *-tivus* (relational adjective), *-torium* (adjective/ instrument/ place) were built on the supine (basic) stem, like the past participle. Italian, as other Romance languages, has inherited from Latin both the process and the derived forms. Some derivatives, and in some cases the past participle itself, underwent semantic drift. Some (ancient) past participles are no longer connected to a verb and remain as independent adjectives (cf. *solito* ‘usual’). Some past participles of still existing verbs did the same (cf. *viso* ‘seen’ → ‘sight’ → ‘eyes’ → ‘face’ in modern Italian) while being replaced in their past participle function by analogic forms (*visto/veduto* ‘seen’). In these cases, derivatives do not display a transparent relation with the past participles of the base verbs. Yet, they maintain formal relations, and these relations allow identifying a basic stem, which is by default identical to the basic stem of the past participle, but can possibly be distinct. This basic stem can be related to other basic stems of the base verb.

In particular, we study the relations between the basic stems of verbal derivatives and the basic stems of the related verb, along with a classification of these relations.

Keywords: computational morphology, paradigms, inflection, derivation, Levenshtein distance.

Ignacio L. Montoya: *A comparison of roots as units of analysis in Modern Hebrew and Spanish: exploring a remnant approach to defining roots*

The root as a morphological unit has been utilized in the description and explanation of many linguistic patterns. The concept of the root, though, is not approached consistently across diverse morphological theories, making cross-linguistic comparison problematic. This paper explores a methodology for comparing roots across languages using what is called a remnant approach: Roots are characterized as the elements that remain after accounting for inflectional and derivational morphology. Such a characterization is preferable to a characterization in which roots are defined *a priori* in terms of a set of properties because it offers a procedure for identifying roots that is independent of the properties that we want to examine. The methodology for identifying roots via a remnant approach is illustrated using data from Modern Hebrew and Spanish, and a comparison of roots identified in this way is offered. Several observations from this comparison are discussed, including, for instance, that roots identified in such a manner have a more characteristic distribution than form or meaning in both Hebrew and Spanish. This and other findings suggest potential parameters that can serve as the basis of a more extensive typology of roots. In addition, the preliminary results of this work offer insights that can inform morphological theory.

Keywords: roots, Hebrew, Spanish, cross-linguistic.

Gergana Popova: *Inflected and periphrastic features: issues of comparison and modelling*

This paper discusses the association of periphrastic constructions, that is syntactic constructions that express grammatical meaning usually realised by inflection, with features. The paper argues that the morphosyntactic non-compositionality in periphrasis and in ‘nested periphrases’ can be modeled by associating periphrastic constructions with feature values that are different from those appropriate for their component parts. The paper then explores the consequences such modeling would have on recent formalizations of periphrasis and points out that this approach would change how we view some ordering problems, as well as

recursion within periphrasis. Taking this alternative view, arguably, also reflects the dual nature of periphrases as expressions of a syntactic nature that occupy cells in a paradigm and gain their meaning from their position in a range of paradigmatic oppositions.

Keywords: features, morphology, periphrasis.

Marcel Schlechtweg, Holden Härtl: *The lexicalization of complex constructions: an analysis of adjective-noun combinations*

The current paper discusses the lexicalization of complex constructions composed of an adjective and a noun. It is argued that compounds/compound-like constructions are more prone to become lexicalized than phrases/phrase-like constructions. The relationship between lexicalization and the cognitive process of memorization represents a key point of our analysis. We report evidence from psycholinguistic studies contrasting compounds/compound-like constructions to phrases/phrase-like constructions either within a single language or across different languages. The results suggest that the former type of constructions show a memorization advantage in comparison to the latter type. These findings support the idea that the two construction types fundamentally differ with regard to their lexicalization.

Keywords: lexicalization, memorization, adjective-noun combinations.

Giuseppina Todaro, Fabio Montermini: *Spatial reduplication in Sicilian: lexicon or grammar?*

This paper proposes a morphological treatment for a particular kind of total reduplication (TR) in Sicilian which does not seem to express any of the meanings generally attributed to this kind of constructs (plurality, emphasis, intensity, iterativity, etc.). Reduplication is considered as a particular subtype of compounding. Compounding itself, however, can be seen either as the outcome of a general cognitive ability to connect two words semantically by simply juxtaposing them (that we call C_1), or as the output of a set of morphological patterns that are codified in the grammar of individual languages (C_2). The same may be claimed to apply to reduplication, under the form of a parallel distinction between a R_1 and a R_2 . Here, Sicilian TRs are considered as examples of R_2 , i.e. morphological derivational (lexeme-forming) strategies, creating either adverbs or adjectives.

The framework adopted is constructionist in the sense of Booij (2010). Sicilian TRs are therefore considered to instantiate generalizations made by speakers on sets of existing complex words with a systematic correlation between form and meaning. In particular, it is proposed to represent these TRs as constructional schemas ($[XX]_a$) which, in their turn, are encompassed into larger constructions involving a semantically and syntactically related entity (in most cases expressed by a NP): $[Y \text{ } \mathfrak{R} \text{ } [XX]_a]$. The three patterns observed can be seen as subcases of a larger construction whose general function is to mark a localisation. More precisely, Sicilian TRs are constructions whose function is to establish a spatial relation between a trajector (T) and a landmark (L). In two of the types, the reduplicated unit can be either a singular (*U picciriddro_T ioca casa_L-casa_L* ‘The child plays all over the house’) or a plural (*U cane_T camina muntagni_L-muntagni_L* ‘The dog walks in/through the mountains’) nominal form. The output of this kind of reduplication is an adverb. In a third type, the

reduplicated unit can only be a plural nominal form, and the output functions as an adjective (cfr. *A cammisa_L è pirtusa_T-pirtusa_T* ‘The shirt is full of holes’).

It is shown that the global semantic interpretation of the constructions in question is conditioned by the semantic features of each participant involved (the trajectory and the landmark) and by pragmatic factors. In particular, possible semantic nuances are: movement or static localization within the boundaries of the landmark, vague or incongruous localization, homogeneous distribution.

Keywords: Reduplication, Spatial Semantics, Lexeme Formation, Sicilian, Construction Morphology

Suspended affixation with derivational suffixes and lexical integrity

Faruk Akkuş
Yale University
faruk.akkus@yale.edu

1. Introduction

The paper aims to contribute to the discussion of suspended affixation and argue that it might have implications for the lexical integrity through the investigation of not-so-commonly-found suspended affixation instances formed with derivational suffixes. In opposition to what has been commonly assumed in the literature, I will argue that these instances of suspended affixation (SA) are ‘rather uncommon to find, quite many to ignore’ in analogy to Kaufmann’s (2014) approach to embedded imperatives as ‘too rare to expect, too frequent to ban’.¹ Examples (1) and (2) illustrate the phenomenon this paper investigates, where a derivational suffix only appears on the rightmost conjunct, but takes scope over both of the conjuncts:

- (1) *Loto-dan kazan-dığı parayı beş lira ve on dolar-lık*
lottery-ABL win-NMLZ-POSS money-ACC five lira and ten dollar-DER
*banknot-lar hal-i-nde boz-dur-du.*²
banknote-PL case-CM-LOC change-CAUS-PASS
‘S/he had her lottery winnings changed into banknotes of 5 liras and 10 dollars.’
- (2) *Sıcak tut-ar-ken dön-üp bak-tır-t-acak bere*
warm keep-AOR-CVB turn-and look- CAUS-CAUS-FUT cap
model-ler-i
model-PL-CM
‘cap models which while keeping you warm will make others turn and look.’

As in (1) and (2), such cases are found both in the nominal and verbal domain. In this paper, I will try to show that such instances are more than just idiosyncrasies, and cannot be accounted for via Wälchli’s (2005) notion of ‘natural coordination’, *pace* Kabak (2007). Accordingly, I argue that this rather productive process calls for an explanation. I also suggest that SA, which has been traditionally considered as peripheral to the discussion of lexical integrity, could in fact constitute a significant piece of evidence for the interaction between syntax and morphology.

The paper is organized as follows: section 2 describes suspended affixation, whereas section 3 looks at the lexical integrity hypothesis and the relation of the suspended affixation to it. Section 4 summarizes the previous approaches to suspended affixation in Turkish, and

¹ Unless otherwise stated, examples of the suspended affixation with derivational suffixes are attested and come from the internet searches.

² The following abbreviations are used: ABL: ablative, ACC: accusative, AL: alienable, AOR: aorist, CAUS: causative, CM: compound marker, COP: copula, CVB: converb, DAT: dative, DER: derivation, FUT: future, GEN: genitive, INST: instrumental, LOC: locative, MOD: modal, NMLZ: nominalizer, NOM: nominative, PART: participle, PASS: passive, PAST: past, PL: plural, POSS: possessive, PROG: progressive, REL: relativizer, SG: singular.

shows how derivational suffixes have been treated, which sets the path for section 5. In section 5, I introduce the instances of suspended affixes constructed with derivational suffixes in both nominal and verbal domain, and argue that previous approaches fail to capture these empirical facts. Section 6 suggests two possible analyses for the phenomenon, and section 7 concludes the paper.

2. Suspended affixation

Lewis (1967: 35) characterizes suspended affixation as when “one grammatical ending serves two or more parallel words”.³ Consider the sentences in (3) and (4), where the (a) sentences illustrate conjoined structures with suspended affixation. Only the final conjunct carries bound morphemes and has scope over the non-final conjuncts, as opposed to the (b) sentences, where both conjuncts are inflected for the same suffixes. It should be noted that discussion is limited to the Turkish coordination morphemes *ve* ‘and’ for verbal and nominal, and *-(y)Ip* for verbal coordination.

- (3) a. [*Zengin ve ünlü*]-*y-dü-m*.
 rich and famous-COP-PAST-1SG
 ‘I was rich and famous.’
 b. *Zengin-Ø-di-m ve ünlü-y-dü-m*
 rich-COP-PAST-1SG and famous-COP-PAST-1SG
 ‘I was rich and famous.’
- (4) a. [*Gid-er, gör-ür ve al-ır*]-*Ø-ız*. (Kabak 2007: 314, example 3)
 go-AOR see-AOR and buy-AOR-COP-1PL
 ‘We go (there), see (it), and buy (it).’
 b. *Gid-er-iz gör-ürüz ve al-ır-Ø-ız*.
 go-AOR-1PL see-AOR-1PL and buy-AOR-COP-1PL
 ‘We go (there), see (it), and buy (it).’

Suspended affixation is also found in the nominal domain, e.g. with the plural (5), and the case endings (6).

- (5) a. [*ev ve okul*]-*lar*
 house and school-PL
 ‘houses and schools’
 b. *ev-ler ve okul-lar*
 house-PL and school-PL
 ‘houses and schools’
- (6) a. [*kitap ve defter*]-*i*
 book and notebook-ACC
 ‘the book and the notebook’

³ Despite its important bearings on the syntax-morphology interface, suspended affixation has been rarely addressed in theoretical literature, and almost all the studies have dealt with suspended affixation formed with inflectional suffixes (e.g. Kornfilt 1996; Kahnemuyipour & Kornfilt 2011; Kenesei 2007; Pounder 2006; Kabak 2007; Johannessen 1998; Erschler 2012). Moreover, as both Stephen Anderson and Greg Stump pointed out to me on separate occasions, such examples could also be considered from the perspective of their implications for the definition of derivation as well.

- b. *kitab-ı ve defter-i*
 book-ACC and notebook-ACC
 ‘the book and the notebook’

As Erschler (2012) points out, the “mirror image” structure is logically possible, but much less common cross-linguistically.

- (7) *s-jə-pçaçe-re* *if’ale-re* *zezaox* (Adyghe, Northwestern Caucasian)
 1SG-AL-girl-and boy-and fight.each.other
 ‘My son and daughter are fighting.’ (Erschler 2012: 154, example 2b)

Note also that suspended affixation can apply not only to words, as Lewis (1967) mentions, but also to units bigger than words, such as phrases, as shown in (8). Here the locative case takes scope over both preceding NPs.

- (8) [*Can’-in* *divan-ı* *ve* *Orhan’-in* *yatağ-in*]-*da* *uyu-du-m*
 Can-GEN couch-3SG and Orhan-GEN bed-3SG-LOC sleep-PAST-1SG
 ‘I slept on Can’s couch and Orhan’s bed.’

Examples like (8) raise problems for lexical integrity, as it is usually understood (Bresnan and Mchombo 1995; Bresnan 2001, see below), although a morphological solution may be possible, as will be dwelt upon later in the paper.

3. Suspended affixation and lexical integrity

This section looks at the several proposals with respect to the Lexical Integrity Hypothesis, and then explores the potential implications of suspended affixation for it.

3.1 Lexical Integrity Hypothesis⁴

Starting from the generative morphology, the Lexical Integrity Hypothesis (LIH) was a widely accepted part of the landscape for morphologists. The LIH appeared in a number of different forms:

- (i) *The Word Structure Autonomy Condition* (Selkirk 1982: 70)
 No deletion or movement transformation may involve categories of both W-structure and S-structure.
- (ii) *The Atomicity Thesis* (Di Sciullo and Williams 1987: 49, cited in Lieber and Scalise 2006)
 Words are “atomic” at the level of phrasal syntax and phrasal semantics. The words have “features”, or properties, but these features have no structure, and the relation of these features to the internal composition of the word cannot be relevant in syntax – this is the thesis of the atomicity of words, or the lexical integrity hypothesis, or the strong lexicalist hypothesis (as in Lapointe 1980), or a version of the lexicalist hypothesis of Chomsky (1970), Williams (1978) and numerous others.

A distinction can be drawn between a Weak Lexicalist Hypothesis and a Strong Lexicalist Hypothesis, the former merely stating that transformations could not look into word structure

⁴ This section draws freely from Lieber and Scalise (2006).

(i.e., derivation and compounding), the latter adding inflection to the domain of the LIH (Spencer 1991).

The notion that words are unanalyzable units persists in Bresnan and Mchombo's (1995: 181) formulation of the LIH.

- (iii) A fundamental generalization that morphologists have traditionally maintained is the *lexical integrity principle*, which states that words are built out of different structural elements and by different principles of composition than syntactic phrases. Specifically, the morphological constituents of words are lexical and sublexical categories – stems and affixes – while the syntactic constituents of phrases have words as the minimal, unanalyzable units; and syntactic ordering principles do not apply to morphemic structures.

It can be seen that despite slight differences in the formulations and the focus of the rationale behind the separation of morphology and syntax, all of these statements of LIH still have in common that they assume a firewall between morphology and syntax, in whatever form syntax takes.

The more recent statements of the LIH include Booij (2005) and Spencer (2005). Following in essence the formulation of LIH of Anderson: “the syntax neither manipulates nor has access to the internal structure of words” (Anderson 1992: 84, cited in Booij 2005: 1) proposes to split the LIH in two parts:

- (iv) a) Syntax cannot manipulate the internal structure of words.
b) Syntax cannot enter into the internal structure of words.

Lieber and Scalise (2006) argue, on the basis of a number of various data which – according to their view – present strong challenges to the LIH, that 0 may be correct, while 0 cannot be. Lieber and Scalise's data include phrasal compounds in English, as shown in (9),⁵ and the Italian *trasporto latte*-type constructions. These data pass the test of insertion/modification, which is traditionally the main test of cohesiveness, although they fail the other tests, or scope in Spanish prefixation, as in (10):

- (9) a *slept all day* look (Lieber and Scalise 2006: 4, example 8)
a *pipe and slipper* husband
over *the fence* gossip

- (10) a. *el* [*ex*-[*futbolista del Barça*]_{NP}] (Lieber and Scalise 2006: 11, example 28)
the *ex*- footballer of Barça
'the former Barça footballer'
b. *comisión* [*pro*-[*legalización de las drogas*]_{NP}]
committee *pro*- legalization of the drugs
'*pro*- drug-legalization committee'

In (10), although phonologically prefixation takes place on the N head of an NP, semantically the prefix affects the whole NP.

⁵ See e.g. Bağrıaçık and Ralli (2013), Göksel (2015) for phrasal compounds in Turkish.

To the examples discussed in Lieber and Scalise (2006), we can add the phrasal derivation observed in Turkish, along with various languages.

- (11) a. [*san-a tap-ıyor-um*]-*cu* *tutum*
 you-DAT adore-PROG-1SG-DER attitude
 Lit: ‘I adore you-ish attitude.’
 b. [[*köpeğ-e evet, kedi-ye hayır*]-*ci*]-*lik*⁶
 dog-DAT yes cat-DAT no-DER-DER
 Lit: ‘yes to dog, no to cat-ism’

Ackema and Neeleman (2004: 11) also mention a case in Quechua where a particular affix attaches to phrases to nominalize them.

Only a look at the cases at hand gives the following picture according to the type of inter-component interaction that they imply:

- (v) Morphology has access to Syntax
 a) syntactic phrases within words (phrasal compounds)
 b) insertion/modification into *trasporto latte* constructions (Italian data)
 c) Turkish and Quechua nominalizations on phrases

Syntax has access to Morphology

Quechua nominalizations: position of verb is dependent on whether VP is nominalized or not.⁷

The examples thus far point to an interaction between syntax and morphology, hence challenge the LIH. Next let us look at a common approach to suspended affixation with respect to the Lexical Integrity.

3.2 The relevance of suspended affixation to the LIH

After noting that the phenomenon *Gruppeninflection* or ‘suspended affixation’ is found in many languages, Spencer (2005) points out that “this is only possible when single words are coordinated, not phrases, and only when the coordinated elements form a ‘natural coordination’ (in an intuitive sense)” along the lines of Wälchli (2005).

Example (8), repeated here as (12), shows that Spencer’s point about only words being coordinated does not cover all the empirical facts:

- (12) [*Can’-in divan-ı ve Orhan’-in yatağ-ın*]-*da* *uyu-du-m*
 Can-GEN couch-3SG and Orhan-GEN bed-3SG-LOC sleep-PAST-1SG
 ‘I slept on Can’s couch and Orhan’s bed.’

Another counterexample can be given from the verbal domain, where two VPs are coordinated:

⁶ Bağrıaçık and Ralli (2013) provide the example in and argue that these three suffixes are word-formation X^{\max} affixes that choose phrase levels, following Ackema and Neeleman (2004).

[_{DP} *karşı* *dağ-ın* *ardındaki* *kasaba*]-*lı*
 opposite mountain.GEN beyond.LOC.RTV town-DER
 ‘from the town beyond the opposite mountain’

⁷ See Lieber and Scalise (2006) for more examples and a detailed discussion of the interaction of different modules of grammar.

- (13) *Leyla [yemek yi-yor ve kitap oku-yor]-du.*
 Leyla food eat-PROG and book read-PROG-PAST.3
 ‘Leyla was having a meal and reading a book.’

I will discuss ‘natural coordination’ in the context of suspended affixation cases in Turkish later in the paper in section 5.1, but it should suffice to say that Lieber and Scalise (2006) state that “[cases of conjunction in English derivation and compounding] constitute a clear violation of the LIH, as do cases of so-called *Gruppeninflection* or ‘suspended affixation’ (Spencer 2005: 83) which seem to constitute a similar phenomenon, albeit concerned with inflection rather than word formation”.

Lieber and Scalise’s remarks represent the general intuition that the phenomenon could in principle be a strong challenge to the LIH, but with the way it is, it can only be peripheral to the discussion of syntax-morphology interaction.

4. Previous treatment of suspended affixation in Turkish⁸

4.1 Kabak (2007)

This article focuses on various constraints on suspended affixation, where Kabak proposes an account of suspended affixation based on the notion of morphological wordhood in Turkish. Investigating the type of material that can be left in nonfinal conjuncts in both verbal and nonverbal coordinate constructions, Kabak argues that suspension of affixes is legitimate if the bare conjunct constitutes a morphological word in Turkish. A morphological word is defined to be a form that is able to occur in isolation.

Kabak points out that although there seems to be a strict constraint on the suspension of derivational morphemes, derivational morphemes can be attached to certain tightly coordinated nouns, which on the surface may look like instances of affix suspension. Consider (14):

- (14) a. *ana (ve) baba-lık* (Kabak 2007: 336)
 mother (and) father-DER
 ‘parenthood’
 b. *ay-yıldız-lı bayrak*
 moon-star-DER flag
 ‘moon-star flag’ (refers to the Turkish flag)
 c. *sarı-kırmızı-lı takım*
 yellow-red-DER team
 ‘team in red and yellow’ (refers to the Galatasaray soccer team)

However, Kabak regards such cases as instances of co-compounds or natural coordination, which express stereotypically conjoined entities in the sense of Wälchli (2005), corroborated by the two well-known instances of antonomasia in (14) and (14). Arguably, such constructions involve coordination of items that are expected to co-occur, and behave as a single conceptual unit with the derivational morpheme attached to it. For these reasons, Kabak argues that they should not be considered as representative of affix suspension.

Kabak concludes that the Turkish morphological system exhibits a split behavior between derivational morphemes and inflectional ones: unlike inflectional morphemes, derivational morphemes cannot have scope over conjuncts and hence they cannot be suspended (see

⁸ Naturally, the literature on suspended affixation in Turkish is more extensive (e.g. Yu and Good 2000; Hankamer 2012), but this paper focuses on those that deal with derivational suffixes.

Kornfilt 2012 for the same argument). This could be attributed to the fact that derivational affixes show closer lexical affinity to the stems that they are attached to. This, he argues, follows from the argument that inflection and derivation correspond to distinct systems in linguistic competence (e.g. Anderson 1992).

4.2 Bozşahin (2007)

Bozşahin gives the following example of coordination in (15), cited in Kornfilt (2012), which is ambiguous between two potential readings:

- (15) *tuz ve limon-luk*
 salt and lemon-container
 a. ‘salt and lemon squeezer’ (Non-SA-reading)
 b. ‘salt shaker and lemon squeezer’ (Apparent SA-reading)

Bozşahin (2007) claims that reading b. shows that SA does not distinguish between derivational and inflectional morphemes, and thus it can apply in the lexicon, too (under the assumption that the suffix *-LIK*, glossed as ‘instrument’ in Bozşahin and as ‘container’ here, is a derivational morpheme merged with the stem in the lexicon).

4.3 Kornfilt (2012)

Although Kornfilt argues that Kabak’s definition of “morphological word” needs revisiting and may not account for all cases in her discussion of suspended affixation, she agrees with Kabak on stipulating that the crucial distinction is syntax versus the lexicon: those affixes that can be “suspended” are syntactic functional heads in phrasal or clausal architecture (i.e. they are merged syntactically). On the other hand, those affixes that are part of the lexical word formation cannot distribute. In other words, SA is a syntactic process that of course applies to syntactic constituents. Thus, only those affixes can be “suspended” that are syntactic heads, i.e. heads of functional projections.

Kornfilt takes Bozşahin’s *tuz ve limonluk* as a case study and discusses it further. She points out that the order of the conjuncts (with the “suspended” ‘instrument’-suffix on *tuz* ‘salt’) eliminates the distributed reading:

- (16) *limon ve tuz-luk*
 lemon and salt-container
 a. ‘lemon and salt shaker’ (Non-SA-reading)
 b. *‘lemon squeezer and salt shaker’ (The (apparent) SA-reading is not available)

She claims that only the “suspended affixation” reading for (16) with the “container” suffix interpreted as distributed over the two conjuncts, is apparent and what actually takes place is that pragmatically *salt* is used in reading B for *salt shaker*.

However, I suggest an alternative explanation, which is in fact in line with what one of the anonymous reviewers’ suggestion of Kornfilt (2012): *tuz* ‘salt’, being uncountable, requires a classifier: the object containing it. Salt would therefore be able to stand for ‘salt shaker’, while lemon, which is not uncountable, would therefore not need a classifier to be interpreted as a definite amount, and would therefore also not be able to stand for ‘lemon squeezer’ on its own. This suggestion predicts that a noun like *biberlik* ‘pepper shaker’, when combined with *tuzluk* ‘salt container’ should have a distributive reading in either order, which turns out to be correct as the google searches confirm:

- (17) *Coghlan's tuz ve biber-lik en iyi fiyat-la Hepsiburada'-da.*
 Coghlan's salt and pepper-DER most good price-INST hepsiburada-LOC
 'Coghlans's salt container and pepper shaker is at Hepsiburada(.com) for the best price!'
- (18) a. *2 ons cam biber ve tuz-luk kap, ücretsiz kargo.*
 2 ounce glass pepper and salt-DER container free shipping
 '2 ounce glass pepper shaker and salt container, free shipping.'
- b. *ön-ümüz-de dur-an biber ve tuz-luk ... garson*
 before-1PL.POSS-LOC lie-REL pepper and salt-DER server
tarafından başka müşteri-ye ver-il-di.
 by other customer-DAT give-PASS-PAST
 'The pepper shaker and salt container in front of us were given to another customer by the server.'

Given that pepper is uncountable similarly to salt, this shows that a purely pragmatic explanation, as that of Kornfilt, since a conjunct, as in (18), targeted by a modifier that modifies the whole phrase, by itself does not express the meaning it gets with the derivational suffix. Therefore, Kornfilt's explanation cannot suffice to account for (17) and (18) even when considering the relative difference in frequency effects for *limon ve tuzluk* and *biber ve tuzluk*. The next section introduces the instances of derivational suffixes that are used in suspended affixation, mainly from Turkish, as well as some examples from other languages.

5. Instances of suspended affixation with derivational suffixes

Although suspended affixation has been considered as peripheral to morphology-syntax interaction, as (Kornfilt and Whitman 2011) argue, it touches on the issue of syntax-lexicon dichotomy. The syntax-lexicon debate roots in the treatment of Japanese causatives. In the early days of the 60s and 70s, causative verbs are formed syntactically (via transformation). The 80s saw the advent of lexicalism, and whether Japanese causatives are formed in the syntax or the lexicon has been controversial.

The example in (19), where suspended affixation happens with the causative suffix, is nowadays widely recognized as a decisive argument against lexical approaches:⁹

- (19) *Hanako-ga Masao-ni [[uti-o soozisuru]-ka [heya-dai-o*
 H.-NOM M.-DAT [[house-ACC clean]-or [room-rent-ACC
haraw]]-aseru koto ni sita
 pay]]-CAUS decided
 'Hanako decided to make Masao clean the house or pay room rent.'
 (Kuroda 2003: 455)¹⁰

In the context of Turkish, as discussed in the previous section, Kabak (2007) and Kornfilt (2012) explicitly argue that the nature of a suffix determines its ability in scoping over conjuncts and that instances with derivational suffixes are not true cases of affix suspension. However, I argue that these explanations fail to capture the wide range of well-formed instances that cannot be reduced to the accounts of co-compounds or pragmatics.

⁹ For a recent semantic analysis of Japanese suspended affixation, see Fukushima (2015).

¹⁰ Kuroda was one of the earliest proponents of syntactic treatment.

5.1 The nominal domain

The following are some of the examples obtained from internet searches from the nominal domain in Turkish (only the relevant parts glossed):

- (20) a. *Pijama, genellikle üst ve alt-lık gibi iki unsur-dan*
 pajama usually top and bottom-DER like two part-ABL
mürekkep ... bir giysidir.
 comprised of a dress
 ‘The pajama is a dress comprised of two pieces, namely top and bottom.’
- b. *Yanınıza gece için kalın alt ve üst-lük ... al-in.*
 side-DAT night for thick bottom and top-DER take-2PL
 Lit: ‘Bring with you a warm top and bottom for the night.’
- (21) *Balyoz konu-su-nda yaz-dığ-ım on-lar ve de*
 sledgehammer topic-POSS-LOC write-NMLZ-1SG.POSS ten-PL and also
on-lar-ca yazı-ya ... (newspaper Radikal)
 ten-PL-DER article-DAT
 Lit: ‘Despite the tens and tens of articles I have written on the Sledgehammer operation...’
- (22) *Loto-dan kazan-dığı parayı beş lira ve on dolar-lık*
 lottery-ABL win-NMLZ-POSS money-ACC five lira and ten dollar-DER
banknot-lar hal-i-nde boz-dur-du.
 banknote-PL case-CM-LOC change-CAUS-PASS
 ‘S/he had her lottery winnings changed into banknotes of 5 liras and 10 dollars.’
- (23) *kitabın giriş, bir ve yedi-nci bölüm-ler-i-ni...*
 book-GEN introduction one and seven-DER chapter-PL-POSS-ACC
 ‘the introduction, first, and seventh chapters of the book...’
- (24) a. *Bütün eğitim çalışmaları boyunca dost ve arkadaş-ça bir*
 all training sessions during buddy and friend-DER an
hava ol-malı-dır
 atmosphere be-must-MOD
 ‘There must be a friendly and intimate environment during the whole training session.’
- b. *... buna uy gundavranmak için arkadaş ve dost-ça*
 this-DAT appropriate behave-NMLZ for friend and buddy-DER
gel-di-k.
 come-1PL-PAST
 ‘Accordingly, we came in a friendly and intimate manner.’

The other derivational suffixes include *-ci*, *-leyin*, *-zede*, *-inci*, etc.

- (25) *İstanbul Valiliği tarafından organize ed-il-en Deprem ve*
 Istanbul governorship by organize do-PASS-REL earthquake and
Afet-zede Anma Yürüyüşü...
 disaster-DER commemoration march
 ‘Earthquake and Disaster-victims Commemoration March organized by the Istanbul Governorship...’

One could argue that these instances may be accountable by Kabak's (2007) explanation, following Wälchli's (2005) notion of 'natural coordination'. However, I argue that as idiosyncratic as they might be, these instances differ from the natural coordination examples Kabak provides for Turkish, for several reasons. First, note that conjuncts in Kabak (2007) have a fixed word order.

- (26) a. *ana (ve) baba-lık* (Kabak 2007: 336)
 mother (and) father-DER
 'parenthood'
 b. **baba (ve) ana-lık*
 father (and) mother-DER
- (27) a. *ay-yıldız-li bayrak*
 moon-star-DER flag
 'moon-star flag'
 b. **yıldız-ay-li bayrak*
 star-moon-DER flag

On the other hand, as examples (20) through (25) illustrate, these coordinations may have a free conjunct order, which poses a problem for a theory that attributes the possibility to their lexicalized nature.

Second, note that as the example (22) shows, the derivational suffix *-LIK* distributes over phrases, not just words, similar to other SA cases formed with inflectional morphemes (cf. (12) and (13)). This stands as a strong challenge to a purely lexical account and has a bearing on the morphology-syntax interaction similar to the Quechua nominalization cases.

Third, strictly speaking, some of the conjuncts are not items that are easily expected to occur together or are necessarily supposed to co-occur always. Accordingly, another item can replace one of the conjuncts, that is, there is no strict rule that allows only the present conjuncts to co-occur. For instance, in the case of (22) one can have a different conjunct, as illustrated in (28). This again does not go well with the argument that these conjuncts are tightly connected.

- (28) ... [*yirmi şekel ve on dolar*]-*lık banknotlar*
 twenty shekel and ten dollar-DER banknotes
 'banknotes of 20 shekels and 10 dollars...'

Fourth, natural coordinations are not expected to allow another conjunct since they are assumed to express stereotypically conjoined entities and to behave as a single conceptual unit with the derivational morpheme attached to it. However, in cases at hand, it is possible to have a third conjunct. Consider (29), where the addition of another conjunct to (23) is possible.

- (29) *kitabın giriş, bir, yedi ve yirmi bir-inci bölüm-ler-i-ni...*
 book-GEN introduction one seven and twenty one-DER chapters-POSS-ACC
 'the introduction, first, seventh and twenty first chapters of the book...'

Finally, maybe as a not very strong point, it could also be said that the possibility of this wide range of derivational suffixes allowed in this operation is not exactly in favor of a lexicon-oriented account.

Therefore, the instances at hand are at best somewhere between the natural coordination cases and fully productive cases of suspended affixation in Turkish.

Moreover, these cases are not compatible with Kornfilt's (2012) understanding of pragmatics either, since for her the possibility of saying *tuz ve limonluk* 'salt shaker and lemon squeezer' is because of the pragmatic use of *salt* for *salt shaker*. Let's take examples in (24). In the example (a) the whole affixed phrase *dost ve arkadaşça* [buddy and friend-Der] functions as an adjective modifying the noun *hava* 'air', therefore, the pragmatic use of *dost* 'buddy' in an adjective function to modify the noun fails here. The same point extends to the (b) example too. The point is essentially that a conjunct by itself does not express the meaning it gets with the derivational suffix.

Brazilian Portuguese is another language where suspended affixation with the derivational suffix *-mente* can be observed, as shown in (30). Note that Brazilian Portuguese also allows free order of conjuncts, although the slight degradation in (30) could be due to phonological reasons.¹¹

- (30) a. *feliz a vagarosa-mente* (Manu Quadros, pers. comm.)
 happy and slow-ly
 'happy and slowly' (as in *He finished his homework happily and slowly*)
 b. *?vagarosa a feliz-mente*
 slow and happy-ly

In addition to the examples in the nominal domain, it is possible to find natural data online in the verbal domain too.

5.2 The verbal domain

Note that example (19) from Japanese, repeated here as (31), is an illustration of suspended affixation with the causative suffix, and has been taken as a strong argument for the syntactic analysis (e.g. Nishiyama 2012).

- (31) *Hanako-ga Masao-ni [[uti-o soozisuru]-ka [heya-dai-o*
 H.-NOM M.-DAT [[house-ACC clean]-or [room-rent-ACC
 haraw]]-aseru koto ni sita
 pay]]-CAUS decided
 'Hanako decided to make Masao clean the house or pay room rent.'
 (Kuroda 2003: 455)

Turkish also has instances of SA constructed with both the causative and passive suffixes, two types of suffix traditionally taken to be derivational in the Turkish literature (Kornfilt 1997; Göksel and Kerslake 2005).

- (32) Causative
 Sıcak tut-ar-ken dön-üp bak-tır-t-acak bere model-ler-i
 warm keep-AOR-CVB turn-and look- CAUS-CAUS-FUT cap model-PL-CM
 'cap models which while keeping you warm will make others turn and look.'

¹¹ In fact, Kayne (2005) suggests that certain derivational suffixes in English such as *-less*, *-ful*, *-ish*, *-y*; also *-th* as in *two hundred and fiftieth*, suggest a strongly syntactic approach. To Kayne's list, one can add *-wise*, as in *in format and content-wise*. The latter is like the mirror image of the *pro-* and *anti-revolution* cases that Lieber and Scalise (2006) deal with.

(33) Causative

Ömür *gerçekten* *yetenekli.* *Hazır cevaplı espriler-i* *ile salonu*
 Ömür really talented witted jokes-POSS with hall

gül-üp kır-dır-dı.
 laugh-and break-CAUS-PAST

‘Ömür is really talented. He cracked up the whole hall with laughter.’

(34) Passive

... *Nice aile-ler, ocak-lar yak-ıp yok ed-il-di.*
 score family-PL home-PL burn-and destroy-PASS-PAST
 ‘Scores of families, homes were burned and destroyed.’

(35) Passive

Depo polis tarafından bas-ıp yık-il-di.
 warehouse police by raid-and destroy-PASS-PAST
 ‘The warehouse was raided and destroyed by the police.’

These examples also speak against a purely lexicon-based account and call for some sort of syntactic involvement.¹²

Bhili of Khandesi provides another instance of suspended affixation in the verbal domain that is formed via a derivational suffix (Grierson 1907: 152). Consider (36), where the conjunctive participle suffix *-san* scopes over two verbs:

- (36) a. *khai-san*
 eat-PART
 ‘having eaten’
 b. *khai-pii-san*
 eat-drink-PART
 ‘having eaten and drunk’

In the next section, I will entertain two possible accounts for this phenomenon.

6. Two possible accounts

One view of morphology might argue that these cases are far too idiosyncratic to find a place in the syntactic system. However, as the arguments in section 5.1 show, I believe, the level of the idiosyncrasy one finds with these instances is not more severe than the idiosyncrasy. In fact, one finds with all sorts of elements that no one would deny are part of the syntactic system, in the sense that they are generated in unique phrase-structural positions and subject to syntactic constraints (Wood 2015). In other words, certain degree of conventionalized use or pragmatics does not rule out the structure, e.g. *attend church* or *going to prom*-type examples where due to conventional use, the definite article is dropped, but still canonically a structure is still assumed.

Rather than taking the path where it is syntax-all-the-way, I assume a system, such as Lieber and Scalise’s (2006) *The Limited Access Principle* or the analysis presented in Ackema

¹² Certain instances of suspended affixation in nominal derivation is observed in Korean as well (e.g. Yoon 2008):

[20-il-ina 21-il]-kkey manna-ca
 20-day-or 21-day-around meet-prop
 ‘Let’s meet on the 20th or the 21st of next month.’

and Neeleman (2004). The two approaches fall within the realm of Minimalist Framework and might give two options. The former allows the interaction between morphology and syntax in a limited way.

The Limited Access Principle (Lieber and Scalise 2006: 21)

Morphological Merge can select on a language specific basis to merge with a phrasal/sentential unit. There is no Syntactic Merge below the word level.

In this system, limited intermodular access may be allowed by virtue of allowing configurations like:

- (37) a. $[[XP] Y]_Y / [Y [XP]]_Y$
 b. $[[XP] [Y]]_Y / [[Y] [XP]]_Y$
 c. $[[XP] Y]_X / [Y [XP]]_X$
 d. $[[XP] [Y]]_X / [[Y] [XP]]_X$

In this system, the morphological merger, together with the Limited Access Principle, would yield the sorts of structures highlighted in (v). It seems possible to place Turkish suspended affixation cases in (22) into the structure in (37), along with Quechua nominalizations.

The other option would be to follow Ackema and Neeleman (2004), who propose that the grammar is constituted by three modules (syntax, semantics and phonology), but each of these modules contains “a submodule that generates phrasal representations and a submodule that generates word-level representations” (2004: 3). The main idea is that morphology is a “set of submodels within these bigger modules” (2004: 6).

Morphology and syntax can thus share common principles, for example, a vocabulary of features and a process of merger, but they can at the same time be based on different principles. Nevertheless, Ackema and Neeleman argue that there can be a number of different types of intramodular interactions between morphology and syntax: first, words and sentences consist of a certain amount of shared vocabulary (certain features, the notion of Merge, etc.); second, word syntax and phrasal syntax are in competition (2004: 9); and finally, the process of insertion works both ways between morphology and syntax: words can of course be inserted into syntactic structures, but it is also possible for phrases to be inserted into words (2004: 10).¹³

The last point in Ackema and Neeleman could account for the Turkish instances, in that affixes choose phrase level.

7. Conclusion

This paper has argued that Turkish (and potentially several other languages) exhibits certain instances of suspended affixation formed with derivational suffixes both in the nominal and verbal domains. I have argued that these instances cannot be reduced to Wälchli's (2005) natural coordination since they differ from the examples Kabak (2007) provides.

This observation undermines a purely pragmatic account, and points to the relevance of derivational suffixes to the interaction of morphology and syntax. I also noted that one could employ the accounts of Lieber and Scalise (2006) or Ackema and Neeleman (2004) in order to give an explanation to such instances.

¹³ A third option not discussed here would be a DM-style analysis. Also Erschler's (2012) phonological deletion account seems applicable.

Acknowledgments

I would like to thank Aslı Göksel, Stephen Anderson, Edit Doron, Anne-Marie Di Sciullo, Mark Aronoff, Aysun Kunduracı for their valuable questions, comments and the audience at MMM10. I also thank Greg Stump for his suggestions. Usual disclaimers apply.

References

- Ackema, P. & A. Neeleman (2004) *Beyond Morphology: Interface Conditions on Word Formation*. Oxford: Oxford University Press.
- Anderson, S. (1992) *A-Morphous Morphology*. Cambridge: Cambridge University Press.
- Bağrıaçık, M. & A. Ralli (2013) N+ N-sI (n) Concatenations in Turkish and the Morphology-Syntax Interface. In: U. Ozge (Ed.), *Proceedings of the 8th Workshop on Altaic Formal Linguistics (WAFL8)* (MIT Working papers in Linguistics 67). MIT Press, 13-24.
- Booij, G. (2005) Construction-Dependent Morphology. *Lingue E Linguaggio* 4 (2): 163-178.
- Bozşahin, C. (2007) *Lexical Integrity and Type Dependence of Language*. Manuscript. METU, Ankara.
- Bresnan, J. (2001) *Lexical-Functional Syntax*. Oxford: Wiley-Blackwell.
- Bresnan, J. & S. A. Mchombo (1995) The Lexical Integrity Principle: Evidence from Bantu. *Natural Language and Linguistic Theory* 13 (2): 181-254.
- Chomsky, N. (1970) Remarks on Nominalization. In: R. Jacobs & P. Rosenbaum (Eds.), *Readings in English Transformational Grammar*. Waltham (MA): Blaisdell, 184-221.
- Di Sciullo, A.-M. & E. Williams (1987) *On the Definition of Word*. Cambridge: Massachusetts: MIT Press.
- Erschler, D. (2012) Suspended affixation in Ossetic and the structure of the syntax-morphology interface. *Acta Linguistica Hungarica* 59: 153-175.
- Fukushima, K. (2015) Compositionality, lexical integrity, and agglutinative morphology. *Language Sciences* 51: 67-85.
- Göksel, A. (2015). Phrasal Compounds in Turkish: Distinguishing Citations from Quotations. *STUF Language Typology and Universals* 68 (3)-*Special Issue: Phrasal compounds from a typological and theoretical perspective* (Issue Editors: C. Trips and J. Kornfilt): 359-394.
- Göksel, A. & C. Kerslake (2005) *Turkish: A Comprehensive Grammar*. London: Routledge.
- Grierson, G. A. (1907) *Linguistic Survey of India*. Vol. IX, Part III. India: Office of the Superintendent of Government Printing.
- Hankamer, J. (2012) Ad-Phrasal Affixes and Suspended Affixation. Talk presented at the *Workshop on Suspended Affixation* (Cornell University, October 26-27, 2012).
- Johannessen, J. B. (1998) *Coordination*. Oxford: Oxford University Press.
- Kabak, B. (2007) Turkish Suspended Affixation. *Linguistics* 45 (2): 311-347.
- Kahnemuyipour, A. & J. Kornfilt (2011) The syntax and prosody of Turkish “pre-stressing” suffixes. In: R. Folli, and C. Ulbrich (Eds.), *Interfaces in linguistics: New research perspectives*. Oxford: Oxford University Press, 205-221.
- Kaufmann, M. (2014) Embedded imperatives across languages: Too rare to expect, too frequent to ban. Talk presented at the *Colloquium Stony Brook* (April 4, 2014; updated on April 11, 2014).
- Kayne, R. (2005) Some notes on comparative syntax, with special reference to English and French. In: G. Cinque & R. S. Kayne (Eds.), *The Oxford handbook of comparative syntax*. New York: Oxford University Press, 3-69.
- Kenesei, I. (2007) Semiwords and affixoids: The territory between word and affix. *Acta Linguistica Hungarica* 54: 263-293
- Kornfilt, J. (1996) On some copular clitics in Turkish. *ZAS Papers in Linguistics* 6: 96-114.
- Kornfilt, J. (1997) *Turkish*. New York: Routledge.
- Kornfilt, J. (2012) Revisiting ‘Suspended Affixation’ and Other Coordinate Mysteries. In: L. Brugé, A. Cardinaletti, G. Giusti, N. Munaro, and C. Poletto (Eds.), *Functional Heads: The Cartography of Syntactic Structures 7*. Oxford: Oxford University Press, 181-196.
- Kornfilt, J. & J. Whitman (2011) Introduction: Nominalizations in Syntactic Theory. *Lingua* 121 (7): 1160-1163.
- Kuroda, S.-Y. (2003) Complex Predicates and Predicate Raising. *Lingua* 113 (4): 447-480.
- Lapointe, S. (1980) *A Theory of Grammatical Agreement*. PhD Thesis. UMass.
- Lewis, G. L. (1967) *Turkish Grammar*. Oxford: Oxford University Press.
- Lieber, R. & S. Scalise (2006) The Lexical Integrity Hypothesis in a New Theoretical Universe. *Lingue E Linguaggio* 1: 7-32.
- Nishiyama, K. (2012) Japanese Verbal Morphology in Coordination. Talk presented at the *Workshop on Suspended Affixation* (Cornell University, October 26-27, 2012).

- Pounder, A. (2006) Broken forms in morphology. In: C. Gurski, and M. Radisic (Eds.), *Proceedings of the 2006 Annual Conference of the Canadian Linguistics Association*. Available at: <http://westernlinguistics.ca/Publications/CLA2006/Pounder.pdf>
- Selkirk, E. (1982) *The Syntax of Words*. Cambridge/Massachusetts: MIT Press.
- Spencer, A. (1991) *Morphological Theory: An Introduction to Word Structure in Generative Grammar*. Oxford: Wiley-Blackwell.
- Spencer, A. (2005) Word-Formation and Syntax. In: P. Štekauer & R. Lieber (Eds.), *Handbook of Word-Formation*. Amsterdam: Springer, 73-97.
- Wälchli, B. (2005) *Co-Compounds and Natural Coordination*. New York: Oxford University Press.
- Williams, E. (1978) Across-the-Board Rule Application. *Linguistic Inquiry* 9 (1): 31-43.
- Wood, J. (2015) *Icelandic Morphosyntax and Argument Structure* (Studies in Natural Language and Linguistic Theory Vol. 90). Switzerland: Springer.
- Yoon, J. HS. (2008) The Lexicalist Hypothesis and Korean Morphosyntax. Talk presented at the 16th *International Circle of Korean Linguistics Conference* (Cornell University, June 26-28, 2008).
- Yu, A. C. & J. C. Good (2000) Morphosyntax of Two Turkish Subject Pronominal Paradigms. In: M. Hirotani et al. (Eds.), *Proceedings of the 30th North Eastern Linguistics Society*. Amherst, MA: GLSA, 759-774.

The Semitic templates from the perspective of reciprocal predicates

Elitzur A. Bar-Asher Siegal
Hebrew University of Jerusalem
ebas@mail.huji.ac.il

1. Introduction

1.1 Background

Among the questions that theories about Semitic verbal morphology aim to answer is the following: what is the relationship between verbs with different but related meanings that share the same phonological root but appear in different templates (cf., *inter alia*, Ornan 1971; Berman 1975; Aronoff 1994; Doron 2003; Arad 2005)? The forms in (1) illustrate the phenomenon of the morphology of templates with the phonological root \sqrt{gdl} , which appears in six different templates:

- (1) \sqrt{gdl} verbal templates:¹
- | | | |
|--------------------------------|-----------------|--------------------|
| a. CaCaC (basic template) | <i>gadal</i> | ‘grow’ |
| b. CiCeC (intensive template) | <i>gidel</i> | ‘raise, cultivate’ |
| b'. CuCaC | <i>gudal</i> | ‘was raised’ |
| c. hiCCiC (causative template) | <i>higdil</i> | ‘enlarge’ |
| c'. huCCaC | <i>hugdal</i> | ‘was enlarged’ |
| d. hitCaCeC (T-template) | <i>hitgadel</i> | ‘become bigger’ |

The goal of this paper is to shed some light on this broader question by considering the so-called “reciprocal verbs”, i.e., verbs with a certain morphological structure that allegedly encodes reciprocal relations. Such verbs often appear in the T-templates (cf. (1d)) across the Semitic languages, thus I would like to examine their relation to other verbs with the same root.

Previous studies on the morphology of the templates pay only little attention to the verbal expressions of reciprocity. Thus, the limited goals of this paper are to see what theories concerning the morphology of the templates in Modern Hebrew would have to account for with respect to these verbs.

1.1 Data and theoretical questions concerning the *reciprocal*-predicate

In most languages with verbal reciprocals, triplets of the type illustrated in (2a-c) are available (Nedjalkov 2007; Behrens 2007). Although this paper focuses on Hebrew, examples from a host of languages could be provided as well. The common assumption is that there is a morpheme related to reciprocity,² which in Hebrew would be the T-template, illustrated in (2) with the root \sqrt{nsk} :

¹ C stands for the consonants of the root.

² In order to avoid, at this point, commitment to the semantic content of the relevant linguistic markers, such morphemes are dubbed REC throughout this paper.

- (2) a. Transitive construction (basic template)
rut niška et miriam
 Ruth kiss.PST.3.F.SG ACC Miriam
 ‘Ruth kissed Miriam’
- b. *Rec*-construction (T-template)
rut ve-miriam hitnašku
 Ruth and-Miriam kiss.REC.PST.3.PL
 i. Collective reading: ‘Ruth and Miriam kissed each other’
 ii. Distributive reading: ‘Both Ruth and Miriam had reciprocal kissing with someone’ (not necessarily the other)
- c. Discontinuous *rec*-construction (T-template)
rut hitnaška im miriam
 Ruth kiss.REC.PST.3.F.SG with Miriam
 Collective reading (only): ‘Ruth and Miriam kissed each other’

Such triplets of sentences raise the following set of questions:

- (3) i. What is the relationship between the *rec*-predicates (predicates of the *rec*-constructions (2b-c)) and the *basic predicates* (predicates of the transitive constructions (2a))? Is it accurate to assume that either derives from the other? And, if so, what are the operations that this derivation involves?
- ii. What is the relation between the *rec*-predicates in (2b) and (2c)? Considering that the relationship between (2b) and (2c) requires also an account for the origin of the distributive reading of the *rec*-construction (2bii).
- iii. *Prima facie* ‘Miriam’ has the same semantic role in (2b) and (2c); however, the question is whether Miriam in sentence (2c) is an argument or an adjunct (the answer to this question determines the argument structure of the *rec*-predicate, and consequently sheds light on its derivation).

All previous analyses of verbal reciprocals assume that *rec*-predicates are at some level of analysis derivatives of more basic predicates. Furthermore, most reciprocals in Hebrew are found in the T-template, and the assumption, within various theories, about verbs in this template is that they are derivative of functions that have as their input either the root or verbs in other templates.

The following table (Table 1) portrays how previous analyses of verbal reciprocals answered the questions in (3). For the sake of brevity, I do not go into the details of each of these proposals. I only outline the various options that already exist in the literature, and additionally I note how I differ from them.

	The relation between (2a) and 2b-c (I)	The relation between (2b) and (2c) (II)	The status of the oblique expression in (2c) (III)	The nature of the <i>rec</i>-predicate
Nedjalkov (2007)	Syntactic derivation (a)>(b)	Extension (b)>(c)	N.A	Reciprocal verb
Dimitriadis (2008)	Semantic operation (a)>(b) same argument structure	(c)>(b) Reflexivization	Argument (one of the two arguments in a symmetric relation)	Symmetric predicate
Siloni (2012)	Lexical operation (a)>(b/c)	Two entries	Phrase unvalued in terms of thematic role	Take a set as an argument, whose members have two thematic roles each
Current proposal	<i>No derivational relation</i>	<i>Two realizations of the same lexical entry</i>	<i>An argument with a lexical role</i>	<i>A predicate with two arguments; not necessarily symmetric</i>

Table 1: Summary of the previous literature on verbal reciprocals.

The current paper proposes an analysis that provides the following answers to the questions in (3):

- (4)
- i. There is no grammatical relation between (2a) and (2b-c), neither morphological nor semantic.
 - ii. The two constructions in (2b) and (2c) are two different syntactic realizations of the same lexical entry/predicate. This is an atom-predicate (see Section 5.2 for a clarification of the term) and therefore the distributive and the collective readings are expected.
 - iii. The oblique phrase is an argument of the predicate.

The differences between the current proposal and the previous ones lies in the fact that I do not consider the *rec*-predicate to be an encoding of reciprocal/symmetric relations and offer, therefore, an alternative portrayal of the relationship between them and the symmetric events they denote. The current paper provides evidence for this analysis, in relation to the morphological form of these verbs (for a more detailed account on the semantics see Bar-Asher Siegal 2015). In the first part of the paper (Sections 2-3) I will still follow the assumption that reciprocity is part of the meaning of the *rec*-predicates, and I will demonstrate the problems of the derivational approaches. Only in Section 4 will I argue against this semantic assumption.

One further note regarding *rec*-predicates is in order. Siloni (2008, 2012) and, similarly, Doron and Rappaport Hovav (2009) consider reciprocalization as a diathetic operation resulting in valency reduction, which is, in essence, very similar to reflexivization, i.e.: an operation of semantic identification of the external θ -role with the internal θ -role of the verb. The difference between reflexives and reciprocals, accordingly, lies in the number of the participants. Reflexives have individuals as participants, while reciprocals have collective sets (Doron 2008: 70.) I will also examine the validity of this claim and consider various reasons for rejecting it.

I will conclude the introduction with some notes concerning the more general literature about the T-template in Hebrew (regardless of the fact that occasionally it encodes reciprocity): Arad (2005) observes that this template is one of the two templates that are marked as non-transitive. As we shall see, it is crucial to examine whether this is merely a syntactic claim, i.e., that verbs in the T-template never have a direct object (which is marked

in Hebrew with the accusative preposition for definite nouns); or whether this claim is also relevant at the semantic level.

Furthermore, it would be relevant to mention that in other theories, the T-template is related in different ways to arity-operations (valence changing operations). According to Siloni (2008), in light of her analysis to the verbal reciprocals (*inter alia* Siloni 2001, 2012), the T-template is, in a sense, a sensor to indicate various types of diathetic operations resulting in valency reduction. For Doron (2003), the T-template is the phonological representation of the head μ . This head operates as a MIDDLE and reciprocals and reflexives fall under this category

The structure of the paper is as follows. Section 2 raises various problems concerning the claim that *rec*-predicates derive from other predicates, and Section 3 demonstrates why reciprocal verbs are not reflexives. In order to support an alternative approach, I will discuss in Section 4 the semantics of *rec*-predicates, and in Section 5 the argument structure of *rec*-predicates, and how it can account for the various readings the *rec*-constructions have. I will conclude, in Section 6, with various remarks concerning *rec*-predicates and the challenges they raise for theories about the Hebrew/Semitic templates.

2. Problems with derivational approaches

While the basic and the *rec*-predicates in Hebrew and in many other languages may have similar phonological roots, there are various reasons to argue that they do not hold any grammatical-derivational relation.

2.1 Deponent³ reciprocal verbs

Cross-linguistically, there are reciprocal verbs that do not have a transitive counterpart, and nonetheless appear as reciprocals. There are even denominative verbs that feed directly the T-template without feeding a transitive template (Kemmer 1988: 30-31; Nedjalkov 2007:14; Knjazev 2007:118-119). This is, for example, the case with the Hebrew verb *lehit'agref* 'to box', a relatively new verb that does not have a transitive equivalent. Obviously, deponent verbs pose a problem for a theory that considers the *rec*-predicate to derive from the basic (transitive) predicate. Admittedly, this is not an insurmountable problem for a derivational approach. Siloni (2001), for instance, considers these verbs as exemplifications of a derivation from "frozen verbs". Such explanations, however, are not ideal from a theoretical standpoint.⁴

2.2 Adjuncts become arguments

If we think about reciprocalization as a function that takes transitive verbs as input and derives *rec*-predicates, the expectation is that the relevant function would operate on the arguments of the transitive-predicate only, since arity-operations should involve only what is included in the argument structure of the predicate in the input. This expectation, however, is not always met in the case of reciprocal verbs. Often, the roles of the sets participating in the reciprocal relation are adjuncts in their transitive counterparts, as can be illustrated with the

³ A deponent verb is a verb that appears in a non-active form, usually the middle or passive, but is active in meaning. Thus, since *rec*-predicates in Hebrew are found in a "middle" morphology (the T-template), and like other deponent verbs such roots have no other active forms, I follow the term found in Kemmer (1988), as this is the term used in the previous works on this phenomenon (I wish to thank Stephen Anderson for the terminological discussion).

⁴ Similarly, according to Doron (2003) and Arad (2005), such verbs may support the Root Hypothesis and be consistent with a claim that *rec*-predicates derive from the root directly with the aid of some *rec*/reflexive functional head.

root \sqrt{ktv} . In the case of the basic predicate, *katav* ‘write’, the recipient is not an argument of the transitive predicate (5a), since a writing does not entail a “recipient”. In the *rec*-predicate (5b), which appears in the T-template (*lehitekatev* ‘to correspond’), the participants are the writers but also the recipients of the writings:

- (5) a. *yosi katav mixtav le-dani ve-dani katav*
 Yosi write.PST.3.M.SG letter to-Dani and-Dani write.PST.3.M.SG
mixtav le-yosi.
 letter to-Yosi
 ‘Yosi wrote a letter to Dani and Dani wrote a letter to Yosi’
- b. *yosi ve-dani hitakatvu*
 Yosi and-Dani write.REC.PST.3.PL
 ‘Yosi and Dani corresponded’

This observation seems to pose a serious theoretical problem for derivational approaches to verbal reciprocals, as it indicates that even if reciprocity is part of the meaning of the verb, it is not a result of a derivation from the transitive predicate.

2.3 Reciprocity already in the transitive verb

In the case of several other verbs, reciprocity is also part of the meaning of the transitive verb, with a symmetric relation between the internal arguments:

- (6) a. *yosi immet et dani im raxel*
 Yosi confront.PST.3.M.SG ACC Dani with Rachel
 ‘Yosi made Dani have a confrontation with Rachel’
- b. *dani ve-raxel hit'amtu*
 Dani and-Rachel clash.REC.PST.3.PL
 ‘Dani and Rachel clashed’
- c. *yosi mizeg et šte ha-xevr-ot*
 Yosi merge.PST.3.MSG ACC two DEF-company-PL
 ‘Yosi Merged the two company’
- d. *šte ha-xevr-ot hitmazgu*
 two DEF-company-PL merge.REC.PST.PL
 ‘The two companies merged’

The verbs in (6a,c) are in the intensive template, those in (6b,d) are in the T-template, as was the case in (2a-b). While in (6) all verbs have a reciprocal component, in (2) only the one in the T-template has. Another difference between these pairs of verbs is that in roots like the one in (6), the sentences with the verbs in the intensive template entail those with verbs in the T-template, which is not the case in verbs like the one in (2). Moreover, there seems to be a more fundamental difference: verbs like the ones in (6) describe some inherently symmetric relations, either of physical situation (*mizeg* ‘merge, incorporate’, *iged* ‘group’, *irev* ‘mingle, blend’, *šidex* ‘put them together’, *xiber* ‘combine’, *ixed* ‘unify’, *hišmid* ‘glue’ and *hifrid* ‘separate’) or of inter-personal relation (*xiten* ‘marry’ and *šidex* ‘put them together, fix people up’). Verbs of the type illustrated in (2) are not inherently symmetric relations. The *rec*-predicates in these roots describe a reciprocal event of asymmetric relations (when “a kisses b”, it does not entail that “b kisses a”; however if “a clashes with b” then necessarily also “b clashes with a”).

These data may lead to think of one of several theoretical options:

- (i) *The derivation can go in both directions: Intensive Template/root => T-Template or T-Template => Intensive Template.*

According to this option, while in (2) the *rec*-predicate derives from the basic, transitive predicate, in (6) the transitive predicate derives from the *rec*-predicate. Accordingly, there is a functional head, which adds the thematic role of agent/cause. This head can take the *rec*-predicate as the input as well (not only the root).

This option poses a theoretical problem to theories according to which the function responsible for adding the thematic role of agent/cause takes only the basic lexical entry/the root as its input; such theories must, at the same time, assume that reciprocity/symmetry is the result of a different operation on the same lexical entry/root.⁵ In fact, Doron (2003, 2008) already notices that some of the verbs in the causative-template must derive from a middle-predicate (in the T-template).⁶

The problem with this option is that, as noted, these verbs have an inherent symmetric meaning. It is unclear in what sense the verbs in the T-template are the results of “derivation” from the root, and not that they are born, so-to-speak, in the T-template, due to their meaning.

- (ii) *The derivation is always in the same direction: Intensive Template/root => T-Template.*

The basic meaning includes the external (causer) argument. Thus, regardless of the level at which this operation takes place, the operation that results with a T-template only indicates the reduction of the external argument. Reciprocity, accordingly, is a type of lexical meaning that is independent of a specific morpheme, and it is not a type of semantic/morphological/lexical derivation.

- (iii) *Non-derivational approach.*

Rec-predicates tend to be in the T-template. It is not a result of any derivation, and it is regardless of whether reciprocity is inherent to the meaning of the predicate or whether the reciprocal verb is related to another asymmetric predicates.

2.4 Morphologically unmarked verbal reciprocals

Finally, verbs can express reciprocal events even without the morphology of T-templates. *Rec*-predicates appear in other templates as well, as is the case in the following verbs: *nilxam* ‘to fight’ (N-template), *rav* ‘to quarrel’ (basic-template) and *soxeax* ‘to converse’ (intensive-template). The last one is significant, since this root appears also in the basic-template *sax* ‘to say, to tell’ and according to Doron (2003, 2008), only roots that appear in a single template may not reflect the semantics involved regularly with the morphology of templates. Thus, while, according to her theory, it might be unnecessary to explain why the meaning of reciprocity is encoded in a verb such as *rav* that appears only in the basic-template, it is still necessary to explain why the verb *soxeax* (intensive-template) has the reciprocal meaning. Thus, such verbs impose a serious problem to a derivational approach, since they show that reciprocals are a class of verbs regardless of their form. Even if one assumes a semantic

⁵ A comparative note: in Akkadian, reciprocity is usually marked with the T-template (indicated by a *t*-infix), whereas the causative template is marked with a *š*-prefix. When a verb has a reciprocal meaning in the T-template, and it has a causative form as well, both morphemes appear. Compare (i) to (ii):

- (i) *mitgurum* ‘to come to an agreement’
šutamgurum ‘to bring to an agreement’
(ii) *našaqum* ‘to kiss’
nitkusum ‘to kiss each other’

These data may support the idea that there are two different relationships between templates. When they both derive from the root only one affix appears, but when one verb is derived from another, both affixes appear.

⁶ There are still some problems for Doron (2003, 2008), namely that the verbs with the additional causer are in the intensive-template and not in the causative-template.

derivation between different verbs, this derivation is not directly reflected in their morphology.

2.5 Preliminary conclusions

When going through the inventory of a language like Hebrew, it becomes clear that the way reciprocal verbs are introduced in the literature, as the triadic constructions (2a-c), is misleading. It is possible to present such a set of constructions for a given phonological root only for the minority of the *rec*-predicates in Hebrew (many of which express physical contact, which involve also emotions). The vast majority of verbs are either deponent or of the types introduced through the examples in (5)-(6). Furthermore, the idea that reciprocal verbs originate through some derivation from a more basic transitive root/verb relies on such triadic verbs. As demonstrated throughout this section, there are strong reasons not to assume such derivation. The facts demonstrated in Sections 2.1 and 2.3, on the one hand, do not need a convoluted explanation, once a derivation is not assumed; and the data presented in Sections 2.2 and 2.4, on the other hand, seem to have no hope in a derivational approach. The observations in Sections 2.1-2.2 weakened the assumption that *rec*-predicates are necessarily related to other predicates in the first place. Furthermore, the data in Sections 2.2-2.3 lead towards an analysis that the reciprocity of these verbs is unrelated to other predicates with the same root, and cannot be a result of a derivation. Finally, Section 2.4 argues that, if reciprocalization is indeed a semantic process, it is not necessarily reflected in the morphology.

Furthermore, the assumption behind the derivational approach is that, from the semantic point of view, the *rec*-predicates are a reciprocal version of the basic-predicates. As we shall see in Section 4, this is not the case, so, in fact, there will be not even a semantic motivation for such approaches.

3. Reciprocalization and reflexivization

Another topic that is often discussed both in the literature about the *rec*-predicates and in the literature about the morphology of the templates is the relationship between reflexives and reciprocals. The reason for this discussion is the fact that, in many languages, the morphology is identical for verbal reciprocals and verbal reflexives, and accordingly they both fall under the category of what many designate as “middle”. Thus, in theories that assume that the verbal templates represent one-to-one functions, the two must be one and the same phenomenon at some level. For Doron (2008), for example, reciprocalization and reflexivization are similar operations of semantic identification of the external θ -role with the internal θ -role of the verb. Accordingly, reflexives and reciprocals differ only in the number of the participants: reflexives have individuals as participants, while reciprocals have collective sets. In the rest of this section, I will demonstrate in which type of languages such a claim is valid, and that this is clearly not the case in Hebrew and other languages with similar verbal morphology.

In the context of languages like French, which convey reciprocity with a clitic anaphoric expression (Siloni 2012), most reciprocal constructions have two possible interpretations: one as reflexive and the other as reciprocal. Thus, for example, (7)-(8) have two readings:

- (7) *Pierre et Jean se sont lavés* [French]
 Pierre and Jean SE be.PRS.PL wash.PASS.PTCP.PL
 a. ‘Pierre and Jean washed (themselves)’
 b. ‘Pierre and Jean washed each other’

- (8) *Pierre et Jean se sont parlés* (+à eux mêmes)
 Pierre and Jean SE be.PRS.PL speak.PASS.PTCP.PL (to REFL)
 a. ‘Pierre and Jean spoke to themselves’
 b. ‘Pierre and Jean spoke to each other’

Only the context determines the choice of the reading. This is not the case in languages like Hebrew. The sentences in (9)-(10) illustrate that verbs are either reciprocal or reflexive, despite the fact that morphologically reflexive actions and reciprocal relations are both often similarly marked with the T-template.

- (9) *yosi ve-dani hitraxšu* [Hebrew]
 Yosi and-Dani wash.REFL.PST.3.PL (<= T-template)
 a. ‘Yosi and Dani washed (themselves)’
 b. ‘*Yosi and Dani washed each other’
- (10) *yosi ve-dani hitnašku* (<= T-template)
 Yosi and-Dani kiss.REC.PST.3.PL
 a. ‘*Yosi and Dani kissed themselves’
 b. ‘Yosi and Dani kissed each other’

If, in Hebrew, reciprocalization is a sub-type of reflexivization, one for a singular subject and the other for a set, it is surprising not to encounter an ambiguity between the two groups of verbs with a plural subject, as is regularly found in the languages of the French type. Similarly, since, according to Siloni, reciprocalization is essentially a similar operation in both types of languages, it is striking that such an ambiguity occurs only in one type of languages. As for the ambiguity of languages like French, Siloni explains that reciprocalization in the French type is productive for all transitive verbs, as is reflexivization, because both are syntactic. Hence, since they share morphology, the ambiguity is necessary. However, it is unclear why such an ambiguity is prevented in languages like Hebrew. Even if the reflexives and the reciprocals derive from different lexical operations, there should have still been room for the occurrence of both derivations with the same root, and we could expect the existence of similar ambiguity in such languages as well.⁷ These data support an analysis according to which reciprocity is related to the meaning of the verbs and is not merely a result of an arity-operation, therefore each verb has only one meaning: either of a reciprocal or of a reflexive. However, as we shall see in the next section, reciprocity in fact is not even part of the meaning of these predicates.

4. The semantics of the *rec*-predicates

In order to be able to discuss the argument structure of *rec*-predicates a preliminary observation regarding the semantics of these verbs is in order. So far, I followed the assumption that part of the meaning of *rec*-predicates is that they denote a reciprocal relation. The goal of this section is to challenge this assumption. This semantic discussion is also relevant for the question of the morphological derivation that we dealt with in the previous

⁷ There are, however, extremely rarely phonological roots that have two meanings in the T-template, one reflexive and the other reciprocal, such as *lehistabex* ‘to entangle’ (which has the two meaning in English as well: ‘to become twisted together’ and ‘to be involved in a complicated situation’) and *lehitxalek* ‘to be divided’ and ‘to slip’. As for the former, I discuss similar verbs in Bar-Asher Siegal (2015). As for the latter, it seems to be related to two independent meanings of this root, one related to the noun *xelek* ‘part’ and the other to *xalak* ‘smooth’. In any event, these examples are the exceptions. I wish to thank Maayan Nidbach for these two examples.

sections. The assumption was that morphological derivation reflects a systematic semantic shift from a regular transitive relation to a symmetric relation. Various theories assume that the derivation brings about a change in the argument structure of the predicates. Once it becomes clear that *rec*-predicates do not encode symmetry, an alternative analysis for the relation between the predicates is due.

Let us begin with Dimitriadis (2008) and Siloni (2012). According to them, *rec*-predicates denote only irreducibly symmetric events, with the following definition:

Definition: A predicate is *irreducibly symmetric* if (a) it expresses a binary relationship, but (b) its two arguments have necessarily identical participation in any event described by the predicate (Dimitriadis 2008: 378).

Despite their claim, these predicates, what we call *rec*-predicates, in point of fact, do not necessarily encode symmetric relations. Bearing in mind that they appear systematically in the discontinuous *rec*-construction in sentences with asymmetric reading like (11), it is clear that a symmetric or even reciprocal meaning is not necessarily entailed (cf. Carlson 1998):

- (11) *yosi hitnašek im ha-karit*
 Yosi kiss.REC.PST.3.M.SG with DEF-pillow
 ‘Yosi kissed the pillow’

While evidently (11) entails two participants, it does not entail that they are either active in the same way or even that both participants are active at all. Assuming that REC denotes reciprocity, one must provide an explanation for the possibility of sentences such as (11). The following three explanations are possible in theory:

- (i) Sentences similar to the one in (11) are representations of a different non-reciprocal predicate (Siloni 2012). However, one can easily illustrate that for each of the *rec*-predicates, in certain contexts, sentences with different levels of agency are available.⁸

⁸ See, for instance, the following examples:

- (i) *spaydermen mit'agref im saq igruf*
 Spiderman box.REC.PST.M.SG with bag.of boxing
 Lit. ‘Spiderman engaged in a “mutual boxing” with a boxing bag’.
 (<http://www.haaretz.co.il/misc/1.980551>)
- (ii) *be-migraš-im im kirot gvoh-im ešar lehitmasev im*
 in-playground-PL with wall.PL tall-M.PL possible throw.REC.INF with
ha-kir ve-la'asot al-av trikim
 DEF-wall and-do.INF on-3.M.SG trick-PL
 ‘In playgrounds with high walls it is possible to play *throw and catch* with the wall and to trick it’
 (<http://forum.vgames.co.il/showthread.php?p=1763593>)
- (iii) *u-me-rov romantika kim'at hitnašakti im eš*
 and-from-much romantics almost kiss.REC.PST.1.M.SG with tree
 ‘Overcome by the romantic atmosphere, I almost kissed (rec) a tree’
 (<http://www.groopy.co.il/forummessage.aspx?messageid=182809>)
- (iv) *oded menase lehit'amet im ha-mašav*
 Oded try.PRS.M.SG confront.REC.INF with DEF-situation
še-nixpa al-av
 REL-force.PASS.PST.3.M.SG on-3.M.SG
 ‘Oded is trying to confront the situation that was forced upon him’
 (http://www.rashut2.org.il/prod_program.asp?pgId=42232&sttsId=2)
- (v) *adif lehizdayen im buba mitnapax-at ve-lo lagaat*
 better have.intercourse.REC.INF with doll blow.up-F.SG and-NEG touch.INF
 ‘It is better to have intercourse with a blow-up doll and not touch’
 (www.shin1.co.il/ya.php?sid=7475664)

Consequently, a theory allowing one entry for all uses of the verbs in the lexicon should be favored.

- (ii) These are metaphorical uses of reciprocalization: *rec*-predicates in sentences like (11) are used metaphorically (Yosi was hugging the pillow as if it were another person). However, if this is indeed a metaphorical reading of a reciprocal construction one may wonder why such metaphorical readings are unavailable with other reciprocal constructions, such as the pronominal ones (11'a) or in the *rec*-construction (11'b) with a plural verb:

- (11') a. #yosi ve-ha-karit nišku exad et ha-šeni
 Yosi and-DEF-pillow kiss.PST.3.PL one ACC DEF-second
 b. #yosi ve-ha-karit hitnašku
 Yosi and-DEF-pillow kiss.REC.PST.3.PL
 Int. 'Yosi and the pillow kissed each other'

Thus, it becomes clear that the option of an asymmetric reading is related to that fact that (11) is in the discontinuous *rec*-construction.

- (iii) The last option to explain the non-reciprocal uses of these predicate is the most trivial one: *reciprocity is not necessarily encoded by the rec-predicates*. *Rec*-predicates are dyadic predicates, with one argument expressing the "agent" (*rec*-er) and the other the "patient" (*rec*-ed). As will become clear, according to this analysis, *rec*-predicates are no longer verbs that encode reciprocity, but a subset of verbs that appear (mostly) in the T-template.

5. The argument structure of *rec*-predicates and the various readings of the *rec*-construction

5.1 The argument structure of *rec*-predicate

One of the leading questions of this paper is the status of the oblique expression in the discontinuous *rec*-construction (cf. (3iii)). I would like to argue that it is an argument of the *rec*-predicate, proposing (12) as the argument structure (=AS) of such predicates:

- (12) *rec*-predicate (*rec*-er, *rec*-ed)

(13) is a reasonable criterion for being an argument of a predicate and for its lexical role (for a justification of this definition see Bar-Asher 2009).

-
- (vi) gam l-i ba lehitxabek im ha-adama hazo
 also to-1.SG come.PST.3.M.SG hug.REC.INF with DEF.land DEM.F.SG
 'I too wish to hug this land'
 (<http://d-spot.co.il/forum/index.php?showtopic=105750>)
- (vii) im ata matxil lehitgofef im ha-ašiš siman
 COND 2.M.SG begin.PRS.M.SG embrace.REC.INF with DEF-plant sign
 še-ata šikor
 REL.2.M.SG drunk
 'If you start embracing a plant, it is a sign that you are drunk'
 (<http://forum.bgu.co.il/Index.php?showtopic=190581>)
- (viii) ve-lex titvakeax im agadot
 and-go.IMP.2.M.SG argue.REC.FUT.2.M.SG with fairy.tale.PL
 'And go argue with fairy tales'
 (<http://www.nelech.co.il/Track/TrackInfo/16>)

- (13) To be an argument of a predicate is to be a function (=of a lexical role), the argument of which is bound with an existential quantifier in every instantiation of the predicate.

When considering the entailments of the *rec*-predicates, and also bearing in mind that these predicates appear systematically in sentences with asymmetric readings of the sort seen in (11)-(14), it is clear that a symmetric or even a reciprocal meaning is not necessarily entailed:

- (14) *yosi hitxabek im ha-karit*
 Yosi hug.REC.PST.3.M.SG with DEF-pillow
 ‘Yosi hugged the pillow’

The entailment, therefore, of all instantiations of this predicate is that there must always be in every hugging at least one hugger and one being hugged (a *hugger* and a *hugged*), and not necessarily two *huggers* and two that are being *hugged*.

- (15) $\forall e [rec-hug' (e) \rightarrow \exists x \exists y (rec-hugger (e, x) \wedge rec-hugged (e, y))]$

Thus, (16) could be proposed as the AS for this predicate:

- (16) a. *rec-hug* (*rec-hugger*, *hugged*)
 b. *rec-hug'* $\rightarrow \lambda y \lambda x \lambda e [rec-hug' (e) \wedge rec-hugger (e, x) \wedge hugged (e, y)]$

In a language like English, without a *rec*-morphology, a lexical description is problematic, as it is hard to distinguish between the transitive-predicate (2a) and the *rec*-predicate (2b-c). Using the equivalent Hebrew vocabulary, *hitxabqut* ‘*rec-hugging*’ involves *mitxabek* (the active participle form of the *rec*-verb) and *mexubak* (the passive participle form of the transitive counterpart verb). This is therefore what stands as the background for the more general proposal in (17) – a repetition of (12) – for the AS of *rec*-predicates:

- (17) *rec*-predicate (*rec-er*, *rec-ed*)

Note that (17) *does not claim that this is the function of the morphology involved with rec-predicates*. It only characterizes the AS of all the predicates that exhibit what we saw in (2b-c). In other words, this is what is in common to all verbs that have the reciprocal collective and distributive readings with plural subjects, and that at the same time may appear in the discontinuous *rec*-construction, where a reciprocal reading is not necessary. As we saw in Section 2.4, such verbs may appear in a variety of templates in Hebrew.

The idea behind this proposal is that, while it is clear that there are irreducibly symmetric events, this does not mean that there are irreducibly symmetric predicates. In other words, *symmetry is a characteristic of an event and not a content of a predicate*. Furthermore, having (17) as the AS, the discontinuous *rec*-construction is accordingly the basic realization of such predicates, as the *rec-er* is represented by the subject and the *rec-ed* by the oblique. Consequently, I will argue for the following two claims:

- (a) The predicates of the *rec*-construction and of the discontinuous *rec*-construction are the same, and represented in (17).
 (b) As for the collective and the distributive readings of the *rec*-construction: the collective reading is a specific reading of these predicates. The two alternative readings are always possible with atom-predicates.

The idea behind these claims is that, considering the following two sets of sentences ((18)-(19)), the sentences in (18c)-(19c) with the plural subjects describe the same events that are described by the sentences in (a+b), and are composed semantically in an identical way. The reciprocal reading of (19c), accordingly, is pragmatic, when (a) and (b) are contextually understood to be a description of the same event:

- (18) a. *John kissed (Jacob)*
 b. *Betty kissed (Marry)*
 c. *John and Betty kissed*

- (19) a. *John kissed (Betty)*
 b. *Betty kissed (John)*
 c. *John and Betty kissed*

When there is a plural subject, it may appear with a reciprocal anaphoric expression (20), and, as the sentence in (21) demonstrates, it is possible for the *rec*-ed to be only an implicit argument, i.e. not overtly expressed:

- (20) *yosi ve-rina hitnašku (exad im ha-šeni)*
 Yosi and-Rina kiss.REC.PST.3.PL (one with DEF-other)
 ‘Yosi and Rina kissed (each other)’

- (21) *rachel hitnašqa ha-yom b-a-pa’am ha-rišona*
 Rachel kiss.REC.PST.3.F.SG DEF-day in-DEF.time DEF-first.F
 ‘Rachel kissed someone today for the first time’ (implied: a reciprocal kissing)

Therefore, according to the current analysis, from a syntactic point of view, sentences with a plural subject (cf. (20)) are like sentences with only one overt argument (in subject position) and additional implicit arguments (cf. (21)).

5.2 The distributive and the collective readings of the *rec*-construction

The analysis presented here consists of two claims. First, the *rec*-predicates are atom predicates, and not set-predicates (cf. Winter’s 2002 typology of predicates).⁹ Second, the collective reading is a particular reading of sentences with plural subjects, and the difference between the distributive and collective readings results from issues related to scope of the event quantification (cf. *inter alia* McCawley 1968; Higginbotham and Schein 1989; Schein 1993 and Lasnik 1995).

Such an approach leads naturally to double readings. If we take a set A, denoted by a plural subject, it may have a distributive reading, as with other dyadic predicates:

- (22) $\forall x \in A \exists e \exists y R_{rec}(e, x, y)$

In such a reading, each member of the set denoted by the subject is the *rec*-er of a different event. As for the collective reading, it is not related to a different syntactic or lexical configuration than the one of the distributive reading. In this reading all members of set A are participating in the same event, without other participants, and all the participants occupying the *rec*-ed position are also members of the set A. Thus, in such a case, *especially when the cardinality of the set A is 2*, there is an irreducibly symmetric event. (23) is a representation of

⁹ A more elaborate version of this paper (Bar-Asher Siegal in prep.) includes support for this claim.

the collective reading (the existential quantifier binds the event variable and has a wider scope), and since (23) entails (22), (23) can be a specific reading of (22).

$$(23) |A| = 2 \quad \exists e \forall x \in A \exists y \in A [(x \neq y \rightarrow R_{\text{rec}}(e, x, y))]$$

According to this hypothesis, it is not the case that the *rec*-predicates are monadic set-predicates, but are dyadic atom-predicates. This is true with every atom-predicate with a plural subject that may be read as a single event or as a plural event (cf. Lakoff and Peters 1969, who claim that all these examples are the result of an ambiguity between sentential and phrasal conjunction), as illustrated by (24):

- (24) *John and Bill went to New York*
 a. *John and Bill, each of them, went separately to New York*
 b. *John and Bill went together to New York*

Therefore, what has been previously analyzed as the most basic manifestation of *rec*-constructions is, in fact, a specific reading of sentences with plural subjects according to which all relations are held in one single event.

5.3 Ramification for the discussion regarding the T-template in Hebrew

As noted earlier, according to Arad (2005), the T-template is marked as non-transitive. In light of the current analysis for the *rec*-predicates, Arad's claim is true only if this is a syntactic observation and it refers to the fact that these verbs never have a direct object. It is not true, however, if it claims that all verbs in the T-template have only one argument at the semantic level. Moreover, it is important to note that the restriction on having a direct object with the *rec*-predicate is not limited to the T-template, it is also true with *rec*-predicates in other templates:

- (25) *yosi rav im rina / *et rina*
 Yosi fight.PST.3.M.SG with Rina / ACC Rina
 'Yosi had a fight with Rina'

- (26) *yosi soxeax im rina / *et rina*
 Yosi converse.PST.3.M.SG with Rina / ACC Rina
 'Yosi held a conversation with Rina'

6. Concluding remarks

6.1 *Rec*-predicates and the challenges for theories regarding the Hebrew/Semitic templates

Section 2 discussed the problems of considering *rec*-predicate as a derivative of the basic-predicate. Section 4 established that reciprocity/symmetry is not part of the meaning of the *rec*-predicates, thus giving further support to the claim that they do not derive from other predicates. Section 5 provided an explanation for the various readings the *rec*-constructions have. Thus, up to this point, we have been seeking to demonstrate that (17) as a *rec*-predicate with one dyadic lexical entry can explain the various constructions involved with *rec*-predicates and their semantics. This leaves us, then, to deal with the following questions:

- (a) What is the difference between *rec*-predicates and transitive predicates?

(b) Why is the *rec*-ed always realized as an oblique and not as a direct object?

Dealing with question (b) is beyond the scope of the current paper (the matter was discussed in Bar-Asher 2009). As for question (a), if we have sentences (2a-c) in mind, *prima facie* the difference is obvious, since (2a) is not reciprocal and (2b-c) are. However, as has been demonstrated with sentences such as (27), reciprocity is not entailed by *rec*-predicates. Thus, the more subtle question concerns the difference between (27) and (28):

(27) *yosi hitxabek im karit*
 Yosi hug.REC.PST.3.M.SG with pillow
 ‘Yosi hugged a pillow’

(28) *yosi xibek karit*
 Yosi hug.PST.3.M.SG pillow
 ‘Yosi hugged a pillow’

I would like to argue that this is a lexical discussion and not a grammatical one; it would then follow that this issue need not be addressed by a theory of *rec*-predicates. In order to justify this claim, it should be mentioned that, as was noted earlier, the introduction of (2a-c) as a triadic construction was misleading. It is incorrect to assume that the *rec*-predicate is a reciprocalized transitive predicate: *rec*-predicates are, instead, *independent* (non-derived) predicates. They are bivalent like other transitive verbs, as their argument structure indicates (cf. (12)-(17)). While speakers of Hebrew may share intuitions about the differences between (27) and (28), such intuitions are similar to other lexical observations that speakers can provide about the subtle nuances concerning the semantic differences between verbs with similar denotations, such as *lehistakel* and *lehabit*, both meaning ‘to look’ / ‘to watch’ in Hebrew. However, we may still say that *lexabek* ‘to hug’ (transitive) and *lehitxabek* ‘to hug’ (*rec*) have different types of hugging events as their denotations. The type of hugging which is “reciprocal” is, accordingly, denoted only by the *rec*-predicate, but it is not a result of some syntactic/semantic/lexical operation, since as we saw throughout this paper there are strong reasons to reject derivational approaches to the *rec*-predicates.

6.2 Returning to the Semitic templates

The non-derivational analysis for the *rec*-predicates proposed throughout this paper is clearly consistent with Aronoff (1994) and Arad (2005), who consider the templates as inflectional/conjugational classes. In this case, it is not even clear if there is a distinguished functional head for the *rec*-predicates (note that we still have to answer the question: why most of them are in the T-templates? A question that was not dealt with here).

Prima facie this analysis challenges the hypothesis put forward by Doron (2003), who considers the templates as a spell-out of syntactic heads, and reciprocity as a type of reflexive/middle morphology. However, if we do accept that *rec*-predicates can still fall under the category of middle (as argued by Kemmer 1993 and others), some amendments must be made to such a hypothesis (see Section 2.3), and some further clarifications for the content of the category of middle must be made.

Acknowledgements

I wish to thank the students of my various advanced seminars at the Hebrew University, in which I presented different parts of this paper, for many insightful questions and comments. In addition, I am grateful to Odelia Ahdout, Nora Boneh, Karen De Clercq, Luka Crnic, Edit Doron, Noam Faust, Danny Fox, Yael Greenberg,

Eitan Grossman, Larry Horn, John Huehnergard, Malka Rappaport Hovav and Aynat Rubinstein for either reading this paper, or discussing with me its content. Finally, I wish to thank the audience of *Reciprocals cross-linguistically*, Berlin November-December 2007, and the audience of the MMM10 for the productive comments I received following the presentation of my paper.

References

- Arad, M. (2005) *Roots and patterns: Hebrew morpho-syntax*. Dordrecht: Springer.
- Aronoff, M. (1976) *Word-formation in Generative Grammar*. Cambridge: The MIT Press.
- Bar-Asher, E. (2009) *A theory of argument realization and its application to features of the semitic languages*. Ph.D Thesis. Harvard University.
- Bar-Asher Siegal, E. A. (2014) Notes on the history of reciprocal NP-strategies in Semitic languages in a typological perspective. *Diachronica* 31: 337-378.
- Bar-Asher Siegal, E. A. (2015) Reciprocal verbs in Hebrew. Paper presented at the Workshop on *Verbs, verb phrases, and verbal categories*, Hebrew University of Jerusalem, March 23-24, 2015.
- Bar-Asher Siegal, E. A. (in prep.) Verbal reciprocal constructions.
- Behrens, L. (2007) Backgrounding and suppression of reciprocal participants: A cross-linguistic study. *Studies in Language* 31: 327-408.
- Berman, R. (1975) Rišumam ha-leksikali šel pealim, šorašim u-vinyanim [Lexical entries for verbs, roots and patterns]. In: R. Nir & B. Fischler (Eds.), *Kišon Amo*. Jerusalem: Council on the Teaching of Hebrew, p. 384-398.
- Carlson, G. (1998) Thematic roles and the individuation of events. In: S. Rothstein (Ed.), *Events and grammar*. Dordrecht: Kluwer, 35-51.
- Dimitriadis, A. (2008) The event structure of irreducibly symmetric reciprocals. In: T. Heyde & Z. J. Dolling (Eds.), *Event structures in linguistic form and interpretation*. Berlin: De Gruyter, 327-354.
- Dimitriadis, A. (2012) An event semantics for the Theta System. In: M. Marelj, T. Siloni & M. Everaert (Eds.), *The Theta System: Argument structure at the interface*. Oxford: Oxford University Press, 308-353.
- Doron, D. (2003) Agency and voice: The semantics of the semitic templates. *Natural Language Semantics* 11: 1-67.
- Doron, D. (2008) The contribution of the template to verb meaning. In: G. Hatav (Ed.) *Modern linguistics of Hebrew*. Jerusalem: Magnes Press, 57-88.
- Doron, E. & M. Rappaport Hovav (2009) A unified approach to reflexivization in Semitic and Romance. *Brill's Annual of Afroasiatic Languages and Linguistics* 1: 75-105.
- Kemmer, S. (1993) *The middle voice: A typological and diachronic study*. Amsterdam: John Benjamins.
- Knjazev, J. P. (2007) Lexical reciprocals as a means of expressing reciprocal relations. In: V. P. Nedjalkov (Ed.) *Reciprocal constructions*. Amsterdam: John Benjamins, 115-146.
- Lakoff, G. & S. Peters. (1969). Phrasal conjunction and symmetric predicates. In: D. A. Reibel & S. A. Schane (Eds.), *Modern studies in English*. Englewood Cliffs: Prentice-Hall INC, 113-142.
- Nedjalkov, V. P. (2007a) Overview of the research: Definitions of terms, framework, and related issues. In: V. P. Nedjalkov (Ed.), *Reciprocal constructions*. Amsterdam: John Benjamins, 3-114.
- Nedjalkov, V. P. (Ed.) (2007b) *Reciprocal constructions*. 5 vols. Amsterdam: John Benjamins.
- Nedjalkov, V. P. & E. Š. Geniušienė (2007) Questionnaire on reciprocals. In: V. P. Nedjalkov (Ed.), *Reciprocal constructions*. Amsterdam: John Benjamins, 379-434.
- Ornan, U. (1971) Binyanim uvsisim, netiyot ugzarot. *Hauniversita* 16: 15-22.
- Siloni, T. (2001) Reciprocal verbs. *Proceedings of the Israeli Association of Theoretical Linguistics* 17. Available at: <http://linguistics.huji.ac.il/IATL/17/>
- Siloni, T. (2008) On the *hitpa'el* template. In: G. Hatav (Ed.), *Modern linguistics of Hebrew*. Jerusalem: Magnes Press, 111-138.
- Siloni, T. (2012) Reciprocal verbs and symmetry. *Natural Language & Linguistic Theory* 30: 261-320.
- Winter, Y. (2001) *Flexibility principles in boolean semantics: The interpretation of coordination, plurality, and scope in natural language*. Cambridge: The MIT Press.
- Winter, Y. (2002) Atoms and sets: A characterization of semantic number. *Linguistic Inquiry* 33: 493-505.

Motivated phonological templates in Sign Language

Gal Belsitzman
University of Haifa
galbel84@gmail.com

Wendy Sandler
University of Haifa
wsandler@research.haifa.ac.il

1. Introduction

One of the most striking and universal characteristics of language is usually taken for granted: the existence of distinct structural levels. All languages have phonology, morphology, syntax, prosody, and semantics, and each level has its own types of forms and its own rules and constraints for combining them. Sign languages – which arise spontaneously within communities of deaf people – are no exception. In fact, sign language linguistics as a field was born as a result of the seminal discovery by Stokoe (1960) that the meaningful level of signs/words is distinguishable from a meaningless level, akin to phonology, which provides its building blocks. This discovery implies that signed and spoken languages are similar in basic ways. It was surprising because signs appear to have iconic form-meaning correspondences, and were therefore assumed to be wholes that could not be broken down into meaningless parts, unlike spoken words, which are divisible into meaningless phonological segments or features. Since Stokoe's work, linguists have gone on to analyze each level of structure in sign languages, and have found numerous similarities between them and spoken languages (Sandler and Lillo-Martin 2006).

Yet sign languages are exceptional in the degree to which their words are iconically motivated, and this high degree of iconicity means that the phonological and morphological levels cannot always be cleanly and discretely separated.¹ Here we will show a unique type of interaction between phonology and morphology in sign languages, suggesting that the physical form that a language takes influences its linguistic form in nontrivial ways. Specifically, we will show, following Lepic et al. (2016), that the availability of two hands in sign languages is exploited in lexical word formation in largely predictable ways, due to iconicity. That is, the phonological structure of certain categories of signs is determined by meaning. We go on to adopt a templatic account that is influenced by morphological templates in Semitic languages, but, in the sign language case only, we show how phonological aspects of lexical templates are determined by meaning.

We begin in Section 2 with a brief overview of sign language phonology, including both one- and two-handed signs, and of morphology, including inflectional morphological templates that have been proposed earlier. In Section 3, we proceed to demonstrate that whether an uninflected lexical sign is one- or two-handed is often determined by particular categories of meaning. We propose some templatic schemata – motivated phonological templates – for different categories of two-handed signs. Despite the fact that formational elements have meaning, such signs are typically analyzed as monomorphemic. This lexical motivatedness blurs the line between morphology and phonology that is usually assumed, and leads us to conclude in Section 4 that the phonological and the morpho-lexical levels of language are not mutually exclusive.

¹ Iconicity does not stop at the level of the sign. There is iconic motivation behind morphology and syntax in sign languages as well (Taub 2001; Wilcox 2004; Perniss, Thompson and Vigliocco 2010; Meir et al. 2013).

2. Phonology and morphology in Sign Languages

The purpose of this section is to demonstrate briefly that the two levels, phonology and morphology, are indeed distinguishable in sign languages.

2.1 Phonology

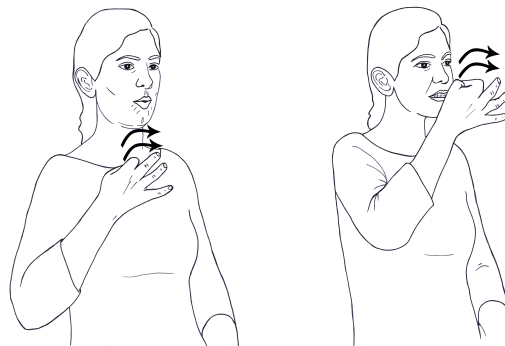
Signs are comprised of three major phonological categories: handshape, location, and movement (Stokoe 1960). Like consonants and vowels of spoken languages, each of these major categories consists of a list of features (Liddell and Johnson 1989; Sandler 1989; Brentari 1998). These elements behave phonologically, in the sense that the constraints on their structure and the rules that manipulate them are related to their form and not to meaning.

Turning first to contrastiveness, Figure 1 shows minimal pairs in Israeli Sign Language (ISL) that differ only in features of (a) handshape, (b) movement, and (c) location.

(a)



(b)



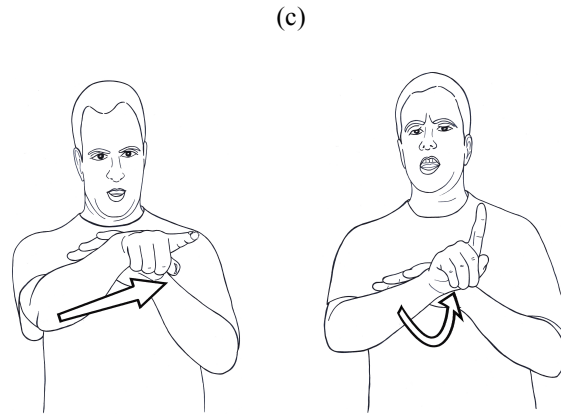


Figure 1: Phonologically distinguished minimal pairs in ISL, (a) MOTHER, NOON - different handshapes (b) WELL-BEING, CURIOSITY - different locations (c) ESCAPE, BETRAY - different movements

Slightly fewer than half of the signs in vocabularies of sign languages use two hands. Of these, there are two basic types (Battison 1978), called balanced and unbalanced (van der Hulst 1993) among other labels. In balanced signs like MEET in Swedish Sign Language (SSL) (Figure 2), the two hands are configured and move symmetrically, and in unbalanced signs like ESCAPE (Figure 1c) above, the nondominant hand serves as a location/place of articulation, while the dominant hand articulates the movement.



Figure 2: Balanced sign, MEET in SSL

The canonical phonological shape of a sign consists of a single hand configuration that moves from one location to another, to manifest a Location-Movement-Location (LML) monosyllable (Liddell 1984; Liddell and Johnson 1986; Sandler 1986, 1987, 1989). The fingers selected that specify a given hand configuration remain constant across a sign, but their position may change, e.g., from open to closed.

In addition to the criterion of contrastiveness, categories and their features are subject to rules based on their form. Processes like assimilation and deletion can affect whole feature categories without regard to meaning, even if the base sign is iconic, in which case the iconicity can be obscured. This means that these formational elements are phonological. Furthermore, if the sign is two-handed, the behavior of each hand is determined by its phonological category membership, as would be expected if meaning is irrelevant. For example, the ISL lexicalized compound THINK+STOP = TAKEN-ABACK (“surprised”) reduces to a canonical, monosyllabic LML form (Sandler 1989, 1999a) by deleting locations and assimilating hand configuration regressively, as shown in Figure 3 (a-c). Crucially, the fact

that the two-handedness of the second sign assimilates regressively, obscuring the form of the first sign of the compound, shows that two-handedness behaves like a phonological element.

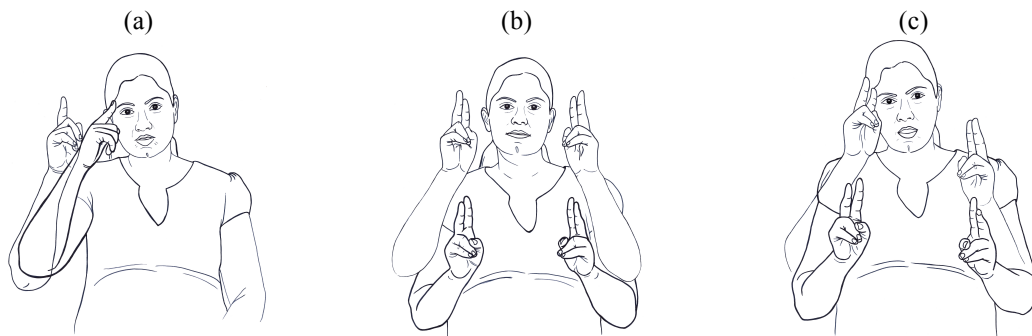


Figure 3: Lexicalized compound in ISL, (a) THINK + (b) STOP = (c) TAKEN-ABACK ("surprised")

2.2 Morphology

Compounding is but one example of sign language morphology, perhaps the richest and most dynamic level of structure in established sign languages. Even though sign languages are young compared to spoken languages – at most 300 years old -- these languages exhibit a range of inflectional and derivational processes, such as verb agreement (Fischer and Gough 1978; Padden 1988; Meir 2002), compounding (Klima and Bellugi 1979; Liddell and Johnson 1986; Sandler 1989), complex classifier constructions (Supalla 1982; Emmorey (ed.) 2003), and temporal aspect inflection (Klima and Bellugi 1979; Sandler 1989, 1990). All of these processes are typically found across sign languages, with similar (but not identical) formal properties, apparently because they are based on iconic images and iconic use of space, while more opaque affixal processes that rely on grammaticalization are fewer and take longer to emerge (Aronoff et al. 2005).

Here we exemplify a particular type of morphological process that is analyzed as templatic: temporal aspect morphology, that is, inflections for aspects such as iterative, durational, etc. There are many other types of morphology, as sketched in the previous paragraph, but we present temporal aspect here because it is templatic, as background for the phonological templates that we will propose for two-handed signs in section 3. Originally described for American Sign Language (ASL) by Klima and Bellugi (1979), and developed by Newkirk (1979, 1981), verbs are inflected for aspects such as habitual, durational, continuative, and iterative. They do so, not by added affixes, but by systematically altering the shape and/or rhythm of the sign's production. The citation form of the sign LOOK-AT and its inflected Durational form are illustrated in Figure 4 (a,b).



Figure 4: ASL sign LOOK-AT (a) citation form and (b) durational form, reproduced with permission from Ursula Bellugi ©

Since the form of these aspects involves gemination of locations or movements and/or changes in the shape of movement, typically from a straight path to an arc, they may be considered comparable to prosodic morphological templates of Semitic languages as proposed in McCarthy (1979, 1981). In that model, the root consonants are associated with templates that account for gemination by double association of the features of the root consonant to C timing slots. Vowel “melodies” vary roughly according to inflectional class, such as active or passive. If a given form requires gemination of the vowel, its features too are doubly associated to V slots. In McCarthy’s model, each ‘tier’ – the consonantal features tier, the vowel features tier, and the CV tier of prosodic form – has the status of a morpheme. Some examples from standard Arabic are shown in Figure 5 (a,b).

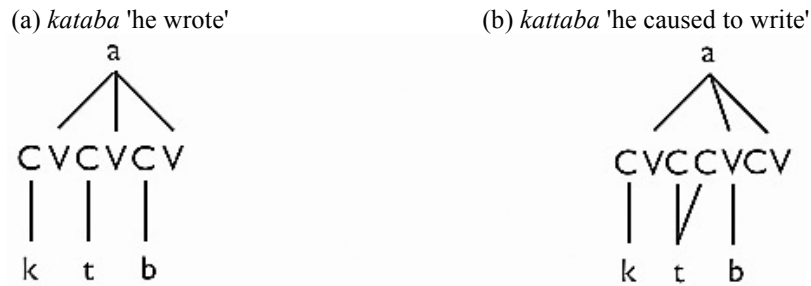


Figure 5: Arabic examples of McCarthy's model, (a) *kataba* and (b) *kattaba*

We have seen that the canonical form of a sign in sign language is LML (roughly comparable to CVC). Under temporal aspect inflection, the basic features of handshape and location stay the same; movements, like vowels, can change their shapes, creating circles, for example, and the prosodic form is altered through gemination of one or more segments. Templates of the citation form of the ASL sign LOOK-AT and its Durational form are shown in Figure 6a and b.² The model schematized here and throughout this paper is from Sandler (1989), inspired by models of CV phonology (Clements and Keyser 1983), autosegmental phonology (Goldsmith 1976) and prosodic templatic morphology (McCarthy 1979, 1981).

The citation form (6a) is produced by configuring the hand with the index and middle fingers extended, fingertips pointed outward, and moving the hand from a position near the body (proximal) to a position farther from the body (distal).³ In (6b), the movement segment is specified for the feature [arc] and the sign is reduplicated, resulting in the circular movement pattern (see Sandler 1990 for details). Temporal aspect morphology of a similar character is found in ISL. For example, the Continuative is formed in ISL by making the final location longer in duration, represented schematically in Figure 7.

² LOOK-AT is an agreement verb, which inflects for subject and object by changing the direction of movement, i.e., by changing the feature specifications of each location (see e.g., Meir 2002). The inflected form can still undergo temporal aspect inflection in the same way as shown in Figure 6, since the aspectual template does not affect the location features, only their durations and whether the sign is reduplicated.

³ The sequential structure shown here, in which locations follow one another separated by a movement, originated in a different model with Liddell (1984); see Sandler and Lillo-Martin (2006, chapter 9) for detailed discussion of sequentiality/linearity in sign language.

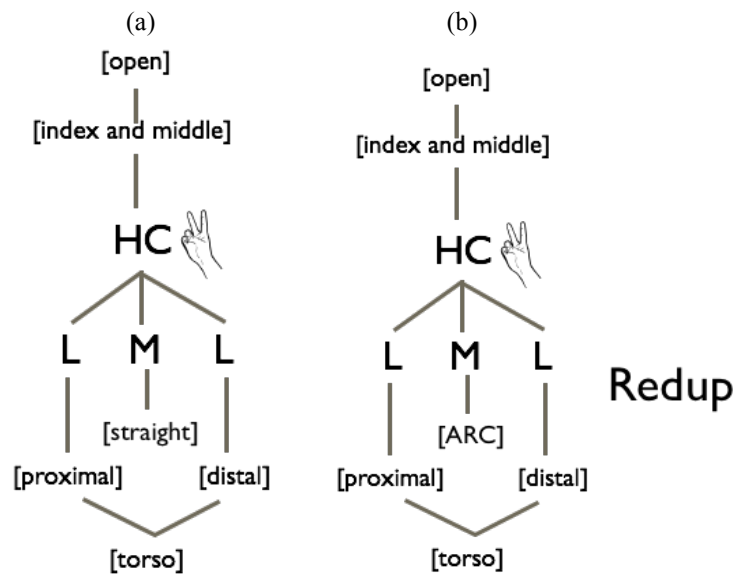


Figure 6: Templates of (a) Partial citation form of ASL LOOK-AT and (b) association to the Durational form (after Sandler 1990). The movement feature in the Durational template is [arc], and the whole sign is reduplicated.

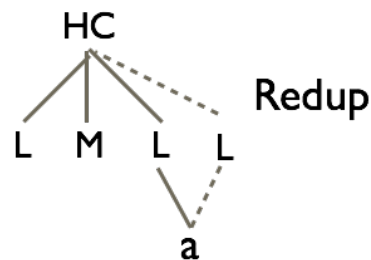


Figure 7: Association of a verb to the Continuative template in ISL (Sandler 1993)

The templates shown involve formal inflectional processes that apply to lexically specified base forms that are comprised of meaningless phonological material. In all of these ways, we find that sign languages are similar to spoken languages in the sense of having distinctly identifiable phonological and morphological levels. In what follows, we demonstrate that certain phonological elements of two-handed signs are semantically motivated, and we propose partially specified phonological templates for these forms.

3. Semantic motivation for phonological features

For many years, sign language researchers were influenced by de Saussure's observation (de Saussure 1983) that the relation between sound and meaning is arbitrary, eschewing iconic motivation in sign language on the assumption that it is 'nonlinguistic'. One reason for treating formational properties such as two-handedness as meaningless is the existence of linguistic constraints and processes that treated them as such, for example, assimilation that affects the nondominant hand as a formational element, together with the dominant hand in symmetrical signs, without reference to meaning (Sandler 1989, 1993). However, in recent years, sign language researchers have begun to view the correspondence between form and meaning in lexical signs as an important clue for understanding linguistic structure (e.g., Taub 2001; Wilcox 2004; Perniss, Thompson and Vigliocco 2010; Meir et al. 2013).

While semantic motivations that underlie phonological elements in sign language lexicons have often been overlooked or ignored, some studies have dealt with them explicitly (e.g., Brennan 1990; Johnston and Schembri 1999; Wilbur 2008; Strickland et al. 2015). Van der Kooij (2002) specifically discusses semantic aspects of unbalanced two-handed signs. She argues that in many asymmetric signs the non-dominant hand is not a phonological element that represents a place of articulation, but in fact a separate morpheme that is semantically motivated. In our analysis, we discuss both balanced and unbalanced two-handed signs, but we do not consider the non-dominant hand to represent a separate morpheme, since its omission from a sign does not leave a morphological unit of any kind. In this section, we present templates of two-handed signs and we identify iconic motivations for several features that have traditionally been treated as meaningless.

3.1 Motivating two-handedness

Broadly speaking, two-handed signs are either balanced or unbalanced.⁴ Figure 8 (a,b) shows representations of these two types of two-handed signs, with examples. Specifically, the representations show that the nondominant hand either acts as a copy of the dominant hand, in balanced signs, or as a place of articulation in unbalanced signs. The details of representations for such signs are different representations in other models of sign language phonology (e.g., Stokoe 1960; Blevins 1993; van der Hulst 1993; Brentari 1998), but in all of these, the elements that are represented are treated as meaningless. We show below that two-handedness is not random, and briefly describe relationships between form and meaning in two-handed signs, following Lepic et al. (2016). It is not only whether a sign is balanced or unbalanced that is motivated; other phonological details such as movement type and contact between the two hands are systematically related to meaning as well, which we illustrate with templates in Section 3.2.

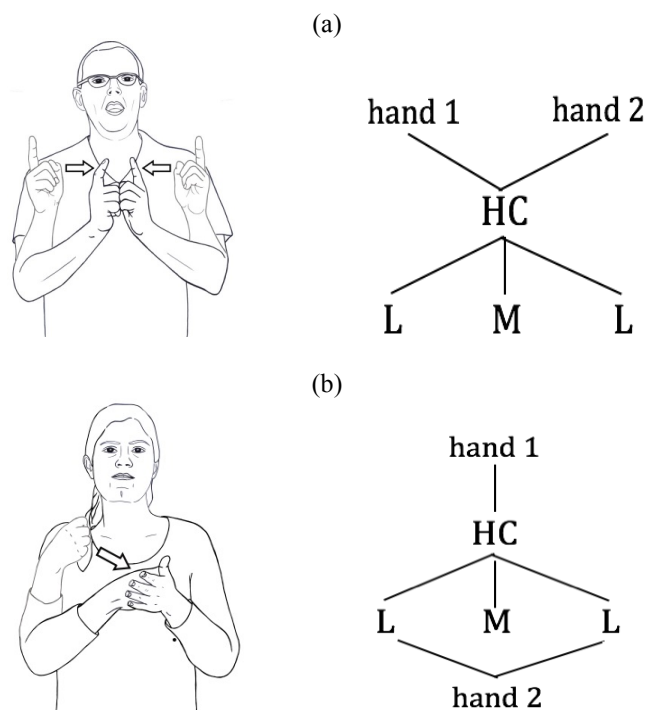


Figure 8: Underspecified templates for two-handed signs, (a) a two-handed balanced template in the SSL sign MEET and (b) a two-handed unbalanced template in the ISL sign END (templates after Sandler 1989, 1993).

⁴ The terms “balanced” and “unbalanced” are from van der Hulst (1993).

As mentioned above, the distribution of signs in the vocabulary of any given sign language is nearly even, about half of the signs are one-handed and the other half, two-handed (Nilsson 2007). Therefore, it may appear as if two-handedness is a chance phenomenon, and that the choice between using one or two hands in a new sign is arbitrary. However, the use of two hands in a sign is often semantically motivated, either underlyingly or by inflection (see references above).

A recent study by Lepic et al. (2016), on which this section focuses, is a cross-linguistic study on three unrelated sign languages – American Sign Language (ASL), Israeli Sign Language (ISL), and Swedish Sign Language (SSL). The study provides comparative data to support the claim that two-handedness in uninflected signs is neither random nor meaningless phonological structure. The study compares 200 lexical items in the three sign languages and finds that the number of cases in which the same concepts were two handed in all three sign languages is highly significant: 33%, where random distribution would be 13%. This study finds that, although the two-handed signs for the same concepts are not necessarily identical in form across the three languages, they draw upon a similar kind of iconic mapping, a similar link between form and meaning. For example, the sign EMPTY is realized in ASL, ISL, and SSL with the non-dominant hand representing a surface or container, and the dominant hand acting upon it, to indicate bareness or lack of content (Figures 9 a-c). As is often the case, the signs are not phonologically identical in the three languages, but the fact that they are two-handed and unbalanced in all three languages is not random. Rather, these aspects of phonological form are motivated by meaning.

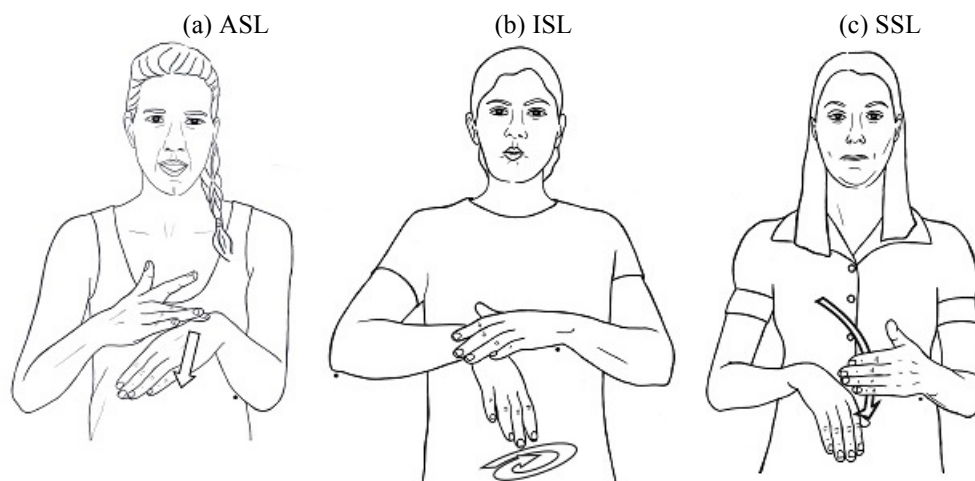


Figure 9: The sign EMPTY in (a) ASL (b) ISL and (c) SSL (from Lepic et al. 2016)

The iconic motivations that underlie the use of two hands in lexical signs typically encode particular types of relationships, shown in (1). Note that (a) and (b) represent relationships that obtain between individual entities, while (c) and (d) represent relationships between component parts of single entities.⁵

(1) Types of relations that motivate two-handedness

- a. *Interaction*: Paired, interacting entities are mapped onto each of the two hands
- b. *Location*: Paired entities and their locations are mapped onto each of the two hands
- c. *Dimension*: Boundaries of an entity's shape/volume are mapped onto the two hands
- d. *Composition*: Component parts of an entity are mapped onto the two hands

⁵ For details of the semantic analysis and illustrations of these relations, see Lepic et al. (2016).

It is not only two-handedness that is often motivated; whether two-handed signs are balanced or unbalanced is often motivated too, as we show in Section 3.2.

The observations in (1) lead to the conclusion that meaning is a key factor, though not the only factor, in predicting two-handedness, a characteristic that was usually treated as strictly phonological. Thus, the templates for two-handed signs are in fact not meaningless – the specification of two hands as a part of the template is itself often motivated by the meaning of the sign.

3.2 Motivation of balanced and unbalanced phonological templates

The two basic forms for two-handed templates correspond to the basic division of two-handed signs – balanced and unbalanced. We now go on to show the relationship between the form of these two templates and meaning.

The two-handed **balanced** template typically represents signs that encode symmetrical relations between similar entities (e.g., SSL MEET, shown in Figure 8a above). By doubly associating the hand configuration (HC) category to each hand, as well as to the same location-movement-location (LML) sequence, the template ensures that both hands have the same handshape and move together. It is known that there are constraints on two-handed signs (Battison 1978). Specifically, if the nondominant hand moves, then it must be configured with the same handshape as the dominant hand and execute similar movement. The phonological template is designed to capture these similarities (Sandler 1989). What now becomes clear is the fact that these similarities are also motivated by the symmetrical semantic relationship between two like entities, in the MEET example, between two people meeting one another.

The two-handed **unbalanced** template typically represents signs that encode an asymmetrical semantic relationship between entities, such as figure-ground (as defined in Talmy 2003; e.g., ISL EMPTY, shown in Figure 9b above; ISL END, Figure 8b). The relationship is clearly reflected in the template so that the non-dominant hand (hand2) represents the *ground* and, as a place of articulation, does not move, and the dominant hand (hand1) represents the *figure* and moves from one location to another. The fact that the nondominant hand does not move in such signs, behaving instead as a location, is motivated by the figure-ground relationship. The constraint on such signs (Battison 1978) is that the nondominant hand must be configured in one of a small number of unmarked handshapes. This too now follows from its status as a more general notion associated with the ground, vis à vis the figure enacting the event.

The form of these templates, we argue, is motivated by the semantics of the two-handed concepts. Are all of the rest of the features in each of these sign types (i.e., features of hand configuration, location, type of movement etc.) provided by the individual lexical representations, or are any of them motivated as well? Lepic et al. (2016) argue that some of them are in fact motivated too, and here we propose that they be included in motivated templates. We focus here on *type of movement* and *contact* in two-handed signs.

3.3. Motivated movement features in balanced signs

In balanced signs, both hands move, and the type of *movement* in such signs may also be specified in the template, in particular, whether the movement is synchronized or alternating. The hands move in a synchronized manner in (1a) and the signs typically represent a relationship denoting synchronized movement of entities, or a symmetrical shape of objects. For example, MEET (Figure 8a above) encodes an event in which two people are approaching each other together, and therefore the movement is synchronized. We repeat the example here, adding the feature [synchronized] (Figure 10).

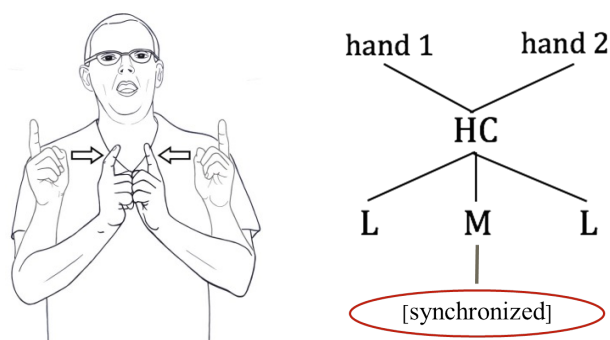


Figure 10: SSL MEET, balanced synchronized template

The feature [alternating movement] is represented in templates to encode a relationship in which one event/entity follows another repeatedly. For instance, in the ISL sign for NEGOTIATE (Figure 11) the hands represent two lines of communication, and the alternating movement shows that the conversation is going back and forth between the participants.

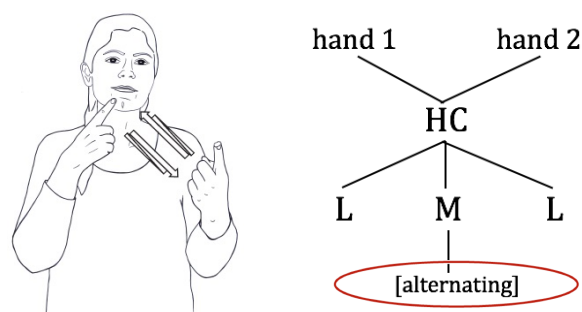


Figure 11: ISL NEGOTIATE, balanced alternating template

3.4. Motivated contact in balanced and unbalanced signs

The contact between the two hands in both balanced and unbalanced signs encodes the spatial relationship/interaction between entities in any of the four relationship types shown in example (1). The contact can be final, as in SSL MEET (Figure 10 above) or ISL END (Figure 8b above), or contact may be initial, as in the SSL sign for SEMESTER, in which the contact notes the beginning of the semester and the movement towards the final location represents the period of time of the semester. Figure 12 repeats the ISL sign END and Figure 13 shows SSL SEMESTER, with the [contact] feature entered into the templates.⁶ The point in the sign at which contact takes place is another example of motivated phonology.

⁶ The fact that the contact may occur on different segments is one of the theoretical motivations for the sequential structure of the model (Sandler 1989).

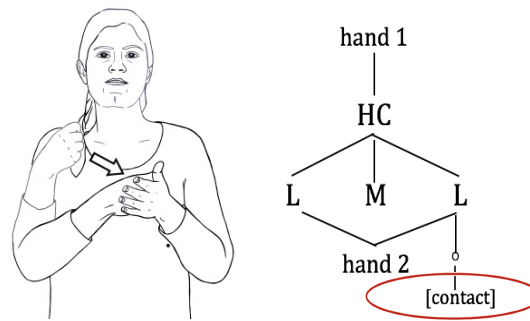


Figure 12: SSL END, unbalanced, final-contact template

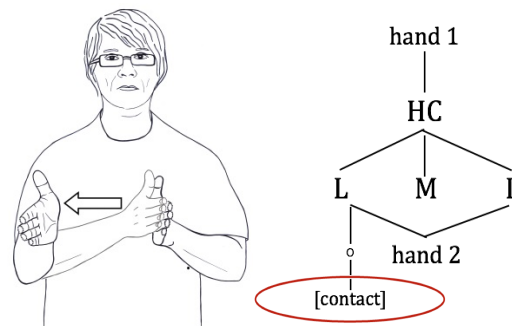


Figure 13: SSL SEMESTER, unbalanced initial-contact template

4. Conclusion: blurring the line between levels of structure

In Section 1, we showed that two-handedness has often been treated as a purely phonological characteristic of signs. The nondominant hand in signs is treated by the phonology as formational elements belonging to the categories of hand configuration or location, both in their specifications, and also in phonological processes, such as assimilation. In various derivations and inflections, the nondominant hand is also referred to without any reference to lexical meaning (see especially Klima and Bellugi 1979; Padden and Perlmutter 1987). However, Lepic et al. (2016) show in detail that two-handedness in a disproportionate percentage of signs in any sign language lexicon is accounted for by meaning. Other aspects of phonological form in these signs are motivated as well, such as the type of two-handed sign, balanced or unbalanced, whether the movement is symmetrical or alternating, and whether and where the dominant hand will contact the nondominant hand. We've made these findings explicit by showing which elements in the abstract templates that have been proposed for signs – previously considered strictly phonological and meaningless – are motivated. In so doing, we see that in sign languages, the line between the phonological and morpho-lexical levels of structure is not as sharp as often assumed. In fact, other aspects of phonological form, such as movement types and locations, are also motivated. Since it has been shown that each of these elements behaves phonologically as well, it is left for future models of sign language structure to incorporate this ambivalence in a theoretically satisfactory way.

In spoken languages, the acoustic-auditory medium limits the extent to which parts of words that are not morphemes can be motivated. At the same time, Bloomfield (1933) demonstrated that this does occur in so-called sound symbolism, such as *sn* in words like *sniff*, *snuff*, *snore*, *snort* all reflecting some kind of breath noise through the nose. In Japanese

mimetics, Hamano (1986) demonstrated that phonetic features correspond to meanings, for example, [-voice] for small/light/fine events, and [+continuant] for continuous movement or shapeless objects.

Sign languages merely show us the extent to which the duality of patterning in human language can be blurred. They also suggest that the paucity of motivation within the phonology of spoken language does not make it insignificant. It is just an accident of modality.

Acknowledgments

We are grateful to all of our consultants, particularly Debi Menashe and Tory Sampson, for sharing their languages with us. Debi Menashe created the sign illustrations presented here. We thank Mark Aronoff, Irit Meir, Hope Morgan, Corrine Occhino-Kehoe, and Carol Padden for their helpful and encouraging comments on this paper. We thank the people who attended the Theoretical Issues in Sign Language Research meeting in London on July 12, 2013, for constructive discussion and criticisms, and participants in the Mediterranean Morphology Meeting in Haifa on September 10, 2015 for valuable questions and suggestions. This work was funded by Israel Science Foundation grant 580/09, National Institutes of Health grant R01 DC006473, and European Research Council Grant 340140.

References

- Aronoff M., I. Meir & W. Sandler (2005) The paradox of sign language morphology. *Language* 81: 301-344.
- Battison, R. (1978) *Lexical borrowing in American Sign Language*. Silver Spring, MD: Linstok Press.
- Bergman, B. (1979) *Signed Swedish*. Stockholm: Liber.
- Bergman B. (1983) Verbs and adjectives: Morphological processes in Swedish Sign Language. In: J. Kyle & B. Woll (Eds.), *Language in sign: An international perspective on sign language*. London: Croom Helm, 3-9.
- Blevins J. (1993) The nature of constraints on the non dominant hand in ASL. In: G. Coulter (Ed.), *Current issues in ASL phonology: Phonetics and Phonology 3*. New York, NY & San Francisco, CA: Academic Press, 43-62.
- Bloomfield, L. (1933) *Language*. New York: Henry Holt.
- Börstell C. (2011) *Revisiting Reduplication: Toward a description of reduplication in predicative signs in Swedish Sign Language*. M.A. thesis. Department of Linguistics, Stockholm University.
- Brennan, M. (1990) *Word formation in British Sign Language*. Stockholm: Stockholm University.
- Brentari, D. (1998) *A prosodic model of sign language phonology*. Cambridge, MA: MIT Press.
- Brentari D. & J.A. Goldsmith (1993) Secondary licensing and the nondominant hand in ASL phonology. In: G. Coulter (Ed.), *Current issues in ASL phonology: Phonetics and Phonology 3*. New York, NY & San Francisco, CA: Academic Press, 19-41.
- Clements, G.N. & S. Keyser (1983) *CV Phonology: A generative theory of the syllable*. Cambridge, MA: MIT Press.
- Crasborn O. (2011) The other hand in sign language phonology. In: M. van Oostendorp, C.J. Ewen, E. Hume & K. Rice (Eds.), *The Blackwell companion to phonology*, Vol. 1. Malden, MA & Oxford, 223-240.
- Emmorey K. (Ed.) (2003) *Perspectives on classifier constructions in sign languages*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Fischer S. & B. Gough (1978) Verbs in American Sign Language. *Sign Language Studies* 7: 17-48.
- Goldsmith J. (1976) *Autosegmental phonology*. Ph.D. thesis. MIT.
- Hamano S.S. (1986). *The sound symbolic system of Japanese*. Ph.D. thesis. University of Florida.
- Hockett C. F. (1960) The origin of speech. *Scientific American* 203: 88-96.
- Hulst H. van der (1993) Units in the analysis of signs. *Phonology* 10: 209-241.
- Hulst H. van der (1996) On the other hand. *Lingua* 98: 121-143.
- Johnston T. & A. Schembri (1999) On defining lexeme in a signed language. *Sign Language & Linguistics* 2(2): 115-185. doi.org/10.1075/sll.2.2.03joh
- Klima, E.S. & U. Bellugi (1979) *The signs of language*. Cambridge, MA: Harvard University Press.
- Kooij, E. van der (2002) *Phonological categories in Sign Language of the Netherlands: The role of phonetic implementation and iconicity*. Utrecht, the Netherlands: LOT.
- Kyle J.G. & B. Woll (Eds.) (1985) *Sign language: The study of deaf people and their language*. Cambridge: Cambridge University Press.
- Lepic R., C. Börstell, G. Belsitzman & W. Sandler (2016) Taking meaning in hand: Iconic motivations for two-handed signs. *Sign Language & Linguistics*, 19(1).

- Liddell S.K. (1984) THINK and BELIEVE: Sequentiality in American Sign Language. *Language* 60: 372-99.
- Liddell, S.K. (2003) *Grammar, gesture, and meaning in American Sign Language*. Cambridge: Cambridge University Press.
- Liddell S.K. & E.R. Johnson (1986) American Sign Language compound formation processes, lexicalization, and phonological remnants. *Natural Language and Linguistic Theory* 8: 445-513.
- Liddell S.K. & E.R. Johnson (1989) American Sign Language: The phonological base. *Sign Language Studies* 64: 197-277.
- McCarthy J. (1979) *Formal Problems in Semitic Phonology and Morphology*. Ph.D. thesis. MIT.
- McCarthy J. (1981) A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* 12: 373-418.
- Meir I. (2002) A cross-modality perspective on verb agreement. *Natural Language & Linguistic Theory* 20: 413-450.
- Meir I., C. Padden, M. Aronoff & W. Sandler (2013) Competing iconicities in the structure of languages. *Cognitive Linguistics* 24(2): 309-343. doi:10.1515/cog-2013-0010.
- Meir, I. & W. Sandler (2008) *A Language in space: The story of Israeli Sign Language*. New York, NY & London: Lawrence Erlbaum Associates.
- Nespor M. & W. Sandler (1999) Prosody in Israeli Sign Language. *Language and Speech* 42(2-3): 143-176. doi:10.1177/00238309990420020201.
- Newkirk D. (1979) *The form of continuative aspect inflection on ASL verbs*. Ms.
- Newkirk D. (1981) *Rhythmic features of inflections in American Sign Language*. Ms.
- Nilsson A.L. (2007) The non-dominant hand in a Swedish Sign Language discourse. In: M. Vermeerbergen, L. Leeson, & O. Crasborn (Eds.), *Simultaneity in signed languages: Form and function*. Current Issues in Linguistic Theory Vol 281. Amsterdam/Philadelphia, PA: John Benjamins, 163-185.
- Padden, C. (1988) *Interaction of morphology and syntax in American Sign Language*. New York, NY & London: Garland Publishing, Inc.
- Padden C. & D.M. Perlmutter (1987) American Sign Language and the architecture of phonological theory. *Natural Language & Linguistic Theory* 5: 335-375.
- Perniss P., R.L. Thompson & G. Vigliocco (2010) Iconicity as a general property of language: evidence from spoken and signed languages. *Frontiers in psychology* 1: 227. doi:10.3389/fpsyg.2010.00227.
- Pfau R. & M. Steinbach (2003) Optimal reciprocals in German Sign Language. *Sign Language & Linguistics* 6(1): 3-42.
- Pfau R. & M. Steinbach (2006) Pluralization in sign and in speech: A cross-modal typological study. *Linguistic Typology* 10(2): 135-182. doi.org/10.1515/LINGTY.2006.006
- Sandler W. (1986) The spreading hand autosegment of American Sign Language. *Sign Language Studies* 50: 1-28.
- Sandler W. (1987) Assimilation and feature hierarchy in American Sign Language. In: *Papers from the Chicago Linguistic Society, Parasession on Autosegmental and Metrical Phonology*. Chicago: Chicago Linguistic Society, 266-278.
- Sandler, W. (1989) *Phonological representation of the sign: Linearity and nonlinearity in American sign language*. Dordrecht: Foris.
- Sandler W. (1990) Temporal aspect and American Sign Language. In: S. Fischer & P. Siple (Eds.), *Theoretical Issues in Sign Language Research*. Chicago: University of Chicago Press, 103-129.
- Sandler W. (1993) Hand in hand: The roles of the nondominant hand in sign language phonology. *The Linguistic Review* 10: 337-390.
- Sandler W. (1999a) Cliticization and prosodic words in a sign language. In: T. Hall & U. Kleinhenz (Eds.), *Studies on the phonological word*. Amsterdam: John Benjamins, 223-255.
- Sandler W. (1999b) Prosody in two natural language modalities. *Language and Speech* 42(2-3): 127-142. doi:10.1177/00238309990420020101.
- Sandler W. (2006) Phonology, phonetics, and the nondominant hand. In: L. Goldstein, D.H. Whalen & C. Best (Eds.), *Papers in laboratory phonology: Varieties of phonological competence*. Berlin: Mouton de Gruyter, 185-212.
- Sandler W. (2012) The phonological organization of sign languages. *Language and Linguistics Compass* 6(3): 162-182. doi:10.1002/lnc3.326.
- Sandler, W. & D. Lillo-Martin (2006) *Sign language and linguistic universals*. Cambridge: Cambridge University Press.
- Saussure, F. de (1983) *Course in General Linguistics*, trans. by Harris, R. Open Court Classics, Chicago, IL.
- Stokoe W.C. (1960) *Sign language structure: An outline of the visual communication system of the American Deaf*. Studies in linguistics: Occasional papers 8. Buffalo, NY: Dept. of Anthropology and Linguistics, University of Buffalo.

- Strickland, B., Geraci, C., Chemla, E., Schlenker, P., Kelepir, M., & Pfau, R. (2015). Event representations constrain the structure of language: Sign language as a window into universally accessible linguistic biases. *Proceedings of the National Academy of Sciences* 112 (19): 5968-5973.
- Supalla T. (1982) *Structure and acquisition of verbs of motion and location in American Sign Language*. Ph.D. thesis. University of California, San Diego.
- Talmy L. (2003) The representation of spatial structure in spoken and signed language. In: K. Emmorey (Ed.), *Perspectives on classifier constructions in sign languages*. Mahwah, NJ: Lawrence Erlbaum, 169-195.
- Taub, S. (2001) *Language from the body: Iconicity and metaphor in ASL*. Cambridge: Cambridge University Press.
- Vermeerbergen M., L. Leeson & O. Crasborn (Eds.) (2007) *Simultaneity in signed languages: Form and function*. Amsterdam/Philadelphia, PA: John Benjamins.
- Wilbur R.B. (2008) Complex predicates involving events, time and aspect: Is this why sign languages look so similar? In: J. Quer (Ed.), *Signs of the time: Selected papers from TISLR 2004*. Hamburg: Signum Verlag, 217-250.
- Wilcox, S. (2004) Cognitive iconicity: Conceptual spaces, meaning, and gesture in signed languages. *Cognitive Linguistics* 15(2): 119-147.

The morphosyntax of definiteness agreement in Neo-Aramaic and Central Semitic

Edit Doron
Hebrew University of Jerusalem
edit.doron@mail.huji.ac.il

Geoffrey Khan
University of Cambridge
gk101@cam.ac.uk

1. Introduction

The question we seek to answer in this paper is: How did the multiple marking of definiteness within the noun phrase develop in Central Semitic? We propose an answer based on the study of Neo-Aramaic, a modern Central Semitic language, and in particular on the process by which the definite article developed in Neo-Aramaic on the basis of its demonstrative pronouns. We suggest that the development in ancient Central Semitic could have paralleled the one in Neo-Aramaic.

We thus argue (contra Pat-El 2009) that definiteness in Semitic originates like in other languages as part of Greenberg's 1978 "demonstrative cycle" of grammaticalizing the demonstrative pronoun as a definite article (cf. Lyons 1999, Gelderen 2007, 2011):

Greenberg's demonstrative cycle

stage I	stage II	stage III	stage IV
demonstrative pronoun	> definite article	> marker of argumenthood	> class marker

2. Multiple marking of definiteness

Languages of the world often mark definiteness on the noun or the determiner (Dryer 2013), but multiple marking of definiteness is less common, and, in particular, definiteness marking of the attributive adjective is not common. One sub-group of Semitic languages, Central Semitic, which includes Arabic, Hebrew, and Aramaic, systematically marks definiteness in a structure with a definite affixal article (DEF), often reconstructed as **hal* or **han*, attached both to the noun and the adjective, either as a prefix, as in (1a), in Classical Arabic and Biblical Hebrew, or a suffix as in (1b), in Classical Aramaic. Indefiniteness is marked in these languages by omission of the article.

- (1) a. prefixal article
[DEF-N DEF-Adj]
b. suffixal article
[N-DEF Adj-DEF]

An example is given for each of the three languages:

- (2) a. prefixal article
Classical Arabic
ʔal-ʔarḍ ʔal-muqaddas-a
DEF-land.F DEF-holy-F
'the holy land'

Biblical Hebrew¹*hā-ʔāreš haq-qadoš-ā*

DEF-land.F DEF-holy-F

‘the holy land’

b. suffixal article

Classical Aramaic*ʔarʕ-ā qaddiš-t-ā*

land.F-DEF holy-F-DEF

‘the holy land’

Below we investigate the diachronic question of the origin of multiple definite articles. This is a question posed for other languages as well, in addition to Semitic languages, which have multiple marking of definiteness, such as German, Yiddish, Norwegian, Swedish, Faroese, Greek, Albanian, Romanian, Bulgarian, and colloquial Slovenian. In French, there is double marking of definiteness in the expression of superlatives (e.g. the double occurrence of the article *la* in the superlative phrase *la terre la plus sainte* ‘the holiest land’).²

Researchers have proposed different accounts for the multiple marking of definiteness. These accounts can be roughly divided into two different types: those which view the multiple marking as representing multiple syntactic phrases (as shown in diagrams (3ai) and (3aii) below), and those which view it as multiple marking of a single phrase (as shown in diagram (3b)).

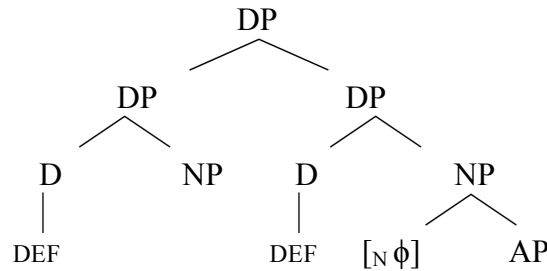
The accounts which view the multiple marking of definiteness as involving multiple nominal phrases come in two variants (3ai vs. 3aii). According to the first variant, notably Lekakou and Szenderői 2012, DEF realizes the syntactic functional head D (determiner). Hence, multiple marking of definiteness involves the multiple occurrence of the syntactic category D. Since D is considered the head of DP (the nominal phrase), the occurrence of multiple D’s reflects the occurrence of multiple DP’s. In other words, a multiple marked DP is actually a complex DP whose daughters are DP’s themselves. The semantic relation between the daughter DP’s is that of close apposition. Moreover, in one of the daughter DP’s, the adjective modifies a null noun. This variant of the multiple-phrase account is shown in (3ai).

¹ On the article in Modern Hebrew see Doron and Meir 2013, 2016.

² The marking of the adjective in the Germanic languages is actually a *weak-strong* marking, and might be unrelated to the definite article marking adjectives in the other languages on this list. Another difference within the list has to do with the obligatoriness vs. optionality of the multiple marking, e.g. the Hebrew *ha-mazon ha-bari ve *(ha-)ta'im* vs. the French *la plus saine et (la) plus délicieuse nourriture*. The languages also differ in which nominal components may be marked for definiteness. For example, in addition to marking nouns and adjectives, as in (i.a) below, Bulgarian marks numerals (i.b) and possessors (i.c) as well:

- (i) a. *xubavata sestra na domakinjata*
beautiful.F.DEF sister.DEF of hostess.DEF
‘the beautiful sister of the hostess’ (Mladenova 2007:30)
- b. *drugite dvete devojki*
other.PL.DEF two.F.DEF girls
‘the other two girls’ (Mladenova 2007:26)
- c. *Naš'te starite dojdoxa.*
our.PL.DEF old.PL.DEF came.3PL
‘Our parents came.’ (Mladenova 2007:45)

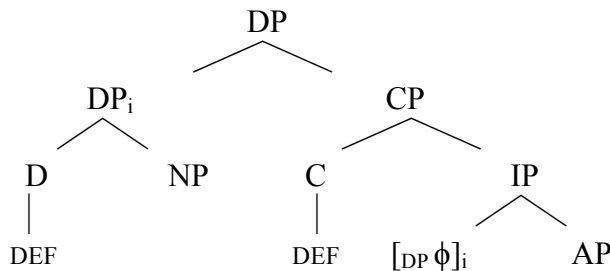
- (3) a. i. multiple-phrase account (DEF realizes a syntactic functional head)
two DPs in close apposition where DEF realizes D



(Lekakou and Szenderői 2012 for Greek)

According to the second variant of the multiple-phrase account, (e.g. Alexiadou and Wilder 1998), DEF realizes either of two syntactic functional heads: D (determiner) or C (complementizer). Hence, the complex DP is viewed as consisting of a DP modified by a relative clause CP. The semantic relation between the daughter DP and CP is that of relative-clause modification. This variant of the multiple-phrase account is shown in (3aii):

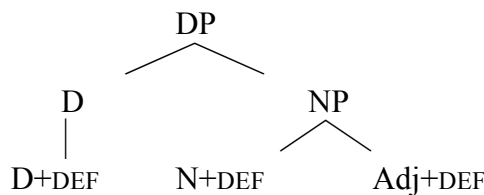
- (3) a. ii. multiple-phrase account (DEF realizes a syntactic functional head)
DP modified by a relative clause where DEF realizes the complementizer C



(Alexiadou and Wilder 1998, Alexiadou 2014 for Greek;
Khan 2008 for Neo-Aramaic)

The second type of account views the multiple marking of definiteness as multiple marking within a single phrase. According to these accounts, DEF is the exponent of definiteness inflection which inflects the various syntactic categories of the DP. In one variant of these accounts, the categories within DP include not only N and Adj but also D (e.g. Delsing 1993). DEF inflection spreads from N to Adj and D. This is shown in (3bi). A second subtype of these accounts only recognizes the categories N and Adj within the nominal phrase (3bii). Such an account is that of Pat-El 2009.

- (3) b. single-phrase account (DEF is an affix attached to syntactic heads)



- i. DEF spreads from N to D and Adj
(Delsing 1993 and most other analyses of Scandinavian)

ii. DEF spreads from Adj to N
(Pat-El 2009 for Semitic)

We follow a version of the single-phrase account (3b), but we would also like to explain how DEF developed into a marker of agreement, assuming the received view that DEF is originally a demonstrative pronoun, an independent phrasal element DP_{DEM} , which, in the course of historical development, was reanalyzed as a D head. We will show how this paved the way to the double attachment of DEF to both N and Adj. The fact that it marks agreement between N and Adj is due to the fact that all its occurrences express the features of a single phrase.

Our view of the single-phrase account for Semitic differs from Pat-El's 2009 account. Pat-El does not share the assumption that DEF in Semitic originates in a demonstrative pronoun. Her arguments against the identification of the definite article as a historical demonstrative include the following. (i) The normal order in Semitic is N-DEM, whereas the article is often placed before the head noun: DEF-N. (ii) The morphological exponent of the definite article in the historically attested Semitic languages, i.e. *ha-*, *han-* (or phonetic variants), does not correspond to any form of an attested independent demonstrative pronoun, rather only to an element that is a deictic prefix to such a pronoun, e.g. Arabic *ha-da* 'look.here-DEM' (= 'this'). (iii) The article exhibits no inflection for gender or number, whereas such inflection is present in paradigms of demonstrative pronouns. Pat-El argues, therefore, that the article began as a deictic/presentative prefix with the form *ha-* or *han-* (cognate to *hinne* in Hebrew), which was used adnominally to nominalize an adjective or mark it as attributive. The article on an attributive adjective then spread to the noun head, e.g.:

- (4) Reconstruction (Pat-El 2009:43)
- a. *han-tāb*
look.here-good 'the good one'
 - b. *kalb han-tāb* > *han-kalb han-tāb*
dog look.here-good 'look.here-dog look.here-good' ('the good dog')

Assuming that the article originates on adjectives rather than nouns allows Pat-El to account for the suffixal nature of DEF in Aramaic as a case of rebracketing:

- (5) N [han Adj] > [N han] Adj

One problem though is that under Pat-El's account, Semitic is different from general language typology in the origin of its definite article. Second, presentative particles are strictly deictic and lack the anaphoric function which is a crucial ingredient of definiteness. We therefore stick to the received view (Rubin 2005), that definiteness in Semitic originates as in other languages, as part of Greenberg's cycle. Greenberg describes the transition to the definite article in terms of the demonstrative pronouns becoming "bleached of deixis by anaphoric uses" (Greenberg 1978:79). The formal properties of the Central Semitic article that Pat-El adduces as arguments against its demonstrative origin, i.e., that it resembles prefixes of attested demonstrative pronouns and that it does not inflect, can be interpreted as the result of structural attrition as a result of grammaticalization. As for her argument relating to the normal syntactic ordering of the demonstrative relative to its head noun in Semitic, it should be pointed out that the ordering in fact exhibits considerable flexibility across the Semitic languages. Most relevantly, in Neo-Aramaic where, as we shall argue, the Greenberg definiteness cycle is taking place (and has been completed in one particular dialect), the demonstrative in question is placed before the head noun.

3. Deictic vs. anaphoric demonstrative pronouns in Neo-Aramaic

In Neo-Aramaic, demonstrative pronouns have a deictic or an anaphoric function.³ Many dialects have distinct forms of the demonstrative pronoun for each of these functions. We discuss two dialects, Barwar (Khan 2008) and Ṭuroyo (Waltisberg 2014).

A well known characteristic of deictic demonstratives, in Neo-Aramaic as in other languages, is that they encode the proximal/ distal contrast. Anaphoric demonstratives do not encode this contrast:

- (6) a. Deictic demonstratives
- | | |
|---------------------|----------------------|
| <u>Barwar</u> | |
| ' <i>awwa kθawa</i> | ' <i>awaha kθawa</i> |
| this book | that book |
| <u>Ṭuroyo</u> | |
| ' <i>u-kθow-ano</i> | ' <i>u-kθow-awo</i> |
| the-book-this | the-book-that |
- b. Anaphoric demonstratives
- | | |
|------------------|--|
| <u>Barwar</u> | |
| ' <i>o-kθawa</i> | |
| 'that/the book' | |
| <u>Ṭuroyo</u> | |
| ' <i>u-kθowo</i> | |
| 'the book' | |

The demonstratives which developed into the definite article are not the deictic but the anaphoric demonstratives. In Barwar, the anaphoric demonstrative '*aw* functions as an embryonic article, typically in clitic form ('*o-*). It is different from the deictic '*awwa*, which has developed historically from attaching the deictic particle (*h*)*a* to '*aw*. In Ṭuroyo, the anaphoric demonstrative **hu* has made the full shift to the status of definite article in its clitic form '*u-*. The independent form of **hu* became *hiye* by the addition of the 3MS suffix *-e* (*hu-e* > *hiye*). '*aw* and *hiye* are anaphoric demonstratives, they function as personal pronouns in particular environments. In most environments, personal pronouns are null in Neo-Aramaic. Continuing topics are generally tracked by null anaphors. Anaphoric demonstratives are predominantly used to track topics that are discontinuative or contrastive (Diessel 1999).

4. Barwar vs. Ṭuroyo: different stages of the demonstrative cycle

In Barwar, the article '*o-* has not yet shifted to the status of definite determiner; that is, Barwar has not yet fully shifted to stage II in Greenberg's cycle. The article is only used to mark *pragmatic* but not *semantic* definiteness (in the terminology of Löbner 1985): the article marks individuals as being unique in the context, e.g. *house* in a context which happens to include a unique house, but does not mark individuals which are unique independently of the the context, i.e. by virtue of their meaning, such as *king*, *sun*, *nose*, *evening*, etc. In Ṭuroyo, '*u-* has already grammaticalized into a determiner:

³ We use the term *anaphoric* to include reference to entities which have been made prominent in the particular discourse in any way, not necessarily by previous mention. Strictly speaking the term is *endophoric*.

- (7) a. Barwar
xoni bnele beθa
 brother.my built house
'o-beθa qurba l-beθa-t malka
the-house near to-house-of king
 'My brother built a house. The house is near the house of the king.'
- b. Turoyo
aḥuni ma 'marle bayto
 brother.my built house
'u-bayto qariwo-yo l-u-bayto d-u-malko
the-house near-COP to-the-house of-the-king
 'My brother built a house. The house is near the house of the king.'

In Barwar, names of kinds do not take the article, which further indicates that the article is still a demonstrative phrase rather than a determiner (Krámsky 1972: 34), but in Turoyo they obligatorily do:

- (8) a. Barwar
 (*'o-) 'arya
 b. Turoyo
 *('u-) 'aryo
 both: 'the lion' (as a kind-name)

We conclude that 'o- in Barwar (and its fem. and pl. counterparts, 'a- and 'an- respectively) is still a phrasal constituent, a demonstrative DP_{DEM}, whereas 'u- in Turoyo (and its fem. and pl. counterparts 'i- and 'a(nn)- respectively) is reinterpreted as D, which moreover is realized as an affix to N. Accordingly, in Barwar, the definite article may be attached to the left of a conjoined noun-phrase, whereas in Turoyo it must be attached to each noun separately:

- (9) a. Barwar
xzayəl-la 'a-yaləxta -w 'isaqθa
 seeing.3MS-OBJ.3PL the.FS-scarf and ring
 'He sees the scarf and ring' (Khan 2008 III A26:9)
- b. Turoyo
hule-la 'i-dasmale 'u 'i-'isqaθo
 gave.3MS-DAT.3FS the.FS-scarf and the.FS-ring
 'He gave her the scarf and the ring.'

In Barwar, 'o- and 'awwa do not co-occur in a single noun phrase, since both are demonstrative phrases. But in Turoyo, the D 'u- cooccurs with the demonstrative DP_{DEM} 'awo:

- (10) a. Barwar
 * 'awaha 'o kθawa
 'that book'
- b. Turoyo
 'u-flaym-awo
 the-boy-that
 'that boy'

Conversely, in Turoyo the D 'u- cannot be modified by a reduced relative clause (introduced by the complementizer -d), whereas the Barwar demonstrative DP_{DEM} 'o- may occur as the head of reduced relative clauses introduced by the complementizer -t:

- (11) a. Barwar
 'o- -t gu-bεθa
 the C in-house
 'the one in the house'
 b. Turoyo
 *'u- -d b-u-bayto
 the C in-the-house

5. Contrastive vs. non-contrastive attributive adjectives

Neo-Aramaic can overtly express contrastive attribution by marking the adjective with the definite article. In Barwar, definiteness marking of the adjective precludes marking of the head-noun, since the article is a demonstrative DP_{DEM} which can only be attached once per noun phrase:

- (12) Barwar
 a. xone diye faqira wewa
 brother of.3MS poor PAST
 šattā-t maθa wewa tiwa 'o-xona faqira
 bottom-of village PAST lived **the-brother poor**
 'His brother was poor... The poor brother lived at the bottom of the village.'
 (Khan 2008 vol 3, A25:1)
 b. 'aw dmixεle xona 'o-gora modi mære
 he slept **brother the-big** what said
 'While he (the youngest brother) slept, what did the eldest brother say?'
 [contrastive]
 (Khan 2008 vol 3, A24:25)

But in Turoyo, the article obligatorily marks the noun in definite phrases, whether or not the adjective is marked as contrastive:

- (13) Turoyo
 a. g-'oyašno b-u-bayto na'im-ano / b-u-bayt-ano na'imo
 FUT-live.1S in-**the-house small-this** in-**the-house-this small**
 'I shall live in this small house.'
 b. 'ono g-'oyašno b-u-bayt-ano 'u-na'imo
 I FUT-live.1S in-**the-house-this the-small**
 hat 'uš b-u-bayt-awo 'u-rabo
 you live.IMP.2S in-**the-house-that the-big**
 'I shall live in this small house. You live in that big house.' [contrastive]

Crucially, contrastive marking of the adjective in Turoyo is only possible in the environment of a demonstrative phrase, as in (13b) above. When the demonstrative phrase is not present, the adjective is not marked by the definite article, cf. (14a), and contrastive marking of the adjective is ungrammatical. In (14b), the modifier *the eldest* can only be interpreted as loose apposition, i.e. *the eldest one*, which is incompatible with contrast:

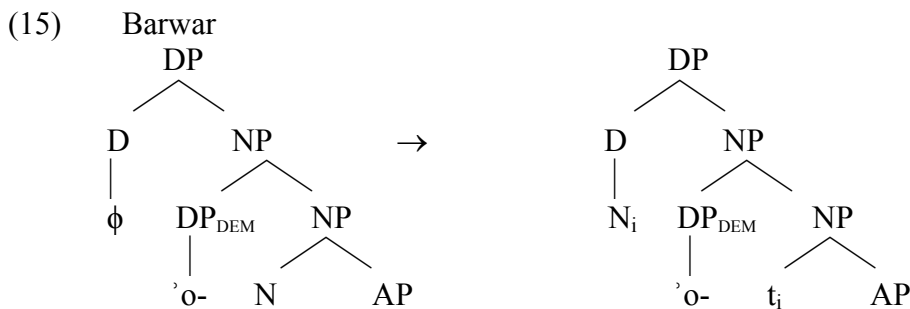
(14) Ṭuroyo

- a. *inaqa d- u-aḥuno na 'imo daməx 'u-aḥuno rabo mən malle*
 while that the-brother young slept **the-brother big** what said
 'While the youngest brother slept, what did the eldest brother say?'
- b. **inaqa d- u-aḥuno na 'imo daməx 'u-aḥuno 'u-rabo mən malle*
 while that the-brother young slept **the-brother the-big** what said
 * 'While the youngest brother slept, what did the brother, the eldest one, say?'

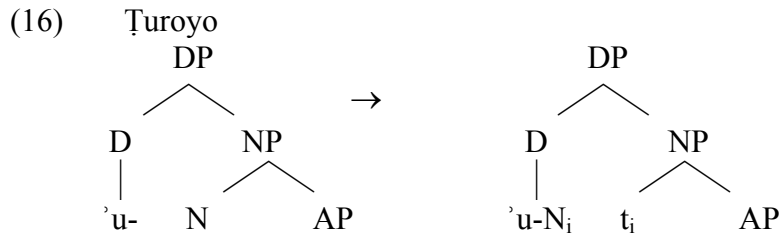
6. The evolution of multiple definiteness marking

In the previous sections we uncovered two crucial factors of the development of multiple definiteness marking in the Neo-Aramaic noun phrase. One factor is the syntactic status of the definite article. Is it a phrasal constituent DP_{DEM} or a lexical head D? In section 4, we showed that the transition from demonstrative pronoun, as in Barwar, to definite determiner, as in Ṭuroyo, corresponds to reanalysis of the phrase DP_{DEM} as the lexical head D. This corresponds to the parallel reanalysis suggested for the Latin demonstrative *ille* by Giusti 2001 and Roberts and Roussou 2003: 131-136. The second factor, discussed in section 5, is the use of the definite article in Neo-Aramaic to mark contrastivity of the attributive adjective. We derive this marking from the reordering of the noun N, or some (extended) projection of N, relative to the determiner D within the noun-phrase. Underlyingly, the noun N intervenes between the determiner D and the adjective Adj. If the noun is raised out of its underlying position, the stranded Adj remains adjacent to D, with no intervening material, resulting in the definite article attaching to Adj. Semantically, the raising of N achieves de-focalization of the noun, and hence contrastive interpretation of the stranded adjective.

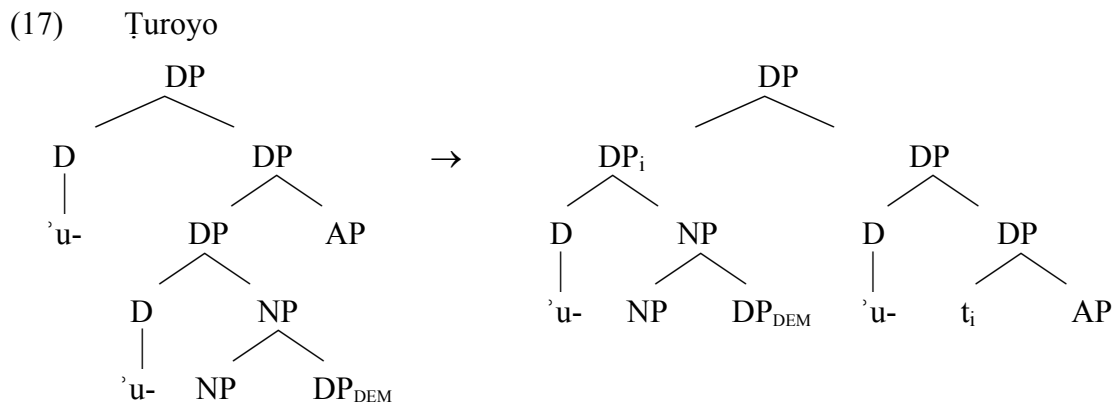
The interaction of these two factors is at the basis of the development of multiple definiteness marking within the noun-phrase in Neo-Aramaic. The simpler case is Barwar, where the article is still a demonstrative DP_{DEM} , and the determiner D is null. If N raises to the null D, semantically marking the attributive AP as contrastive, it allows the attachment of the article 'o- to the AP:



In Ṭuroyo, the article is a D head. Raising of N in simple noun phrases without demonstrative phrases, such as in (14) above, does not alter the relative order of N and the article, and thus does not result in contrastive marking:



In noun phrases containing a demonstrative DP_{DEM} , such as example (13) above, what makes AP contrastive is the de-focalization of an extended projection of the noun N, i.e. the constituent $[_{DP} D [_{NP} NP DP_{DEM}]]$. Raising this constituent allows the phonological attachment of D to AP:



7. Conclusion

We have shown how the Neo-Aramaic dialects progress along the transition from the first to the second stage of Greenberg’s cycle, where demonstrative pronouns turn into definite articles. In the less progressive dialect, the article is still an anaphoric demonstrative pronoun which has not yet turned into a determiner. Depending on whether the attributive adjective is contrastive or not, the article attaches either to the adjective or the noun, never to both. In the more progressive dialect, the anaphoric demonstrative pronoun has already turned into a determiner, and it may attach to both noun and adjective within the same noun phrase. Assuming that the ancient Central Semitic development might have followed the same path that we uncovered in Neo-Aramaic, our analysis suggests how the double marking of definiteness might have come about.

Bibliography

- Alexiadou, A. (2014) *Multiple Determiners and the Structure of DP*. Amsterdam: John Benjamins.
- Alexiadou, A. & C. Wilder (1998) Adjectival modification and multiple determiners. In: A. Alexiadou & C. Wilder (Eds.), *Possessors, Predicates and Movement in the Determiner Phrase*. Amsterdam: John Benjamins, 303–332.
- Delsing, L.O. (1993) *The Internal Structure of Noun Phrases in the Scandinavian Languages: A Comparative Study*. Ph.D. thesis. University of Lund.
- Diessel, H. (1999) *Demonstratives: Form, Function, and Grammaticalization*. Amsterdam: John Benjamins.
- Doron, E. & I. Meir (2013) Amount Definites. *Recherches Linguistiques de Vincennes* 42: 139-165.
- Doron, E. & I. Meir (2016) The Impact of Contact Languages on the Degrammaticalization of the Hebrew Definite Article. In: E. Doron (Ed.), *Language Contact and the Development of Modern Hebrew*. Leiden: Brill, 281-297.

- Dryer, M.S. (2013) Definite Articles. In: M.S. Dryer, & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available at: <http://wals.info/chapter/37>.
- van Gelderen, E. (2007) The Definiteness Cycle in Germanic. *Journal of Germanic Linguistics* 19: 275–308.
- van Gelderen, E. (2011) *The Linguistic Cycle: Language Change and the Language Faculty*. Oxford: Oxford University Press.
- Giusti, G. (2001) The birth of a functional category: from Latin *ille* to the Romance article and personal pronoun. In: G. Cinque & G. Salvi (Eds.), *Current Studies in Italian Syntax: Essays Offered to Lorenzo Renzi*. Amsterdam: North Holland, 157-171.
- Greenberg, J.H. (1978) How does a language acquire gender markers? In: J.H. Greenberg, C.A. Ferguson & E.A. Moravcsik (Eds.), *Universals of Human Language, Volume 3: Word Structure*. Stanford: Stanford University Press, 47–82.
- Khan, G. (2008) *The Neo-Aramaic Dialect of Barwar*. Leiden: Brill.
- Krámský, J. (1972) *The Article and the Concept of Definiteness in Languages*. The Hague: Mouton.
- Lekakou, M. & K. Szenderői (2012) Polydefinites in Greek: ellipsis, close apposition and expletive determiners. *Journal of Linguistics* 48: 107-149.
- Löbner, S. (1985) Definites. *Journal of Semantics* 4: 279–326.
- Lyons, C. (1999) *Definiteness*. Cambridge: Cambridge University Press.
- Mladenova, O.M. (2007) *Definiteness in Bulgarian: Modelling the Processes of Language Change*. Berlin: Mouton de Gruyter.
- Pat-El, N. (2009) The Development of the Semitic Definite Article: A Syntactic Approach. *Journal of Semitic Studies* 54:19–50.
- Roberts, I. & A. Roussou (2003) *Syntactic Change: A Minimalist Approach to Grammaticalization*. Cambridge: Cambridge University Press.
- Rubin, A.D. (2005) *Studies in Semitic Grammaticalization. Harvard Semitic Series*. Winona Lake, IN: Eisenbrauns.
- Vincent, N. (1997) The emergence of the D-system in Romance. In: A. van Kemenade & N. Vincent (Eds.), *Parameters of Morphosyntactic Change*. Cambridge: Cambridge University Press, 149–169.
- Waltisberg, M. (2014) *Syntax des Tuoyo*. Habilitationsschrift. Philipps-Universität, Marburg.

Telicity makes or breaks verb serialization

Kazuhiko Fukushima
Kansai Gaidai University
 kaz.zisin@gmail.com

1. Introduction: Lexical V-V compounds and telicity

Lexical V-V compounds have attracted much theoretical attention (inter alia, Kageyama 1993; Matsumoto 1996; Nishiyama 1998; Himeno 1999; Fukushima 2005/2007; Yumoto 2005). However, much research addresses the matters of “argument-synthesis”, i.e. how arguments of component verbs are/are not inherited or realized in the argument structure of the compound. Some non-exhaustive examples of such compounds are given in (1) reflecting a descriptive type classification such as cause, manner, etc.¹ For example, the subject arguments of *odor* ‘dance’ and *tukare* ‘get.tired’ are matched and inherited as the subject argument of the compound *odori-tukar* in (1a).

(Im)possible patterns of argument-synthesis have been identified and have given accounts of various sorts which are not reviewed here.² Anticipating the exposition below, I indicate the aspectual types of component verbs in the examples here and below. A standard telicity test (cf. Vendler 1967 and Dowty 1979), like using duration expressions like *itizikan* ‘for an hour’ and *itizikan-de* ‘in an hour’ suffices to determine (unmarked) telicity of verbs. For example, the adverbs give different results regarding telicity: *itizikan*/**itizikan-de odor* ‘dance for an hour/in an hour’ or *itizikan-de*/**itizikan tukare* ‘become tired in an hour/for an hour’. Also, simply “telicity” is employed here due to (i) stative and activity being atelic and (ii) achievement and accomplishment being telic. Actually, the absence of an incremental theme (or VP, a crucial factor for accomplishment), which is unavailable for lexical word-formation, renders accomplishment rather irrelevant to lexical word formation (see note 5).³

(1) a. Cause/resultative compounds:

Hanako-ga odori-tukare-ta. (atelic-telic)
 Hanako-NOM dance-get.tired-PAST
 ‘Hanako got tired from dancing.’
obore-sin ‘drown(telic)-die(telic), i.e. die from drowning’,
sini-tae ‘die(telic)-get.extinct(telic), i.e. become extinct by dying’, etc.

b. Manner compounds:

Ziroo-ga gohan-o tabe-nokosi-ta. (atelic-telic)
 Ziroo-NOM rice-ACC eat-leave-PAST
 ‘Ziroo left rice after eating (some).’
koroge-oti ‘roll(atelic)-fall.down(telic), i.e. fall down rolling’,
taore-kakar ‘fall(telic)-cover(telic), i.e. cover (something) by falling onto (it)’, etc.

c. Coordinating (dvandva) compounds:

Taroo-ga naki-saken-da. (atelic-atelic)
 Taroo-NOM cry-scream-PAST
 ‘Taroo cried and screamed.’

¹ The type-classification here, though convenient for illustration, is not directly relevant to this paper, since the aspectual properties of the component verbs are the central concern.

² For the intricacies of the compounds and theoretical issues arising from them, see the sources cited above.

³ The lexical (as opposed to syntactic) nature of “lexical” V-V compounds is demonstrated by Kageyama (1993).

hikari-kagayak ‘shine(atelic)-glitter(atelic), i.e. shine and glitter’,
itari-tuk ‘reach(telic)-arrive(telic), i.e. arrive-reach’, etc.

Adapting the assumptions of the Optimality Theory, Fukushima (2007) investigates the productivity of such compounds, but the perspective is argument-centered as well.

Compared with the situation above, the contribution of the aspectual (telicity) or the temporal properties of the component verbs has received rather sporadic attention. In this connection, a classic temporal account of V-V compounds in Chinese and Japanese is found in Li (1996). Additionally, there have been several observations and accounts focusing on telicity/temporal properties like Matsumoto (1996), Hasegawa (2000), Yumoto (2005), and Asao (2007) (see below for additional remarks on these predecessors). This paper is another contribution to the exploration of the role of telicity of V_1 and V_2 in lexical V-V compound formation. In particular, (i) it investigates (im)possible V_1 - V_2 combinations by looking at the aspectual properties of component verbs, (ii) it shows that V_1 cannot be telic (achievement/accomplishment) *unless* V_2 is also telic (given that V_2 is the head); any other patterns are possible: i.e. atelic-atelic, atelic-telic, telic-telic, and (iii) (im)possible aspectual combinations are shown to be a consequence of aspectual composition for V-V compounds based on the classification of verbal telicity by Dowty (1986).

2. Telic verbs as spoilers

What is described here regarding the aspectual properties of V_1 and V_2 should be considered to be an additional necessary condition for lexical V_1 - V_2 compound formation. In that sense, other constraints, such as proper argument-synthesis in particular, have to be satisfied independently. Accordingly, examples employed in this paper are constructed by observing these non-aspectual constraints noted by researchers mentioned above.

The current observation is simple and straightforward: telic V_1 (non-head) is an unacceptable component (i.e. a “spoiler”) as long as V_2 (head) fails to be telic as (2) demonstrates.

- (2) a. **Taroo-ga terebi-o naosi-tukat-ta.*
 Taroo-NOM TV-ACC repair(telic)-use(atelic)-PAST
 ‘(Int.) Taroo repaired and (then) used a TV.’ or
 ‘(Int.) Taroo used a TV by repairing it.’
- b. **hiroge-ur* ‘spread(telic)-sell(atelic), (Int.) sell after spreading (merchandise) or sell by spreading (merchandise)’
- c. **koware-nokor* ‘break(telic)-remain(atelic), (Int.) remain after going out of order’
- d. **taosi-fum* ‘knock.down(telic)-step.on(atelic), (Int.) step on after knocking (something) down’ or ‘(Int.) step on by knocking (something) down’, etc.

All the examples here involve telic verbs as V_1 and atelic verbs as V_2 . Actually, excluding instances where component verbs are used non-literally/figuratively, in 1157 examples of V-V compounds in Tagashira and Hoff (1986), there is one potential counter-example, namely, *kati-hokoru* ‘win-boast’. Though there are a few more apparent/potential counter-examples to be mentioned below, the generalization seems to be solid, demanding an explanation as to why it holds. It is interesting to note that a contraposition version of (2d), *fumi-taos* ‘step.on(atelic)-knock.down(telic)’, turns out to be an actual/legitimate compound, suggesting that argument-synthesis –since two semantically identical verbs are involved– is simply one of the factors (i.e. a necessary but not sufficient condition) determining the outcome.

We also note, as shown in (3) for example, that there is no intrinsic inconsistency in the combination of *naos* and *tukaw* or *hiroge* and *ur* with the former (temporally) preceding the latter and being construed as a (temporally preceding) factor/manner for the action indicated by the latter.

- (3) a. *Taroo-ga terebi-o [naosi sosite tukat-ta].*
 Taroo-NOM TV-ACC repair CONJ use-PAST
 ‘Taroo repaired and used a TV.’
- b. *Taroo-ga syoohin-o [hiroge sosite ut-ta].*
 Taroo-NOM merchandise-ACC spread CONJ sell-PAST
 ‘Taroo spread and sold merchandise.’

3. A proposal to capture the telicity restriction: a compositional approach

Given the empirical exposition and generalization offered above, let me outline the current account of the role of telicity found in lexical V-V compound formation.

3.1 Assumptions and aspectual composition

The account proposed here follows the aspectual characterization of predicates (verbs) due to Dowty (1986) given in (4):⁴

- (4) a. A predicate is stative (atelic) iff it follows from the truth of a sentence ϕ to which the predicate gives rise to is true at an interval I that ϕ is true at all subintervals of I .
- b. A predicate is activity (atelic) iff it follows from the truth of a sentence ϕ to which the predicate gives rise to is true at an interval I that ϕ is true at all subintervals of I down to a certain limit in size.
- c. A predicate is achievement/accomplishment (telic) iff it follows from the truth of a sentence ϕ to which the predicate gives rise to is true at an interval I that ϕ is false at all subintervals of I .

According to (4a,b) both statives and activities can have subevents that can also be classified as stative and activity (both atelic), respectively. For the latter, the size of such subevents has to satisfy a certain size limit. For example, just lifting one leg slightly may not qualify as the act of walking. Basically, for example, we can “chop up” the state of knowing one’s own name and still have a state where he/she knows his/her own name. Or an activity of swimming for five minutes qualifies as a subevent of swimming for 10 minutes and the latter of swimming for one hour, etc.

Due to (4c), the same story is inapplicable to achievement and accomplishment (both telic) –these do not have proper subevents. For example, there is no subevent of dying or arriving somewhere. One could have been involved in a sub-process (like pounding nails and sawing, i.e. activity) of building a house but that does not automatically mean that one completed an event of building a house. While the fact that someone was building a house for five days does not entail that a house was built, the fact that someone was swimming for an hour does entail that swimming indeed took place.

⁴ Dowty originally classifies sentences as stative, activity, achievement, and accomplishment types. His definitions are adapted here for defining predicates instead. Though formalization of the proposal is possible along the lines of the algebra of events framework of Bach (1986) and Krifka (1998) –both of which draw heavily on Link’s (1983) mereological approach in the nominal domain– or Filip (2008) for that matter, I am not going to pursue such a direction in this paper.

Based on the aspectual properties of predicates above, I propose the following:

- (5) a. A V-V compound represents a single event belonging to a single aspectual type with subevents attributed to V_1 and V_2 .
 b. When V_1 and V_2 differ in telicity, either (i) or (ii) holds. (i) If the head (usually V_2) is telic, a termination-point (distinct from an inception-point) is imposed on the interpretation of the non-head (usually V_1). Or (ii) if the head is atelic, the truth-at-all-subintervals requirement is imposed on the interpretation of the non-head.
 c. Otherwise, the aspectual type the whole compound is identical to those of V_1 and V_2 .

Some remarks are due regarding (5). First, while the term “aspectual type” encompasses telic and atelic, the former is a category with some width. According to Krifka (1998), telic predicates are either “quantized” (“strongly” telic, so to speak) or just telic (“weakly” telic). Quantized predicates are telic and have no proper sub-part other than itself. For example, *die* is a predicate with no proper sub-part (e.g. no proper sub-part of it counts as dying). In contrast, atelic *run* (being “cumulative” and not quantized) comes with proper sub-parts (e.g. two instances of running together count as running). *Run* is atelic but can be rendered telic by, for example, adding *for an hour* (see Krifka’s proof to this effect). Thus an atelic predicate like *run* is compatible with both atelic and telic readings. The introduction of a termination-point (distinct from an inception-point) in (5b-(i)) captures the weaker notion of telic-ness and the falsity-at-all-subintervals requirement in (5c) above the stronger one.⁵

Second, Japanese is generally morphologically head-final and V-V compounds are no exception. Thus (5) mostly picks V_2 as the determiner of the aspectual property of the whole compound. However, the formulation in (5) is not biased a priori regarding headedness. The reason is that there are some instances where both V_1 and V_2 are the heads (i.e. dvandva compounds) like *naki-sakeb* ‘cry-yell’. In some other cases V_2 behaves as if it is a non-head adverbial modifier as in *mi-oros* ‘look-lower, i.e. look down’ where V_1 is the semantic head (Matsumoto 1996). For the former type, due to (5c), the compound inherits the (identical) telicity of V_1 and V_2 . For the latter, V_1 has to be designated as the head for the purpose of aspectual composition.

Third, since the formation of these V-V compounds takes place in the lexicon (see note 3), the process depends on the kind and amount of lexical information available to the lexical process. For example, as mentioned above in connection with the distinction between achievements and accomplishments, the latter requires a VP with an incremental theme to qualify as such. But such information is not available for lexical word formation –no structure like VP is available yet. To accommodate this situation, we proceed in the following way: we apply the traditional telicity tests (like the one mentioned at the onset of this paper) and criteria like (4) to verbs and determine their “basic” telicity. Given the possibility that some verbs can eventually be construed as accomplishment, it may be necessary to appeal to a post-lexical semantic adjustment in some cases. For example, when forming a VP with an incremental theme, *tabe* ‘eat’ (basically activity) is quantized (i.e. turned accomplishment) in *hitotu-no ringo-o tabe* ‘eat an apple’.

According to (5), there are two cases to consider: when the telicity of the respective component verbs is different, and when it is identical. In the latter case, the whole compound simply inherits the telicity of the components as mentioned above in (5c) –a situation that is expected and hardly remarkable. On the other hand, if we encountered a literal dvandva V-V

⁵ Filip (2008) goes farther and claims that “accomplishment” verbs are atelic lexically. They qualify as telic only when their incremental theme arguments are present (in a VP).

compound with different telicity from that shared between V_1 and V_2 , that would be truly surprising.

The strength of (5) is (b-i) and (b-ii) covering the second case above. When the head is telic (b-i), the non-head is interpreted as having a termination point, i.e. treated as giving rise to a culminating (telic) event. This does not go against (4a,b), since both state and activity qualify as state and activity, respectively, even when they are terminated at one point or another (recall Krifka's conception of a "weakly" telic predicate). In contrast, when the head is atelic (b-ii), the non-head must be interpreted as satisfying (4a,b) but that goes against (4c) due to the fact that there is no subinterval where "strongly" telic (quantized) predicates can qualify as such.

3.2 Demonstration

Some well-/ill-formed V-V compounds are repeated in (6).

- (6) a. Matching telicity:
obore-sin (1a) 'drown(telic)-die(telic), i.e. die from drowning'
hikari-kagayak (1c) 'shine(atelic)-glitter(atelic), i.e. shine and glitter'
- b. Atelic-telic combination:
odori-tukare (1a) 'dance(atelic→telic)-get.tired(telic)'
tabe-nokos (1b) 'eat(atelic→telic)-leave(telic), i.e. leave (food) after eating'
- c. Telic-atelic combination:
**naosi-tukaw* (2a) 'repair(*telic→atelic)-use(atelic), (Int.) use after/by repairing'
**koware-nokor* (2c) 'break(*telic→atelic)-remain(atelic), (Int.) remain after having been broken'

The ones in (6a) are those with V_1 and V_2 matching in terms of telicity. They pose no problem due to (5c). Those in (6b) are just fine as well. For example, for *odori-tukarer*, according to (5b-(i)), a termination point is imposed on the verb *odor* (activity) but the action denoted by this verb can be terminated at some point without going against (4b) (up to a certain limit in size). The same goes for examples like *omoi-itar* 'think-reach' with a stative verb as V_1 and achievement V_2 . Now, for those in (6c), the story is a bit different. For instance, since V_2 *tukaw* is atelic, *naos* (telic) in **naosi-tukaw* is supposed to be interpreted as having sub-intervals where the action of fixing (of V_1) holds due to (5b-(ii)). But this contradicts the demand of (4c) above.

4. What others have said

In this section we briefly review what other researchers have said about temporal/aspectual combinations of verbs in lexical V-V compounds. I am not commenting on their proposals regarding other dimensions.

In this domain, a classic approach in this domain is found in Li (1993). The concept of "temporal iconicity" is central to his theory about Chinese/Japanese resultative V-V compounds. He claims that the temporal ordering of subevents e_1 and e_2 of e must be directly reflected in the surface linear order of the elements denoting the subevents. As we have seen above, however, this seems too restrictive. Though the account holds up as far as resultative ones like *odori-tukarer* (1a) go, there are other types of V-V compounds where the concept of result is irrelevant. One example is *koroge-oti* (1b) where V_1 is most naturally construed as manner. So it is either Li's account is restricted to resultatives and not having anything to say about the manner type (or others), or the term "resultative" is extended to include manner that

is relevant for bringing about a certain result. If the latter choice is made, the ill-formed examples in (2) would be mysterious. As we know from (3) above, there is nothing wrong about *naos* ‘fix’ temporally preceding *tukaw* ‘use’ in **naosi-tukaw*.

Another temporal account is found in Matsumoto (1996) where the “coextensiveness condition” is designated as one of the necessary factors in V-V compound formation. The condition states that the main component event (i.e. the one expressed by the head) must be temporally coextensive with (i) the subordinate component’s (i.e. non-head’s) event itself, or (ii) its result or effect, or (iii) an intention to execute or actualize it. For his account to be viable, first, the notion of “coextensiveness” has to be clarified. Does it mean that for example, two events overlap (completely or partially) or can occur adjacent to each other (with/without any temporal gap in between)? The latter consideration is relevant to examples like *tabe-nokosu* (1b). Second, a precise definition of “an intention to execute or actualize it” has to be offered. Would such an intention need to be expressed by a component verb or can be inferred from other information relevant to the sentence? In any case, conceptual unclarity has to be addressed so that the consequences of the condition can be tested adequately.

Hasegawa (2000) suggests a syntactic account employing the abstract predicate Res (for “result”). This abstract predicate is invisible in English but heads its own projection. It raises and attaches to a verb and creates a resultative counterpart of the verb. In Japanese the predicate is absent as an independent element, but telic verbs function as Res does in English. One such verb is *nob* ‘flatten’ (intransitive) which is raised and combined with *tataki* ‘hit’ to render a resultative compound *tataki-nob*. The compound finally combines with *as(ita)* (i.e. a “transitivizer”) as in (7b,c) for (7a).

- (7) a. *Hanako-ga kinzoku-o tataki-nobasi-ta.*
 Hanako-NOM metal-ACC hit-flatten-PAST
 ‘Hanako flattened the metal by hitting it.’
 b. [_{VP} Hanako [_V [_{VP} kinzoku-o [_V [_{VP} taira-ni [_{VR} nob]] tataki]] as(ita)_{V+Tr}]]
 c. [_{VP} Hanako [_V [_{VP} kinzoku-o [_V [_{VP} taira-ni *t_i*] *t_j*]] [*tataki-nob_i*]_j as(ita)]]

Being similar to Li’s account above, the focus is on resultatives; consequently, the coverage of Hasegawa’s account is restricted to the cases where V₂ is telic. Since telic verbs can be V₁ as well, e.g. *obore-sin* in (1a) but not all such cases can give rise to a legitimate compound, e.g. **obore-nagare* ‘(Int.) drown(telic)-float(atelic)’, there has to be something more said to extend her coverage.

What is observed by Yumoto (2005) comes very close to the current observation. Her supposition can be labeled as the “avoid-telic-V₁ condition”. She states that (i) for the dvandva type, V₁ and V₂ are of the same telicity, and (ii) for the modifying (cause/manner) type, it is very rare to find telic verbs as V₁. Though her observation seems to be on the right track, V-V compounds with telic V₁ are not that difficult to find, e.g. *obore-sin*, *sini-tae* in (1a), and *itari-tuku* in (1c). As was made clear above, however, not all instances with telic V₁ are acceptable, in particular, the ones with atelic V₂. Though Yumoto’s approach is correct in focusing on telicity of the component verbs (contra Matsumoto 1986), the scope of investigation needs to be broadened to include the aspectual properties of both V₁ and V₂. Thus the current proposal can be considered to be an extension based on Yumoto’s lead.

Finally, again based on the notion of temporal iconicity, Asao (2007) suggests another temporal condition -the “no complex V₁ without a complex V₂ condition”- which states that in V-V compound formation, if V₁ has a subordinate state of affairs, then so must V₂. Employing Lexical Conceptual Structure (LCS) to establish “complexity”, he assumes that the left-hand side of the predicated CAUSE (i.e. a superordinate state of affairs) of LCS is temporally prior to the one of the right-hand side of it (i.e. a subordinate state of affairs). The

event attributed to V_2 must not precede the one attributed to V_1 due to temporal iconicity. If V_1 alone has a subordinate state of affairs, it would be referred to after the state of affairs presumably of V_2 with the superordinate state of affairs of V_1 and simple state of affairs of V_2 being identified (though this point is not made explicit by Asao). But this goes against the temporal iconicity supposition above.

In any case, to make his condition viable, Asao classifies verbs and designates “causative transitives” (e.g. *naos* ‘repair’ and *otos* ‘drop’), “unaccusatives” (e.g. *make* ‘lose’), and “reflexive unergatives” (e.g. *ik* ‘go’) as the complex type with CAUSE and a subordinate state of affairs. For example, *ik* is defined as [x ACT] CAUSE [BECOME [x BE-AT y]]. But “non-causative transitives” (e.g. *os* ‘push’) and “unergatives” (e.g. *yorokob* ‘be.pleased’ and *nak* ‘cry’) are the simplex type without CAUSE (or a subordinate state of affairs). So examples like **naosi-yorokobu* ‘(Int.) repair-be.pleased’ and **make-naku* ‘(Int.) lose-cry’ are predicted to be impossible, where V_1 s are complex and V_2 s are simplex.

This outcome, however, crucially depends on the way verbs are defined, in particular, with or without CAUSE (plus a subordinate state of affairs). In this connection, Asao’s definition seems to be arbitrary. *Naos* is supposed to come with CAUSE since there is an agent (or an agentive event) that causes repairing but so should *yorokob* or *nak* since there is a factor/agent (or an event as such) responsible for pleasing and crying. True, there may not be an agent with a definite intention in the latter, but then there is no need for such an intentional agent for *otos* either –we may drop things unintentionally– which is defined as complex by Asao. Without independent evidence regarding what counts as being complex or simplex, the empirical consequences of the account cannot be tested adequately.

5. Some apparent counter-examples

Let me comment on apparent/potential counter-examples, though they are few. The first one is *yake-nokor* ‘burn(telic)-remain(atelic), i.e. remain after having burned’ which seems to be going against the aspectual generalization about compounding above. Two cases have to be considered here: (i) when “burning” and “remaining” concern a single object, and (ii) when “burning” and “remaining” are applicable for separate objects (as pointed out by Hiroaki Tada, p.c.). Under (i) burning is incomplete (i.e. without culmination) to the extent that some portion of the object escapes burning. In such a situation, V_1 *yake* is not telic but rather used as being equivalent to *moe* ‘be.on.flame’ (atelic). That is contrasted with *yake-oti* ‘burn(telic)-crumble(telic), i.e. crumble after having been burned’ where burning is complete or at least to the extent that the original state of the object burned cannot be deciphered. In the second case, the object is not affected by burning at all, e.g. a house escapes a town fire. This is similar to *Ie-ga ne-sizumar-u* ‘The house becomes quiet due to (its occupants’) falling asleep’ where V_1 is irrelevant to the subject *ie-ga*.

As mentioned above, *kati-hokoru* ‘win(telic)-boast(atelic)’ found in Tagashira and Hoff (1986) is a potential counter-example along with *kati-nokoru* ‘win(telic)-remain(atelic)’ (by Saeko Urushibara, p.c.). In these instances, what regularly appears to be a telic verb *kat* seems to be reinterpreted as an atelic verb. Evidence for this comes from ambiguity when the suffix *-iru* is employed. *Katte-iru* can be ambiguous; it could mean, for example, in a context of competitive sport ‘be winning’ (progressive) or ‘have won (and have been in that state)’ (perfective). Thus the verb in question can be construed as either atelic or telic. The same situation does not obtain with strongly telic (quantized) verbs like *sin* ‘die’, for example *sinde-iru* exclusively indicates a resulting perfective state.

Another potential problem is *sini-isog* ‘die(telic)-hurry(atelic), i.e. hurry to die’. As the translation suggests, this example may have a syntactic control structure where V_1 (or VP) serves as a complement of V_2 . We note that unlike the typical examples of V-V compounds

seen above, V_1 in this case does not function as cause, manner/means, or any other “modifier” for V_2 . Nor is the compound double-headed (i.e. not dvandva). In fact the following VP coordination seems to be fine with only one V_2 *isog* taking within its semantic scope a VP complement made up of two smaller VPs.

- (8) *Taroo-ga* [_{VP} [_{VP} *zairyoo-o kai*] (*sosite*) [_{VP} *seihin-o uri*]]-*isoi-da*.
 Taroo-NOM material-ACC buy and product-ACC sell-hurry-PAST
 ‘Taroo hurried to [buy materials and sell products].’

6. Conclusion

This paper has investigated the role of telicity in lexical V-V compound formation. The current observation is that the aspectual interaction between V_1 and V_2 needs to be taken into account. More specifically, telic V_1 is a spoiler if V_2 is not telic as well. To capture this generalization, aspectual composition was proposed based on Dowty’s conception of verb aspect. If valid, aspectual composition can be counted among the conditions for lexical V-V compound formation. This paper is also a contribution to research in not only the aspectual properties of V-V compounds but also the nature of aspectual composition of complex eventualities.

Acknowledgements

Many thanks go to the audience of MMM10 for their input. In addition, I would like to acknowledge individual contributions of Hee-Don Ann, Daisuke Bekki, Anna Maria Di Sciullo, Edit Doron, Nobuko Hasegawa, Toshio Hidaka, Claudio Iacobini, Takane Ito, Taro Kageyama, Hans Kamp, Hideki Kishimoto, Jong-Bok Kim, Nikos Koutsoukos, Fred Landman, Francesca Masini, Kunio Nishiyama, Naoyuki Ono, Chris Piñón, Masaki Sano, Yasuhiro Shirai, Sayaka Sugiyama, Hiroaki Tada, Yukinori Takubo, Anna Maria Thornton, Satoshi Uehara, Saeko Urushibara, Atsuko Wasa, Masaaki Yamanashi and Yoko Yumoto. I would also like to acknowledge the late Emmon Bach’s contribution to linguistics, in particular, the role of semantics in morphology. He was a great friend the interactions with whom helped to enrich my perception of formal semantics. Remaining inadequacies that may be found herein are my own fault.

References

- Asao, Y. (2007) Imi-no kasane-awase-to-site-no Nihongo hukugoo dooshi [Japanese compound verbs as superimpositions of meaning]. *Kyoto University Linguistic Research* 26: 59-75.
- Bach, E. (1986) The algebra of events. *Linguistics and philosophy* 9: 5-16.
- Dowty, D. (1979) *Word meaning and Montague Grammar*. Dordrecht: Reidel.
- Dowty, D. (1986) The effects of aspectual class on the temporal structure of discourse: semantics or pragmatics? *Linguistics and philosophy* 9: 37-61.
- Filip, H. (2008) Events and maximalization. In: S. Rothstein (Eds.), *Theoretical and crosslinguistic approaches to the semantics of aspect*. Amsterdam: John Benjamins, 217-256.
- Fukushima, K. (2005) Lexical V-V compounds in Japanese: lexicon vs. syntax. *Language* 81: 568-612.
- Fukushima, K. (2007) On the type-wise productivity of lexical V-V compounds in Japanese: a thematic proto-role approach. *Gengo Kenkyu* [Linguistic Research] 134: 119-140.
- Hasegawa, N. (2000) Resultatives and language variations: Result phrases and V-V compounds. In: M. Nakayama & C. J. Quinn Jr. (Eds.), *Japanese/Korean Linguistics* 9. Stanford: CSLI, 269-282.
- Himeno, M. (1999) *Hukugoodooshi-no kozo-to imiyoo* [Structure and semantic usage of compound verbs]. Tokyo: Hitsuji.
- Kageyama, T. (1993) *Bunpoo-to gokeisei* [Grammar and word-formation]. Tokyo: Hitsuji.
- Krifka, M. (1998) The origins of telicity. In: S. Rothstein (Ed.), *Events and grammar*. Dordrecht: Kluwer, 197-235.
- Li, Y. (1993) Structural head and aspectuality. *Language* 69: 480-504.
- Matsumoto, Y. (1996) Complex Predicates in Japanese: a syntactic and semantic study of the notion ‘word’. Stanford: CSLI.

- Nishiyama, K. (1998) V-V compounds as serialization. *Journal of East Asian Linguistics* 7: 175-217.
- Tagashira, Y. & J. Hoff (1986) *Handbook of Japanese Compound Verbs*. Tokyo: Hokuseido.
- Vendler, Z. (1967) *Linguistics in philosophy*. Ithaca, NY: Cornell University Press.
- Yumoto, Y. (2005) *Fukugodoshi/haseidoshi-no imi-to togo: mojuru keitairon-kara mita Nichieigo-no doshi keisei* [The semantics and syntax of compound verbs/derived verbs: verb-formation in Japanese and English viewed from a modular morphological perspective]. Tokyo: Hitsuji.

The role of stem frequency in morphological processing

Hélène Girardo	Serena Dal Maso	Sabrina Piccinin
<i>CNRS & University of Toulouse</i>	<i>University of Verona</i>	<i>University of Verona</i>
helene.girardo@univ-tlse2.fr	serena.dalmaso@univr.it	sabrina.piccinin@univr.it

1. Introduction

In usage-based models, it is generally acknowledged that the frequency of use of a lexical item influences its representation in long term memory, its organization in lexical storage and its processing mechanism during lexical access, since repeated exposures to an item lead to its entrenchment in the minds of the speakers (e.g., Bybee 2003; Bybee and Beckner 2010; Croft and Cruse 2004; Tomasello 2003). When it comes to morphologically complex words, processing mechanisms and lexical access are crucially determined by the relative frequency of the derived word and of its morphological components (both the stem and the affix). The results of thirty years of psycholinguistic research, mostly from lexical decision tasks, have demonstrated that, since frequent derived words have a strong lexical representation and are therefore highly entrenched in the speakers' minds, they are more likely to be processed holistically, i.e., as whole words. On the contrary, low frequency derived words made up of frequent morphological components would be more prone to be processed through a parsing process, i.e., by the activation of their morphological constituents. Starting from these results, an intense debate has sparked (as will be discussed in § 1.1): more precisely, researchers do not unanimously agree on the role of the stem during the processing of derived words. According to some studies (e.g., Taft and Foster 1975; Taft 1979), every morphologically complex form would be decomposed and consequently accessed through the stem whereas other researchers argue that this would only happen under specific conditions (Caramazza, Laudanna and Romani 1988).

An interesting contribution to this debate has come from psycholinguistic studies exploiting the masked priming experimental design combined with a lexical decision task (see §.1.2). What distinguishes this experimental protocol is the fact that, since the time given to access the derivative (usually used as a prime) is limited, a more sophisticated picture of the automatic and unconscious processing mechanisms of derived words is provided. Focusing on the very early stages of lexical access, this methodology highlights different nuances of frequency effects in comparison with traditional experimental tasks (lexical decision tasks with no priming). More precisely, it could be argued that the data coming from the latter might be evidence of a post-lexical effect, and therefore might reflect the organization of the lexicon in the mind, rather than the mechanisms involved during access to it.

In order to explore in more detail the role of the stem in the processing of morphologically complex words, we conducted a masked priming experiment on Italian, where we manipulated the frequency of the stem, while keeping constant that of the derived prime. More precisely, we contrasted stems which were on average twice as frequent as the derivatives (e.g., *trasferimento* 'transfer' (112) vs. *trasferire* 'to transfer' (284); target/prime ratio: 2,53) to stems which were on average half as frequent as the derivatives (e.g., *motivazione* 'motivation' (98) vs. *motivare* 'to motivate' (44); prime/target ratio: 2,22).

On the one hand, an influence of the stem frequency in the magnitude of priming effects would point towards a decompositional interpretation, i.e., a segmentation process at a

prelexical level could be hypothesized. On the other hand, if the stem frequency was found to have no effect, this would mean that access occurs through the whole form. A similar finding would not necessarily mean denying the role of the stem frequency in the organization of the lexical architecture; rather, it would be consistent with a view which hypothesizes that the mechanisms involved in lexical access are crucially different from the criteria of organization involved in lexical representation.

2. Background

2.1. Frequency effects in lexical decision tasks

In traditional psycholinguistic research, frequency effects have been observed mainly by means of visual lexical decision tasks (from now on LDT), i.e., experiments where the subjects are asked to decide whether the letter string they are briefly exposed to is a word of their language or not. Generally, these experimental studies oppose derived or inflected words of comparable surface frequency, but crucially differing in their stem frequency (high vs. low). In this kind of studies, when reaction times (RTs) are found to be a function of the stem frequency, this is considered as evidence of the fact that word recognition implies the activation of the stem.

Research on frequency effects has flourished after the seminal – though controversial – proposal put forward by Taft and Forster (1975) and Taft (1979), according to which affixes (precisely, prefixes and inflectional suffixes) would be ‘stripped off’ their stem prior to lexical access, which proceeds from the stem.

As for Italian, Burani and Caramazza (1987) investigated derived suffixed forms (verbal roots combined with highly productive suffixes such as *-mento*, *-tore*, *-zione*) by opposing stimuli matched for whole-word frequency, but differing in root frequency (Exp. 1), to stimuli matched for root frequency but differing in whole-word frequency (Exp. 2). Their aim was to find out whether reaction times in lexical decisions were determined by the frequency of the root-morpheme or rather by the frequency of the whole-word and thus to verify Taft and Forster’s proposal. Since their results indicated that RTs were influenced by both root and whole-word frequencies (faster RTs were obtained for items containing a high frequency root in Exp. 1 and for higher whole-word frequency items in Exp. 2), the authors suggested that the access procedure crucially operates with both whole-word and morpheme access units. This kind of results led them to elaborate a ‘dual-route’ model, i.e., the Augmented Addressed Morphology Model (AAM, proposed in Caramazza, Laudanna and Romani 1988), which includes both prelexical morphological computation and whole-word lexical representation of complex words. According to this model, the activation of an access unit corresponding to the whole word is a faster and more efficient procedure for familiar words, while the activation of morpheme access units is exploited for the recognition of items which do not have a global lexical representation, i.e., novel words, neologisms or pseudo-words. To sum up, what this model shares with Taft’s is the assumption that there is a lexical component where the words are decomposed into morphemes. However, while in Burani and Caramazza’s activation model access to lexical representations for previously experienced words takes place via a whole word procedure, Taft proposes a serial architecture where the search process is always conducted via the root.

Frequency effects have been observed also in French by Colé, Beauvillain and Segui (1989), who similarly considered derived words matched for surface frequency but differing in their cumulative root frequency (e.g., *jardinier* ‘gardener’, containing a high frequency root, vs. *policier* ‘policeman’, containing a low frequency root). Since a clear cumulative root effect was observed only for suffixed words (but not for prefixed ones), Colé and colleagues suggest that only the former are accessed through decomposition via the root.

However, it is in Schreuder and Baayen's Race Model that frequency is acknowledged to play a crucial role in determining processing mechanisms: in this model morphological parsing would not be limited to unknown or new complex words (as in the AAM), but would be extended also to less frequent words, which contain commonly used roots (Schreuder and Baayen 1995 and Baayen and Schreuder 1999 for the evolution into the Parallel Dual-Route Model).

In the same line of reasoning, Burani and Thornton (2003) conducted a study on the interplay between the frequency of the root, the frequency of the suffix and the whole-word frequency in processing Italian derived words. More precisely, in Exp. 3, they considered low frequency suffixed words which differed with respect to the frequency of their morphemic constituents. As expected, the results showed that lexical decisions were faster and more accurate when the derived words included two high-frequency constituents (e.g., *pensatore* 'thinker') and slowest and least accurate when both constituents had low frequency (e.g., *luridume* 'filth'). Interestingly, when the derived words included only one high-frequency constituent (either the root or the suffix), the lexical decision rate was found to be a function of the frequency of the root only, irrespective of suffix frequency. The authors conclude that access through activation of morphemes is beneficial only for derived words with high frequency roots, while lexical decision latencies to suffixed derived words are a function of their surface frequency when they contain a low frequency root.

To sum up, frequency effects have been considered as a diagnostic for determining whether an inflected or derived form is recognized through a decompositional process that segments a word into its morphological constituents or through a direct look-up of a whole-word representation stored in lexical memory. Frequency has therefore played a crucial role in the debate which opposed full parsing models, which assume a prelexical treatment of the morphological constituents with a consequent systematic and compulsory segmentation of all complex words (Taft and Forster 1975, Taft 1979), and full listing models, which defend a non-prelexical processing of the morphological structure and a complete representation of all morphologically complex words (McClelland and Rumelhart 1981).

2.2. Frequency effects and masked priming

More recently, frequency effects have been investigated using the masked priming experimental paradigm associated with a LDT, which gives a slightly different picture of such effects on word recognition. This experimental paradigm is usually employed to verify the existence of a facilitation effect between a derivative word (i.e., the prime word) and its stem (i.e., the target word). The advantage of this technique is that, since the prime is presented on the screen for a very short time (usually between 45 and 60ms), subjects are not aware of its presence and any observed facilitation cannot therefore be considered to derive from a conscious recognition of the prime-target relationship. Moreover, as the activation of the prime is constrained by strictly determined time limits, this technique allows for observation of the very early stages of lexical access, during which other bits of information possibly encoded in the lexicon are not yet available to the subjects.

Despite the introduction of this experimental methodology, the dichotomy between decompositional and holistic approaches has not been reconciled yet and the results of masked priming studies have not been univocally interpreted. We shall now briefly consider the main contributions to this debate. Giraudo and Grainger (2000) in a series of masked priming experiments investigated the role of surface and base frequencies using French material. More precisely, they manipulated the surface frequencies of derivatives used as primes for the same target (HF *amitié* 'friendship' - *ami* 'friend'; LF *amiable* 'friendly' - *ami* 'friend'). They found an interaction between priming effects and the prime surface frequency (Exp.1), but no effect for the base frequency. Experiments 1 and 3 demonstrated that the

surface frequency of morphological primes affects the size of morphological priming: high surface frequency derived primes showed significant facilitation relative to orthographic control primes (*amidon* ‘starch’ - *ami* ‘friend’), whereas low frequency primes did not. The results of Experiment 4 revealed, on the contrary, that cumulative root frequency does not influence the size of morphological priming on free root targets. Suffixed word primes facilitated the processing of free root targets with low and high cumulative frequencies. These data suggest that during the early processes of visual word recognition, words are accessed via their whole form (as reflected by surface frequency effects) and not via decomposition (since the base frequency did not interact with priming).

It is worth noticing that in the traditional experiments carried out with the masked priming protocol, the usual prime-target configuration is one where the target word (i.e., a stem) is the most frequent member of the pair and thus the easiest-to-activate member of the morphological paradigm (e.g., *darkness/ dark*). The reverse situation has rarely been investigated, yet it is potentially very useful in order to understand the processing mechanisms, mainly with respect to the parsing/listing dichotomy. Interestingly, Voga and Giraud (2009) reversed the traditional prime-target configuration and, considering French inflected forms, used the less frequent item as target word (e.g., *montons* ‘we mount’) and the most frequent one as prime (e.g., *monter* ‘to mount’) (Exp. 1b). Moreover, in order to better verify the strength of lexical representations, they also considered the ‘environment’ of the prime, i.e., its competitors, the number of words sharing the same letters of the stem without having a morphological relationship with it. Morphological priming was expected to be more efficient for verbs with a small pseudo-family (i.e., a limited number of antagonists). Their results confirmed that high frequency targets (usually the base or infinitive forms of inflectional paradigms) have a low threshold of activation and are therefore the easiest-to-activate form of the paradigm, independently of their antagonists (size of the pseudofamily). However, when the targets are not the easiest-to-activate forms because of their low frequency, the effect of pseudo-family size emerges, as the competition which takes place at the lexical-orthographic level influences what happens at the morphological level. In other words, it seems that the presence of many antagonists at the lexical-orthographic level interferes with the processing of low frequency targets, leading to the absence of morphological effects. According to the authors, this is the reason why morphologically related forms with no antagonists at the lexical level can facilitate the processing of the target, while forms with many pseudo-relatives fail to do so.

Another piece of evidence against the decompositional hypothesis comes from the study conducted by Orihuela and Giraud (submitted), which considers the effects of the relative frequencies of complex primes and their base target opposing the configuration with high frequency primes / low frequency targets to the configuration with low frequency primes / high frequency targets in French. Their results reveal that, relative to both the orthographic and unrelated conditions, morphological priming effects emerge only when the surface frequency of the primes is higher than the surface frequency of the targets. Again, these data contradict the prediction of the classical decomposition hypothesis, according to which the reverse effects would be expected.

The interpretation of frequency effects with respect to psycholinguistic models, however, remains very controversial. McCormick, Brysbaert and Rastle (2009) defend a completely different position, in favour of an obligatory decomposition of all kinds of stimuli (even for the non-morphologically structured ones). They carry out a masked priming experiment manipulating the frequency of the prime, thus comparing high frequency, low frequency and non-word primes. Their hypothesis is that if morphological decomposition is limited to unfamiliar words, as predicted by the horse-race style of dual-route models, then priming should be limited to the last two conditions. On the contrary, if morphological decomposition

is a routine, an obligatory process applying to all morphologically structured stimuli should occur in all three conditions. The results show that the priming effect observed with high-frequency primes is equivalent to the one observed with low-frequency primes and with non-word primes. Such findings seem to confirm the claim that a segmentation process is not restricted to low-frequency words or non-words, as assumed by horse-race models (e.g., Caramazza et al. 1988; Schreuder and Baayen 1995).

3. The present study

3.1. Method

3.1.1 Participants

38 Italian native speakers, 16 males and 22 females, aged 17 to 55 years (mean age: 29,39), with normal or corrected-to-normal vision, participated in the experiment. All of them had high-school or university education and took part in the experiment voluntarily.

3.1.2 Stimuli and design

We selected 80 critical items to be used as target stimuli for this experiment and divided them in two subsets: 40 of them were high frequency words (mean frequency: 238,23) and 40 were low frequency words (mean frequency: 42,30). For each stimulus we created a morphological, an unrelated and an orthographic prime. Since we were interested in the effects of the manipulation of the stem frequency, the frequency of the morphological primes was kept constant in the two subsets (mean frequency in the HF subset: 87,75; LF subset: 95,33). The orthographic and the unrelated primes were matched as well (mean frequency in the HF subset: orthographic: 92,78; unrelated: 90,63; mean frequency in the LF subset: orthographic: 136; unrelated: 96). Moreover, all of the three primes were matched for letter length (mean length in the HF subset: morphological: 10,90; orthographic: 8,03; unrelated: 9,15; mean length in the LF subset: morphological: 11,13; orthographic: 8,03; unrelated: 9,43). We made sure that the ratio between the frequencies of primes and targets was comparable in the two subsets: in other words, in the HF subset the target was 2,71 as frequent as the prime, while in the LF one the prime was 2,25 as frequent as the target. Frequency values were extracted from the CoLFIS database. The targets were either verbs or nouns, with respectively deverbal nouns and denominal adjectives as morphological primes. We created four experimental lists in which the targets were rotated across the four priming conditions by means of a Latin square design, so that participants saw each target only once in one of the possible four conditions. The experimental design is summarized in the tables below:

CONDITIONS	PRIME	FREQ.	HF TARGET	FREQ.	TARGET/PRIME RATIO
Identity	<i>trasferire</i>		<i>trasferire</i>		
Morphological	<i>trasferimento</i>	112	<i>trasferire</i>	284	2,53
Orthographic	<i>trasparenza</i>		<i>trasferire</i>		
Unrelated	<i>sacrificio</i>		<i>trasferire</i>		

CONDITIONS	PRIME	FREQ.	LF TARGET	FREQ.	PRIME/TARGET RATIO
Identity	<i>motivare</i>		<i>motivare</i>		
Morphological	<i>motivazione</i>	98	<i>motivare</i>	44	2,22
Orthographic	<i>motorino</i>		<i>motivare</i>		
Unrelated	<i>orizzonte</i>		<i>motivare</i>		

Table 1: Experimental design (HF vs LF targets)

3.2. Procedure and apparatus

The experiment was run on a PC using the DMDX software (Forster and Forster 2003). Each trial consisted of three visual events: the first was a forward mask made up of a series of hash marks that appeared on the screen for 500ms. The mask was immediately followed by the prime, which appeared on the screen for 50ms. The target word was then presented and remained on the screen until participants responded. Both primes and targets were presented in lowercase; however, they differed in size and font. Prime stimuli were presented in Arial 12 and target stimuli in Cambria 16. Participants were seated 50 cm from the computer screen and were instructed to decide as quickly and accurately as possible whether the target stimuli they saw were words or not, by pressing the appropriate buttons on the keyboard. They were not aware that a prime word was presented. After 20 practice trials, participants received the 160 items in two blocks.

4. Results

Correct response times (RTs) were averaged across participants after excluding outliers (RTs >1500ms, 0,65% of the data). Results are presented in Table 2. An ANOVA was performed on the data with prime type factor (identity, morphologically related, orthographic and unrelated controls) and frequency factor (high frequency targets and low frequency targets) as within-participant factors. The list factor was included as a between-participant factor in order to extract any variance associated with this variable and turned out to be not significant ($F < 1$). A Latin Square design was used in the present experiment; therefore, as recommended by Raaijmakers, Schrijnemakers and Gremmen 1999, we did not perform separate subject and item analyses, but only a F1 statistic test.

The analysis of RTs latencies showed a main effect for Frequency, $F(1,37)=43.17$, $p < .0001$ and Prime Type, $F(3,111)=29.53$, $p < .0001$. The interaction of frequency by prime was not significant, $F(3,111)=1.04$, $p > .10$. Main comparison differences (significant at $p < .05$) are indicated in Table 2. An analysis of the error rates showed no main effect (All Fs < 1).

		Net priming effects						
	Prime type	RTs (SD)	Errors	I-M	O-M	I-U	M-U	O-U
Low Frequency targets	Identity	660 (124)	2 (0,52%)	+7		+47*		
	Morphological	653 (121)	4 (1,05%)		+24*		+44*	
	Orthographic	677 (107)	12 (3,15%)					+20*
	Unrelated	697 (113)	10 (2,63%)					
High Frequency targets	Identity	623 (113)	4 (1,05%)	-2		+57*		
	Morphological	625 (111)	4 (1,05%)		+33*		+55*	
	Orthographic	658 (123)	3 (0,78%)					+22*
	Unrelated	680 (115)	9 (2,36%)					

Table 2. Reaction Times and SD (in milliseconds), error rates (in %) for lexical decisions to word targets in each frequency and priming condition, with net priming effects relative to the identity, orthographic and the unrelated prime conditions. * : $p < .05$

We found significant morphological effects for both HF and LT targets: morphologically related primes yielded faster RTs with respect to the unrelated condition, independently of the prime/target frequency ratio. Moreover, morphological primes did not differ from the identity primes, suggesting comparable magnitudes of morphological facilitation. These equivalent effects indicate that the activation strength of the morphological primes is the same because their surface frequency was equal. This is in line with the results of Giraudo and Grainger

(2000) according to which the priming effect is a function of the surface frequency of the prime (as a whole-word) and not of its stem frequency. Finally, both morphological primes differed from their orthographic controls (+29 ms on average), while both orthographic controls induced faster RTs than the unrelated primes (+21 ms on average). This indicates that while orthographic primes produced mere formal facilitation effects, morphological primes induced genuine morphological effects reflecting form and meaning correlations between primes and targets.

As expected, overall we obtained faster RTs for HF targets, as their activation threshold is lower than the threshold of LF targets.

5. Discussion

The present data show that full morphological priming effects are obtained whatever the frequency of the targets (high or low). Accordingly, the frequency of the base contained in the derived primes (e.g., *trasferire* in *trasferimento*) does not interfere with morphological facilitation: primes whose base had a high frequency did not induce stronger facilitation than primes with a low frequency base. As a consequence, contrary to a decompositional approach of lexical access to complex words, the prior presentation of a complex prime whose stem has a high surface frequency does not accelerate the access to its lexical representation relative to primes whose stem frequency is low.

Moreover, in the present experiment both morphological primes induced full priming (i.e., same priming effects for morphological and identity condition) and both differed significantly from orthographic and unrelated controls. Hence, morphological primes are as efficient for the recognition of their base target as the repetition of the target itself. This suggests that morphological families are not accessed through stem activation but through the activation of each of its members and without being segmented into stem + affix. Such findings replicate the results found in Giraudo and Grainger study (Exp. 4) in which the cumulative frequency of the targets (high vs. low) was manipulated with French materials. The results of their Exp. 4 showed that cumulative root frequency did not influence the size of morphological priming on free root targets. Suffixed word primes facilitated the processing of free root targets with low and high cumulative frequencies. Thus, a manipulation designed to modulate the involvement of morphological processing (according to a generic dual-route model) did not lead to variations in morphological priming.

The present data do not, however, contradict the results found by Colé et al. (1989) in French or by Burani and Caramazza (1987) and Burani and Thornton (2003): rather, they are not in line with their interpretation. These authors found a root frequency effect which has been interpreted as evidence in favour of the activation of the stem during lexical access. Since we did not find a base frequency effect using masked primes, this suggests that lexical access does not take place via the automatic activation of the stem of suffixed words.

The important difference between these studies and ours resides indeed in the experimental designs employed to examine morphological processing. While Burani et al. used a simple lexical decision task, we combined it with a masked priming paradigm. While the former provides information on the time needed to consciously process a visual target, the latter, constraining the duration of the prime exposure, allows for the observation of very early stages of unconscious processing. Moreover, this paradigm does not examine the entire process of word recognition as the LDT does, but explores a transfer of activation between a prime and a target. More precisely, it questions whether the information extracted from the prime representation during a very reduced time is able to spread towards other representations thanks to links or shared representations. As a consequence, we can say that in the LDT the observed RTs reflect an overall process of lexical access resulting from two

types of activation: the time required to activate the lexical representation depends on its lexical/surface frequency and on the activation of its synagonists (i.e., morphologically related words that facilitate the recognition of the target as demonstrated by the robust family size effect, see De Jong, Schreuder and Baayen 2000). Masked priming effects, on the other hand, depend much more on the lexical frequency of the prime (that determines its activation threshold) given the fact it is presented for a limited time. Hence, masked priming corresponds to a small window in the overall process of word recognition. Accordingly, this paradigm is much better suited to constrain the process engaged while accessing the mental lexicon. This, however, does not mean that priming effects only correspond to pre-lexical effects, otherwise morphological priming effects would not differ from orthographic priming effects. This significant difference has been extensively demonstrated in the literature on masked morphological priming (see Baayen 2014 for a review) and there is no doubt that masked priming also reflects lexical effects. We would like to highlight here that the prime exposure duration is a parameter that constrains the access to lexicon.

Therefore, we claim that our data clearly show that the stem frequency does not interfere with the access to the mental lexicon. Moreover, morphological priming effects reveal that, as soon as a lexical representation is activated within the mental lexicon, such a representation automatically triggers the activation of all its family members. The result of the overall activation of the morphological family is revealed in those LDT experiments in which it has been observed that both the lexical and the base frequency determine the recognition latencies of suffixed words. Only models that consider the word as the main unit of analysis, be it morphological (e.g., Giraudo and Voga 2014) or not (e.g., Baayen et al. 2011), are able to account for our data and the previous findings.

References

- Baayen, R.H. & R. Schreuder (1999) War and Peace: Morphemes and Full Forms in a Noninteractive Activation Parallel Dual-Route Model. *Brain and Language* 68: 27-32.
- Baayen, R. H., P. Milin, D. Filipovic Durdevic, P. Hendrix & M. Marelli (2011) An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review* 118: 438-482.
- Baayen, R.H. (2014) Experimental and psycholinguistic approaches to studying derivation. In: R. Lieber & P. Stekauer (Eds), *Handbook of derivational morphology*. Oxford: Oxford University Press, 95-117.
- Beauvillain, C. (1996) The integration of morphological and whole-word form information during eye-fixation on prefixed and suffixed words. *Journal of Memory and Language* 35: 801-820.
- Bertinetto, P. M., C. Burani, A. Laudanna, L. Marconi, D. Ratti, C. Rolando & A.M. Thornton (2005). Corpus e Lessico di Frequenza dell'Italiano Scritto (CoLFIS). http://linguistica.sns.it/CoLFIS/CoLFIS_home.htm.
- Booij, G. (2010) *Construction Morphology*. Oxford: Oxford University Press
- Burani, C. & A. Caramazza (1987) Representation and processing of derived words. *Language and Cognitive Processes*, 2, 217-227.
- Burani, C. & A.M. Thornton (2003) The interplay of root, suffix and whole-word frequency in processing derived words. In: R. H. Baayen & R. Schreuder (Eds.), *Morphological Structure in Language Processing*. Berlin: Mouton de Gruyter, 157-208.
- Bybee, J. (1988) Morphology as lexical organization. In: M. Hammond & M. Noonan (Eds.), *Theoretical morphology*. San Diego: Academic Press, 119-141.
- Bybee, J. & J. Scheibman (1999) The effect of usage on degrees of constituency: the reduction of *don't* in English. *Linguistics* 37(4): 575-596.
- Bybee, J. (2003) Mechanisms of change in grammaticization: The role of frequency. In: B. D. Joseph & J. Janda (Eds.), *The Handbook of Historical Linguistics*. Oxford: Blackwell, 602-623.
- Bybee, J. (2007) *Frequency of use and the organization of language*. Oxford: Oxford University Press.
- Bybee, J. & C. Beckner (2010) Usage-based Theory. In: B. Heine & H. Narrog (Eds.), *The Oxford Handbook of Linguistic Analysis*. Oxford: Oxford University Press, 827-856.
- Caramazza, A., A. Laudanna & C. Romani (1988) Lexical access and inflectional morphology. *Cognition*, 28: 297-332.

- Colé, P., C. Beauvillain & J. Segui (1989) On the representation and processing of prefixed and suffixed derived words: A differential frequency effect. *Journal of Memory and Language* 28: 1-13.
- Croft, W. & D.A. Cruse (2004) *Cognitive Linguistics*. Cambridge: Cambridge University Press.
- De Jong, N. H., R. Schreuder & R.H. Baayen (2000) The morphological family size effect and morphology, *Language and Cognitive Processes*, 15: 329-365.
- Forster, K.I. & C. Davis (1984) Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10: 680-698.
- Forster, K.I. & J.C. Forster (2003) DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, 35(1): 116-124.
- Giraud, H. & J. Grainger (2000) Effects of prime word frequency and cumulative root frequency in masked morphological priming. *Language and Cognitive Processes*, 15 (4/5): 421-444.
- Giraud, H. & M. Voga (2014) Measuring morphology: the tip of the iceberg? A retrospective on 10 years of morphological processing. *Carnets de Grammaire*, 22: 136-167.
- McClelland, J. L. & D.E. Rumelhart (1981) An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, 88: 375-407.
- McCormick, S.F., M. Brysbaert & K. Rastle (2009) Is morphological decomposition limited to low-frequency words? *The Quarterly Journal of Experimental Psychology*, 62: 1706-1715.
- Orihuela, K. & H. Giraud (submitted) Reversing the surface frequencies of prime-target morphologically related pairs: implications for the decompositional and the holistic approaches of lexical access. *Proceedings of the Network 2015 Conference, Word Knowledge and Word Usage*, March 30-April 1, 2015, Pisa, Italy.
- Raaijmakers, J.G.W., J.C.M. Schrijnemakers & F. Gremmen (1999) How to deal with “The Language-as-a-Fixed-Effect Fallacy”: Common misconceptions and alternative solutions. *Journal of Memory and Language*, 41:416-426.
- Schreuder, R. & R.H. Baayen (1995) Modeling morphological processing. In: L.B. Feldman (Ed.), *Morphological aspects of language processing*. Erlbaum: Hove, 131-154.
- Taft, M. (1979) Recognition of affixed words and the word frequency effect. *Memory & Cognition*, 7: 263-272.
- Taft, M. (1994) Interactive-activation as a framework for understanding morphological processing. *Language and Cognitive Processes*, 9: 271-294.
- Taft, M. & K.I. Forster (1975) Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 14: 638-647.
- Tomasello, M. (2003) *Constructing a language: A Usage-Based Theory of Language Acquisition*. Cambridge: Harvard University Press.
- Voga, M. & H. Giraud (2009) Pseudo-family size influences processing of French inflections: Evidence in favor of a supralexicale account. In: F. Montermini, G. Boyé & N. Hathout (Eds.), *Selected Proceedings of the 6th Décembrettes: Morphology in Bordeaux*. Somerville, MA: Cascadilla Proceedings Project, 148-155.

“Romanes eunt domus”: where you can go with Latin morphology. Variation in motion expression between system and usage

Claudio Iacobini
University of Salerno
ciacobini@unisa.it

Luisa Corona
University of Salerno
lcorona@unisa.it

1. Introduction

The title of this paper alludes to a famous scene in the film *Monty Python's Life of Brian* in which Brian daubs an anti-Roman slogan (*Romanes eunt domus*) on the walls of Governor Pontius Pilate's palace in Jerusalem (cf. <https://www.youtube.com/watch?v=IIAdHEwiAy8>). The centurion who catches him corrects Brian's sloppy grammar, and orders him to write out a hundred times the correct form *Romani ite domum* (better would be *redite domum*) on the palace walls. The comical dialogue of the forced lesson reveals the richness of morphological resources available for the encoding of motion expression in Classical Latin.

In this paper, besides outlining the way Classical Latin encodes motion events, we will show that, although Latin displays at the system level a wide array of linguistic resources characterizing Satellite-Framed languages, the actual usage of the strategies employed in motion encoding significantly differs from what is expected in a typical Satellite-Framed language.

These findings, resulting from a detailed corpus-based analysis, lead us to formulate a remark and a hypothesis:

- (i) an investigation limited to the resources offered by the morphological and the lexical system is not sufficient to provide a proper typological classification of a language, since the resources available at the system level may be not consistently employed in actual usage. As a consequence, typologies of languages as a whole are generalizations that can be useful only at a very broad level of classification; a fine-grained typological classification should primarily refer to the actual usage of constructions employed to encode event types (cf. Croft et al. 2010; Verkerk 2014, 2015; Ibarretxe-Antuñano 2015);
- (ii) the preferred ways in which Latin encodes dislocational motion (e.g. simple Path, relative scarcity of manner verbs) allow us to hypothesize a possible way of typological change in preferred motion encoding from Satellite- to Verb-Framed strategies (as occurred in the transition from Latin to the Romance languages).

The paper is organized as follows. Section 2 summarizes the basic tenets of the classification of motion expressions we used to evaluate the linguistic strategies deployed by Classical Latin. In Section 3 we briefly present our corpus and methodology of data analysis. In Section 4 the most important morphological means Latin makes available for motion encoding are listed. Section 5 reports the results of our corpus-based analysis on the preferred strategies of motion encoding in Latin. In the conclusions, we put forward some methodological considerations and we refer to the usage preferences in Latin motion encoding which may have favoured the emergence of Verb-Framed strategies in the passage to the Romance languages.

2. The linguistic classification of motion encoding

In this section, we briefly illustrate some key concepts of the typological classification of motion event lexicalization patterns upon which there is general consensus among scholars, and which have been used here to assess the linguistic strategies deployed by Classical Latin in the encoding of motion events.

Space and Motion are basic concepts in human cognition: therefore, in the last decades, they have been widely studied in cognitive linguistics and, above all, in linguistic typology. These notions, highly represented in every human language, can be easily used for large-scale comparison in a cross-linguistic perspective. In this work, we will focus on dislocational expression.

A motion event can be decomposed into four major components: Figure, Manner, Path, and Ground. In the expression *Brian walked into the room*, *Brian* refers to the Figure, *walked* to the Manner of motion, *into* to the Path, and *the room* to the Ground component. Among these, the most important component is Path, because motion essentially consists in a change of location, cf. also Bohnemeyer (2003), Grinevald (2011). As a consequence, the defining criterion for the categorization of a language is the identification of the linguistic element encoding the Path (e.g. the main verb or an element different from the verb). According to Talmy’s (2000) macro-typology, which divides languages into two groups, Ancient Greek, Latin, Slavic and Germanic languages are classified as Satellite-Framed because they typically express Path outside the verb root, in elements called “satellites” (e.g., adverbs, prefixes, post-verbal particles, etc.), and the Manner of motion in the verb (e.g. Latin *ad-curro*, *ex-curro*, *per-curro* vs. English *run in/out/across*, etc.). Verb-Framed languages lexicalize the Path component in the verb, whereas Manner is optionally expressed as an adjunct (as in Spanish *entrar corriendo* ‘enter running’). Romance languages, Hebrew, Turkish are classified in this group.

The linguistic resources each language makes available for the lexicalization of motion are highly intertwined with speakers’ attention to different aspects of a same motion event (cf. Slobin 2006). Satellite-Framed languages allow speakers to describe both Manner and Path frequently and in detail. As a consequence, Satellite-Framed languages are characterised by a rich and expressive Manner of motion verb lexicon, and by the possibility to attach more than one Path segment to a single verb. On the other hand, Verb-Framed languages hardly describe Manner unless it is discursively important (as a consequence, Manner of motion verb lexicon is limited and general), and at most one Path element is added to the verb. Instances of Path complexity in Satellite-Framed languages are listed in the examples from (1) to (3).

- (1) English (from Slobin 2005)

He ran out of the house, across the field, into the forest.

- (2) Polish (from Fortis and Vittrant 2011)

chłopiec wy-biegł z morza na plażę
 boy(M).NOM out-ran from sea(N).GEN to beach(F).ACC
 ‘The boy ran out from the sea to the beach’

- (3) Jakalte’k Popti (from Craig 1993)

sirnih-ay-toy sb’a naj sat pahaw b’et wichen
 threw-down-away REFL.3SG PRON.3SG in.front cliff into gully
 ‘He threw himself away over the cliff into the gully’

According to Slobin (1996), the main distinction to be made with respect to Path complexity is between minus-ground and plus-ground expressions. The former are cases where the verb is alone (English *to slip*, Latin *labor*) or with a satellite (English *slip down*, Latin *delabor*), and the latter are those containing one (or more) extra Path element(s) (English *he slips down from the cliff*; Latin *summo delabor Olympo* ‘I descend from the top of Olympus’ Ovid *Metamorphoses* I.212).

The situation where the typological differences between Satellite- and Verb-Framed languages are most noticeable is the linguistic encoding of boundary-crossing. Boundary-crossing occurs when there is an explicit change of spatial configuration, as in *Brian ran out of the cage into the arena*. Manner verbs tend to be blocked in Verb-Framed languages in such situations (due to the lack of dedicated satellites or other available means to express the crossing of a boundary), and a verb encoding Path is used instead (cf. Aske 1989; Filipović 2007). In the expression of non-boundary-crossing, Manner of motion verbs can be used in both Verb- and Satellite-Framed languages. Verb-Framed languages may use verbs with definite end-states or origin in the encoding of motion events (e.g. Italian *Corsero verso casa* ‘(they) ran towards the house’; French *Le poisson a nagé vers la rive* ‘The fish swam towards the river bank’), but this happens less often than in Satellite-Framed languages.

3. Corpus and methodology

Our analysis is based on a systematic scrutiny of two texts of different kinds belonging to two authors of the Golden Age: Caesar’s account written in prose of the Gallic Wars (published in 58-49 BC) *De bello gallico* and Ovid’s *Metamorphoses*, an early imperial poem in dactylic hexameters (published in 8 AC).

From these works, we extracted the contexts describing dislocational events and excluded all metaphorical uses of motion verbs and all senses of particles and prepositions not related to the direction of motion.

The linguistic encoding of motion events has been classified and analyzed by means of a coding grid developed starting from Fortis and Vittrant’s (2011) proposal of typology of constructions. Our grid combines both morpho-syntactic and semantic information and allows intra- and cross-linguistic data comparability (for an in-depth description of the grid cf. Iacobini, Corona, Buoniconto and De Rosa 2016; some works issued by using this methodological tool are Corona 2015; De Pasquale 2015; Iacobini et al. in press). We have identified four main *loci* for Path encoding, corresponding to two grammatical categories (Noun and Verb) and two functional categories (Adverbial and Adnominal). We analyzed the lexical features of the verbs (i.e. basic motion verbs: *eo* ‘to go’; verbs denoting caused motion: *duco* ‘to lead’; Manner verbs: *curro* ‘to run’, *vagor* ‘to roam’; Path verbs, within which we have distinguished source-oriented *linquo* ‘to leave’ from goal-oriented *venio* ‘to arrive’), as well as the semantics and the distribution of the spatial prefixes, the role of the prepositional phrases in conveying directional meaning, and more in general the way in which spatial information is distributed in the sentence.

4. Latin as a “classical” Satellite-Framed language

At the system level (i.e., at the level of morphological and lexical means the language makes available) Latin can be considered a typical Satellite-Framed language (cf. Corona 2015; Iacobini and Corona 2016). In the next three sections (4.1-4.3) we will briefly show the linguistic resources available (at the system level) in the Latin language for the expression of

the three main conceptual components that can be associated with Path: spatial orientation, deictic anchoring, and boundary-crossing encoding.

4.1 Spatial orientation

Spatial orientation, i.e. the oriented line covered by a moving Figure, is normally explicitly expressed through specifically dedicated linguistic means and is conditioned by the frame of reference the speakers of each language prefer to use.

The notion of frame of reference was introduced by Levinson (1996, 2003) in relation to the coordinate system that languages adopt to identify the location of an object. Levinson and his collaborators (cf. Brown and Levinson 2000; Levinson and Wilkins 2006) distinguish three main frames of reference: the relative, the intrinsic, and the absolute. Each frame of reference in a language can be associated with distinct linguistic expressions. In the intrinsic frame the location of an object is defined in relation to specific properties of the Ground; in the relative frame the axes of the human body of the perceiver are the main point of reference for the location of an object; in the absolute frame the location of an object is defined in relation to arbitrary fixed bearings (for instance, cardinal directions), or to bearings considered salient by a community of speakers, such as seacoast, upriver/downriver, uphill/downhill, human settlements.

Latin preferentially adopts the relative frame of reference. Spatial orientation can be expressed by a rich system of prefixes and prepositions (and also, albeit marginally, by spatial cases governed by the verb). In Table 1, we provide a schematic inventory of the main means deployed by Latin to encode spatial orientation on both the vertical and the horizontal axes.

Case (assigned by the preposition or by the prefixed verb)	HORIZONTAL AXIS		VERTICAL AXIS	
	Prefix	Preposition	Prefix	Preposition
ACCUSATIVE	<i>ante-</i>	<i>ante</i>		
	<i>post-</i>	<i>post</i>	<i>in-</i>	
	<i>inter-</i>	<i>inter</i>		
	<i>praeter-</i>	<i>praeter</i>	<i>e(x)-/de-</i>	<i>de</i>
ABLATIVE	<i>prae-</i>	<i>prae</i>		
	<i>pro-</i>	<i>pro</i>		
	<i>re(d)-</i>			
	<i>retro-</i>	<i>retro</i>		
ACCUSATIVE/ABLATIVE			<i>sub-</i>	<i>sub</i>
			<i>super-</i>	<i>super</i>
			<i>subter-</i>	<i>subter</i>
	Other: <i>dextra / dexter, -tera, -terum</i> ‘right’ <i>sinistra / sinister, -tera, -terum</i> ‘left’ <i>longe / procul</i> ‘far / from afar’ <i>propre</i> ‘close / closely’ <i>porro</i> ‘before’		Other: <i>insuper</i> ‘upwards’/ <i>desuper</i> ‘from above’ <i>subtus</i> ‘below, downwards’	

Table 1. Main means for spatial orientation encoding in Latin.

4.2 Deictic anchoring

Spatial deixis can be defined as the lexicalized information of the position of the speaker / scene, in which “scene” refers to the location where narrative attention is focused (cf. Lyons 1981: 170). In Classical Latin – as in most of the languages in which deictic anchoring is expressed – there is a distinction between ‘venitive’ and ‘andative’ deixis (i.e. ‘towards the speaker’ and ‘in the opposite direction to the speaker’). The most common way to express deixis is by the adverbs *obviam* ‘in the way (of)’, *adversus* ‘facing, opposite, against’, *contra* ‘in front of, on the other side, against’ (Cuzzolin 2010). Deixis can also be expressed by morphological means through adverbs derived from deictic demonstrative pronouns, cf. Table 2.

Demonstrative	GOAL		SOURCE		MEDIAN Segment	
HIC, HAEC, HOC	<i>huc</i>	towards here	<i>hinc</i>	from here	<i>hac</i>	through here
ILLE, ILLA, ILLUD	<i>illuc</i>	towards there	<i>illinc</i>	from there	<i>illac</i>	through there
ISTE, ISTA, ISTUD	<i>istuc</i>	towards there (2 nd person oriented)	<i>istinc</i>	from there (2 nd person oriented)	<i>istac</i>	through there (2 nd person oriented)

Table 2. Latin deictic demonstrative adverbs.

Differently from the Romance languages, the lexical opposition between the Latin verbs *eo* ‘to go’ and *venio* ‘to come’ is related to actionality, rather than to the spatial relationship between the speaker and the source of motion. *Venio* indicates a movement culminating in the reaching of a goal, that is to say, entailing an arrival. On the other hand, *ire* indicates a neuter movement, from both a deictic and a goal-reaching point of view (cf. Ricca 1993).

4.3 Boundary crossing

Boundary-crossing (called Conformation in Talmy’s 2000 terminology) is the conceptual category that refers to the linguistic encoding of the events where there are one or more boundaries to be crossed on the way of the moving Figure.

As shown by Slobin’s works (cf. Slobin 2004 amongst others), the crossing of a boundary has cognitive and linguistic salience in the encoding of motion events: the languages which belong to the Satellite-Framed type generally encode this category by items in functional opposition (English *in* vs. *into*: *I run in the park* vs. *I run into the park*), in contrast with Verb-Framed languages (Italian *in*: *Corro nel parco*) which lack dedicated linguistic items.

Grammatical descriptions of Classical Latin identify a functional opposition – linked to the boundary-crossing category – expressed by prefixes and prepositions taking either the accusative or the ablative in the encoding of Source and Goal.

In Goal expression, the preposition *ad* + accusative encodes no-final contact or spatial coincidence between Figure and Ground. Motion ending with the Figure contained in the Ground is expressed by *in* + accusative, as illustrated in Figure 1.



Figure 1. Goal encoding in Latin.

Other means for the encoding of Goal attested in our corpus are: i) simple dative with verbs indicating approach such as *adpropinquo* ‘to come near, approach’ (*muro oppidi portisque adpropinquarent* ‘They approached the wall and the doors of the city’, Caesar, *De bello gallico* VII.47.3); ii) simple accusative in crystallized uses, such as transitive verbs (*invado* ‘to invade’); or iii) verbs made transitive by their prefix (*Scythiam septemque triones horrifer invasit Boreas* ‘The horrible Boreas invaded the Scythia and the Northern region’, Ovid, *Metamorphoses* 1.64-65).

Latin can further articulate Source depending on the initial position of the Figure with reference to the Ground. In this respect, the Figure can be contained within the Ground, stand in a relation of contact with the Ground, or be in generic proximity to the Ground. The spatial meanings of such relations are respectively named *elative*, *delative* and *ablative*, from the Latin prepositions and homonymous prefixes *ex*, *de* and *ab*, which govern the ablative case.

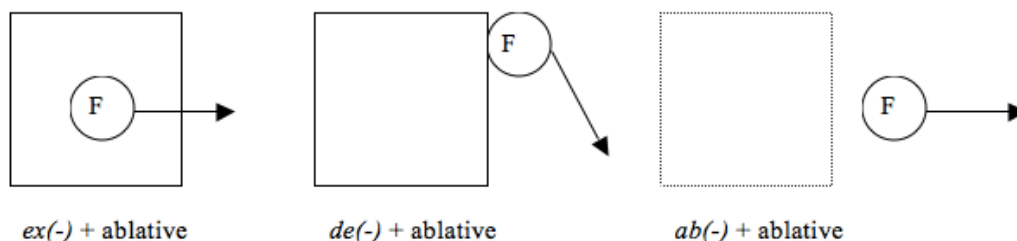


Figure 2. Source encoding in Latin.

Other types of source encoding attested in our corpus are: i) bare ablative in its original spatial value ‘motion away from a place’ (*canis fluit unda capillis* ‘water (lit. wave) streams from his grey hair’, Ovid, *Metamorphoses* I.266); ii) accusative with transitive verbs indicating abandonment (*ubi deseruit madidos septemfluus agros Nilus* ‘where the seven-steamed Nile left the wet fields’, Ovid, *Metamorphoses* I.422).

As an interim conclusion of this section, we can confirm that, at the system level, Classical Latin can be classified according to Talmy’s (2000) typology as a Satellite-Framed language. The prefixed verbs listed in Table 3 can be considered as another piece of evidence of the typological classification of Latin, since they show the rich array of spatial prefixes applied to a Manner of motion verb (*curro* ‘to run’) taken from the dictionary entries of the Oxford Latin Dictionary.

However, as we will see in the next section, our corpus-based analysis shows that, differently from a typical Satellite-Framed language, Latin tends to express simple Path and to blur the encoding of boundary crossing. Moreover, spatial prefixation of Manner verbs is much more limited than expected.

<i>adcurro</i>	‘to run or hurry to or up to’
<i>antecurro</i>	‘to run in front of’
<i>circumcurro</i>	‘to run or extend round’
<i>concurro</i>	‘to assemble at a run or in haste, hurry together’
<i>decurro</i>	‘to run down, hurry down’
<i>discurro</i>	‘to run off in several directions’
<i>excurro</i>	‘to run or rush out’
<i>incurro</i>	‘to rush or charge (at), make an attack (on) ’
<i>intercurro</i>	‘to run or hasten (from one place to another) ’
<i>introcurro</i>	‘to run or speed inside’
<i>occurro</i>	‘to run or hurry to meet’
<i>percurro</i>	‘to run, move quickly over or through’
<i>praecurro</i>	‘to run in front of others’
<i>procurro</i>	‘to run forward or ahead’
<i>recurro</i>	‘to run or hurry back’
<i>succurro</i>	‘to run or move quickly (under) ’
<i>supercurro</i>	‘to come up (to a person) at a run’
<i>transcurro</i>	‘to travel rapidly, hurry, run, etc., across (from one place to another) ’

Table 3. Prefixed verb with *curro* ‘to run’ (from the entries of the Oxford Latin Dictionary).

5. Latin as a “non-classical” Satellite-Framed language

Until very recently, we knew very little about motion expression in Classical Latin, as more attention was paid to Late Latin. The literature on the subject is very limited in number and quite recent: cf. Baldi (2006); Ferrari and Mosca (2010); Meini and McGillivray (2010); Brucale and Mocciano (2011); Brucale, Iacobini and Mocciano (2011); Corona (2015); Iacobini and Corona (2016); and the strongly theoretical committed analysis by Acedo Matellán (2010).

With respect to boundary-crossing encoding, our analysis shows an asymmetric behaviour of Goal and Source. We found a quite systematic use of *in* when the Figure performs a boundary-crossing movement entailing an actual entrance, as in (4) and (5), while *ad* is used to describe a generic arrival (6).

- (4) Caesar, *De bello gallico* VI.37.2
in castra in-rumpere conantur
 into camp(N.PL).ACC into-burst:PRS.INF begin(DEP):PRS.3PL
 ‘They begin to burst into the camp’
- (5) Caesar, *De bello gallico* VII.53.3
exercitum in castra re-duxit
 army(M).SG.ACC in camp(N.PL).ACC back-bring:PFV.3SG
 ‘[Caesar] brought back the army in the camp’
- (6) Caesar, *De bello gallico* V.22.2
cum ad castra venissent
 when to camp(N.PL).ACC arrive:SBJV.PPFV.3PL
 ‘When they arrived at the camp’

The same functional opposition does not hold for Source expression. The distinction on the basis of the relative vs. ablative value is blurred, with the consequence of an overlap, and an overgeneralization of Source elements. As the example in (7) shows, boundary-crossing at starting points of motion events can be encoded in the adnominal locus not only by the expected preposition *ex* but also by *de*, even in those cases in which the Ground (*pars sepulta*) is characterized by boundaries that are crossed by the Figure (*scorpius*).

- (7) Ovid, *Metamorphoses* XV.370
de parte sepulta scorpius ex-ibit
 from part(F).SG.ABL hidden:PTCP.PRF.F.SG.ABL scorpion(M).SG.NOM out-go:FUT.3SG
 ‘From the inner a scorpion will exit’

The functional overlap of prepositions expressing Source is even more evident in examples like the one in (8), where the departure from a place which cannot be conceived as a container (*tumulus*) is expressed by the preposition *ex*. Similar cases of the use of prepositions in Ancient Greek have been explained as a result of the so-called “Weakening of the Container Metaphor” (cf. Luraghi 2003: 315), a metaphor implying that the Ground corresponds to a space physically delimited by boundaries and that the Figure initially or finally coincides with a portion of said Ground.

- (8) Caesar, *De bello gallico* II.27.4
ex tumulo tela in nostros conicerent
 from hill(M).SG.ABL dart(N).PL.ACC against POSS.F.1PL.ACC throw:SBJV.IPFV.3PL
 ‘From the hill they threw darts against our soldiers’

In our corpus, we have also noticed the reverse phenomenon, that could be called extension or “Strengthening of the Container Metaphor”: in examples like the one in (9) the Source (*Menapii*) and the Goal (*nostri* ‘our army’) are expressed respectively by the prepositions *ex* and *in* (instead of the expected *ab* and *ad*) implying that populations and armies seem to be considered as containers, since they are perceived as homogeneous entities within which a Figure can be metaphorically included.

- (9) Caesar, *De bello gallico* VI.9.1
Caesar postquam ex Menapiis
 Caesar(M).NOM after out.from Menapi(M.PL).ABL
in Treveros venit
 into Treveri(M.PL).ACC arrive:PFV.3SG
 ‘Caesar, after he arrived from the Menapii[’s] to the Treveri[’s territories]’

Perhaps more compelling for an in-depth analysis of the preferred strategies of motion encoding employed in Latin are the data concerning the expression of complex Path and Manner of motion. It is worth recalling that the accumulation of Ground expressions around a Manner verb is a key characteristic of Satellite-Framed languages.

The most striking data arising from our analysis concern: i) the relative scarcity of Manner of motion verbs with respect to the typological expectations (cf. (10)-(24)); and ii) the almost complete lack of complex Path expression (cf. (25)-(28)).

Latin directional prefixes are equally distributed between Manner and other motion verbs with regard to the number of both types and tokens (if not overbalanced towards directional verbs). The cases in which a prefix is the only element used to add directional meaning to a Manner base (cf. (10)) are few in number.

- (10) Ovid, *Metamorphoses* XV.739
scinditur in geminas partes
 split:PRS.3SG.PASS into dual:F.PL.ACC branch(F).PL.ACC
circum-fluus amnis
 around-flowing river(M).NOM.SG
 ‘The river, flowing around, splits into two branches’

Even in the cases in which a Manner verb is available in the Latin lexicon, oftentimes this is not used if Manner is easily inferable from other elements in the clause: the verbs *navigo* ‘to sail’ and *fluo* ‘to flow’ might have been used in the examples in (11) and (12) instead of *transeo* and *eo*.

- (11) Caesar, *De bello gallico* III.11.2
si (...) navibus flumen trans-ire conentur
 if ship(M).PL.ABL river(N).SG.ACC through-go:PRS.INF try(DEP):SBJV.PRS.3PL
 ‘if they had tried to cross the river’

- (12) Ovid, *Metamorphoses* I.111
flumina lactis iam flumina
 river(N).PL.NOM milk(N).SG.GEN by.now river(N).PL.NOM
nectaris ibant
 nectar(N).SG.GEN go:IMPFV.3PL
 ‘By now, rivers of milk and nectar flowed’

As shown by Brucale, Iacobini and Mocciaro (2011), from which the following examples are taken, the cases in which the prefix is highly meaningful in Path expression are represented not by Manner verbs but by the set of verbs derived from *eo* ‘to go’, in which the very general meaning of the verbal base probably constitutes the reason for the prefix to play such a crucial role (13). Besides expressing direction, the prefix can also modify the argument structure of the base verb, for example licensing a direct object as in (14).

- (13) Phaedrus, *Fabulae* 4.23.7
red-ire in patriam
 back-go:PRS.INF into homeland(F)C.F.SG
voluit cursu pelagio
 want:PRS.3SG way(M).SG.ABL of.the.sea:M.SG.ABL
 ‘Homewards by shipping he wants to return’

- (14) Tacitus, *Germania* 21.1
proximam domum non invitati ad-eunt
 nearest:ACC.F.SG.SUP house(F).SG.ACC not invite:PTCP.PRF.M.PL.ABL to-go:PRS.3SG
 ‘and without invitation they go to the next house’

Other examples of the combination of directional prefixes with the verb *eo* taken from our corpus are provided in (15) and (16).

- (15) Caesar, *De bello gallico*, I.33.3
in provinciam ex-irent
 into province(F).SG.ACC out-go: SBJV.IPFV.3PL
 ‘going forth into the province’

(16) Caesar, *De bello gallico*, II.27.3

ascendere altissimas ripas sub-ire
 ascend:PRS.INF highest:ACC.F.PL.SUP bank(F).PL.ACC up-go:PRS.INF
iniquissimum locum
 disadvantageous:ACC.M.SG.SUP place(M).SG.ACC
 ‘ascending the highest banks, and coming up to a very disadvantageous place’

Moreover, a prefixed Manner base easily fades or loses its Manner meaning, as in the case of the verbs derived from *gradior* ‘walk’, as shown by the examples (17)-(19), taken from the Packard Humanities Institute 5 corpus. In (17) the unprefixated verb *gradior* is used in its original meaning (as can be noticed from the contrast with other Manner verbs), while in (18) and (19), prefixes convey a directional meaning (*e-*, *in-*) and the Manner value of the base verbs ‘to walk’ and ‘to fly’ is lost. This situation could suggest that, in contrast with prototypical Satellite-Framed languages, Manner information is not so salient in Latin motion events, and can be easily blurred by a more prominent directional meaning.

(17) Cicero, *De natura deorum*, II.122.4

alia animalia gradiendo alia serpendo
 some:N.PL.NOM animal(N).PL.NOM walking some:N.PL.NOM crawling
ad pastum accedunt alia volando
 to food(N).SG.ACC approach:PRS.3SG some:N.PL.NOM flying
alia nando
 some:N.PL.NOM swimming
 ‘some animals approach their food by walking, some by crawling, some by flying, some by swimming’

(18) Quintilianus, *Institutio Oratoria* XII.proem.4

tanta atque ita instructa nave
 such_big:F.SG.ABL and so well_found:F.SG.ABL ship(F).SG.ABL
hoc mare in-gressus
 this:N.ACC.SG sea(N).ACC.SG inside-gone:PTCP.PRF.DEP.NOM.M.SG
 ‘[Cicero] though the ship of such size and so well found in which he entered this sea’

(19) Caesar, *De bello gallico*, III.28.3

subito ex omnibus partibus silvae e-volaverunt
 suddenly out.from all:F.PL.ABL part(F).PL.ABL forest(F).SG.GEN out-rush:PFV.3PL
 ‘[they (sc. the enemies)] suddenly rushed out from all parts of the forest’

As already pointed out by Brucale (2011), not infrequent are those cases where the Manner meaning can be expressed by a verbal item different from the main directional verb. Among the constructions in which Manner is expressed outside the main verb, one of the most frequent is the one exemplified in (20) and (21), in which Manner is expressed in a nominal adjunct.

(20) Ovid, *Metamorphoses* II.772

passu=que incedit inerte
 step(M).SG=and move.forward:PRS.3SG lazy:M.SG.ABL
 ‘he moves forward by lazy steps’

- (21) Ovid, *Metamorphoses* II.772
in mare lassatis decidit alis
 into sea(N).SG.ACC fatigue:PTCP.PRF.F.PL.ABL fall.down:PRS.3SG wing(F).PL.ABL
 ‘(the bird) falls down with his fatigued wings into the sea’

In Classical Latin, there are also attestations of constructions linking a directional verb to the present participle of a Manner verb (*properanti* < *propero* ‘to act with haste, hurry, be quick’; *festinans* < *festino* ‘to act hurriedly, make haste’), which clearly prelude to the Romance construction “main verb + gerund (of Manner verb)” (cf. (22) and (23), from Brucale, Iacobini and Mocciaro 2011).

- (22) Sallust, *de Catilinae coniuratione* 57.3
igitur, [...], castra propere movit ac sub ipsis
 therefore camp(N.PL).ACC immediately move:PRS.3SG and under DEM:F.PL.ABL
radicibus montium consedit, qua illi
 root(F).PL.ABL mountain(M).PL.GEN sit:PFV.3SG where DEM:M.SG.DAT
de-scensus erat in Galliam properanti
 down-rise:M.SG.NOM be:IPFV.3SG Into Gaul(F).SG.ACC hasten:PTCP.PRS.M.SG.DAT
 ‘he immediately broke up the camp, and took his post at the very foot of the hills, at the point where Catiline's descent would be, in his hurried march into Gaul’

- (23) Phaedrus, *Fabulae* 3.19.9
hominem inquit quaero, et ab-iiit
 man(M).SG.ACC say:PRF.3SG look.for:PRS.1SG and away-go:PRF.3SG
festinans domum
 hasten:PTCP.PRS.M.SG.NOM house(F).SG.ACC
 ‘He answered briefly, as he ran, “Fellow, I’m looking for a man.”’

Directional prefixes can also combine with Path verbs, as in (24), where Path is already expressed in the verb root, and the prefix does not add any further directional information.

- (24) Ovid, *Metamorphoses* II.684-5
incustoditae pylios memorantur
 unwatched:F.PL.NOM of.Pylos.M.PL.ACC remind:PRS.3PL.PASS
in agros processisse boves
 into land(M).PL.ACC advance:PFV.INF kine(F).PL.NOM
 ‘his kine, forgotten, strayed away to graze over the plains of Pylos’

Differently from what can be expected from a typical Satellite-Framed language (and despite the available resources), Classical Latin tends to avoid the expression of multiple Path.

The most frequent pattern of motion event encoding in Classical Latin is the one in which there is a so-called semantic congruence (cf. Borrillo 1998) between the portion of Path encoded in the verbal locus by the prefix and the one encoded in the adnominal locus. The directional information is distributed in the sentence, and prefixes and prepositions in the same clause generally encode the same Ground, either lexically coinciding (25) or (less frequently) differing (27)-(28).

- (25) Caesar, *De bello gallico* II.8.5
suas copias ex castris
 POSS:F.PL.ACC force(F.PL).ACC from.out camp(N.PL).ABL

e-ductas *instruxerant*
 out-bring:PTCP.PRF.F.PL.ACC draw.up:PPFV.3PL
 ‘[the enemies] drew up their forces which they had brought out of the camp’

(26) Ovid, *Metamorphoses* I.569-70

Peneos *ab* *imo* *effusus*
 Peneus(N).NOM from bottom.M.SG.ABL out.flow:PTCP.PFV.PASS.M.SG.NO
Pindo
 Pindus(M).ABL
 ‘[the river] Peneus, flowing from the bottom of the Pindus’

(27) Ovid, *Metamorphoses* I.608

delapsa=que *ab* *aethere* *summo*
 glide.down:PTCP.PFV.F.SG.NOM=and from heaven(M).SG.ABL topmost:M.SG.ABL
constitit *in* *terris*
 stand.upon:PFV.3SG into earth(F).PL.ABL
 ‘From the dome of heaven she glided down and stood upon the earth’

Our corpus presents some instances of complex Path. In such cases, Source is preferentially encoded by the prefix, while Goal occupies the adnominal locus (and therefore is expressed in a more detailed way). Complex Paths are normally restricted to Grounds that are in contact (i.e., physically or conceptually understood as joined or close together in a spatial continuum), and in which only one boundary-crossing occurs (cf. Bohemeyer 2003; Filipović 2010).

(28) Ovid, *Metamorphoses* III.67

totum *de-scendit* *in* *ilia* *ferrum*
 all.N.SG.NOM down-rise:PRS.3SG into innards(N).PL.ACC sword(N).SG.NOM
 ‘the entire sword descends into the innards’

(29) Ovid, *Metamorphoses* VIII.796-97

sub-vecta *per* *aera* *curru*
 up-carry:PTCP.PFV.F.SG.ABL through air(N).PL.ACC chariot(M).SG.ABL
 ‘carried up through the air by a chariot’

To sum up, although the partial character of our corpus does not allow final generalizations, the examples presented thus far clearly indicate the presence in Latin usage of non-prototypically Satellite-Framed strategies.

The main tendencies displayed by Latin may be summed up in the following points:

- simple Path (no cumulation of satellites);
- encoding of the same portion of Path in the satellite and in the adnominal locus;
- bleaching of directional meaning of some prefixes (re-interpretation as actional markers when associated with Manner verbs);
- non-widespread use of Manner of motion verbs;
- emergence of Manner expression outside the main verb.

6. Conclusions

Even though Latin can be classified as a Satellite-Framed language on the basis of the available morphological means, our corpus analysis has revealed that it does not fall neatly within this typological group if preferred strategies are investigated at the usage level.

Studies addressing the problem of the presence of strategies that do not fit in with one typology are quite recent and, according to Nikitina (2013: 186) “little is known about the specific factors that determine the choice of a strategy in particular cases”.

Our research supports the need to investigate this issue together with the claim that “[t]he study of texts in motion event typology is crucial. Only by considering texts can we explore how encoding is shaped by language use. Particularly important for encoding is frequency of occurrence and frequency of co-occurrence” of spatial elements (Wälchli and Sölling 2013: 110).

At a methodological level, our main result is the neat distinction between system and usage: a rich set of morphological means is not a sufficient condition for assessing the preferred strategies of encoding.

At the descriptive level, we brought new findings on the encoding of motion in Classical Latin.

Finally, we suggest that our findings may shed new light on the possible pathways of typological change. We believe that the limited use of Manner verbs in motion expression, the lack of complexity of Path in dislocational motion encoding, together with the semantic congruence of Path expressed in the prefixed verb and the prepositional phrase we found in Latin may constitute conditions paving the way for the typical expression of dislocation motion in Verb-Framed languages, in which the function of indicating the direction is carried out by the verb, whereas prepositional phrases tend to express neutral meanings with respect to the static / dynamic distinction.

References

- Acedo Matellán, V. (2010) *Argument structure and the syntax-morphology interface. A case study in Latin and other languages*. Ph.D. thesis. Universitat de Barcelona.
- Aske, J. (1989) Path predicates in English and Spanish: A closer look. In: K. Hall (Ed.), *Proceedings of the Fifteenth Annual Meeting of the Berkeley Linguistics Society*, 1-14.
- Baldi, Ph. (2006) Towards a history of the manner of motion parameter in Greek and Indo-European. In: P. Cuzzolin & M. Napoli (Eds.), *Fonologia e tipologia lessicale nella storia della lingua greca, Atti del VI Incontro Internazionale di Linguistica Greca*. Milano: Franco Angeli, 13-31.
- Bohnmeyer, J. (2003) The unique vector constraint: The impact of direction changes on the linguistic segmentation of motion events. In: E. Van der Zee & J. Slack (Eds.), *Representing direction in language and space*. Oxford: Oxford University Press, 86-110.
- Borillo, A. (1998) *L'espace et son expression en français*. Paris: Ophrys.
- Brown, P. & S. C. Levinson (2000) Frames of spatial reference and their acquisition in Tenejapan Tzeltal. In: L. Nucci, G. Saxe & E. Turiel (Eds.), *Culture, thought, and development*. Mahwah: Erlbaum, 167-197.
- Brucale, L. (2011) Manner of motion verbs in Latin. Paper presented at *Historical-Comparative Linguistics in the 21st Century*, University of Pavia, 22-25 September 2011.
- Brucale, L., C. Iacobini & E. Mocciaro (2011) Typological change in the expression of motion events from Latin to Romance languages. Paper presented at *The 44th Annual Meeting of the Societas Linguistica Europaea*, Universidad de la Rioja – Logroño, 8-11 September 2011.
- Brucale L. & E. Mocciaro (2011) Continuity and discontinuity in the semantics of the Latin preposition *per*: a cognitive hypothesis. In: J. Helmbrecht & E. Verhoeven (Eds.), *Linguistic typology and language universals (Sprachtypologie und Universalienforschung)*. Special issue of *STUF* 64(2): 148-169.
- Corona, L. (2015) *Gli eventi di moto in diacronia. Variazione e continuità dal latino all'italiano*. Ph.D. thesis. Università di Salerno.
- Craig, C. G. (1993) Jakaltek directionals: their meaning and discourse function. *Languages of the World* 7(2): 23-36.

- Croft, W., J. Barðdal, W. B. Hollmann, V. Sotirova & C. Taoka (2010) Revising Talmy's typological classification of complex event constructions. In: H. C. Boas (Ed.), *Contrastive studies in Construction Grammar*. Amsterdam: John Benjamins, 201-235.
- Cuzzolin, P. (2010) How to move towards somebody in Plautus' comedies: Some remarks on the adverb *obuiam*. In: B. R. Page & A. D. Rubin (Eds.), *Studies in classical linguistics in honor of Philip Baldi*. Leiden-Boston: Brill, 7-21.
- De Pasquale, N. (2015) The “classical” way to encode motion. Path and manner expression in Ancient Greek. Paper presented at the *International Colloquium of Ancient Greek Linguistics*. Rome, 23-27 March 2015.
- Ferrari, G. & M. Mosca (2010) Some constructions of path: From Italian to some classical languages. In: G. Marotta, A. Lenci, L. Meini & F. Rovai (Eds.), *Space in language: Proceedings of the Pisa International Conference*. Pisa: Edizioni ETS, 317-338.
- Filipović, L. (2007) *Talking about motion: A crosslinguistic investigation of lexicalization patterns*. Amsterdam: John Benjamins.
- Filipović, L. (2010) The importance of being a prefix. Prefixal morphology and the lexicalization of motion events in Serbo-Croatian. In: V. Hasko & R. Perelmutter (Eds.), *New approaches to Slavic verbs of motion*. Amsterdam: John Benjamins, 247-266.
- Fortis, J. M. & A. Vittrant (2011) L'organisation syntaxique de l'expression de la trajectoire: vers une typologie des constructions. *Faits de Langues. Les cahiers* 3: 71-98.
- Grinevald, C. (2011) On constructing a working typology of the expression of path. *Faits de Langues. Les cahiers* 3: 43-20.
- Iacobini, C., A. Buoniconto, L. Corona & N. De Pasquale (in press) How should a “classical” Satellite-Framed language behave? Path encoding asymmetries in Ancient Greek and Latin. In: S. Luraghi, T. Nikitina & C. Zanchi (Eds.), *Space in diachrony*. Amsterdam-Philadelphia: John Benjamins.
- Iacobini, C. & L. Corona (2016) L'espressione della direzione del moto dal latino classico all'italiano antico. In: M. Fruyt, G. Haverling & R. Sornicola (Eds.), *Actes du XXVII^e Congrès international de linguistique et de philologie romanes (Nancy, 15-20 juillet 2013), Section 2 : Linguistique latine/linguistique romane*. Nancy, ATILF: 87-100 (<http://www.atilf.fr/cilpr2013/actes/section-2.html>).
- Iacobini, C., L. Corona, A. Buoniconto, & A. De Rosa (2016) A grid for decoding motion encoding. Manuscript. University of Salerno.
- Ibarretxe-Antuñano I. (2015) Going beyond motion events typology: The case of Basque as a verb-framed language. *Folia Linguistica* 49(2): 307-352.
- Levinson, S. C. (1996) Language and space. *Annual Review of Anthropology* 25: 353-382.
- Levinson, S. C. (2003) *Space in language and cognition*. Cambridge: Cambridge University Press.
- Levinson, S. C. & D. P. Wilkins (Eds.) (2006) *Grammars of space: Explorations in cognitive diversity*. Cambridge: Cambridge University Press.
- Luraghi, S. (2003) *On the meaning of prepositions and cases. The expression of semantic roles in Ancient Greek*. Amsterdam: John Benjamins.
- Lyons, J. (1981) *Language and linguistics. An introduction*. Cambridge: Cambridge University Press.
- Meini, L. & B. McGillivray (2010) Between semantics and syntax: spatial verbs and prepositions in Latin. In: G. Marotta, A. Lenci, L. Meini & F. Rovai (Eds.), *Space in language: Proceedings of the Pisa International Conference*. Pisa: Edizioni ETS, 384-401.
- Nikitina, T. (2013) Lexical splits in the encoding of motion events from Archaic to Classical Greek. In: J. Goschler & A. Stefanowitsch (Eds.), *Variation and change in the encoding of motion events*. Amsterdam: John Benjamins, 185-202.
- Ricca, D. (1993) *I verbi deitici di movimento in Europa: una ricerca interlinguistica*. Firenze: La Nuova Italia.
- Slobin, D. I. (1996) Two ways to travel: Verbs of motion in English and Spanish. In: M. Shibatani & S. A. Thompson (Eds.), *Grammatical constructions: Their form and meaning*. Oxford: Clarendon Press, 195-220.
- Slobin, D. I. (2004) The many ways to search for a frog: Linguistic typology and the expression of motion events. In: S. Strömquist & L. Verhoeven (Eds.), *Relating events in narrative. Vol. 2. Typological and contextual perspectives*. Mahwah: Lawrence Erlbaum Associates, 219-257.
- Slobin, D. I. (2005) Relating narrative events in translation. In D. Ravid & H. B.-Z. Shyldkrot (Eds.), *Perspectives on language and language development: Essays in honor of Ruth A. Berman*. Dordrecht: Kluwer, 115-129.
- Slobin, D. I. (2006) What makes manner of motion salient? Explorations in linguistic typology, discourse, and cognition. In M. Hickmann & S. Robert (Eds.), *Space in languages: Linguistic systems and cognitive categories*. Amsterdam: John Benjamins, 59-81.
- Talmy, L. (2000) *Toward a cognitive semantics: Typology and process in concept structuring. Vol. 2*. Cambridge: The MIT Press.
- Verkerk, A. (2014) Diachronic change in Indo-European motion event encoding. *Journal of Historical Linguistics* 4(1): 40-83.

- Verkerk, A. (2015) Where do all the motion verbs come from? The speed of development of manner verbs and path verbs in Indo-European. *Diachronica* 32(1): 69-104.
- Wälchli, B. & A. Sölling (2013) Building typology bottom-up from text data in many languages. In: J. Goschler & A. Stefanowitsch (Eds.), *Variation and change in the encoding of motion events*. Amsterdam: John Benjamins, 77-114.

Morphology: the base processor

Aysun Kunduracı
Yeditepe University
 kunduraca@gmail.com

Aslı Göksel
Boğaziçi University
 gokselas@boun.edu.tr

1. Introduction

Since the early 1980s, the Lexical Integrity Hypothesis (LIH) has been one of the cornerstones of morphological research that delineates the division between morphology and syntax. It is a set of principles that distinguish morphology from syntax and this division, and the constraints that lead to this division have been a significant tool for the arguments towards morphology as a separate component.

Among its various characterizations, there are two properties that seem related but are distinct:

- (i) Syntax cannot manipulate morphological structure (Lapointe 1980 among others): this first characterization implies that syntactic rules do not create lexemes and word forms and that syntactic outputs are not equivalent to morphological outputs.
- (ii) Syntax cannot enter word-internal structure (a.k.a. No Phrase Constraint, Botha 1984): the second characterization can be rephrased as the ban against syntactic items in word structure.

In this paper, we hold the view that the first property above is sufficient to characterize a word-formation component separate from a phrase-structure component, and the violation of the second property above does not pose a threat to this. We propose a system which excludes syntax from creating morphologically complex structures, but allows syntactic outputs within words. We first show that syntax cannot create lexemes and word forms, and that various well-known data support the implications of (i) but not of (ii), thus leading to a model of autonomous morphology (section 2). Next, we lay down what an autonomous morphology component does (section 3).

2. The limitations on syntactic operations in word structure

In this section, we summarize some well-known reasons discussed in the literature that argue against syntax-based word formation.

2.1 Syntax does not predict fixed affix order; fixed order despite ambiguity

The ordering of affixes may be fixed, irrespective of the presumed position of these functional projections in syntax.

2.1.1 Headless relative clauses

The various syntactic and semantic functions of the plural and possessive affixes in headless relative clauses in Turkish do not yield variable affix order. The sequence *V-relativizer-plural-possessive* is fixed, yet this word form is four-ways ambiguous as shown in (1),

2.1.3 The position vs. the scope of negation

The position of the negative marker in Turkish and Mongolian is fixed although it has scope ambiguity (from Göksel 1993):

- (5) a. *oku-t-ma-dı-m* vs. b. **oku-ma-t-tı-m*
 read-CAUS-NEG-PST-1.SG
 ‘I didn’t make him read’

(5a) has two interpretations as below, despite the fixed order of the negative marker:²

- (i) not > cause > read: I did not cause him to read
 ≠ ≠
 (ii) cause > not > read: I caused him not to read

2.1.4 The interpretation of voice morphology

In Kinande, the order of an applicative and reflexive is fixed, although the construction is ambiguous (from Alsina 1999):

- (6) *hum-ir-an* vs. **hum-an-ir*
 hit-APP-REFL
 (i) ‘hit X for each other’
 (ii) ‘hit each other for X’

2.2 Syntax does not predict variable affix order; variable order despite single interpretation

In Turkish, the plural (agreement) marker, *-lar*, and the past marker, *-DI*, have variable position, the sole difference being differences in register.

- (7) a. (*onlar*) *gid-iyor(-lar)-dı* (...Per.Num+Tense)
 they go-PROG(-PL)-PST
 ‘they were going’
 b. (*onlar*) *gid-iyor-du(-lar)* (...Tense+Per.Num)
 they go-PROG-PST(-PL)
 ‘they were going’

There are also inter paradigm differences between the behavior of 1st/2nd and 3rd Person in ordering tense, person/number values following aspect:

- (8) a. *gid-iyor-mu-sun-uz* (...Asp+Q+Per+Num)
 go-PROG-QP-2-PL
 ‘are you going’
 b. *gid-iyor-lar-mı* (...Asp+Num+Q)
 go-PROG-PL-QP
 ‘are they going’

² See also Li (1990) and Kelepir (2000) for similar observations.

2.3 Syntax does not predict slot competition or syncretism

There is an abundance of examples from position class languages where order is purely morphological. Below we present some of these.

In Pazar Laz, *-t* shows PL of subject, object, or both without changing position (from Öztürk and Pöchtrager 2011):

- | | | | |
|--------|------------------------------------|-------------|----------------------|
| (9) a. | <i>t'k'va</i> | <i>ma</i> | <i>ce-m-ç-i-t</i> |
| | you.PL.ERG | I.DAT | PV-P.1-beat-PST.2-PL |
| | 'you(PL) beat me' | | |
| b. | <i>si</i> | <i>şk'u</i> | <i>ce-m-ç-i-t</i> |
| | you.SG.ERG | we.DAT | PV-P.1-beat-PST.2-PL |
| | 'you(SG) beat <u>us</u> ' | | |
| c. | <i>t'k'va</i> | <i>şk'u</i> | <i>ce-m-ç-i-t</i> |
| | you.PL.ERG | we.DAT | PV-P.1-beat-PST.2-PL |
| | ' <u>you</u> (PL) beat <u>us</u> ' | | |

In Georgian *v-* (1.SUB) and *g-* (2.OBJ) are mutually exclusive as Anderson (1982: 597 example 18) shows:

- | | | | |
|---------|----------------|-------------------|---------------------|
| (10) a. | <i>v-xedav</i> | b. <i>g-xedav</i> | c. <i>g-xedav-s</i> |
| | 'I see him' | 'I see you' | 'he sees you' |

In Nimboran, the same slot can be occupied by values that belong to different features, and a single item can block affixes in more than one slot (cf. Inkelas 1993: 589):

- | | | | | | |
|------|------|---------------|----------|----------|-------|
| (11) | 0 | 1 | 2 | 3 | |
| | Root | PL.SUB | DUAL.SUB | MASC.OBJ | |
| | | PL.OBJ | PART | | |
| | | -----Dur----- | | | |

These are a few of the examples cited in the literature, examples that show that affixes of different paradigms, person, number, voice morphology, negative, interrogative can be fixed, irrespective of syntactic and semantic function, or free, again, irrespective of semantic and syntactic function. If syntax determined affix positions and combinatorial conditions, we would expect invariable parallelism between syntactic operations and morphological elements and we would not need an autonomous morphological component. In the face of the examples above, the only logical conclusion is that we need a component that forms words.

3. Autonomous morphology: scope and function

We claim that an autonomous morphological component is responsible for the following:

- (i) Creating candidate-lexemes regardless of their complexity/simplicity (cf. Kunduracı 2013),
- (ii) Using and restricting a variety of base types in morphological operations (cf. Pounder 2000) (These base types are not only roots and stems but also word forms, syntactic phrases, and prosodic phrases),
- (iii) Creating bases only for morphological well-formedness, e.g. concatenative purposes (specific stem forms to enable derivation or inflection),
- (iv) Organizing paradigmatic relations (slots).

3.1 Autonomous morphology creates candidate lexemes

As discussed in previous literature, Turkish possessive phrases (12a) and Noun-Noun compounds (12b) show a superficial similarity: both involve the *-(s)I* suffix in the head element, which functions as the possessive marker (POSS) in possessive phrases and as the compound marker (CM) in compounds. Two of the well-known structural differences between a possessive phrase and a compound are reversibility and separability: unlike compounds, possessive phrases can be reversed and their constituents can be separated (12c). We assign such properties of compounds to their structural status: compounds must be candidate-lexemes created by autonomous morphology (cf. Kunduracı 2013) unlike possessive phrases.

- (12) a. [*çocuğ-un kitab-ı*] *nerede* (Possessive Phrase)
 child-GEN book-POSS where
 ‘Where is the child’s book?’
 b. [*çocuk kitab-I neredede*] (Compound)
 child book-CM where
 ‘Where is the child book (childrens’ book)’
 c. *kitab-I neredede çocuğ-un/*çocuk* __
 book-POSS where child-GEN
 ‘Where is the child’s book?’

3.2 Autonomous morphology operates on and restricts its own bases

Bases for operations may be chosen according to lexical, semantic, phonological, prosodic, syntactic or purely morphological properties. However, it is always morphology that determines such conditions on bases (of morphological operations). Turkish and Yakut display differences in word size, i.e. there is a limit that, once it is reached, auxiliaries are used for further concatenations (cf. Göksel 1998):

- (13) Turkish
 a. **var-mış-acağ-ı*
 arrive-PRF-FUT.NZ-3.POSS
 Int.: ‘that (s)he/it will have arrived’
 b. *var-mış ol-acağ-ı*
 arrive-PRF AUX-FUT.NZ-3.POSS
 ‘that (s)he/it will have arrived’
 Yakut
 c. **si-i-bit-im*
 eat-AOR-PST-1.SG
 Int.: ‘I had been eating’
 d. *si-i olor-but-um*
 eat-AOR AUX-PST-1.SG
 ‘I had been eating’

The auxiliary verb in (13b) is necessary for the combination of PRF and FUT.NZ in Turkish. This combination is not possible within the same word morphologically, as shown in (13a). Yakut also displays a similar case (13c, d).

3.3 Autonomous morphology may use syntactic and prosodic units as a base

It is well-known that morphology can select a stem or a word as its base (Ralli 2013: 79, cited in Bağrıaçık et al. frth., *i.a.*):

- (14) Greek
 a. [WORD [STEM [STEM STEM] –CM– [STEM STEM]]-INFLECTION]
 b. [WORD [STEM STEM] –CM– [WORD STEM-INFLECTION]]

It is also known that units in the lower levels of the Prosodic Hierarchy (Nespor and Vogel 1986) can be selected as bases for reduplication (Inkelas and Zoll 2005). What becomes clearer is that morphological operations can also involve higher prosodic units (e.g. Phonological Phrase) and syntactic phrases as bases or part(s) of bases. We present examples for both types in the next sections.

3.3.1 Prosodic Phrases as bases

Reduplication of a prosodic base for emphasis, a derivational process outputting emphatic adverbial constructions, is attested in Turkish and Greek. In Turkish, all phonological phrases can be doubled and located in the post-verbal position, creating emphatic adverbs, which is a derivational process (cf. Göksel, Kabak and Revithiadou 2013):

- (15) { [[Ev-e] } *gid-iyor-uz*] {ev-e} !
 home-DAT go-PROG-1.PL house-DAT
 ‘We’re going to HOME’

Such an operation occurs, even if the Phonological Phrase is part of an exocentric compound:

- (16) {*eşek arı-sı*} *bu eşek!*
 donkey bee-CM this donkey
 ‘This is a wasp!’

These are not cases of copy and elide, as neither the elided part nor the remnant is a syntactic constituent: they belong to an exocentric (lexicalized) compound.³

In a similar vein, prosodic units inside morphological units (here affixes) serve as bases for higher level prosodic operations:

- | | |
|--|--|
| <p>(17) a. <u>Presentational Focus</u>
 <i>ye-miş-lér-di</i>
 eat-PRF-3.PL-PST
 ‘They’d eaten (it).’
 (from Sebüktekin 1984)</p> | <p>b. <u>Contrastive focus</u>
 <i>ye-MİŞ-ler-di</i>
 eat-PRF-3.PL-PST
 ‘They HAD eaten (it).’</p> |
| <p>(18) a. <u>Presentational Focus</u>
 <i>ya-slán-in</i> L%
 lean.back-2.IMP
 ‘Lean back.’
 (from Göksel and Güneş 2013)</p> | <p>b. <u>Comma intonation</u>
 <i>YASlan-in</i> H%
 lean.back-2.IMP
 ‘Lean back...’</p> |

In (17b), a high (proposition) level prosodic operation operates on a base which is a word form and focal stress targets a single morpheme, distinguishing the construction from the one with presentational focus (17a). In (18b), comma intonation, which again operates on propositions, targets the first syllable of a word form.

³ We follow Kunduracı (2013) in that we consider lexicalized compounds to undergo the same process of morphology that also outputs novel compounds and that there is one additional, process in the case of lexicalized compounds, i.e. entering the lexicon with or without a new meaning.

3.3.2 Syntactic phrases as bases

Turkish phrasal compounds host bona fide phrases:

- (19) a. { { [ev-e gid-ecek] } {düşünce} }-si
 home-DAT go-FUT thought-CM
 ‘the thought that he is going to go to home’
 b. { { [nerede mi ol-duğ-u] } {soru} }-su
 where QP be-NZ-3.POSS question-CM
 ‘the question about where he is’

The nonhead in this compound is a phrase, not a quotation of type N. Quotations and phrases in such compounds can be distinguished syntactically, semantically and prosodically in Turkish (see Göksel 2015). As in Kunduracı (2013), we consider that such forms involve a syntactic item as the first base, which undergoes a morphological operation, i.e. compounding, and that this does not mean that the compounding operation is syntactic; the two elements are still inseparable:

- (20) { {XP} {Y} }
 {syntactic unit} {lexeme stem} = morphological output

(20) represents the structure of a phrasal compound: XP represents the first base, which is complex, i.e. a syntactically created element, and Y, the second base, which is a lexeme stem. These two are compounded by means of a morphological process outputting morphologically complex forms. In this way, not only the fact that one of the compounded elements is syntactic but the fact that the compounded elements are not separable can be accounted for as well.

3.3.3 Support for phrasal bases from nonconcatenative morphology

The examples we have given so far involved concatenative morphology. In this section, we show that nonconcatenative morphology may also build on phrasal bases.

In Karajá, a clausal structure undergoes a nonconcatenative morphological process: base modification in the formation of a relative clause. Note that the only exponent of the relative clause morphology here is tonal change (from Ribeiro 2006):⁴

- (21) a. [d̥ori ɔ-d-ɔ-ɔrɔ=d-e]
 white 3-CTPT-INTR-go.ashore=CTPT-IMPRF
 ‘The white man came ashore.’
 b. d̥ori ɔ-d-ɔ-ɔrɔ=d-é
 white 3-CTPT-INTR-go.ashore=CTPT-IMPRF.REL
 ‘the white man that came ashore’

3.4 Autonomous morphology may create bases only for combinatorial purposes

In Turkish, there are cases where affixes are necessary only to provide bases for further morphological derivations. In (22), the expression in (a) does not mean ‘something with some

⁴ Interrogative morphology is also known to involve base modification (intonational change), see Cheng and Rooryck (2000), Göksel et al. (2009).

piece of wrapping’, and the expression in (b) does not mean ‘something with (some piece of) building’, which would be the expected meaning if the affix *-i* were interpreted independently. Rather, here the nominalizer (NZ) is clearly not added for semantic reasons:

- (22) a. *sar*(-i)-li*
wrap-NZ-*II*
‘wrapped’
- b. *yap*(-i)-li*
do-NZ-*II*
‘made/bodied’

The form *sarı-*, for example, is not (necessarily) an existing lexeme with the meaning ‘wrapping’, but is a novel form required to allow the further operation above.

Thematic vowels in many languages serve this purpose too. In Latvian, a thematic vowel (*-i* below) is necessary when producing word forms, as shown by Haspelmath and Sims (2010).

- (23) a. *gulb-i-m*
swan-TV-DAT
‘to (the) swan’
- b. **gulb*

Note that the same applies to compound markers that have no meaning but function (cf. Ralli 2008; Kunduracı 2013).

3.5 Autonomous morphology organizes paradigms

Kunduracı (2013) proposes that Turkish N-N compounds, like (24a), are produced in a word-formation paradigm and that the compound marker (CM) is in a paradigmatic relation with certain derivational affixes, as exemplified in (24b). It is important that the compound marker would be required semantically, considering the meaning of the compound *elma ağac-ı* ‘apple tree’, which is identical in both (a) and (b), whereas (b) lacks the compound marker. This case is assigned to a paradigmatic relation involving compounding and derivational affixations:

- (24) a. *elma ağac*(-ı)*
apple tree-CM
‘a/the apple tree’
- b. *elma ağaç(*-ı)-li*
apple tree-*CM-*li*
‘with an/the apple tree’

Another paradigmatic point to make is for inflection: although the 1st person possessive marker, *-(I)m* would be expected semantically, as shown below, it is not affixed after the 2nd person possessive marker, *-(I)n*, due to, again, a paradigmatic slot competition:

- (25) a. *ben-im [sen-in resm-in]-*im*
I-GEN you-GEN picture-2.POSS-*1.POSS
for ‘your picture that I have’
- (from Kunduracı 2013)

In this section, we hope to have shown that morphology arranges word size and restrictions on word structure. It may use stems, syntactic constituents, and prosodic constituents as a base. This means that not only stems but also phrasal and prosodic constituents (i.e. constituents of morphology-external components) can be input to morphological operations such as affixation, compounding, reduplication, base modification, and this spans derivational and inflectional morphology (contra arguments in weak lexicalism, Aronoff 1976; Anderson 1982, 1992). Morphology also organizes paradigms.

An autonomous morphological system then can explain:

- (i) obvious differences between syntactic and morphological outputs
- (ii) morphological complexes
- (iii) morphologically required stem forms
- (iv) formal similarities between derivational and inflectional processes
- (v) paradigms⁵

4. Conclusions

Morphological principles and syntactic principles (i.e. operations, inputs, outputs) are not identical (as in Aronoff 1976, 1994; Anderson 1982, 1992; Zwicky 1984, 1992; Di Sciullo and Williams 1987; Spencer 1991; Beard 1995; Stump 2001; Ackema and Neeleman 2007; Di Sciullo 2009). Morphology works on bases and either provides lexemes for the lexicon or provides new bases, which are not lexemes, for further operations. In terms of concatenation conditions and outputs, morphology seems less flexible than syntax. But in terms of the inputs, morphology seems more flexible than syntax (wider variety of input/base types).

There are post-syntactic morphological operations as well as pre-syntactic ones (cf. Aronoff 1976; Anderson 1992; Lieber and Scalise 2006). However, this does not lead to split morphology (cf. Booij 1994, 1996; Kunduracı 2013). There is one autonomous morphology (as in Aronoff 1994; Sadock 2012; Pounder 2000; Göksel 2006; Kunduracı 2013 among others).

We also consider two alternatives for the autonomous morphology and its operations: there could be (i) simultaneous representations of syntax and morphology (cf. Ackema and Neeleman 2004; Sadock 2012), or (ii) feeding between syntax and morphology, but without a strict order (cf. Lieber and Scalise 2006; Kunduracı 2013). The next task would then be to decide what kind of data could be used in understanding this.

Going back to the Lexical Integrity Hypothesis, we hope to have shown that syntax cannot manipulate morphology, nor determine the output conditions of morphology, i.e. syntax cannot create words, but can supply constituents for words.

Acknowledgements

We thank TÜBİTAK (grant number: 114C064) for supporting this study.

References

- Ackema, P. & A. Neeleman (2004) *Beyond morphology: interface conditions on word formation*. Oxford: Oxford University Press.
- Ackema, P. & A. Neeleman (2007) Morphology ≠ Syntax. In: G. Ramchand & C. Reiss (Eds.), *The handbook of linguistic interfaces*. Oxford: Oxford University Press, 325-352.
- Alsina, A. (1999) Where is the mirror principle? *The Linguistic Review* 16: 1-42.
- Anderson, S. R. (1982) Where is morphology? *Linguistic Inquiry* 13: 571-612.
- Anderson, S. R. (1992) *A-morphous morphology*. Cambridge: Cambridge University Press.
- Aronoff, M. (1976) *Word formation in Generative Grammar*. Cambridge, Massachusetts: MIT Press.
- Aronoff, M. (1994) *Morphology by itself: stems and inflectional classes*. Cambridge: MIT Press.
- Bağrıaçık, M., Göksel, A. & A. Ralli (frth.) The effects of language contact on compound structure: the case of Phrasiot Greek.
- Beard, R. (1995) *Lexeme-morpheme base morphology*. Albany: State University of New York Press.
- Booij, G. (1994) Against Split Morphology. In: G. Booij & J. van Marle (Eds.), *Yearbook of morphology 1993*. Dordrecht: Kluwer, 27-50.

⁵ Although productivity is also an important task of morphology, it is not discussed in the present study.

- Booij, G. (1996) Inherent versus Contextual Inflection and the Split Morphology Hypothesis. In: G. Booij & J. van Marle (Eds.), *Yearbook of morphology 1995*. Dordrecht: Kluwer, 1-16.
- Botha, R. (1984) *Morphological mechanisms*. Oxford: Pergamon Press.
- Bresnan, J. & S. A. Mchombo (1987) Topic, pronoun, and agreement in Chichewa. *Language* 63: 741-782.
- Di Sciullo, A. M. (2009) Why are Compounds a Part of Human Language? A View from Asymmetry Theory. In: R. Lieber & P. Štekauer (Eds.), *The Oxford handbook of compounding*. Oxford: Oxford University Press, 145-177.
- Di Sciullo, A. M. & E. Williams (1987) *On the definition of word*. Cambridge: MIT Press.
- Cheng, L. & J. Rooryck (2000) Licensing wh-in-situ. *Syntax* 3: 1-19.
- Göksel, A. (1993) What's the Mirror Principle? *SOAS Working Papers in Linguistics and Phonetics* 3: 168-205.
- Göksel, A. (1998) Word Length. In: G. Booij, A. Ralli & S. Scalise (Eds.), *Proceedings of the first Mediterranean Morphology Meeting*. Patras: University of Patras, 190-200.
- Göksel, A. (2006) Pronominal participles in Turkish and Lexical Integrity. *Lingue e Linguaggio* 5: 105-125.
- Göksel, A. (2015) Phrasal Compounds in Turkish: Distinguishing Citations from Quotations. In: C. Trips & J. Kornfilt (Eds.), *Phrasal compounds (Language typology and universals, STUF 68)*. Berlin: De Gruyter, 359-394.
- Göksel, A., Kelepir, M. & A. Üntak-Tarhan (2009) Decomposition of Question Intonation: The Structure of Response Seeking Utterances. In: J. Grijzenhout & B. Kabak (Eds.), *Phonological domains: universals and deviations*. Berlin, New York: Mouton de Gruyter (Interface Explorations), 249-286.
- Göksel, A. & G. Güneş (2013) Expressing Discourse Structure & Information Structure within the WORD. Poster presented at the *9th Mediterranean Morphology Meeting*, Dubrovnik, 15-18 Sept.
- Göksel, A., Kabak, B. & A. Revithiadou (2013) Prosodically constrained non-local doubling. *The Linguistic Review* 30: 185-214.
- Haspelmath, M. & A. D. Sims (2010) *Understanding morphology* (2nd Ed). London: Hodder Education.
- Inkelas, S. (1993) Nimboran position class morphology. *NLLT* 11: 559-624.
- Inkelas, S. & C. Zoll (2005) *Reduplication: doubling in morphology*. Cambridge: CUP.
- Kelepir, M. (2000) Scope of Negation: Evidence from Turkish NPIs and Quantifiers. *Proceedings of GLOW Asia II*, Nagoya: Nanzan University, 213-231.
- Kunduracı, A. (2013) *Turkish noun-noun compounds: a process-based paradigmatic account*. Ph.D thesis. University of Calgary.
- Lapointe, S. (1980) *The theory of grammatical agreement*. Ph.D thesis. University of Massachusetts.
- Li, Y. (1990) X⁰-binding and verb incorporation. *Linguistic Inquiry* 21: 399-426.
- Lieber, R. & S. Scalise (2006) The Lexical Integrity Hypothesis in a new theoretical universe. *Lingue e Linguaggio* 5: 7-32.
- Nespor, M. & I. Vogel. (1986) *Prosodic phonology*. Dordrecht: Foris.
- Öztürk, B. & M. Pöchtrager (2011) *Pazar Laz*. München: LINCOM.
- Pounder, A. (2000) *Processes and paradigms in word-formation morphology*. Berlin: Mouton de Gruyter.
- Ralli, A. (2008) Compound markers and parametric variation. *Language Typology and Universals (STUF)* 61: 19-38.
- Ralli, A. (2013) *Compounding in Modern Greek*. Dordrecht: Springer.
- Ribeiro, E. R. (2006) Subordinate Clauses in Karajá. *Boletim do Musu Paraense Emílio Goeldi. Ciências Humanas* 1: 17-47.
- Sadock, J. M. (2012) *The modular architecture of grammar*. Cambridge: Cambridge University Press.
- Sebüktekin, H. (1984) Turkish Word Stress: Some Observations. In: A. Aksu-Koç & E. Taylan (Eds.), *Proceedings of the second International Conference on Turkish Linguistics*. Istanbul: Boğaziçi Üniversitesi Yayınları, 295-307.
- Spencer, A. (1991) *Morphological theory*. Oxford: Blackwell.
- Stump, G. (2001) *Inflectional morphology*. Cambridge: Cambridge University Press.
- Zwicky, A. M. (1984) Reduced words in highly modular theories: Yiddish Anarthrous locatives reexamined. *Ohio State University Working Papers in Linguistics* 29: 117-126.
- Zwicky, A. M. (1992) Some Choices in the Theory of Morphology. In: R. Levine (Ed.), *Formal grammar: theory and implementation*. Oxford: Oxford University Press, 327-371.

Lexical blends and lexical patterns in English and in American Sign Language

Ryan Lopic
University of California, San Diego
 rlepic@ucsd.edu

1. Introduction

Lexical blends are words that have been coined through the fusion of parts of other words, and lexical blending is the creative process through which new blends are made. Some blends, like *brunch* and *motel*, are widely known and well-established, while others, like *cronut* 'a kind of sweet pastry', from *croissant+donut*, and *mansplaining* 'patronizing explanation', from *man+explaining*, are relatively novel creations. Lexical blends therefore provide an excellent opportunity to examine the relationship between how conventional words are formed and how novel words are created¹. Accordingly, this paper seeks to answer two questions about the role of lexical blends in morphology. First, what do blends and blending reveal about the morphological systems of individual languages? Second, when taken seriously, what consequences do blends hold for theories of morphological structure in human language? To answer these questions, this paper presents examples of lexical blends from English, as a representative spoken language, and American Sign Language (ASL), as a representative sign language. I argue that, both in speech and in sign, lexical blending is an analogical process that both exploits and creates paradigmatic relationships among whole words. The examples presented here also suggest that lexical blend constructions can be strikingly similar across languages, even those with quite different structural properties.

Perhaps more than other neologisms, lexical blends are highly salient as new words in English (cf. Metcalf 2002). Speakers seem to hold a wide range of opinions about the suitability of individual blends and of the process of blending. For example, two "pop-linguistics" articles from 2013 speak out against the (perceived) increasing popularity of lexical blending (Figure 1). Citing examples like *bridezilla*, *manscaping*, *chillax*, and *staycation*, and lamenting that many blends are simply not that funny, the authors conclude that lexical blending is a newly popular phenomenon, a viral trend that, they hope, will quickly pass.

¹ Following Halle (1973), many studies of word-formation seek to procedurally build up complex words from smaller parts, and so "word-formation" refers to the process through which a target word like *transformational* is formed from the elements *trans-*, *form*, *-at*, *-ion*, and *-al*. This view of word-formation does not explicitly distinguish between the formation (=derivation) of an established, conventional word and the formation (=creation) of a novel, previously-un-encountered word. Separate terminology is needed to distinguish between different senses of "word-formation". Zwicky and Pullum (1987) and Miller (2014) therefore distinguish "plain/core" morphology from "expressive" morphology, and Ronneberger-Sibold (1999), uses the more neutral terms "word-formation" and "word-creation" to make a similar distinction. Here I use "word-formation" to refer to the description of static sub-lexical structure in established words, and "word-creation" to refer to the process of coining a new word.

THE GOOD WORD LANGUAGE AND HOW WE USE IT MARCH 6 2013 5:37 AM

Please Do Not *Chillax*

Adjoinages and the death of the American pun.

By Simon Akam



I am appalled by *Bridezillas*. I should make it clear that I have never seen an episode of the reality show. I hate *Bridezillas* for one simple reason: *Bride* does not rhyme with *god*. Ergo, *Bridezillas* is not a functioning pun.

The point is significant because *bridezilla* appears to be symptomatic of a wider malaise: the death of the American pun, replaced by something grosser, dumber, uglier. Examples abound: Take one of the most read websites in the world, *Wikipedia*, a "pun" on *encyclopedia* that shares nothing but its suffix. Or *techpreneur*, the loathsome fusion of *technology* and *entrepreneur*.

Likewise *mansplain*, a coinage popular with Internet feminists that adroitly glosses a man addressing a woman in a condescending fashion (e.g., "Akam mansplains that *mansplain* is not a functioning pun.") but is still not a functioning pun. *Manscaping*, the removal of all or part of male body hair, is better—there is at least assonance between the vowel sounds in *man* and *land*—but as a pun it remains perilously


7.8k 0

http://www.slate.com/articles/life/the_good_word/2013/03/chillax_wikipedia_and_bridezilla_are_not_puns_against_adjoinages.html

BROW BEAT SLATE'S CULTURE BLOG JULY 19 2013 11:29 AM

Sharknado, Cronut ... Is This the Summer of the Neolexic Portmanteau?

By L.V. Anderson



101 38 0

Dominique Ansel's witless pastry.

Photo illustration by Andrew Burton/Getty Images

http://www.slate.com/blogs/browbeat/2013/07/19/sharknado_cronut_and_the_summer_of_the_neolexic_portmanteau.html

Figure 1: Two pop-linguistics articles discussing blends in English: "Please do not *chillax*: Adjoinages and the death of the American pun" (left) and "*Sharknado*, Cronut ... Is this the summer of the neolexic portmanteau?" (right)

However, the process of blending two words together to create a new one, and even the meta-linguistic judgment that blending is a new phenomenon, are themselves not that new. Linguists have been discussing the novelty of blending for at least a century: In the prefatory note to her dissertation, for example, Pound admits that the most interesting section is likely to be the one "dealing with the present-day vogue of blend formations", noting that "it seems time that specific attention be called to the contemporary popularity of blends" (1914: ii). Another example can be found in Bryant's (1974) discussion of lexical blends in *American Speech*, the title of which is the simple declaration that "blends are increasing". It is also relatively easy to find examples of historical lexical blends in the Oxford English Dictionary online, for example *foolosopher*, from *fool*+*philosopher*, suggesting that, regardless of whether blends are truly on the rise, the mechanism of blending has been used to create new words in English since at least the Early Modern period:

- (1) a. *Suche men..that in deede are archdoltes, and woulde be taken yet for sages and philosophers, maie I not aptely calle them foolelosophers* (1549)
- b. *What stand yee idle my fooleosophers* (ca. 1600)
- c. *A fine foolosopher!* (1694)

<http://www.oed.com/view/Entry/72672>

In the domain of morphology, the enduring perception that lexical blends are unpredictable novelties seems to have precluded a systematic analysis of lexical blending as a productive morphological process. Many accounts view lexical blending as a marginal, peripheral, or extra-grammatical word-formation process, or even deny the reality of blending outright (e.g., Scalise 1984; Zwicky and Pullum 1987; Spencer 1991; Marantz 2013). However, recent work has shown that though lexical blend structure is probabilistic and gradient, rather than deterministic and categorical, it is indeed conditioned by prosodic and semantic considerations (e.g., Bat-El 2000; Gries 2004; Renner, Maniez, and Arnaud 2012; Arndt-

Lappe and Plag 2013; Bauer, Lieber, and Plag 2013). It seems that lexical blending can therefore be counted among the phenomena that an adequate theory of human language should be expected to address.

Beyond whether or not it is possible to correctly predict the formal structure of individual blends, one aspect of lexical blending that is of particular interest for morphology is the fact that some lexical blends can come together to form families of related blend words, giving rise to new morphological patterns (see Berman 1961; Lehrer 1998; Kemmer 2003; Booij 2010; Lepic 2015). The classic example concerns the word *Watergate* which, through reanalysis, has come to serve as the basis for a number of *gate* words, including *nipplegate*, *deflategate*, and *gamergate*. These new words all have in common that they reanalyze and repurpose the *gate* from *Watergate* in order to name political scandals and pop culture controversies.

A less well-established set of examples involves the recent blend *cronut*, already mentioned above. The *cronut* is a hybrid pastry with characteristics of both a *croissant* and a *donut*. Bakers hoping to cash in on the popularity of the *cronut* have also created their own "knockoff" pastries with corresponding blend names, including the *dossant*, from *donut+croissant*; the *crullant*, from *cruller+croissant*; the *churron*, from *churro+macaron*; and the *cronot*, from *cronut+not*. These words have all apparently been formed on analogy to the original blend *cronut*, combining and recombining words to create a product that is quite similar to, yet legally distinct from, their source of inspiration.

The remainder of this paper analyzes this tendency for certain lexical blends to serve as a template from which other blends may be created, by drawing on examples from English and ASL. Lexical blends have not yet, to my knowledge, been described in a sign language. As I suggest in section 2, this may be because many blends fit into the morphological patterns of sign languages so seamlessly that they are not even remarkable as new words to the same extent that blends are in English. In section 3, I argue that the facts about lexical blend families in English and in ASL can be straightforwardly accounted for under the theory of Construction Morphology (Booij 2010), which anticipates the notion of analogical motivation in morphology.

2. The structure of blends in English and in ASL

Though English and ASL are typologically quite different languages, in part because they are expressed in distinct perceptual modalities (see Sandler and Lillo-Martin 2006), both languages contain words that can be analyzed as having been created through the recombination of parts of other words. In addition to the examples already mentioned above, in English, examples of lexical blends include *sharrows* 'marks indicating lanes to be shared by motorists and cyclists', from *share+arrows*; *webinar* 'an online lecture or class', from *web+seminar*; and *glamping* 'luxurious accommodations at scenic vacation destinations', from *glam(orous)+camping*. Following Arndt-Lappe and Plag (2013) and Lepic (2015), these blends can be analyzed as incorporating the segmental material of one word into the overall prosodic/segmental frame provided by the other. This can be formally represented as in Figure 2, where the words *share* (Figure 2a) and *arrows* (Figure 2b) are aligned so as to share the stressed syllable ['eɪ.] in the resulting blend *sharrows* (Figure 2c).

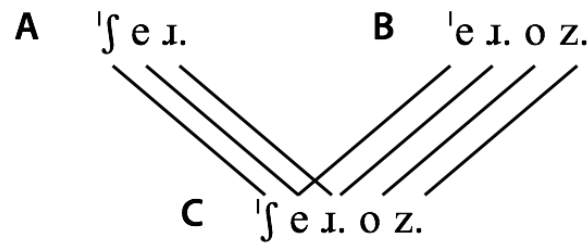


Figure 2: Phonological elements of the English words (a) *share* and (b) *arrows* combine to form the blend (c) *sharrows*.

In ASL, blends also incorporate phonological sub-constituents from one word into the overall prosodic frame provided by another.² Some examples are the sign glossed as TRIPPING³ 'to be on drugs', from TRAVEL⁴+INVENT⁵; HEARING-MINDED⁶ 'to uncritically embrace the values of the hearing majority', from THINK⁷+HEARING⁸ (see Padden and Humphries 1988; Wilcox 2000); and even one of the handful of variant signs for MORPHOLOGY, from WORD⁹+MEANING¹⁰. In each of these examples, the resulting blend sign retains the overall movement pattern of one of its constituent signs, yielding blend signs that are segmentally and prosodically indistinguishable from simplex signs (see Sandler 1989, 1999 for discussions of canonical sign length and complexity).

As an illustration of this point, consider the (partial) forms of the signs glossed as INVENT, TRAVEL, and TRIPPING, shown in Figure 3. Like the sign INVENT (Figure 3a), the blend TRIPPING (Figure 3c) is a one-handed sign in which the hand first makes contact with the temple, and then moves up and forward, away from the head. Like the sign TRAVEL (Figure 3b), the blend sign TRIPPING is formed with the hand configured in what is known as a "bent-V" handshape, with the index and middle fingers extended and slightly bent. The blend sign TRIPPING therefore matches the overall prosodic shape of the sign INVENT, however incorporating the handshape of the sign TRAVEL.

² In sign language linguistics, there is no agreed-upon or practical notational system for transcribing the forms of signs. Instead, it is typical to refer to signs with glosses that approximate their meanings, and to supplement these sign glosses with still images of signs and impressionistic descriptions of their forms. This poses a serious descriptive and representational challenge, especially for readers unfamiliar with a sign language. Wherever possible, I have provided links to videos of individual signs posted online, to help bridge this gap.

³ Lapiak, Jolanta. "trip (drug)". *Handspeak*. <http://www.handspeak.com/word/search/index.php?id=2498>

⁴ Lapiak, Jolanta. "trip". *Handspeak*. <http://www.handspeak.com/word/search/index.php?id=2497>

⁵ Lapiak, Jolanta. "make up". *Handspeak*. <http://www.handspeak.com/word/search/index.php?id=1332>

⁶ Lapiak, Jolanta. "hearing-minded". *Handspeak*. <http://www.handspeak.com/word/search/index.php?id=1010>

⁷ Lapiak, Jolanta. "think". *Handspeak*. <http://www.handspeak.com/word/search/index.php?id=2201>

⁸ Lapiak, Jolanta. "hearing". *Handspeak*. <http://www.handspeak.com/word/search/index.php?id=1001>

⁹ Lapiak, Jolanta. "word". *Handspeak*. <http://www.handspeak.com/word/search/index.php?id=2422>

¹⁰ Lapiak, Jolanta. "meaning". *Handspeak*. <http://www.handspeak.com/word/search/index.php?id=1364>

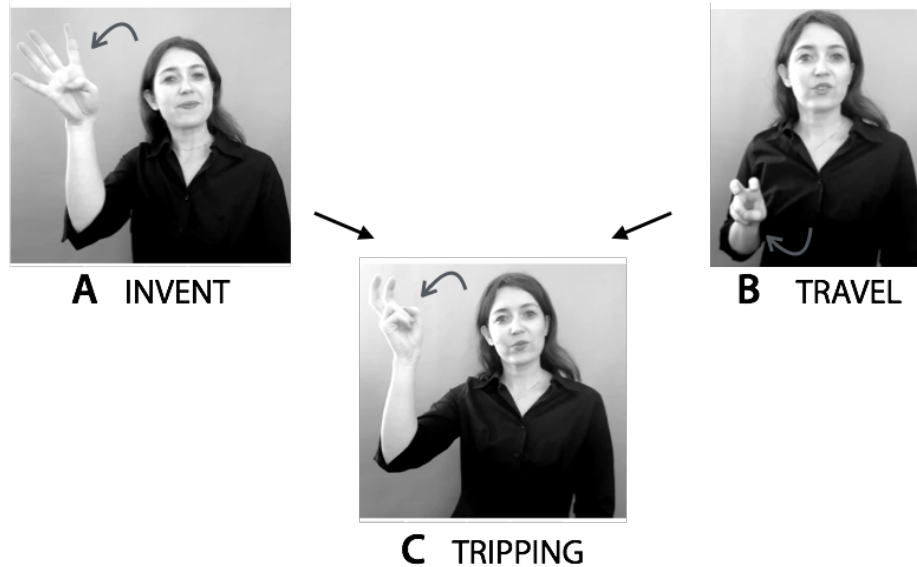


Figure 3: The ASL signs (a) INVENT and (b) TRAVEL combine to form the blend (c) TRIPPING.

These correspondences between INVENT, TRAVEL, and the blend TRIPPING can be formally represented as in Figure 4, where the appropriate values for the formational parameters of handshape, location, and movement are listed for each sign. Similar to the representation of the blend *sharrows* in Figure 2, this representation captures the fact that INVENT (Figure 4a) and TRIPPING (Figure 4c) are articulated with similar movement patterns, and that TRAVEL (Figure 4b) and TRIPPING are signed with the same handshape.

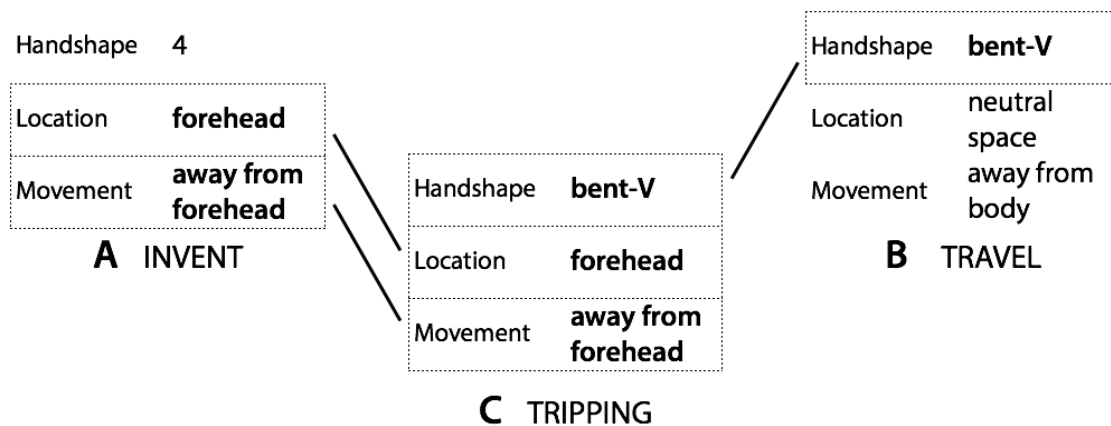


Figure 4: Phonological elements of the ASL signs (a) INVENT and (b) TRAVEL combine to form the blend (c) TRIPPING.

In ASL as in English, then, we can identify examples of words that have been created through the combination of sub-parts of other words. Beyond these singleton blends, in both languages we can also find families of blends that all share aspects of form and meaning in common. An example in English involves words containing *(a)licious*, for example *bubblicious*, *snugglicious*, *cougarlicious*, *divalicious*, *hunkalicious*, and *bootylicious*. These words are all relatively new creations that combine an existing word with the element *(a)licious*, which adds the connotation of being 'deliciously or extremely X'. Similarly, in ASL, a recurring word-formation pattern, relevant for the previously-mentioned example

HEARING-MINDED, involves changing the location of an existing sign so that it is instead articulated at the forehead. This has the effect of adding 'the mind as the site of cognition' to the original sign's meaning, as in COMMIT-TO-MEMORY, from WRITE-DOWN¹¹, MENTAL-SCAR, from SCRATCH¹², and WEAK-MINDED¹³, from WEAK¹⁴. These blend words in English and in ASL can be considered members of small lexical families because they have all changed an existing word in similar ways, whether affixing the element *(a)licious* or changing their place of articulation to the forehead, to create new word with a corresponding change in meaning, as is represented in Figure 5.

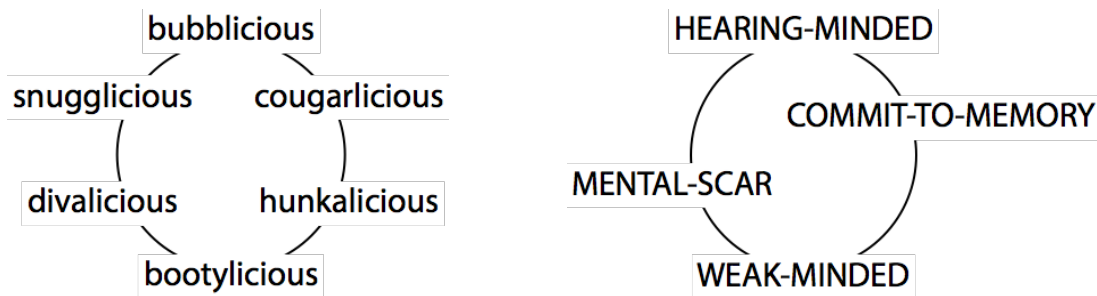


Figure 5: A family of related blends in English (left) and in ASL (right)

It seems obvious, even though the meaning of *(a)licious* differs from the meaning of the English word *delicious*, that the use of *(a)licious* to form new words is the result of repeated lexical blends involving the word *delicious*. With the ASL pattern, however, it is not entirely clear which particular sign the forehead location has been extracted from. There are many conventional ASL signs that are signed at the forehead and relate to 'cognition' (Frishberg and Gough 1973/2000; Meir, Padden, Aronoff, and Sandler 2013), for example THINK, KNOW¹⁵, and WONDER¹⁶. Accordingly, and owing to the simultaneous, non-concatenative structure of many ASL signs (Emmorey and Corina 1990; Aronoff, Meir, and Sandler 2005), any of these signs could be the potential source for the creation of a blend sign like COMMIT-TO-MEMORY.

Though both languages contain families of blends, the English and ASL examples seem to differ in that, in English, we see a pattern where repeated blending of the same word, *delicious*, has led to the association of a particular meaning with an element of form, *(a)licious*, that was not previously meaningful in its source word. In ASL, in contrast, blending seems to take advantage of elements which have already achieved some "meaningful" status, by virtue of the fact that they recur in groups of related signs. Thus, the fact that the forehead location is selected in the formation of signs like WEAK-MINDED is neither arbitrary nor surprising; the forehead location is available to be factored out of conventional ASL signs, and it can be put to work making new signs.

¹¹ Lapiak, Jolanta. "write down". *Handspeak*. <http://www.handspeak.com/word/search/index.php?id=2647>

¹² Lapiak, Jolanta. "scratch". *Handspeak*. <http://www.handspeak.com/word/search/index.php?id=4083>

¹³ Lapiak, Jolanta. "feeble-minded". *Handspeak*. <http://www.handspeak.com/word/search/index.php?id=5956>

¹⁴ Lapiak, Jolanta. "weak". *Handspeak*. <http://www.handspeak.com/word/search/index.php?id=2365>

¹⁵ Lapiak, Jolanta. "know". *Handspeak*. <http://www.handspeak.com/word/search/index.php?id=1207>

¹⁶ Lapiak, Jolanta. "wonder". *Handspeak*. <http://www.handspeak.com/word/search/index.php?id=2420>

3. Analogical and schematic blend-formation

In English, families of new blends can lead to the formation of new affix-like elements, while in ASL, new words typically result from changing the form of an existing sign so as to join an already-conventionalized family of words. Interestingly, though many individual novel blends are unlikely to go on to become established, conventional lexical items, speakers can, and often do, produce and interpret novel blends. This suggests that, for any group of related attested blends, at least some speakers will have identified an abstract pattern that generalizes over families of blend words that they have encountered. My proposal is that at the morphological level, blending takes advantage of proportional analogy, which then results in the formation of small patterns, or *constructions*, that can be used to form new words. Here, I adopt a very general construction-theoretic view of morphology to provide a formal account for the development of these patterns (following e.g., Jackendoff 1975; Bochner 1993; Goldberg 1995, 2006; Booij 2010; Jackendoff 2013; Bauer, Lieber, and Plag 2013). In the morphological constructions discussed here, elements of form are paired with elements of meaning, with some aspects of the construction left unspecified, or *schematic*, in ways that allow them to be extended to create new words.

An illustrative example can be seen with the English word *bromance*. As an established but recently coined blend, *bromance* has served as the template for the creation of a number of other *bro* words, including *brototype* 'a prototypical bro', from *prototype*; *brocabulary* 'the language of bros', from *vocabulary*; *brogrammers* 'men who program together', from *programmers*; and even *bromanteaux* 'blends containing bro', from *portmanteaux*. Though any one of these blends may have been formed on the basis of an analogical extension of the relationship between *romance* and *bromance*, the aspects of form and meaning shared among all of them can be described using a constructional template, as in Figure 6. In this representation, proportional analogy eventually gives way to a construction as a more abstract pairing of form and meaning.

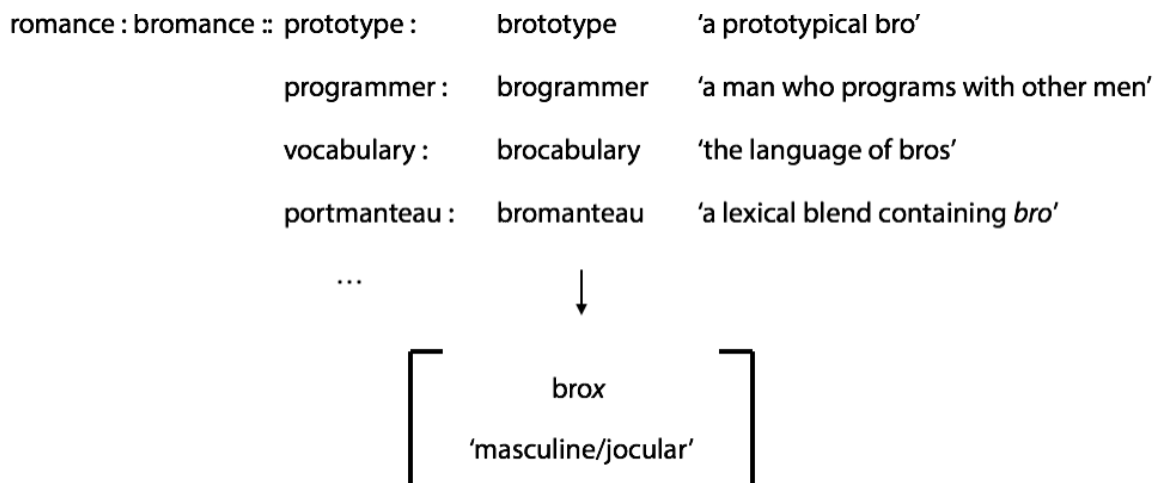


Figure 6: The relationship between analogical and schematic *bro* words

The pattern involving *bro* is notable for being fun and jocular, but another, slightly more serious example in English is *splain(ing)*, which is used to denote a condescending explanation from a position of privilege, as in the pair of examples in (2).

- (2) ...a junior colleague in another department, who is both black and of Caribbean origins, likes to **mansplain** to me about how I *must* wear a suit or I will not be taken seriously. I am thus in the bizarre position of **whitesplaining** to him that I, indeed, as a rich white lady, can get away with being tweedy and disheveled because students will accept that from me as an expected full professor costume...

<http://whatever.scalzi.com/2013/10/30/why-i-wear-what-i-do/>

Here as well, with *mansplain* and *whitesplaining*, we see that the relatively more-established blend *mansplaining* serves as the basis for the formation of other blends referring to other kinds of privilege. This pattern has been extended not only to form *whitesplaining* 'white condescension', but also yields other new words, including *straightsplaining* 'heterosexual condescension', *cissplaining* 'cisgendered condescension', and *geeksplaining* 'over-explaining to assumed non-experts'. As with *bro*, the construction-theoretic analysis of this group of words is that they reflect that at least some speakers of English have generalized an abstract constructional pairing of meaning and form, such that words ending in *splaining* can refer to a patronizing kind of explanation from a position of privilege, resulting in a constructional template that pairs the form *splaining* with the meaning 'patronizing explanation'.

In Construction Morphology (Booij 2010), constructions serve two grammatical functions. The first is that they are a description of how known words are formed, that is, of the relationship between meaning and form that can be observed in actually-occurring lexical items. Second, they serve as a template for producing or interpreting novel words. In English, then, a word like *anniversary*, or even a phrase like *Military Industrial Complex*, when re-used to create a set of blend derivatives, can serve as the source for the innovation of new affix-like elements with meanings like 'commemorative milestone' and 'suspiciously corporate', respectively. These constructions can then be deployed to create other new words (Figure 7). This analysis suggests that in at least some domains of English derivational morphology, we have a cumulative gradient shift from things that look more like canonical blends to things that look more like derivational affixes; the development of affix-like elements from blends follows a transition from more analogical formations to more schematic ones.

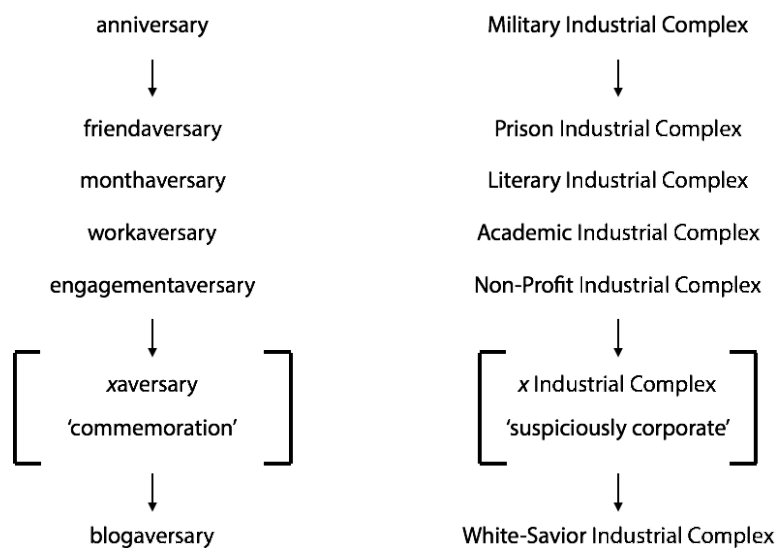


Figure 7: The creation of new blends from schematic constructions

However, in American Sign Language, which makes relatively infrequent use of segmental and concatenative morphology (Fernald and Napoli 2000; Aronoff, Meir, and Sandler 2005), we instead observe ambiguity between things that look like blends and things that look like typical non-concatenative morphology: In ASL, families of blends seem to exploit and systematize existing, partially motivated pairings of form and meaning. Beyond the example that we have already seen, in which the forehead location is systematically reused among signs relating to cognition and mental processes, another quite productive word-formation construction concerns the practice of *initialization* (Lepic 2015). Initialization is a conventional system for borrowing words from English into ASL, and is driven by English/ASL diglossia in the American Sign Language community. Initialization is also facilitated by the practice of fingerspelling, which pairs ASL handshapes with written English letters (Padden 1998; Fernald and Napoli 2000; Brentari and Padden 2001).

Initialized signs in ASL can be identified as signs whose handshapes correspond, via the conventions of fingerspelling, to the initial letters of their English translations. An example is the sign MATH¹⁷, which combines the movement and location of the ASL sign FIGURE-OUT¹⁸ with the "M"-handshape from fingerspelling in order to create a sign that blends aspects of ASL FIGURE-OUT with English *math*. MATH and FIGURE-OUT also belong to a somewhat large family of signs; these signs, including TRIGONOMETRY¹⁹, ALGEBRA²⁰, CALCULUS²¹, and GEOMETRY²², all have in common that they are signed with the same movement pattern, that their handshapes correspond to fingerspelled English letters borrowed from English words, and that they denote 'a kind of calculation'. We can therefore hypothesize that these signs have led to the abstraction of a construction where a particular configuration of location and movement can be used to create or interpret previously unseen initialized signs relating to 'calculation'.

In initialized signs, two different kinds of constructions, one based on shared movements and a second based on shared handshapes, together describe the structure of existing signs. These constructions also provide a template for producing or interpreting novel words. As an example, another lexical family in ASL provides the basis for a morphological construction for signs denoting 'groups of people', abstracted from the ASL sign GROUP²³ and established initialized signs like FAMILY²⁴, TEAM²⁵, and CLASS²⁶. These 'group' signs are all articulated with the same tracing movement, however their handshapes differ. Another morphological construction describes the structure of "U-initialized" signs, abstracted from signs like UNIVERSE²⁷, UNIVERSITY²⁸, and UNCLE²⁹. These signs have in common only that they are initialized: they articulated with a "U"-handshape, and that they correspond to concepts for which there is a "U-initial" word in English, but their movement and location features differ.

The construction that has been abstracted from signs denoting 'groups of people' has its movement pattern specified, and its handshape left schematic, while the construction for "U-

¹⁷ Signing Savvy. "math". <https://www.signingsavvy.com/sign/MATH/801/1>

¹⁸ Signing Savvy. "figure out". [https://www.signingsavvy.com/sign/FIGURE OUT/6603/1](https://www.signingsavvy.com/sign/FIGURE%20OUT/6603/1)

¹⁹ Signing Savvy. "trigonometry". <https://www.signingsavvy.com/sign/TRIGONOMETRY/5737/1>

²⁰ Signing Savvy. "algebra". <https://www.signingsavvy.com/sign/ALGEBRA/871/1>

²¹ Signing Savvy. "calculus". <https://www.signingsavvy.com/sign/CALCULUS/1073/1>

²² Signing Savvy. "geometry". <https://www.signingsavvy.com/sign/GEOMETRY/1391/1>

²³ Signing Savvy. "group". <https://www.signingsavvy.com/sign/GROUP/1432/1>

²⁴ Signing Savvy. "family". <https://www.signingsavvy.com/sign/family>

²⁵ Signing Savvy. "team". <https://www.signingsavvy.com/sign/TEAM/421/1>

²⁶ Signing Savvy. "class". <https://www.signingsavvy.com/sign/CLASS/75/1>

²⁷ Signing Savvy. "universe". <https://www.signingsavvy.com/sign/UNIVERSE/840/1>

²⁸ Signing Savvy. "university". <https://www.signingsavvy.com/sign/UNIVERSITY/457/1>

²⁹ Signing Savvy. "uncle". <https://www.signingsavvy.com/sign/uncle>

initialized" signs, in contrast, requires a specific handshake, but its movement is schematic. These constructions are therefore complementary, each specifying a phonological value that is left schematic in the other; what happens when they combine to form a new sign? Its form is predictable, taking the movement and location that are observed in signs like GROUP, and combining them with the U-handshake from signs like UNIVERSE, as Figure 8 shows.

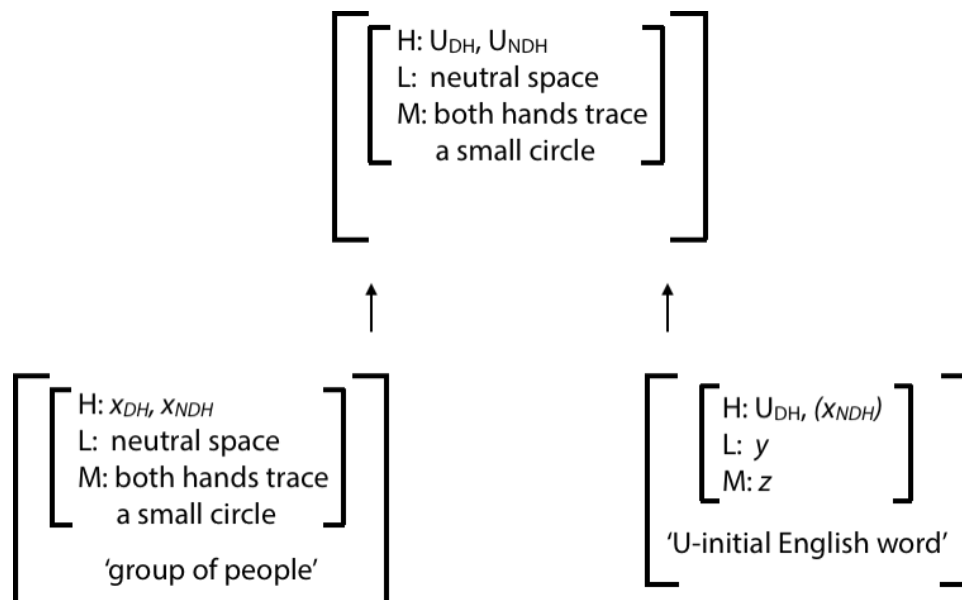


Figure 8: The creation of a new blend from schematic constructions in ASL

However, in seeing this particular initialized sign for the first time, the only inference that can be made about its meaning, in the absence of any discourse context, is that it refers to some kind of "U-group" in English, perhaps a *unit*, or a *union*, or something done in *unison*. Indeed, the fact that this sign means UNION³⁰ in ASL is something that must be learned. This is precisely because the U-handshape does not mean 'union', but rather, one of the functions of this handshape is to represent the fingerspelled letter U in certain types of signs. This fact is also captured by the representation in Figure 8: With the sign UNION, we have the unification of two complementary configurations of meaning and form which potentiate, rather than determine, the correct interpretation of a particular novel or previously-un-encountered sign.

While initialized signs present an illustrative example, it is not only initialized signs that demonstrate that ASL signers are aware of, and can systematically deploy, partially schematic configurations of meaning and form that can be formalized as constructions. The benefit of positing morphological constructions is that they also account for other examples of signs that have changed their form in order to join an existing family of signs. An example is the abstraction of the "horns" handshape from the signs MOCK³¹, STUCK-UP³², and IRONY³³, which all share an implied 'negative' aspect of meaning. Recognition of this pattern among conventional ASL signs licenses the formation of a construction that can be deployed to coin another negative sign, OVER-IT, made by changing the handshape of the already-existing sign BORED³⁴ to the "horns" handshape. Returning to an example we have already seen, we can also analyze signs like THINK, KNOW, and WONDER as licensing the abstraction of a

³⁰ Lapiak, Jolanta. "union". *Handspeak*. <http://www.handspeak.com/word/search/index.php?id=3835>

³¹ Signing Savvy. "mock". <https://www.signingsavvy.com/sign/MOCK/3908/1>

³² Lapiak, Jolanta. "snob". *Handspeak*. <http://www.handspeak.com/word/search/index.php?id=2002>

³³ Signing Savvy. "irony". <https://www.signingsavvy.com/sign/IRONY/1598/1>

³⁴ Signing Savvy. "bored". <https://www.signingsavvy.com/sign/BORED/50/1>

sign UN-UNDERSTAND, articulated with a downward movement and used to describe something that was once understood now becoming incomprehensible. Another example, already anticipated in Figure 9, involves inverting the downward movement of the ASL sign OPPRESS⁴⁰ to form a possible sign for PRIVILEGE, as recently observed in an ASL video log posted online⁴¹. The conventional sign OPPRESS is signed with the dominant hand pushing the non-dominant hand downward, while in the nonce sign PRIVILEGE, this configuration is reversed such that the dominant hand lifts the non-dominant hand upward (Figure 10).



Figure 10: The conventional sign OPPRESS (left) has downward movement, and the related neologism PRIVILEGE (right) has upward movement

4. Conclusion

The main questions addressed in this paper concerned the consequences that lexical blends hold for the morphological systems of individual human languages, as well as the consequences that lexical blends hold for the development of adequate morphological theories. Here I have suggested that in English, lexical blending is a productive word-creation process that can, in certain cases, lead to the creation of elements that resemble affixes, for example, the reanalysis of *splaining* in *explaining* and *mansplaining*, or of *aversary* in *anniversary* and *monthaversary*. In ASL, in contrast, lexical blends reveal when nascent patterns in the language have become systematically organized in a way that allows signers to deploy them in the production of new signs. These examples included the formation of COMMIT-TO-MEMORY by changing the location of WRITE-DOWN to match that of signs like THINK and KNOW, and also the creation of the sign OVER-IT by changing the handshape of BORED to match that of the signs IRONY and MOCK.

Crucially for this description of lexical blends in English and in ASL, a construction-theoretic lexicon treats morphological patterns as emergent phenomena that have been abstracted over whole words as learned pairings of meaning and form. Lexical blends in general provide support for this view because they are necessarily made from existing whole words, rather than more theoretically familiar roots and affixes. ASL derivational morphology in particular also provides support for this construction-theoretic view of the lexicon, because ASL morphology is overwhelmingly non-concatenative (Sandler 1989; Fernald and Napoli 2000; Aronoff, Meir, and Sandler 2005), and, as we have seen, lends itself well to the analogical blending operation described above. The facts about lexical blends in both English and in ASL reveal then, that not only are individual words made up of smaller identifiable

⁴⁰ Signing Savvy. "oppress". <https://www.signingsavvy.com/sign/OPPRESS/1979/1>

⁴¹ <https://youtu.be/P071B5sPCvg?t=1m20s>

parts, but they themselves are also parts that participate in larger patterns known as lexical families. Under this view, the construction-theoretic lexicon is one consisting of whole-part relations all the way down.

References

- Arndt-Lappe, S. and Plag, I. (2013). The role of prosodic structure in the formation of English blends. *English Language and Linguistics* 17(3): 537–563.
- Aronoff, M., Meir, I., and Sandler, W. (2005). The paradox of sign language morphology. *Language* 81(2), 301–344.
- Bat-El, O. (2000). The grammaticality of 'extragrammatical' morphology. *Extragrammatical and marginal morphology, Munich: LINCUM*, 61-81.
- Bauer, L., Lieber, R., and Plag, I. (2013). *The Oxford Reference Guide to English Morphology*. Oxford University Press.
- Berman, J. M. (1961). Contribution on blending. *Zeitschrift für Anglistik und Amerikanistik* 9: 287–281.
- Bochner, H. (1993). *Simplicity in Generative Morphology*. Mouton de Gruyter.
- Booij, G. (2010). *Construction Morphology*. Oxford University Press.
- Booij, G. and Masini, F. (2015). The role of second order schemas in the construction of complex words. In L. Bauer, L. Körtvélyessy, and P. Štekauer (Eds.) *Semantics of Complex Words*. Springer. 47–66.
- Brentari, D. and Padden, C. (2001). Native and foreign vocabulary in American Sign Language: A lexicon with multiple origins. In D. Brentari (Ed.) *Foreign Vocabulary: A Cross-Linguistic Investigation of Word Formation* (87–119). Lawrence Erlbaum.
- Bryant, M. (1974). Blends are increasing. *American Speech* 49: 163–184.
- Emmorey, K. and Corina, D. (1990). Lexical recognition in sign language: Effects of phonetic structure and morphology. *Perceptual and motor skills* 71(3): 1227–1252.
- Fernald, T. and Napoli, D. J. (2000). Exploitation of morphological possibilities in signed languages: Comparison of American Sign Language with English. *Sign Language and Linguistics* 3(1): 3–58.
- Frishberg, N. and Gough, B. (2000). Morphology in American Sign Language. *Sign Language and Linguistics* 3(1), 103–132. (Originally published 1973).
- Goldberg, A. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. University of Chicago Press.
- Goldberg, A. (2006). *Constructions at work: The nature of generalization in language*. Oxford University Press.
- Gries, S. Th. (2004). Shouldn't it be *breakfunch*? A quantitative analysis of the structure of blends. *Linguistics* 42(3), 639–667.
- Halle, M. (1973). Prolegomena to a theory of word formation. *Linguistic Inquiry* 4(1): 3–16.
- Jackendoff, R. (1975). Morphological and semantic regularities in the lexicon. *Language* 51: 639–671.
- Jackendoff, R. (2013). Constructions in the parallel architecture. *The Oxford Handbook of Construction Grammar*, 70-92.
- Kemmer, S. (2003). Schemas and lexical blends. In Thomas Berg et al. (eds.), *Motivation in Language: From Case Grammar to Cognitive Linguistics. A Festschrift for Günter Radden*. Philadelphia: Benjamins. 69–97.
- Klima, E. and Bellugi, U. (1979). *The Signs of Language*. Harvard University Press.
- Lepic, R. (2015). *Motivation in morphology: Lexical patterns in ASL and English*. PhD thesis. University of California, San Diego. <http://escholarship.org/uc/item/5c38w519>
- Lehrer, A. (1998). Scapes, holics, and thons: The semantics of English combining forms. *American Speech* 73(1), 3-28.
- Marantz, A. (2013). No escape from morphemes in morphological processing. *Language and Cognitive Processes* 28(7): 905–916.
- Metcalf, A. (2002). *Predicting New Words*. Houghton Mifflin.
- Meir, I., Padden, C. A., Aronoff, M., and Sandler, W. (2013). Competing iconicities in the structure of languages. *Cognitive Linguistics* 24(2): 309–343.
- Miller, D. G. (2014). *English Lexicogenesis*. Oxford University Press.
- Padden, C. A. (1998). The ASL Lexicon. *Sign Language and Linguistics*. 1(1): 39–60.
- Padden, C. A. and T. Humphries. (1988). *Deaf in America: Voices from a Culture*. Harvard University Press.
- Pound, L. (1914). *Blends: Their relation to English Word Formation*. Swets and Zeitlinger.
- Renner, V., Maniez, F., and Arnaud, P. (Eds.). (2012). *Cross-disciplinary perspectives on lexical blending*. De Gruyter.
- Ronneberger-Sibold, E. (1999). On useful darkness: loss and destruction of transparency by linguistic change, borrowing, and word creation. In G. Booij and J van der Marle (Eds.) *Yearbook of Morphology* (97–120). Kluwer Academic Publishers.

- Sandler, W. (1989). *Phonological representation of the sign: Linearity and nonlinearity in American Sign Language*. Walter de Gruyter.
- Sandler, W. (1999). Cliticization and prosodic words in a sign language. In T. A. Hall and U. Kleinhenz (Eds.) *Studies on the Phonological Word* (223–254). John Benjamins.
- Sandler, W, and D. Lillo-Martin. (2006). *Sign Language and Linguistic Universals*. Cambridge University Press.
- Scalise, S. (1984). *Generative Morphology*. Foris.
- Spencer, A. (1991). *Morphological Theory*. Blackwell.
- Wilcox, P. (2000). *Metaphor in American Sign Language*. Gallaudet University Press.
- Zwicky, A. and G. Pullum. (1987). Plain morphology and expressive morphology. *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*, 330-340.

Grammaticalization is not the full story: a non-grammaticalization account of the emergence of sign language agreement morphemes

Irit Meir
University of Haifa
imeir@univ.haifa.ac.il

1. Introduction

Of all the processes that induce language change, grammaticalization stands out as the only process that introduces new grammatical categories into the language. While other processes may introduce new forms, new morphemes or allomorphs, or cause the disappearance of morphemes, grammaticalization is considered the only process (or assembly of processes working together) that may cause the emergence of a new grammatical *category*.

Grammatical categories, also referred to as inflectional, functional or morpho-syntactic categories, have two facets: they are *conceptual* categories that are manifested *morphologically* in a given language. In order for a new category to arise, a novel bonding between a conceptual category and its morpho-syntactic manifestation has to be created. For example, the creation of forms such as *shew* and *snew* as the past tense forms of *show* and *snow* in the Norfolk dialect of English (Trudgill 2009) cannot count as cases of the emergence of a new category, since past tense already exists in the language; it is only a new form that is introduced. But a change in the meaning of a verb such as *will* from volition to future tense marker (in English and many other languages, see Heine and Kuteva 2002:310-311), does introduce a new category, that of future tense.

The fact that only grammaticalization can cause the emergence of new grammatical categories in a language has been pointed out by Meillet, the linguist who coined the term 'grammaticalization' (1912). Meillet describes two basic core mechanisms for language change – analogy and grammaticalization. However, he points out that there is a fundamental difference between the two: "While the analogy can renew the detail of the forms, but often leaves untouched the overall plan of the grammatical system, the 'grammaticalization' of certain words creates new forms, introduces categories for which there was no linguistic expression, and transforms the whole of the system." (1912:133). According to Meillet, analogy can introduce new forms into the language, but not new categories.

The rich body of research on grammaticalization from Meillet until today has revealed a vast range of changes in languages all over the globe that can be regarded as grammaticalization, the change from a lexical item to a more grammatical item, eventually leading (in some cases) to the creation of a new grammatical category. Heine and Kuteva (2002) compiled a detailed data-set of about 400 attested grammaticalization processes on the basis of a study of about 500 languages.

Grammaticalization has been so widely attested and studied, that it is regarded as the main mechanism for the development of grammatical "stuff" in a language, resonating Meillet's view. In a recent position paper by Beckner et. al (2009), summarizing their view of language as a complex adaptive system, the dominant role of grammaticalization in creating the grammar of a language is emphasized: "Given that grammaticalization can be detected as

ongoing in all languages at all times, it is reasonable to assume that the original source of grammar in human language was precisely this process.” (Beckner et al. 2009: 9).

Kiparsky (2012) holds a different view regarding the relationship between grammaticalization and analogy. He suggests that they are not orthogonal to each other, but rather that grammaticalization is a special type of analogy, one that is not exemplar-based; it is a specific kind of analogy driven by general principles (UG constraints or rules). He further suggests that exemplar-based analogy can, in rare cases, give rise to new categories, "provided they are built from simpler ones in conformity with existing combinatoric patterns of the language." (Kiparsky 2012: 9). However, the major process for introducing new categories in the language is still grammaticalization, as it does not need to have an existing model in the language in order to start the process.

This paper suggests that Meillet's and Kiparsky's assumptions are too strong, in that new forms and new categories can arise not by means of grammaticalization of free words, or by exemplar-based analogy (in specific cases), but rather by reanalysis of a set of monomorphemic words into multi-morphemic, consisting of a base (stem) and inflectional morphemes. These monomorphemic words have a shared phonological component and a shared meaning component. By comparing the different words in this set, the shared phonological stretch is reanalyzed as associated with the shared meaning component, and thus a new form-meaning element is created, a new morpheme. This novel morpheme can then attach to other bases and become a productive morpheme in the language.

As an illustration, consider a hypothetical language in which a few words end in /ko/ and they all happen to denote entities that are inherently red: blood, lips, a rose, Cardinal (bird), a tomato, etc. Speakers of that hypothetical language may identify the shared phonological element and the shared meaning component, and start to associate /ko/ with red color. Eventually, /ko/ may be reanalyzed as the segment representing red, thus creating a new morpheme in the language, *-ko*, a suffix denoting red color. The next stage would be for *-ko* to attach to other bases, to indicate that they are red. Thus, 'house+ *-ko*' would mean a red house, 'fruit+*-ko*' could mean a red apple, and so on. If the language did not have a grammatical category of color, the emergence of this morpheme introduces not only a new form to the language, but also a new category.

This hypothetical example gives the flavor of the process that will be described below. This process is not grammaticalization, since it involves the fission of an existing word into two parts, rather than the fusion of elements into one word. However, it is not analogical, since it is not based on a pre-existing model in the language.¹ Furthermore, it cannot be said to eliminate unmotivated grammatical complexity of idiosyncrasy, as analogical changes often do; if anything, it may be regarded as creating more complexity in the language, as it creates complex words out of non-complex ones. Yet it is based on comparison across a set of words and on consistent form-meaning pairing.

As we demonstrate below, processes along these lines have been attested in derivational morphology, but they are rare in inflectional morphology. But we show that at least in one case, it is responsible for the creation of a new morphological category. The category in question is the class for agreement verbs in Israeli Sign Language (ISL), a class of verbs that denote transfer and are inflected for agreement with their subject and (indirect) object. This class of verbs exists in many other sign languages too, but the diachronic evidence for its emergence presented here is from ISL.

The paper is structured as follows: we first describe the class of sign language agreement verbs (section 2), and present diachronic evidence for the different stages of its emergence in ISL, from mono-morphemic verbs into multi-morphemic words (section 3). We then analyze

¹ This contrasts with Kiparsky's (2012: 10) view that the fission of one word into two is always exemplar-based, occurring only by analogy to specific existing constructions.

the process involved (section 4), and discuss some phenomena in spoken languages (section 5) which are similar yet are derivational, and therefore do not lead to the emergence of a new grammatical category.


2. Sign language agreement verbs

Agreement verbs in sign languages constitute a class of verbs that are inflected for agreement with the verb's arguments. This class is one of three verb classes of sign languages, identified first by Padden (1988) for American Sign Language (ASL) and then attested in many other sign languages. The two other classes are *plain verbs* and *spatial verbs*. These three classes differ in their agreement patterns: plain verbs do not mark agreement with the verb's arguments at all, and spatial verbs mark agreement with spatial referents, that is, locations. Agreement verbs mark agreement with the verb's subject and (indirect) object arguments by a special mechanism, described below, of copying the locational features of the arguments' referential indices onto the verb's location slots. This tri-partite verb classification is semantically grounded (Meir 2001, 2002). Agreement verbs denote transfer, whether concrete (as in GIVE, SEND) or abstract (as in TEACH, HELP).² Spatial verbs denote motion in space, and plain verbs are defined negatively, as not involving transfer or motion. Many plain verbs denote psychological and emotional states.

Verb agreement in sign languages is based on their referential system, in which referential indices are realized by specific loci in space. In sign languages, the referential features of nominals in a clause are associated with discrete locations in space, called 'R(eferential)-loci' (Lillo-Martin and Klima 1990). If the referent is present in the signing situation, the R-locus is towards the actual location of the referent. If the referent is not present, it is assigned a point in the signing space (provided that other NPs have not already been assigned that point). This association is achieved by signing a noun and then pointing to, or directing the gaze towards, a specific point in space.³ Once an R-locus has been established for a specific referent, subsequent reference to that locus is equivalent to pronominal reference; that is, pointing again to that locus has the function of referring back to the NP associated with it. These R-loci are therefore regarded as the visual manifestation of the pronominal features of the nominals in question (see e.g., Meier 1990; Lillo-Martin and Klima 1990; Janis 1992; Neidle et al. 2000; Rathmann and Mathur, 2002; Lillo-Martin and Meier 2011. Liddell 2003, presents a different view of these loci).

In addition to pronouns, verbs which inflect for agreement, the class of agreement verbs, also make use of the system of R-loci: their beginning and end points are determined by the R-loci of their grammatical arguments. The initial point of the verb is associated with the R-locus of the subject argument, and the final location with that of the object's R-locus (Padden, 1988).⁴ Agreement verbs can be regarded as verbs whose beginning and end points are not lexically specified, but rather contain a variable whose phonetic value is filled by copying the location features of the R-loci of the verb's arguments (Meir 1998a). According to this

² Not all agreement verbs are prototypical verbs of transfer, as pointed e.g. by Quer and de Quadros (2008). I attribute this to the grammaticization of this class of verbs (Meir 2012). However, in all sign languages that have agreement verbs, this class includes a core set of typical verbs of transfer.

³ Localization of referents may also be achieved by signing the noun itself in a specific location in space, if the sign is not body-anchored. For example, the sign CHILD is signed by placing a  handshape facing downwards in neutral space. If the signer places his/her hand to the right or to the left of the signing space, this location may serve as an R-locus for the particular child introduced into the discourse.

⁴ This description of the mechanism of sign language verb agreement is oversimplified. The more precise generalization is that the verb's initial and final locations are associated with its source and goal arguments, and the facing of the palm is towards the grammatical indirect object. For a fuller description and analysis, see Meir (2002).

analysis, which we adopt here, the agreement *markers* are the location values of the beginning and end points of the verbs, since these values encode features of the arguments (their R-loci) on the verb.⁵

However, not all agreement verbs behave the same regarding these agreement markers. Some verbs agree only with one argument, the object, while their initial point is anchored to the body and is invariable (Pizzuto 1986 for Italian Sign Language; Engberg-Pedersen 1993 for Danish Sign Language; Meir et al. 2007 for ISL and ASL). Additionally, signers do not always use inflected forms of agreement verbs; they might use a non-inflected form of a verb (called also citation form; in these forms the movement of the verb is from the signer's body outwards), or they may inflect only the final point of the sign, while the initial point is anchored to the signer's body. This variable behavior has been attested in various sign languages (e.g. Padden 1988 for ASL; Engberg-Pedersen 1993 for Danish Sign Language; Meir 1998a, 2010, 2012, Padden et al. 2010 and Meir et al. 2013 for ISL; de Beuzeville et al. 2009 for Australian Sign Language, and Schembri et al. 2016 for British Sign Language). Some researchers found diachronic differences in the use of these different forms of agreement verbs, to which we now turn.

3. The diachronic development of Israeli Sign Language verb agreement

Sign language verb agreement, then, is a complex grammatical system as it involves systematic encoding of syntactic roles, as well as the referential features of the arguments on the verb. How did this system develop? How did the class of agreement verbs arise in a given sign language?

Luckily, sign languages are relatively young, and so diachronic developments may be easier to trace since many of them might be quite recent. We present here a diachronic study of Israeli Sign language, a language that evolved with the emergence of the Israeli Deaf community about 85 years ago, in a language-contact situation. The members of the first generation came from different backgrounds, both in terms of their country of origin, and in terms of their language. A few were born in Israel, and some of them went to the school for the deaf in Jerusalem that was founded in 1932, but the majority were immigrants who came to Israel from Europe (Germany, Austria, France, Hungary, Poland), and later on from North Africa and the Middle East. Some of these immigrants brought with them the sign language of their respective communities. Others had no signing, or used some kind of home sign.⁶ Today, the Deaf community numbers about 10,000 members, spanning over four generations, from the very first generation, which contributed to the earliest stages of the formation and development of the language, to the fourth generation, that has acquired and further developed the modern language as a full linguistic system.

By studying the language of different generations of signers in the community, we can trace diachronic developments in the language (based on Labov's 1963 '*apparent time*' construct, which assumes that differences in the linguistic productions of different age groups in a language reflect diachronic changes in that language). The three groups of ISL signers who participated in the study are presented in Table 1.

⁵ There are further complications regarding the characterization of agreement verbs, which cannot be described here for the sake of brevity. See e.g. Meir (1998a), Meir et al. (2007, 2013) and Lillo-Martin and Meier (2011) for discussion.

⁶ For a description of the history of the Deaf community in Israel and the development of ISL, see Meir and Sandler (2008).

<i>Group</i>	<i>Generation</i>	<i>Age</i>	<i>N</i>	<i>characteristics</i>
1	First	65+	13 (7 male, 6 female)	<ul style="list-style-type: none"> • came from a variety of linguistic backgrounds • not exposed to a unified linguistic system
2	Second	45-65	10 (5 male, 5 female)	<ul style="list-style-type: none"> • had linguistic models for ISL
3	Third	25-45	8 (4 male, 4 female)	<ul style="list-style-type: none"> • exposed to ISL from their early childhood, • six are native ISL

Table 1: Participants in the study, according to groups

3.1 Task

In order to elicit verb forms from the participants, we used a set of 30 video clips, each depicting a single event (Sandler et al. 2005). Of these, relevant to our study are nine clips depicting transitive or di-transitive events with two human participants (see Appendix). The verbs in these clips have the potential of becoming agreement verbs in a sign language, since agreement verbs are always transitive and have two human/animate arguments. Participants were asked to view the clips and describe the event in each clip to another ISL signer. The verbs in all the responses obtained from the signers were analyzed as to whether they displayed full agreement, that is, agreeing with two arguments (the subject and the object), with one argument (the object) or did not agree at all.

3.2 Results

The analysis of the signed productions of the different groups shows that there is a noticeable difference between the two older groups and the youngest group (Figure 1): while signers of the youngest group used agreeing forms in 72.5% of their verb productions, signers in the two older groups used agreeing forms in less than 40% of their responses. Furthermore, in the two older groups, fully agreeing forms are very rare: only 5%-6% of their verb forms mark agreement with two arguments. The majority of forms inflected for agreement in these groups mark agreement with only one argument. In Group 3, fully agreeing forms appear in 45% of the responses.

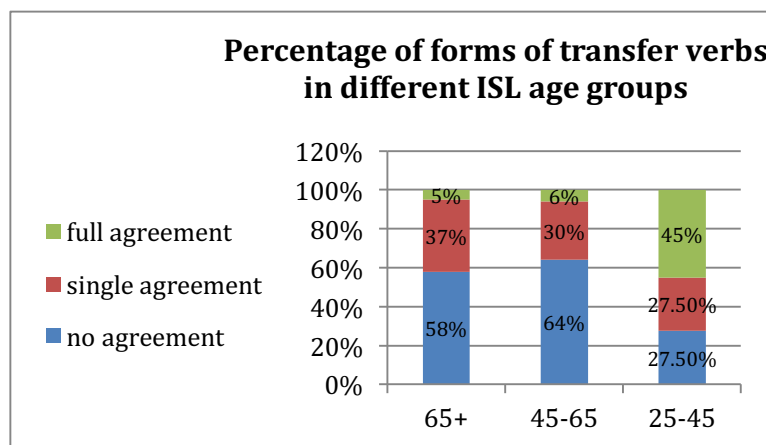


Figure 1: Percentage of forms of transfer verbs in the three ISL groups

Though Groups 1 and 2 look very similar, there is an interesting difference in the form of the single-agreeing verbs. A single-agreeing verb can be articulated in the *Z axis*, that is, the axis from the signer towards a potential addressee, or it can be directed towards locations in space that are to the right or to the left of the signer (the *diagonal axis*). As shown in Figure 2, there is a marked difference between Group 1 and the other groups: in Group 1, almost half of the single agreeing forms are signed on the *Z axis*. In Groups 2 and 3, their percentage is 17% and 10% respectively. This means that in Group 1, many more verbs are confined to the *Z axis*, and make more restricted use of the horizontal plain around the signer's chest. As we shall see shortly, this difference in form is important for understanding the diachronic development of agreeing forms.

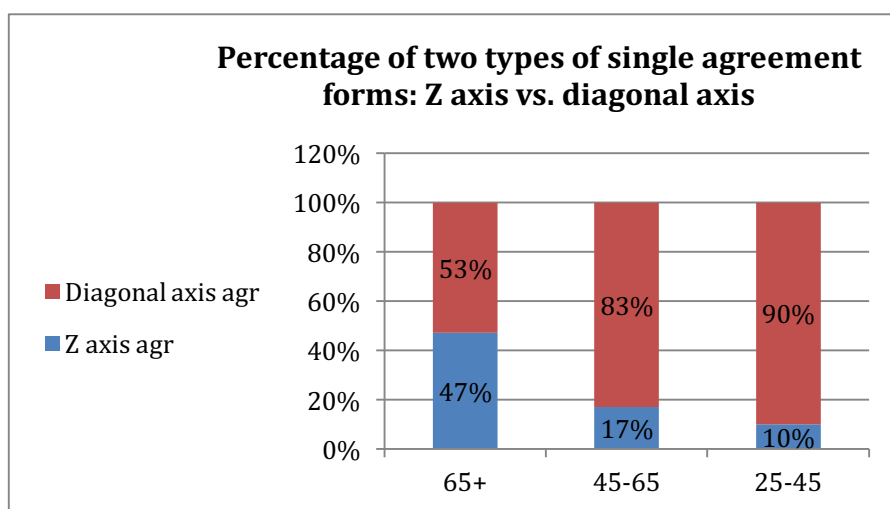


Figure 2: Use of the *Z axis* vs. diagonal axis in single agreeing forms in the three age groups

The results above indicate that fully agreeing verbs were very rare in the first generation of signers, and that the verb forms were mainly confined to the *Z axis*. In the second generation, verbs started to be directed at various points of the signing space, but fully-agreeing forms were still rare. In the third generation, fully agreeing forms became wide-spread, and most verb forms take advantage of the plane around the signer, liberating themselves from the *Z axis*. Yet how these changes occur? And what triggered them? A closer examination of the responses of signers from the different groups reveals several stages that led to the emergence of fully agreeing forms.

3.3 Stages in the development of verb agreement

Stage 1: Verbs of transfer as plain verbs. In Groups 1 and 2, 60% of the responses contain uninflected forms. These verbs of transfer behave like plain verbs; their initial and final points are not determined by the R-Loci of the verb's arguments, but rather stay stable in all the verb's productions, and their movement is from the signer outwards, on the *Z axis*. Figure 3 shows the verb GIVE signed by a Group 1 signer, when describing a clip in which a woman is giving a shirt to a man.



Figure 3: A non-inflecting form of the transfer verb GIVE

Stage 2: Re-analysis of end point. Some signers in Group 1, however, produce forms that can be regarded as an initial step towards marking agreement. For example, in a clip depicting a man throwing a ball to a girl, one signer signed the following:

- (1) I FATHER, FEMALE CHILD_{Z->2} I THROW_{0->2}
 ‘I am the father, the child is there, I throw (to the child)’

In this response, the signer places the sign CHILD in a location in space right in front of her (notated by the Z->2 subscript), as if the child were the addressee. She then directs the verb THROW towards the point in space where she had previously localized the child. This verbal form is quite similar to the uninflected forms described in stage 1. There are two main differences, though. First, the signer explicitly localizes the argument (CHILD) in front of her; second, the verb is signed as if directed towards the child. Such a form, then, shows the buds of the sign language agreement system: a verb is directed towards a location in space associated with an argument of the verb. The end point of the sign is being reanalyzed as a morpheme, encoding a feature of the verb’s argument. But the verb is still articulated on the Z axis. These forms can be analyzed as agreeing with the object argument, but only when the object argument is the addressee or construed as the addressee. Three signers out of the thirteen in Group 1 used this form in two thirds or more of their responses.

Stage 3: from axis to plane. In the next step towards an agreement system, the verb is no longer restricted to the Z axis. Some signers, mostly from Group 2, produce forms in which the sign’s initial location is on the signer’s body, and its end point is directed towards a spatial locus associated with the object argument. Crucially, the location associated with the argument is not on the Z axis, but rather to the right or to the left of the signer; the verb’s ‘loose end’ can be directed to any location associated with an argument in the signing plane. Such forms can be regarded as verbs marking agreement with the (indirect) object (Figure 4).⁷ The reanalysis of the verb’s final location as a morpheme marking agreement with an argument has been completed. This morpheme is not restricted to a specific point in space on the Z axis, but can take any value of an R-locus associated with the verb’s argument. Such forms are found in the responses of signers from all three groups, but they become much more prevalent as the language matures (see Figure 2 above).

⁷ Thanks to Ann Senghas for this point.



Figure 4: The transfer verb GIVE directed towards the R-locus of the (indirect) object argument

Stage 4: Re-analysis of the verb's initial location. The final stage of the emergence of verb agreement is obtained when the verb movement's heretofore fixed and body-anchored initial point can be articulated at a place off the body; it is reanalyzed as the subject argument marker. When such a reanalysis occurs, the verb might be said to have left the body; it is no longer body-anchored. A verb form marked for agreement with two 3rd person referents moves from one location in space to another, often on the left-right spatial axis (see Figure 5). Both end-points, then, are analyzed as morphemes, and the verb becomes multi-morphemic.



Figure 5: A fully agreeing form of the verb GIVE, signed on the X axis

Further developments: Once this mechanism of agreement (associating the verb's initial and final points with the R-loci of the arguments) is established for a group of verbs with specific semantic characterization, namely, verbs of transfer, other verbs may also adopt this morphological mechanism, and become agreement verbs. These verbs may share some, but not all of the semantic attributes of verbs of transfer, e.g. they are transitive and involve two human participants, but the transfer sense is less obvious. For example, verbs of communication, such as PHONE or FAX, involve two human participants, as do verbs of transfer, and also the act of communication. In some languages (i.e. ASL and ISL) these signs were originally plain verbs, but have become agreement verbs over the years. Similarly, verbs of saying, such as TELL, ASK, ANSWER, TELL-A-STORY, and ASL SAY-NO-TO are also agreement verbs. Other verbs that became agreement verbs in ISL are VIDEOTAPE (someone), HATE/DETEST (someone) and STAY-AWAY-FROM/AVOID. As the formal mechanism becomes established in a language, the semantic basis for the category becomes more opaque, and the grammatical characteristics of the elements become more prominent.

4. The diachronic mechanism underlying the development of sign language verb agreement

The preceding section described the chain of changes that occurred in verbs of transfer in ISL, from plain verbs that are not marked for agreement at all to fully agreeing forms. What is the mechanism underlying this chain of changes?

The crucial change is in stage 2, where verbs denoting transfer were reanalyzed as multi-morphemic; their final location was reanalyzed as directed towards the addressee, that is, a morpheme marking agreement with the addressee. I suggest that this reanalysis was triggered by two factors: (i) a resemblance of form and meaning components in words denoting transfer in the manual modality; and (ii) an increase in the use of the signing space for reference indexing.

When depicting a transfer event in gestures, the hands typically move outwards from the signer's body, as if tracing the transfer of an entity from one possessor (represented by the signer's body) towards another person (the recipient, who is conceived of as being located in front of the signer). One end of the sign is at the signer's body, and the other end is in space, away from the body. This 'loose end' of the verbs is crucial here: when a language develops systematic use of space for referential purposes, the 'loose end' lends itself more easily towards reanalysis; it is reanalyzed as a morpheme encoding the R-locus associated with the object (recipient) argument. After one end point undergoes such reanalysis, the other end point, the one close to the signer's body, may also be reanalyzed in a similar way, as encoding properties of the argument associated with the signer's body, the subject argument (Meir et al. 2007, 2013). Verbs of transfer share a *meaning component* and a *physical formational component*: they denote the transfer of an entity from one possessor to another, and their form consists of a path movement between the signer's body and space. The two endpoints (first the spatial end corresponding to the object and then the body-anchored end corresponding to the subject) lend themselves quite easily to reanalysis: they become morphemes, encoding person features of the two possessors.

What makes verbs of transfer special in sign languages is that they share not only meaning (as they do in spoken languages as well), but also a specific iconic form: a path movement from or towards the signer's body. It is this packaging of shared meaning and form that makes verbs of transfer in the manual modality amenable to reanalysis, eventually leading to the creation of a morphological class. Sign language agreement morphemes, then, came into being not by being grammaticalized from free words, but rather by being carved out of a set of existing words that share a form and a meaning component.

Though this process is not based on analogy to existing forms in the language, it is dependent on the language using space for referential purposes. And indeed, we find that concomitant with an increase in the use of agreeing forms, there is also an increase in the use of space for localizing referents in the younger groups. As evident in Figure 6, signers in Group 2 provided spatial information more often than Group 1 signers, and signers in Group 3 provide spatial information in 40% of their responses, and localize referents in about 60% of their responses. The increased use of space, then, together with the form-meaning packaging of verbs of transfer, gave rise to the reanalysis of the endpoints as agreement markers, thus creating a new grammatical category in the language.

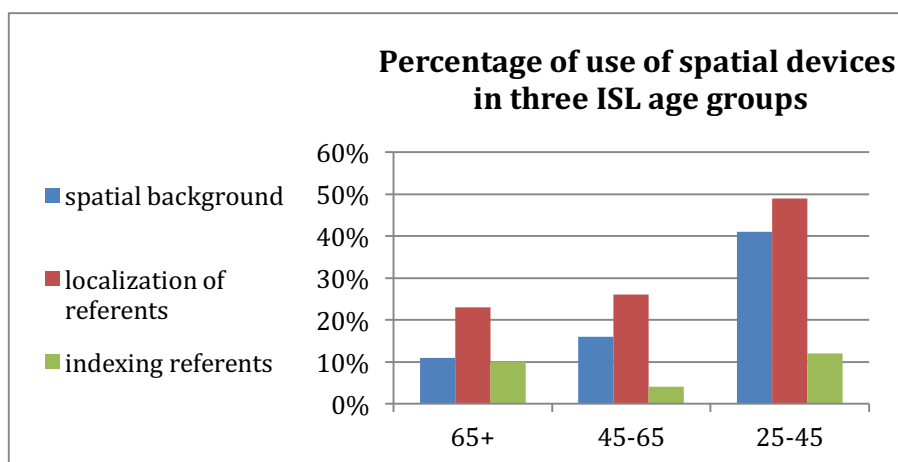


Figure 6: Percentage of reference to spatial background, localization and indexing of referents in the three age groups

5. A non-grammaticalization mechanism for morpheme emergence

The analysis of the emergence of agreement morphemes in ISL presented above shows how morphemes can come into existence by carving off or detaching phonological segments from the words they occurred in, thus granting them morphemic status. In the case of ISL verbs of transfer, not only a new morpheme emerged, but a new grammatical category, that of agreement verbs. Has such a process been attested in spoken languages too?

The carving of a morpheme from a non-complex word has been attested in several types of phenomena, mainly derivational, resulting in a novel morpheme, but not a novel grammatical category in the language. I briefly describe these phenomena here.

- (i) **Backformation:** In backformation a mono-morphemic word is reanalyzed as multi-morphemic because of phonological similarity to existing morphemes in the language. The word [*sculptor*], a mono-morphemic word to begin with, was reanalyzed as [*sculpt+or*], that is, as containing the agentive suffix *-or* and the verbal base *sculpt*, by analogy to the existing verb-noun pairs such as *terminate- terminator*, and the phonological identity of the sequence /or/ to the morpheme *-er/or* that exist in the language (Booij 2005: 40-41). The morpheme itself is not new in the lexicon (though the verb is); the novelty is the morphemic status of this sequence in an initially monomorphemic word.
- (ii) **The creation of productive splinters:** The sequence *-oholic*, extracted from the word *alcoholic*, now attaches to a number of words, such as *workaholic*, *foodoholic*, *golfaholic*, *chocoholic*, *shopaholic* (Lepic 2015). In this case there is no analogy to existing morphemes in the language, as the sequence *-oholic* did not exist as an independent unit in the language. On the contrary; it is unique in the language, and therefore unequivocally associated with the word *alcoholic* and its meaning 'a person addicted to alcoholic drinks'. By detaching *-oholic* from *alc-*, the more general meaning of 'being addicted to' is detached from the specific cause of addiction, and a new suffix is created, *-oholic*, which attaches to a noun X, and whose meaning is 'a person addicted to X'. This process created a new derivational suffix in the language. Similarly, the sequence *-tron* in Hebrew was extracted from *te'atron* ('theatre'), and signals a theatre group (Mirkin 1968), and the sequence *-fie* in Dutch was extracted from *selfie*, denoting specific self-photos of X taken by X (Audring and Jackendoff 2014).

- (iii) Hebrew morphologized derivational suffixes: In Hebrew there is a set of derivational suffixes, whose origin is attributed to non-morphemic subparts of certain nominal vocalic patterns in the language. The agentive suffix *-an*, which appears in words such as *psantran* ('pianist', *psanter* ('piano')+*-an*), *yeynan* ('wine maker', *yayin* ('wine')+*-an*), has its origins in the sequence *-an*, which is part of the vocalic pattern *CaCCan* (Ornan 1983: 30-33, 2001: 138; Schwarzwald 2002 (vol. 2): 20 and references there). Similarly, the Hebrew suffix *-ut*, deriving abstract nouns, was morphologized from words formed from roots whose third radical is *w/y*, and denote abstract nouns, such as *d^emut* ('image', root *d.m.w.*), *p^edut* ('salvation', root *p.d.y*) (Bolzky and Schwarzwald 1992: 52).
- (iv) Literal alliterative agreement: This phenomenon refers to cases where agreement markers attached to agreement targets are not affixes but rather part of the noun radical. In Bainuk (Sauvageot 1967), an Atlantic (Niger-Congo) language, prefixed nouns fall into eleven genders (each containing a matched pair of singular and plural prefixes) and show agreement through the prefix, which appears on pronouns, demonstratives, and adjectives, as exemplified by the noun *gu-sɔl* 'tunic' in Table 2. Nonprefixed nouns, which are often loanwords, trigger one of two sorts of gender agreement in both singular and plural: either a default prefix *a-* or an affix copy of the first CV of the noun stem (Sauvageot can find no reason for which method a given noun chooses). It is the latter mechanism, exemplified by the noun *kata:ma* 'river' in Table 2, that is relevant for our discussion here.⁸ What is copied onto the agreement targets is the sequence */ka/*, the first CV of the noun stem, which becomes an agreement affix.

Noun	class	Singular	Plural	Agreement environments
'tunic'	7/8 (prefixed)	gu-sɔl	ha-sɔl	gu-sɔl gu-fɛr 7-tunic 7-white 'white tunic'
'river'	0/0 (non-Prefixed)	kata:ma	kata:ma-ã	<u>k</u> ata:ma-ŋɔ in- <u>k</u> a river-DEF this-CV 'this river' <u>k</u> ata:ma-ã <u>k</u> a-nak-ã river-PL CV-two-PL 'two rivers'

Table 2: Agreement with prefixed and non-prefixed nouns in Bainuk

What all the four cases described above have in common is that morphemes arise by taking a phonological sequence in existing words and reanalyzing it as a unit, a morpheme. The reanalysis may be triggered by analogy to existing forms and constructions in the language, as with the case of *-er* in *babysitter*, or the case of *ka* in the literal alliterative agreement in Bainuk. It can also be triggered by a strong connection between specific forms and specific meanings, either in a single word, as in *alcoholic*, or in a group of words such as those sharing the sequences, *-an* and *-ut* in Hebrew.

The case of the emergence of ISL agreement morphemes is very similar to the case of the Hebrew affixes. A group of words that share a meaning component, namely an event of transfer, also shares a form: a path movement on the signer-addressee axis, going to or from the signer. Since points in space have a morphemic status in the R-loci system, the part of the verb that moves towards a point in space is reanalyzed as a morpheme. Unlike the Hebrew

⁸ For a deeper discussion of literal alliterative agreement, see Aronoff et al. 2005.

morphemes and the productive blends, the ISL agreement morphemes denote the referential features (phi-features) of the verb's arguments. As such, they contain more grammatical (rather than lexical) content, and they form a group of morphemes rather than an individual morpheme. Therefore they form a new grammatical category, not only new forms. While such a process has not been described for spoken languages, to the best of my knowledge, the emergence of agreement verbs in ISL shows that it is a possible way for a language to acquire new grammatical categories, in addition to the much more wide-spread and very-well attested mechanism of grammaticalization.

Acknowledgements

The research was funded by the Israel Science Foundation grant number 553/04 to Irit Meir. Figures are © of the Sign Language Laboratory, University of Haifa.

References

- Aronoff, M., I. Meir & W. Sandler (2005) The paradox of sign language morphology. *Language* 81: 301-344.
- Audring, J. & R. Jackendoff (2014) The texture of the lexicon. ACLC Colloquium, University of Amsterdam, November 7.
- Beckner, C., R. Blythe, J. Bybee, M. H. Christiansen, W. Croft, N. C. Ellis, J. Holland, J. Ke, D. Larsen-Freeman & T. Schoenmann (2009) Language is a complex adaptive system: Position paper. *Language Learning* 59: Suppl. 1: 1-26.
- de Beuzeville, L., T. Johnston, & A. Schembri (2009) The use of space with indicating verbs in Australian Sign Language: A corpus-based investigation. *Sign Language and Linguistics* 12(1): 52-83.
- Bolozky, S. & O. R. Schwarzwald (1992) On the derivation of Hebrew forms with the +*ut* suffix. *Hebrew Studies* 33: 51-69.
- Booij, G. (2005) *The grammar of words*. Oxford: Oxford University Press.
- Engberg-Pedersen, E. (1993) *Space in Danish Sign Language*. Hamburg: Signum-Verlag.
- Heine, B. & T. Kuteva (2002) *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Janis, W. (1992) *Morphosyntax of the ASL verb phrase*. Ph.D. Dissertation, SUNY Buffalo.
- Kiparsky, P. (2012) Grammaticalization as optimization. In: D. Jonas, J. Whitman & A. Garrett (Eds.), *Grammatical Change: Origins, Nature, Outcomes*. Oxford: Oxford University Press.
- Labov, W. (1963) The social motivation of a sound change. *Word* 18: 1-42.
- Lepic, R. (2015) *Motivation in Morphology: Lexical Patterns in ASL and English*. PhD dissertation, University of California, San Diego.
- Liddell, S. (2003). *Grammar, Gesture, and Meaning in American Sign Language*. Cambridge: Cambridge University Press.
- Lillo-Martin, D. & R. Meier (2011) On the Linguistic Status of 'Agreement' in Sign Languages. *Theoretical Linguistics* 37: 95-141.
- Lillo-Martin, D., & E. S. Klima (1990) Pointing out differences: ASL pronouns in syntactic theory. In S. D. Fischer, & P. Siple (Eds.), *Theoretical Issues in Sign Language Research, volume 1: Linguistics*. Chicago: University of Chicago Press, 191-210.
- Meier, R. P. (1990) Person deixis in American Sign Language. In S. D. Fischer, & P. Siple (Eds.), *Theoretical Issues in Sign Language Research, volume 1: Linguistics*. Chicago: University of Chicago Press, 175-190.
- Meillet, Antoine. (1912) L'évolution des formes grammaticales. *Scientia* (Rivista di Scienza) 12, No. 26, 6. *Linguistique historique et linguistique générale*, 130-48. Paris: Librairie Ancienne Honoré Champion, 130-148.
- Meir, I. (1998a) *Thematic Structure and Verb Agreement in Israeli Sign language*. PhD dissertation, Hebrew University, Jerusalem.
- Meir, I. (1998b) Syntactic-Semantic Interaction in Israeli Sign Language verbs: The Case of Backwards Verbs. *Sign Language and Linguistics* 1: 3-33.
- Meir, I. (2001) Motion and Transfer: The analysis of two verb classes in Israeli Sign Language. In: V. Dively (Ed.): *Signed Languages: Discoveries from international research*. Washington D.C.: Gallaudet University Press, 74-87.
- Meir, I. (2002) A cross-modality perspective on verb agreement. *Natural Language and Linguistic Theory* 20: 413-450.

- Meir, I. (2010) The Emergence of Argument Structure in Two New Sign Languages. In M. Rappaport Hovav, E. Doron & I. Sichel (Eds.), *Syntax, Lexical Semantics and Event Structure*. Oxford: Oxford University Press, 101-123.
- Meir, I. (2012) The evolution of verb classes and verb agreement in signed languages. *Theoretical Linguistics* 38(1-2): 145-152.
- Meir, I., C. Padden, M. Aronoff & W. Sandler (2007) The body as subject. *Journal of Linguistics* 43: 531-563.
- Meir, I., & Sandler, W. (2008) *A language in Space: The Story of Israeli Sign Language*. New York: Lawrence Erlbaum Associates.
- Meir, I., W. Sandler, C. Padden & M. Aronoff (2013) Competing Iconicities in the Structure of Languages. *Cognitive Linguistics* 24(2): 309-343.
- Mirkin, R. (1968) Le-hašlamat ha-xase ba-milonim (Filling in gaps in dictionaries). *Lešonenu la'am* 19(3): 72-76.
- Neidle, C., J. Kegl, D. MacLaughlin, B. Bahan & R. G. Lee (2000) *The syntax of American Sign Language: Functional Categories and Hierarchical Structure*. Cambridge, MA: MIT Press.
- Ornan, U. (1983) Tsurat ha-mila be-'ivrit keycad (How do we build a Hebrew word?), in M. Bar-Asher et al. (Eds.), *Hebrew Language studies presented to Zeev Ben-Hayyim*. Jerusalem: Magnes Press, 13-42.
- Ornan, U. (2001) *The Final Word*. Haifa: University of Haifa Press.
- Padden, C. (1988) *Interaction of Morphology and Syntax in American Sign Language*. New York: Garland.
- Padden, C., I. Meir, M. Aronoff & W. Sandler (2010) The grammar of space in two new sign languages. In D. Brentari (Ed.), *Sign Languages: A Cambridge Survey*. New York: Cambridge University Press, 573-595.
- Pizzuto, E. (1986) The verb system in Italian Sign Language. In B. Tervoort (Ed.), *Signs of Life: Proceedings of the Second European Congress on Sign Language Research*. Amsterdam: The Dutch foundation for the Deaf and Hearing Impaired Child, The University of Amsterdam and the Dutch Council of the Deaf, 17-31.
- Rathmann, C. & G. Mathur (2002) Is Verb Agreement the Same Cross-modally? In: R. Meier, K. Cormier & D. Quinto-Pozos (Eds.), *Modality and Structure in Signed and Spoken Languages*. Cambridge: Cambridge University Press, 370-404.
- Quadros, R. de & J. Quer. (2008) Back to Back(wards) and Moving on: On Agreement, Auxiliaries and Verb Classes. In: R. M. de Quadros (Ed.), *Sign Languages: Spinning and Unraveling the Past, Present, and Future. 9th Theoretical Issues in Sign Language Research Conference, Florianopolis, Brazil, December 2006*. Petropolis: Editora Arara Azul. [Available at: www.editora-arara-azul.com.br/EstudosSurdos.php].
- Sandler, W., I. Meir, C. Padden & M. Aronoff (2005) The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences* 102(7): 2661–2665.
- Sauvageot, S. (1967) Note sur la classification nominale en Bainouk. In *La classification nominales dans les langues Négro-Africaines*. Paris: CNRS, 225-236.
- Schembri, A., J. Fenlon & K. Cormier (2016) "Agreement" verbs in sign languages: are we missing the point? A talk presented at the 38th DGFS conference, Konstanz, Germany.
- Schwarzwald, O. R. (2002) *Studies in Hebrew Morphology*. Tel Aviv: The Open University of Israel.
- Trudgill, P. (2009) Sociolinguistic typology and complexification. In G. Sampson, D. Gil & P. Trudgill (Eds.), *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 98-109.

Appendix: List of clips used in the study

1. A woman *looks* at man
2. A girl *pulls* a man
3. A woman *pushes* a girl
4. A man *taps* a girl (on the shoulder)
5. A woman *gives* a shirt to man
6. A woman *takes* a pair of scissors from a girl
7. A man *shows* a picture to a woman
8. A man *throws* a ball to a girl
9. A girl *feeds* a woman yogurt

A paradigmatic analysis of the Italian verbal derivation

Fabio Montermini
CLLE-ERSS, CNRS
& Université de Toulouse II Jean Jaurès
fabio.montermini@univ-tlse2.fr

Matteo Pascoli
Università di Bologna
matteo.pascoli2@unibo.it

1. Introduction

According to the classification established by Stump (2001), Word-and-Paradigm theories of inflection can be classified as inferential-realizational. In these models, in fact, morphological exponents are not listed in the lexicon along with the corresponding morphosyntactic properties. Rather, they are inferential in that exponents are introduced by rules relating inflected forms with the corresponding lexemes, and are selected by their morphosyntactic properties; they are realizational in that morphosyntactic properties are not added to the word by an exponent, but select the exponents that realize them.

It has been observed (cf. *inter alios* Maiden 1992; Pirrelli and Battista 2000) that, on verbal stems that display allomorphy, alternations show a surprisingly regular distribution, which is not dictated by either the phonological context or the semantic homogeneity. This regularity reflects the organization of the verbal paradigm, or the set of all inflected forms for each lexeme, into *morphomes*, i.e. purely formal entities independent from morphosyntactic features (Aronoff 1994).

In the last twenty years, there has been much interest in the study of the *paradigmatic distribution of allomorphy*, or the way in which variation between forms (the traditional “irregularity”) of a paradigm rests on regular patterns.

Practically, these studies aim at analyzing the paradigmatic structure of inflection, i.e. to decompose the paradigm into zones where forms are realized on possibly distinct basic stems, and to examine the formal relations (at a phonological level) between these basic stems, looking for predictability chains allowing to handle both regular and irregular lexemes.

Moreover, it has been shown that, in Romance languages, the formal alternations that apply to inflection are the same that function in derivation (cf. Bonami *et al.* 2009 on French). In the present work, we examine the formal relations between some verbal derivatives in Italian and the basic stems of the related verbs with the goal of extending the study of paradigmatic distribution to derivation.

In Latin, deverbal derivatives in *-tionem* (event/result), *-torem* (agent/instrument), *-tura* (event/result), *-tivus* (relational adjective), *-torium* (adjective/instrument/place) were built on the supine (basic) stem, the same that was used to form the past participle. Italian, like other Romance languages, has inherited from Latin both the process and the derived forms.

(1)	event/result, N	<i>frittura</i> ‘frying’/‘fried food’	<i>friggere</i> ‘fry’, P.P. <i>fritto</i>
	event, N	<i>chiusura</i> ‘closure’	<i>chiudere</i> ‘close’, P.P. <i>chiuso</i>
	place, N	<i>pensatoio</i> ‘thinking place’	<i>pensare</i> ‘think’, P.P. <i>pensato</i>
	instrument, N	<i>rasoio</i> ‘razor’	<i>radere</i> ‘shave’, P.P. <i>raso</i>
	(place, N)/A	<i>uditorio</i> ‘audience’/‘auditory’	<i>udire</i> ‘hear’, P.P. <i>udito</i>
	A	<i>illusorio</i> ‘illusory’	<i>illudere</i> ‘deceive’, P.P. <i>illuso</i>
	converted, N	<i>raccolto</i> ‘harvest’	<i>raccogliere</i> ‘pick up’, P.P. <i>raccolto</i>
	converted, N	<i>riassunto</i> ‘summary’	<i>riassumere</i> ‘sum up’, P.P. <i>riassunto</i>

Some derivatives, and in some cases the past participle itself, underwent semantic drift. Some (ancient) past participles are no longer connected to a verb and remain in the language as independent adjectives (cf. *solito* ‘usual’). Some past participles of existing verbs underwent the same phenomenon (cf. *viso* ‘seen’ → ‘sight’ → ‘eyes’ → ‘face’ in modern Italian), while being replaced in their past participle function by analogical forms (*veduto/visto* ‘seen’). In these cases, derivatives do not display a transparent relation with the past participles of the base verbs. Yet, they maintain formal relations, and these relations allow identifying a basic stem, which is by default identical to the basic stem of the past participle, but can possibly be distinct. This basic stem can be related to other basic stems of the base verb.

2. Paradigmatic relations in Italian

A number of recent contributions on the inflectional morphology of Romance languages, and in particular of Italian, showed that, when a lexeme displays allomorphy, stems can be formally connected in predictable ways (cf. Pirrelli and Battista 2000; Montermini and Boyé 2012). Lexemes are identified by a “stem space”, i.e. a collection of stems bearing specification about the paradigm area they fill. The links connecting each stem are the same for all lexemes in the language, and some stems are more closely linked than others (cf. Montermini and Boyé 2012 on Italian for details). Under this view, regularity can be viewed as a gradient phenomenon: a lexeme is regular when all its stems are connected in a predictable way; it is irregular when at least one of its stems is unconnected with the rest of the stem space, and thus must be stored independently. Figures 1 and 2 show a portion of the stem space for a regular (traditional 1st class) verb in Italian, *ammirare* (‘admire’). S1 corresponds to the simplest stem in the paradigm (used, among others, to form the present 2PL, the imperfect indicative and subjunctive); S7 corresponds to the stem used to form the past participle, and S9 corresponds to the stem used to form the derivatives we referred to above. As the representation we propose shows, we consider that for *ammirare* the three stems are connected by predictable functions. This means that a speaker of Italian is able to recover any form in the paradigm (including derivatives) from any other.

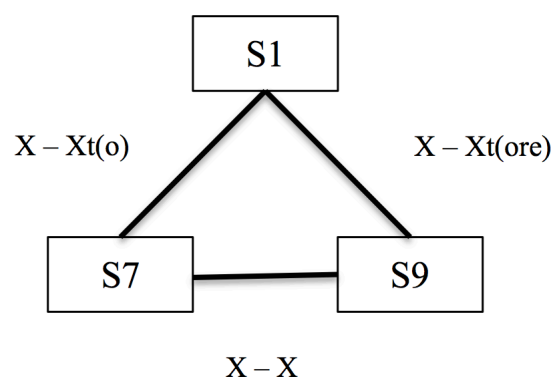


Figure 1: connections between S1, S7 and S9 for a regular verb

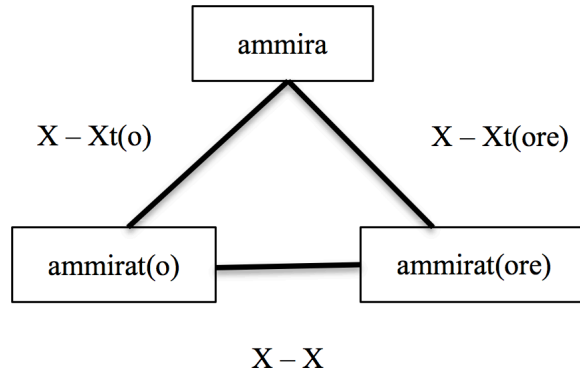


Figure 2: connections between S1, S7 and S9 for a regular verb (*ammirare*)

On the contrary, for an irregular verb, we consider that more than one stem must be stored for a lexeme. It is the case, for instance for *distruggere* (‘destroy’) (Figure 3).

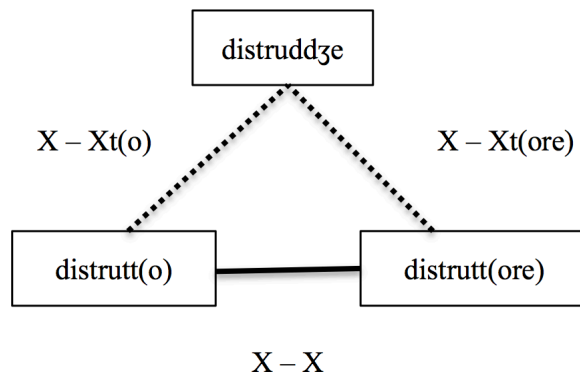


Figure 3: connections between S1, S7 and S9 for an irregular verb (*distruggere*)

As we see, for *distruggere* S7 and S9 are still connected by a predictable function (in this case an identity function), but in any case they are linked to S1. As the examples in (2) show, a S9 ending in [utt] can in fact correspond to S1 having different forms. Thus, we cannot consider that a predictable connection exists between the stems [distruddze] and [distrutt].

(2)	S1 S9	Glosses	
	<u>_uddze</u> (<i>distruggete</i> _{2.PL.PRES.IND})	<u>_utt</u> (<i>distruttore</i>)	‘destroy’/‘destroyer’
	<u>_ui</u> (<i>costruite</i> _{2.PL.PRES.IND})	<u>_utt</u> (<i>costruttore</i>)	‘build’/‘constructor’
	<u>_dutfè</u> (<i>producete</i> _{2.PL.PRES.IND})	<u>_dutt</u> (<i>produttore</i>)	‘produce’/‘producer’
	<u>_romp</u> (<i>interrompete</i> _{2.PL.PRES.IND})	<u>_rutt</u> (<i>interruttore</i>)	‘interrupt’/‘switch’

Some verbs display a structure which is even more complex. For a verb like *produrre* (‘produce’), for instance, neither of the stems under consideration is connected to the others by predictable functions (Figure 4).

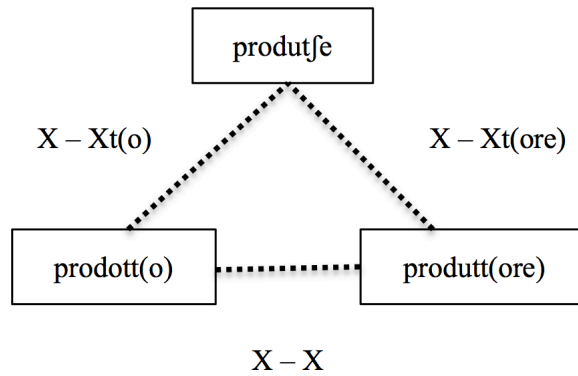


Figure 4: connections between S1, S7 and S9 for an irregular verb (*produrre*)

In synchrony, the stem connections that can be identified for verbs have thus the function of facilitating the recovery of unknown forms by speakers; in diachrony, it can be shown that in some cases paradigms have been restructured following the same lines. Verbs in *_struire* (Figure 5), for instance *costruire* ‘build’, *istruire* ‘educate’, etc., display a S9 which is etymological (cf. Latin *construo*, P.P. *constructum*), and thus unconnected to the rest of the paradigm, whereas the S7 has been remodelled in order to be homogeneous with the rest of the paradigm, and is thus connected to the S1.

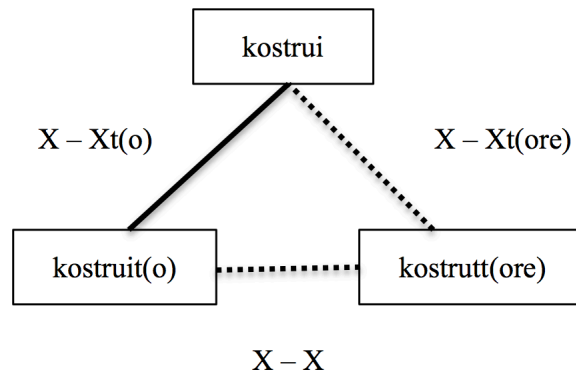


Figure 5: connections between S1, S7 and S9 for a partially regularized verb (*costruire*)

Conversely, for a verb like *scoprire* (‘discover’) the S7 is etymological (from Lat. *operio*, P.P. *opertum* ‘cover’ by double prefixation), whereas the S9 has been remodelled on the rest of the paradigm (Figure 6).

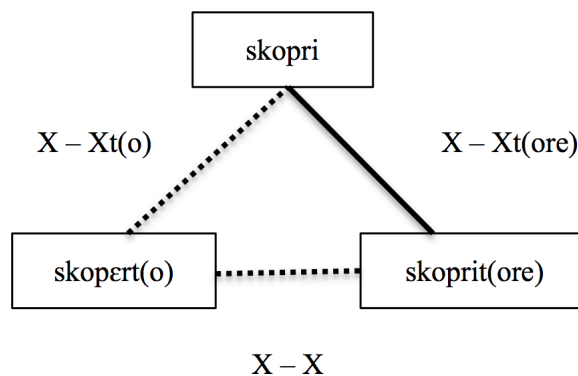


Figure 6: connections between S1, S7 and S9 for a partially regularized verb (*scoprire*)

3. Finding relation between verb stems

We constructed a corpus of 55 forms for 139 Italian verbs. The verb dataset includes all the conjugation models of the GRADIT dictionary (De Mauro 1999). The considered forms include the 46 strictly verbal synthetic forms, 6 participial forms, and 3 deverbal derivatives (-*tivo*, -*tore*, -*zione*), when they exist in the same dictionary for the corresponding verbal lexeme. For the purpose of the present research, the derivatives are considered as part of the verbal paradigm.

We tested a computational technique that could perform two tasks:

- (i) for each paradigm cell, separation of the invariant part which is common to all lexemes (we will call this the “model”) from the part specific to each lexeme (we will call this the “stem” in a broad sense);
- (ii) for each lexeme, collection of the morphophonetic relations between each couple of stems identified in (i).

As an example, the first task should compute the forms *aboliamo*, *capiamo*, *cogliamo*, *dirigiamo*, *siamo*, *vediamo* (“we abolish”, “we understand”, “we pick”, “we direct”, “we are”, “we see”), and give as result: ‘X+iamo; X=*abol, cap, cogl, dirig, s, ved*’ (the actual computations were performed on phonetically transcribed strings). The second task should take for example two stems for the lexeme DIRIGERE “to direct”, [di'rig] and [di'ridʒ] and automatically give the output ‘replace [g] with [dʒ] in word final position’.

To perform the tasks described above, we examined the *string distance measure* algorithms. These algorithms aim at quantifying the difference between two given strings, or sequences of symbols.

For instance, one of the distance measures is computed by using the Hamming algorithm (Hamming 1950). This algorithm counts the symbols of the two strings that are different in each given position:

a	l	b	e	r	o
a	l	b	e	r	i
√	√	√	√	√	X
Hamming distance: 1					

Table 1: Hamming distance between *albero* and *alberi*

a	l	b	e	r	o		
v	i	l	l	e	t	t	e
X	X	X	X	X	X	X	X
Hamming distance: 8							

Table 2: Hamming distance between *albero* and *villette*

However, such a simple algorithm is not adequate for our needs. Let us take for instance the strings *abcdefgh* and *bcdefgh*:

a	b	c	d	e	f	g	h
b	c	d	e	f	g	h	
X	X	X	X	X	X	X	X
Hamming distance: 8							

Table 3: Hamming distance between *abcdefgh* and *bcdefgh*

Intuitively, these strings are very similar, much more than the strings *albero* and *villette* in Table 2, for which we get the same result. There exists, however, a different algorithm which can deal with such cases: the Levenshtein algorithm can find similarities and differences in two strings even when they are in different positions (Levenshtein 1966).

The Levenshtein method can analyse the transformation of a string into another by decomposing it into a sequence of basic operations. These basic operations are Insertion (a symbol occurs in the second string but is absent from the first string), Deletion (a symbol occurs in the first string but is absent from the second one), Substitution, (two symbols in the same position are different in the two strings), No-change (the same symbol occurs in the same position in the two strings).

The Levenshtein algorithm is relatively complex and is based on the construction of a rectangular matrix with as many columns as the length of the first string plus one, and as many rows as the length of the string word plus one.

Let us take a simple example to illustrate how the algorithm works for finding the transformation of the string [kaste] into the string [kare]. First, we prepare a matrix, in the way described above, and we fill the first row and the first column with consecutive numbers starting from 0, like in the example below (Figure 7):

•	•	k	a	s	t	e
•	0	1	2	3	4	5
k	1					
a	2					
r	3					
e	4					

Figure 7: setup of the Levenshtein computation

Then, we fill the other cells, from top to bottom and from left to right, with a simple computation: if the symbols labeling the column and the symbol labeling the row coincide, we copy the value found in the cell above at the left of the new cell (Figure 8):

•	•	k	a	s	t	e
•	0	1	2	3	4	5
k	1	0				
a	2					
r	3					
e	4					

Figure 8: cell relative to corresponding symbols

If the symbols labeling the column and the row are different, we take the lowest of the values in the above, left, and above left cells (relative to the next empty cell), add 1, and write the resulting value in the new cell (Figure 9).

•	•	k	a	s	t	e
•	0	1	2	3	4	5
k	1	0	→1			
a	2					
r	3					
e	4					

Figure 9: cell relative to different symbols

Once all the cells in the matrix have been filled, the Levenshtein measure corresponds to the value in the bottom right cell (2 in the case in point, Figure 10):

•	•	k	a	s	t	e
•	0	1	2	3	4	5
k	1	0	1	2	3	4
a	2	1	0	1	2	3
r	3	2	1	1	2	3
e	4	3	2	2	2	2

Figure 10: finishing the computation

This method of computation is satisfying because it can detect invariant symbols even if they do not occupy the same positions in the two strings, like “e” in the example above.

For our research, however, it was necessary that the output of the computation corresponded to a transformation rather than to a single numeric value. A crucial point is that the distance measure is minimal and unique, but not so the set of transformations that brought to it. Let us examine the matrix in the example above; two different transformation sequences in (3) give the same distance measure, i.e. 2:

- (3) N (k); N (a); S (r,s); D (t); N (e)
 N (k); N (a); D (s); S (r,t); N (e)

To take another example, let us see the transformations of [sedevo] into [sjedo] (corresponding, respectively, to the 1SG of the imperfect and to the 1SG of the present of *sedere* ‘sit’):

•	•	s	e	d	e	v	o
•	0	1	2	3	4	5	6
s	1	0	→1	→2	3	4	5
j	2	1	→1	→2	3	4	5
e	3	2	→1	→2	→2	3	4
d	4	3	2	→1	→2	→3	4
o	5	4	3	2	2	3	→3

Figure 11: Levenshtein algorithm computed for *sedevo* and *siedo*

In this case, three possible paths give the result 3:

- (4) a. N(s); D(e); S(d,j); N(e); S(v,d); N(o)
 b. N(s); S(e,j); D(d); N(e); S(v,d); N(o)
 c. N(s); I(j); N(e); N(d); D(e); D(v); N(o)

So, we developed a modified implementation of the Levenshtein algorithm, that gives exactly the output above. Additionally, the transformation paths can be simplified, by grouping together consecutive similar transformations:

- (5) d. N(s); S(ed,j); N(e); S(v,d); N(o)
e. N(s);I(j);N(ed);D(ev);N(o)

Of course not all these transformation chains are morphologically significant, like for instance (5d) above. We will see how some of these are peculiar to single form pairs of single lexemes, while others are common to many lexemes.

Note, moreover, that the algorithm in question works promisingly also for non-concatenative inflection. For instance, given as input the two Arabic forms *faʕaltu* ‘I did’ and *saʕimtu* ‘I hated’, the algorithm gives the following output (“+” stands for the lexeme-variant matter, to be replaced by the lexeme’s specific stem for that paradigm cell, and “/” suggests the discontinuity in the stems of Arabic verbs):

- (6) *faʕaltu* → *saʕimtu*:
operations: S(f,s); N(a); S(ʕ,ʔ); S(a,i); S(l,m); N(t); N(u)
model: +a+tu
stems: f/ʕal, s/ʕim

4. Models and stems

By applying the algorithm described above to all lexemes contained in the corpus for each paradigm cell, we obtained the models in Table 4:

IND.PRES.1S	+	IND.FUT.2S	+r'aj	COND.PRES.3S	+r'ebbe
IND.PRES.2S	+	IND.FUT.3S	+r'a	COND.PRES.1P	+r'emmo
IND.PRES.3S	+	IND.FUT.1P	+r'emo	COND.PRES.2P	+r'este
IND.PRES.1P	+v'amo	IND.FUT.2P	+r'ete	COND.PRES.3P	+r'ebbero
IND.PRES.2P	+te	IND.FUT.3P	+r'anno	IMP.2S	+
IND.PRES.3P	+no	CONG.PRES.1S	+	IMP.2P	+te
IND.IMPF.1S	+o	CONG.PRES.2S	+	GERU	+ndo
IND.IMPF.2S	+i	CONG.PRES.3S	+	INF	+re
IND.IMPF.3S	+a	CONG.PRES.1P	+v'amo	PART.PRES.S	+nte
IND.IMPF.1P	+v'amo	CONG.PRES.2P	+v'ate	PART.PRES.P	+nti
IND.IMPF.2P	+v'ate	CONG.PRES.3P	+no	PART.PASS.M.S	+o
IND.IMPF.3P	+ano	CONG.IMPF.1S	+ssi	PART.PASS.M.P	+i
IND.PREM.1S	+	CONG.IMPF.2S	+ssi	PART.PASS.F.S	+a
IND.PREM.2S	+sti	CONG.IMPF.3S	+sse	PART.PASS.F.P	+e
IND.PREM.3S	+	CONG.IMPF.1P	+ssimo	D_TIVO.S	+ivo
IND.PREM.1P	+mmo	CONG.IMPF.2P	+ste	D_TORE.S	+ore
IND.PREM.2P	+ste	CONG.IMPF.3P	+ssero	D_ZIONE.S	+j'one
IND.PREM.3P	+o	COND.PRES.1S	+r'ej		
IND.FUT.1S	+r'o	COND.PRES.2S	+r'esti		

Table 4: invariants for each paradigm cell

Furthermore, it happens that for some paradigm cell sets, all the lexemes have the same stem, or the same stem set in case of overabundance, like in the example below (Table 5):

<i>cell</i>	<i>model</i>	AMARE	CRESCERE	FARE
IND.IMP.F.1P	+v'amo	ama	krejfe	fatfe
IND.IMP.F.2P	+v'ate	ama	krejfe	fatfe

Table 5: consistent identity of stems for some groups of paradigm cells

Thus, only one paradigm cell from the set has to be included in further computations, and we will refer to that set as *metacell*.

5. Finding transformation rules

At this point, our interest lay in extracting a list of possible relations occurring between the stems of a lexeme. The modified Levenshtein Algorithm we implemented was applied to every pair of stems for all lexemes, each representing a metacell.

Table 6 shows the relations between the stem of the INDICATIVE PRESENT 1S and all the other stems (the figures indicate the number of lexemes for which the relation is valid):

CONG.PRES.1S	122 (87%)	N;S(a,o)	ABOLIRE, ACCENDERE, ADDURRE, ...
CONG.PRES.1S	10 (7%)	N;S(i,o)	ABBANDONARE, AMARE, CONSUMARE, ...
CONG.PRES.1S	2 (1%)	N;S(ia,o)	DARE, STARE
CONG.PRES.1S	2 (1%)	N;S(i,j,o)	CAMBIARE, COPIARE
CONG.PRES.1S	2 (1%)	N;D(j);N;S(da,ggo)	POSSEDERE, SEDERE
CONG.PRES.1S	2 (1%)	N;I(j);N;S(gga,do)	POSSEDERE, SEDERE
CONG.PRES.1S	1 (0%)	N;S(va,bbo)	DOVERE
CONG.PRES.1S	1 (0%)	N;S(ia,ono)	ESSERE
CONG.PRES.1S	1 (0%)	N;S(abbja,o)	AVERE
CONG.PRES.1S	1 (0%)	N;S(bba,vo)	DOVERE
CONG.PRES.1S	1 (0%)	N;S(appja,o)	SAPERE
IND.PRES.3P	120 (86%)	N	ABOLIRE, ACCENDERE, ADDURRE, AFFIGGERE, ...
IND.PRES.3P	12 (8%)	N;S(a,o)	ABBANDONARE, AMARE, CAMBIARE, ...
IND.PRES.3P	4 (2%)	N;S(an,o)	AVERE, DARE, SAPERE, STARE
IND.PRES.3P	2 (1%)	N;D(j);N;S(d,gg);N	POSSEDERE, SEDERE
IND.PRES.3P	2 (1%)	N;I(j);N;S(gg,d);N	POSSEDERE, SEDERE

Table 6: relations of the IND.PRES.1S stem

Table 7 below shows the relations between the stem of the nominal derivative *-zione* and some other stems:

D_TIVO.S	11 (40%)	N;S(t,ts)	ABOLIRE, ASSOLVERE, ASSUMERE, DARE, ...
D_TIVO.S	8 (29%)	N	ALLUDERE, CONCEDERE, DIFENDERE, EMETTERE, LEDERE, ...
D_TIVO.S	7 (25%)	N;S(tt,ts)	ADDURRE, CONDURRE, COSTRUIRE, DIRIGERE, ...
D_TIVO.S	1 (3%)	N;S(tt,ss)	CONNETTERE
D_TORE.S	11 (28%)	N;S(tt,ts)	ADDURRE, CONDURRE, COSTRUIRE, DIRIGERE, ...
D_TORE.S	10 (26%)	N;S(t,ts)	APPARIRE, ASSOLVERE, ASSUMERE, CONSUMARE, ...
D_TORE.S	7 (18%)	N	AGGREDIRE, CONCEDERE, DIFENDERE, FONDERE, ...
(...)			
PART.PASS.M.S	9 (16%)	N;D(');N;S(t,ts)	ABOLIRE, ASSUMERE, CONSUMARE, ...
PART.PASS.M.S	7 (12%)	N;D(');N;S(tt,ts)	AFFLIGGERE, CUOCERE, DISTRUGGERE, ...
PART.PASS.M.S	7 (12%)	N;D(');N	AFFIGGERE, ALLUDERE, CONNETTERE, ...
PART.PASS.M.S	3 (5%)	N;S('ett,ets)	DIRIGERE, LEGGERE, PREDILIGERE
(...)			
IND.PRES.2P	7 (12%)	N;D(');N;I(ts)	ABOLIRE, APPARIRE, CONSUMARE, DARE, ...
IND.PRES.2P	5 (8%)	N;S('a,ats)	CONSUMARE, DARE, FARE, MASTICARE, ...
IND.PRES.2P	4 (7%)	N;S(d'e,s)	ACCENDERE, ARDERE, DIFENDERE, TENDERE
IND.PRES.2P	3 (5%)	N;S(d'e,ss)	CEDERE, CONCEDERE, POSSEDERE

Table 7: relations of the -ZIONE stem

6. Conclusion and future work

With this work, we researched an automatic method of identifying verbal stems and finding their morphophonetic relations. An ad-hoc implementation of the Levenshtein algorithm seems to give the expected results on a corpus of Italian verbal forms representing all the conjugation models. Moreover, the same procedure can be conducted by introducing to the corpus also the forms of deverbal derivatives, in order to include them in the paradigmatic analysis of inflection.

Future work should undoubtedly include a method for constructing chains of predictability, based on the computation of the conditional entropy of the relations found, coupled with frequency data from a corpus.

References

- Aronoff, M. (1994) *Morphology by Itself: Stems and Inflectional Classes*. Cambridge, MA-London: The MIT Press.
- Bonami, O., Boyé, G. & F. Kerleroux (2009) L'allomorphie radicale et la relation flexion-construction. In: B. Fradin, F. Kerleroux, & M. Plénat (Eds.), *Aperçus de morphologie du français*. Saint-Denis: Presses Universitaires de Vincennes, 103-125.
- De Mauro, T. (Ed.) (1999) *Grande dizionario italiano dell'uso*. Torino: Utet.
- Hamming, R. W. (1950) Error Detecting and Error Correcting Codes. *Bell System Technical Journal* 29 (2): 147-160.
- Levenshtein, V. I. (1966) Binary Codes Capable of Correcting Deletions, Insertions and Reversals. *Soviet Physics Doklady* 10: 707-710.
- Maiden, M. (1992) Irregularity as a Determinant of Morphological Change. *Journal of Linguistics* 28(2): 285-312.
- Montermini, F. & G. Boyé (2012) Stem relations and inflection class assignment in Italian. *Word Structure* 5(1): 69-87.
- Pirrelli, V. & M. Battista (2000) The Paradigmatic Dimension of Stem Allomorphy in Italian Verb Inflection. *Italian Journal of Linguistics* 12: 307-379.
- Stump, G. T. (2001) *Inflectional Morphology: A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.

A comparison of roots as units of analysis in Modern Hebrew and Spanish: exploring a remnant approach to defining roots

Ignacio L. Montoya
City University of New York
 imontoya@gradcenter.cuny.edu

1. Introduction

The root as a morphological unit has been utilized in the description and explanation of a multitude of linguistic patterns. In most accounts of Semitic languages, for example, abstract, non-concatenative roots serve as the formatives around which the system is structured (e.g., Harris 1941; McCarthy 1981; Arad 2005), and in Indo-European historical linguistics, most of the lexicon has been reconstructed in terms of roots that correspond to the template *CeC, or a predictable variant (Forston 2004).

The concept of the root, though, is not approached consistently across diverse morphological theories. In some models, for example, the root is a meaningful unit (McCarthy 1981), and in others, simply a unit of form (Aronoff 1994). Even in the treatment of Semitic languages, where the root has long played a role (Aronoff 2013), the root as a necessary grammatical unit has recently been questioned (e.g., Bat El 1994; Ussishkin 2005). Moreover, across different linguistic traditions, the construct of the root has been conceptualized in distinct ways, making it difficult to make cross-linguistic comparisons. Considering the treatment of roots in Semitic, Indo-European, and Austronesian, Blust (1988) speculates that “the principal common denominator in [the various uses of the term root] appears to be an accidental coincidence in terminology” (2). In fact, cross-linguistic comparisons of word-internal morphological constituents such as roots and stems are not widely represented in the literature.

Given this scarcity of cross-linguistic comparisons of roots, the present study considers one of the problems involved with cross-linguistic investigations, that of defining roots in such a way that meaningful comparisons can be made. To that end, this paper explores a methodology for identifying roots across languages that addresses the concern expressed by Blust (1998) by ensuring that we are comparing like units across different languages. This methodology will be applied to the morphological systems of Modern Hebrew (henceforth Hebrew) and Spanish, which then allows for us to consider the extent to which roots are comparable across the two languages.

This paper is organized as follows. In Section 2, two approaches for characterizing roots are presented: the positive approach, in which roots are defined in terms of properties they possess or functions they perform, and the negative approach, in which roots are characterized as the element that remains after all morphological processes have been accounted for. The negative, remnant approach will be adopted in this paper, and will be discussed in greater detail in Section 3. In Section 4, the process of identifying roots will be presented and applied to Hebrew and Spanish. The elements yielded from the application of this procedure will be described, as will problematic issues that merit further investigation. Section 5 offers a comparison of the properties of roots identified via the remnant approach. Section 6 concludes

with a discussion about how this approach might be used to develop a more extensive typology of roots and about potential theoretical implications of the findings of this paper.

2. Operationalizing roots

Broadly speaking, roots may be characterized in one of two ways, which may be termed positive and negative characterizations. The positive characterization involves conceptualizing roots in terms of the properties they possess or the functions they perform, both of which are determined prior to the analysis of a specific language. In the introductory textbook *Language Files*, for example, a root is defined as “the free morpheme or bound root in a word that contributes most semantic content to the word, and to which affixes can attach” (Mihalicek and Wilson 2011: 703). As this glossary entry exemplifies, a root is defined in terms of what it is (i.e., a type of morpheme) and what it does (i.e., contributes semantic content and serves as the base of affixation). In this characterization, roots are building blocks, units that are selected, or the starting points of morphological processes.

The negative characterization, on the other hand, involves conceptualizing roots in terms of what remains after inflectional and derivational morphology has been accounted for (Blevins 2006). This characterization is represented in a different introductory textbook. In Robins’ (1964) *General Linguistics: An Introductory Survey*, the root is defined as “the part of a word structure which is left when all the affixes have been removed” (206). More recently, it has similarly been described as “what is left when all morphological structure has been wrung out of a form” (Aronoff 1994: 40). In this characterization, words are the starting point and roots are the remnants of analysis.

Broadly speaking, morphological analyses described as morpheme-based (following Stewart 2008), such as Distributed Morphology (Halle and Marantz 1994), adopt a positive characterization, whereas those that are word/lexeme-based, such as A-Morphous Morphology (Anderson 1992), adopt a negative characterization. The present paper adopts a negative characterization consistent with word/lexeme-based models.

3. Remnant approach to roots

Though certainly in many instances both the positive and the negative approaches identify the same element as a root, this paper is based on the assumption that the negative, remnant characterization is better suited for answering questions about cross-linguistic comparisons of roots. One of the problems with the positive approach is that it entails having a set of properties or functions for roots defined *a priori*. If we take a positive approach to characterizing roots, then we run the risk of either (a) having used different characteristics in different languages, as Blust (1988) suggests has happened, thereby hindering meaningful cross-linguistic comparisons, or (b) deciding on an *a priori* set of properties or functions that may not actually be appropriate for certain languages. In contrast, the negative, remnant approach provides a procedure for identifying roots that is independent of particular properties. Given that it can be applied in the same way across different languages to yield what can be considered like units, this procedure for identifying roots facilitates meaningful cross-linguistic comparisons. Across different languages, we would be comparing the part of the word that expresses no grammatical information. Furthermore, the fact that roots are not identified based on their properties or functions then allows for a more empirical exploration of the properties and functions of those units.

Though implications for morphological theory using this approach are considered in Section 6, the procedure for identifying roots developed in this paper is not based on very many strong theoretical claims about morphological structure, though, certainly, no linguistic

analysis is truly pre-theoretical. One of the main assumptions of this paper, for example, is that the word is a relevant and relatively stable cross-linguistic unit of analysis. This follows work in what may be broadly called a Word and Paradigm orientation (e.g., Anderson 1992). Though the construct *word* is not a fundamental unit in certain approaches (e.g., the Distributed Morphology framework of Halle and Marantz 1994), a great deal of literature supports the notion of the word as a foundational morphological unit (e.g., Robins 1959; Aronoff 1976; Dixon and Aikhenvald 2002; Blevins 2016). Given the emphasis on the surface grammatical word as a starting point of analysis, therefore, it would be reasonable to say that this approach is better aligned with a word-based model, though it is not necessarily inconsistent with a morpheme-based model, depending on how such a model characterizes the grammatical word.

4. Identifying roots

Adopting the remnant approach, one would identify roots by starting with a surface grammatical word. The next step would be to remove all of the exponents of grammatical properties that are expressed in that word. This includes both inflectional elements (i.e., those that participate in relating forms within paradigms), such as person, number, and tense, and derivational properties (i.e., those that involve relationships between lexemes), such as adjectivization, causative, and diminutive markers.

One consequence of this approach is that, though there are debates in the literature regarding the status of, for example, theme vowels in Spanish and templates (a.k.a., *binyanim*) in Hebrew, this process would remove all such exponents regardless of whether they are inflectional or derivational since they are in any case associated with grammatical features. Once this unit, which we will call the root, is determined in this way, we can then start to examine the properties that the root possesses and the roles it plays in the linguistic system.

We should also note that a similar procedure can be applied to identify the stem, which can be considered the element that remains after inflectional exponents have been removed. The stem merits a much fuller discussion of its own, and conclusions regarding the nature of the stem would also inform us further regarding the nature of the root. However, such an investigation is beyond the scope of this paper.

4.1 Identifying roots in Hebrew

Using the remnant approach to identify roots in verbs in Hebrew is relatively straightforward, unlike the more problematic nominal system, which is discussed below. Identifying the root in a verb involves removing the exponents of person, number, gender, tense, mood, voice, and inflectional class. A key component of Hebrew verbal structure is the *binyan*, also known as the template. This paper assumes Aronoff's (1994) treatment of *binyanim*, in which they constitute both inflectional and derivational elements, though the overall analysis would not change if *binyanim* were considered either strictly inflectional or strictly derivational. Exponents of inflectional class are the markers of the *binyan* (i.e., conjugation class, which is expressed as a prosodic template) that a surface word belongs to. In the verbal system, this procedure yields exactly the kind of element that one would expect: an abstract, discontinuous, consonantal sequence – the traditional Semitic root. The following examples illustrate the application of this procedure:

(1) *fomeret* 'she guards' → *fomer* → *f-m-r*

(2) *hitlabef* 'he dressed himself' → *l-b-f*

In (1), starting with the surface form *fomeret* ‘she guards’ and removing the gender exponent – the suffix *-et* – produces *fomer*. Since the fact that this verb is in the present tense is indicated by the vowels *o* and *e*, these vowels are also removed, yielding the root *f-m-r*. The same process yields the same kind of tri-consonantal unit in (2), where the surface word is *hitlabef* ‘he dressed himself’. Of note in this example is the fact that one could parse the exponents of the various features in different ways. For example, one might consider that the initial *h* by itself marks past tense since it is found in the past tense form of at least one other inflectional class, or one could treat the whole prefix *hit* as a unit associated with inflectional class. In either case, though, all of the elements except *l-b-f* provide grammatical, as opposed to lexical, information. For example, the distinction between prefixal consonants and root consonants can be seen in McCarthy’s (1979) treatment of the prefixes as occupying a tier distinct from the root tier. With regard to vowels, in some cases, such as in (1), they mark tense and in others, such as in (2), they mark inflectional class. Moreover, in some cases, the *binyan* itself can also be considered derivational (Aronoff 1994). Regardless of the parsing, though, all of these elements are exponents of one grammatical feature or another, so they are all removed to yield the root.

With regard to identifying nouns and adjectives in Hebrew using the remnant approach, removing the inflectional exponents is much more straightforward than dealing with the derivational exponents. The inflectional exponents are readily identifiable; these include gender and number suffixes, as well as the possessive markers and markers of definiteness, whether these are considered affixes or clitics. The complications for nominal forms arise, however, as a result of issues regarding nominal derivational patterns, what are referred to as the *mishkalim*. The following examples serve to illustrate:

- (3) *jatsran* ‘manufacturer’ → *j-ts-r* (cf. *jitser* ‘he manufactured’; *dabran* ‘talkative’, *diber* ‘he talked’)
- (4) *fulxan* ‘table’ → *fulxan* (no relation to *falax* ‘he sent’)

In some cases, a word indeed seems to clearly exhibit derivational morphology that can be removed. For example, in (3), the noun *jatsran* ‘manufacturer’ seems straightforwardly derived from its verbal counterpart, *jitser* ‘he manufactured’, with *j-ts-r* as its root, in the same way that *dabran* ‘talkative’ is derived from the verb *diber* ‘he talked’ and shares the root *d-b-r*.

On the other hand, a word such as *fulxan* ‘table’ in (4) has no derivational morphology associated with it. What might appear to be a tri-consonantal root in that word, namely *f-l-x*, is indeed a root associated with a verb. However, the verb in this case, *falax* ‘he sent’ is not related to *fulxan*, either synchronically or diachronically. An alternative analysis would be to claim that words such as *fulxan* are not based on roots. However, if we are defining the root as the element that remains after all inflectional and derivational information is removed, then we must conclude that the root of *fulxan* is essentially the non-decomposable surface word itself. This observation that nominals in Hebrew appear to have two different kinds of roots was noted in Arad’s (2005) Distributed Morphology analysis of Hebrew roots in which she speaks of two different kinds of roots: consonantal and syllabic.

4.2 Identifying roots in Spanish

Applying the remnant approach to surface words in Spanish yields elements that have been referred to as roots by some and stems by others, which supports the idea that the distinction

between roots and stems is complicated and suggests that further work on how to treat the stem is critical for understanding how these two elements interact.

As we did for Hebrew, we start our treatment of Spanish by first removing the inflectional exponents, such as tense, aspect, person, and number, and then removing derivational exponents. Both of these processes are relatively straightforward. Theme vowels, though, which are sometimes considered to be empty morphs, potentially pose some complications. However, the analysis of this paper follows that of others in which theme vowels are treated as markers of inflectional classes (Harris 1991; Aronoff 1994). Thus, given that they are exponents of grammatical properties, they are removed in order to determine the root.

As in Hebrew, another issue that can arise has to do with multiple potential segmentations. For example, a variety of segmentations can be justified for a verb like *bailaríamos* ‘we would dance’ in (5): *bail-ariamos*, *bail-aria-mos*, *bailar-íamos*, *bail-ar-ía-mos*. In all cases, though, the root is the same, since the competing segmentations all involve grammatical exponents.

(5) *bailaríamos* ‘we would dance’ → *bail-*

(6) *desordenar* ‘to disarrange’ → *desorden-* → *orden-*

The example in (6) illustrates how a derivationally complex word is stripped of its inflectional morphology to yield a stem and then stripped of its derivational morphology to yield a root.

4.3 Interim conclusion and problems to be resolved

The discussions in Sections 4.1 and 4.2 involve relatively unproblematic examples. They are meant to illustrate the principles involved in the remnant approach and to demonstrate how this procedure in its most straightforward application leads to analyses that are different from some other accounts. For example, forms such as *fulxan* in (4) are not generally regarded as roots, in large part because they do not have the properties that are adopted prior to analysis as the defining properties of roots in Semitic (i.e., their abstract, discontinuous nature). In a remnant approach, though, such forms are like the canonical tri-consonantal roots, such *f-m-r*, *l-b-f*, and *j-ts-r* in (1-3), in that they too are the parts of the word that contain no grammatical information. There is a way, therefore, in which the two types are like entities and it is worth considering the kinds of conclusions that can be reached in treating them as such. In addition, given that the Spanish roots in (5) and (6) are also operationalized as the parts of the word that do not convey any grammatical information, roots in Hebrew and roots in Spanish determined using a remnant approach are also like entities, and a cross-linguistic comparison of the two has the potential to yield meaningful insights.

It is important to note, though, that the present analysis serves as simply a preliminary treatment of the topic. As such, it raises many issues that need further development. One of the issues to be resolved in the application of the remnant approach involves cases where some kind of morphological structure seems apparent but it is problematic to treat that structure as either derivational or inflectional. In Spanish, for example, one of those cases involves Latinate roots, which have a parallel in English (cf. Aronoff 2007). If we take a word such as *recibir* ‘to receive’ we see that it seems to be composed of two elements, *re-* and *-cibir*, both of which recur in other other words in which they play a similar role (e.g., *reducir* ‘to reduce’, *percibir* ‘to perceive’). However, these relationships cannot be straightforwardly categorized as either clearly inflectional or derivational and are perhaps best considered as simply diachronic artifacts that are no longer active synchronically. In Hebrew, this issue is evident with certain kinds of nominal patterns, or *mishkalim*. In the word *misrad* ‘office’, for

example, we can recognize a morphological element, the *miCCaC* template, that is evident in words such as *mikdaf* ‘temple’ or *mivxan* ‘test’ (cf. *kidef* ‘he sanctified’ and *baxan* ‘he examined’); however, it is difficult to argue that in *misrad* that template is an exponent of a grammatical or semantic feature. Other issues that should be considered are how to deal with cases of root allomorphy in which a basic form is not clearly motivated. Both Spanish and Hebrew offer examples of this (e.g., involving diphthongization in the former and spirantization in the latter).

5. Comparison of roots in Spanish and Hebrew

Once roots have been identified in a language, we can then ask questions about the properties and roles of roots in that language. Thus, rather than assuming that, for instance, roots in Hebrew are abstract, discontinuous, consonantal forms and then concluding that they are not evident in certain nominal forms, we instead include those nominal forms in our calculations of roots and then ask whether roots in Hebrew have a particular shape. In this section, those kinds of questions are asked for Hebrew and Spanish, and the answers for each language are compared.

One of the questions that is commonly asked with regard to roots is whether they are morphemic (i.e., units of meaning), or whether they are purely units of form. Can a stable lexical meaning be associated with a root? In Hebrew, the answer to that depends on whether we look within a lexeme or across lexemes. Within a lexeme, roots do indeed seem to be associated with a stable lexical meaning, as in (7), in which all forms are clearly associated with a meaning of LEARN.

- (7) *lamad* ‘he learned’, *lomedet* ‘she learns’, *nilmad* ‘we will learn’

This is not necessarily the case across lexemes, though, as illustrated by (8).

- (8) *kibed* ‘he offered refreshment’, *hixbid* ‘he burdened’, *kaved* ‘liver’

Indeed, as has been noted in the literature, trying to find a common meaning in certain sets of words that share a common root, such as those in (8), is theoretically and empirically very problematic (Aronoff 2007).

In Spanish, the same kind of stability within a lexeme is evident. In contrast to Hebrew, though, significant stability is also found across lexemes. The nature of the morphological system supports fewer instances of completely opaque meanings. That is in large part because roots do not tend to be as frequently shared across lexemes. For example, whereas one root can be associated with several conjugation classes in Hebrew, this is not the case in Spanish. Furthermore, Spanish tends to have more transparent derivational morphology, unlike the complex system of nominal derivational patterns, the *mishkalim*, of Hebrew.

Another question that can be asked about roots is whether they have a distinctive characteristic shape in a particular language. In Hebrew, the answer to that is, in large part, yes. With regard to all verbal and many nominal words, roots have a readily identifiable prosodic shape: the canonical tri-consonantal sequence or one of a limited set of variants (e.g., roots with four or more consonants). These sequences are abstract in the sense that they are not pronounceable in isolation. As noted earlier, though, there is an asymmetry between nouns and verbs such that abstract roots cannot be identified in certain nominals, such as the previously mentioned *fulxan* ‘table’, in (4), from which no additional morphological structure can be removed. Therefore, though abstract canonically tri-consonantal roots can be identified

for many lexical words, the root system in Hebrew is not uniform with regard to the shape of roots.

In Spanish, roots do not have a particularly distinctive form, neither with verbal nor with nominal words. Indeed, roots exhibit a wide range of shapes, including single segment roots such as the /d/ in *dar* ‘to give’ and multisyllabic roots such as that of *experimental* ‘to experience’. Roots in Spanish are also not particularly abstract; they resemble surface words and other morphological elements, with the exception that roots do not necessarily follow coda constraints in which, for example, clusters are prohibited (e.g., the root *experiment-* of *experimental*).

If characteristic form is one property that can potentially be associated with roots, characteristic distribution within larger grammatical or prosodic units is another. In this case, both Hebrew and Spanish have roots that are readily identifiable distributionally. In Hebrew, abstract roots are interdigitated with vocalic templates in a predictable manner, such that they are readily easy to recognize within a word. With both verbs and nominals, the distributional relationship between affixes and roots is consistent. For example, number morphology always occurs after a root, most future affixes are before the root, and so forth. In Spanish, roots also have a diagnostic distribution. For instance, they always occur before inflectional morphology and in a predictable relationship to derivational morphology. Based on these two languages, then, it seems that one hypothesis worth exploring further is whether cross-linguistically distribution is a more reliable identifier of roots than form.

In some theoretical approaches, roots have been described as a-categorical. Therefore, one of the questions that can be asked is whether roots identified using the remnant approach can also be described as a-categorical, or whether they are instead reliably associated with a particular lexical category. With the notable exception of the aforementioned non-abstract nominal roots in Hebrew, which are consistently associated with nouns and adjectives, one cannot reliably predict whether a given root in Hebrew or in Spanish will be associated with, for instance, a noun or verb or both. For example, in both (9) and (10), the root *g-d-l* in Hebrew and the root *camín-* in Spanish respectively can be associated with a verb, a noun, or an adjective.

(9) *gadal* ‘he grew’, *godel* ‘size’, *gadol* ‘large’ → *g-d-l*

(10) *caminar* ‘to walk’, *camino* ‘path’, *camínable* ‘walkable’ → *camín-*

Notably, though, some exceptions to this generalization can be found. For example, in Hebrew, *kise*, which is another example of a non-abstract nominal root, is reliably associated only with a noun. In general, however, it does seem to be the case that in these two languages roots can be a-categorical.

In addition to considering the properties of roots, we can also ask about the roles that roots play in a language: How useful is the root as an analytical tool? Does it facilitate descriptions of patterns in a language with regard to morphology, phonology, syntax, or semantics? Such patterns include distributions, processes, and constraints, and the root may serve as a domain, a trigger, a target, etc.

One of the most commonly cited patterns involving roots in Semitic languages involves consonant co-occurrence restrictions. This refers to the observation that roots in surface words can have neither identical nor homorganic first and second consonants nor, with very few exceptions, homorganic second and third consonants (Greenberg 1950; McCarthy 1981; Bat El 2003). On the other hand, we do find many examples of identical second and third root consonants, and, though less frequently, identical and homorganic first and third consonants, as in (11).

- (11) *ʔided* ‘he encouraged’, *nigen* ‘he played’; **dideʔ*, **nineg*

These patterns can be described using the construct *root*, which indicates that the root can indeed be a useful tool for describing the pattern.

Two phonological processes in Hebrew that can be described using the root are sibilant metathesis and the blocking of post-vocalic spirantization. Sibilant metathesis occurs in a very specific domain: with the *hitpa’el* conjugation class when a sibilant occurs in the first root position. To illustrate, in *lehifstalet* ‘to dominate’, which is related to *liflot* ‘to control’, the *t* of the prefix *lehit-* metathesizes with the first root consonant, the *ʃ* (cf. *lehitlabef* ‘to dress oneself’). Spirantization is a common process in Hebrew whereby *b*, *p*, and *k* become fricatives after vowels. This process, however, does not apply in all cases. One of the instances where it fails to apply is in certain conjugation classes, those which historically contained geminates. This can be seen in the alternation between *xover* ‘he joins’ where post-vocalic spirantization does apply and *mexaber* ‘he connects’ where the process is blocked. The root is a useful construct for describing this exception: Post-vocalic spirantization fails to apply to the second root consonant of certain conjugation classes. Though both of these processes may potentially be described without the root, this unit offers elegant and efficient statements of the linguistic patterns.

Roots are also useful in characterizing certain morphological alternation classes. Aronoff (2007) observes that, for example, some words with *n*-initial roots drop the *n* in certain cases and some other words do not. This is illustrated by (12) and (13), which have words associated with the roots *n-s-ʃ* and *n-h-g*, respectively.

- (12) a. *pa’al* conjugation class: *nasa* ‘he traveled’, *jisa* ‘he will travel’
 b. *hif’il* conjugation class: *masia* ‘he transports’
- (13) a. *pa’al* conjugation class: *nahag* ‘he drove’, *jinhag* ‘he will drive’
 b. *hif’il* conjugation class: *manhig* ‘he leads’

In the third singular masculine past, ‘travel’ is *nasa* and ‘drive’ is *nahag*. In the future, the *n* drops for ‘travel’, *jisa*, but not for ‘drive’, *jinhag*. Thus, these two words belong to two different alternation classes. What we can observe is that when in other conjugation classes where the *n* might potentially drop, the root behaves the same. For example, in the *hifil* conjugation class, the *n* also drops for the word associated with the first root but not for the word associated with the second root: We find *masia* ‘he transports’ versus *manhig* ‘he leads’. Thus, these morphological alternation classes are root-based.

In contrast to the variety of patterns that can be described using roots in Hebrew, fewer such patterns are available in Spanish. One pattern that can be described using roots, however, involves diphthongization. In words where the alternation occurs, the stressed syllable of the root is diphthongized (Hualde 2005), as illustrated in (14) and (15).

- (14) *pienso* ‘I think’, *pensar* ‘to think’
- (15) *duermo* ‘I sleep’, *dormir* ‘to sleep’

The first syllable in the first word of each pair is stressed but not the first syllable in the second word of each pair. The domain of diphthongization is what has been identified as the root using the remnant approach; thus the root can potentially be used in the description of this morpho-phonological process. Overall, though, fewer patterns in Spanish seem to rely on a root for their description. Comparing similar patterns across Hebrew and Spanish reveals

that the kinds of patterns in Hebrew that can be described by roots cannot similarly be described in Spanish. For example, though both Hebrew and Spanish exhibit morphological alternation classes such as those exemplified in (12) and (13) for Hebrew, such patterns in Spanish make reference to a lexeme and not to a root (Bermúdez-Otero 2013). Thus, the root is a more useful tool for describing linguistic patterns in Hebrew than in Spanish.

6. Conclusions

With regard to taking preliminary steps towards a typology of roots, one of the biggest contributions of this work is the development of a methodology – what is referred to as the remnant approach – for identifying roots across distinct languages such that meaningful comparisons can be made. This methodology has the advantage of employing a characterization of roots that is independent of properties or functions and that therefore can be used to explore properties and functions in a cross-linguistic context.

In addition, applying this process to Spanish and Hebrew also illuminates some potential dimensions on which a more extensive typology can be based. Roots can be described on a scale of abstraction based on how close they are to surface forms. Some roots, such as those in Hebrew are abstract, and some, such as those in Spanish, are more concrete. In addition, abstractness can be considered with regard to both form and meaning. Distribution is another criterion that can be used for describing roots. Distribution of roots is predictable in both Hebrew and Spanish. Another dimension that can be used to describe roots is their prosodic shape, which may or may not be consistent for a given language (e.g., roots in Hebrew verbs have a characteristic discontinuous consonantal shape but Spanish roots can be very variable). We can also ask how uniform a system is for a particular language. As the asymmetry between nouns and verbs in Hebrew demonstrates, it is possible for roots in one lexical category to behave differently than in another. Another typological question to ask is whether certain kinds of roots in a language are associated with a particular lexical category. Though this does not appear to be generally the case in Hebrew and Spanish, the asymmetry in Hebrew between nouns and verbs hints at the possibility that this might be more clearly the case in other languages. Finally, it might also be fruitful to consider the kinds of patterns – distributions, constraints, processes – in which a root plays a role, and to ask the question: How useful is the root in describing the patterns of a language?

Though the purpose of this study was primarily to explore the implications of defining roots using the remnant approach, the preliminary findings present potential theoretical implications. For example, in terms of the status of roots – whether they are morphemes or purely formal – the results of this exploration support other literature on the topic that suggests that roots cannot straightforwardly be thought of as morphemes, if a morpheme is defined as a minimal unit of meaning. In addition, this work speaks to the importance of not taking for granted characterizations of roots when offering generalizations beyond a single language or language family. It is likely the case that those prior characterizations are based on a set of language-specific properties.

Indeed, these findings also speak to maintaining a certain level of caution against defining the root, and potentially other linguistic units, on the basis of properties in a limited set of languages. Though roots across languages certainly are expected to share some properties, we can also see that potentially few and potentially no properties across languages for roots will converge when we look at a much larger set of languages.

References

- Anderson, S. R. (1992) *A-morphous morphology*. Cambridge: Cambridge University Press.
 Arad, M. (2005) *Roots and patterns: Hebrew morpho-syntax*. Dordrecht: Springer.

- Aronoff, M. (1976) Word formation in generative grammar. *Linguistic Inquiry Monographs Cambridge, Mass.*, (1), 1-134.
- Aronoff, M. (1994) *Morphology by itself: Stems and inflectional classes* (Vol. 22). The MIT Press.
- Aronoff, M. (2007) In the beginning was the word. *Language* 83(4): 803-830.
- Aronoff, M. (2013) The roots of language. In S. Cruschina, M. Maiden & J.C. Smith (Eds.), *The boundaries of pure morphology*. Oxford: Oxford University Press, 161-180.
- Bat-El, O. (1994) Stem modification and cluster transfer in Modern Hebrew. *Natural Language and Linguistic Theory* 12: 571-593.
- Bat-El, O. (2003) Semitic verb structure within a universal perspective. In J. Shimron (Ed.), *Language processing and language acquisition in a root-based morphology*. Amsterdam: John Benjamins, 29-59.
- Bermúdez-Otero, R. (2013). The Spanish lexicon stores stems with theme vowels, not roots with inflectional class features. *International Journal of Latin and Romance Linguistics* 25(1): 3-103.
- Blevins, J.P. (2006) Word-based morphology. *Journal of Linguistics* 42: 531-573.
- Blevins, J.P. (2016) *Word and Paradigm Morphology*. Oxford: Oxford University Press.
- Blust, R.A. (1988) *Austronesian root theory: An essay on the limits of morphology* (Vol. 19). Amsterdam: John Benjamins Publishing.
- Dixon, R.M.W., & Aikhenvald, A.Y. (2002) Word: a typological framework. In R.M.W. Dixon & A.Y. Aikhenvald (Eds.). *Word: A cross-linguistic typology*. Cambridge University Press, 1-41.
- Fortson, B. (2004) *Indo-European language and culture: an introduction*. Oxford: Oxford University Press.
- Greenberg, J.H. (1950) The patterning of root morphemes in Semitic. *Word* 6: 162-181.
- Halle, M., & Marantz, A. (1994) Some key features of Distributed Morphology. *MIT working papers in linguistics* 21 (275): 88.
- Harris, Z. (1941) Linguistic structure of Hebrew. *Journal of the American Oriental Society* 61(3): 143-167.
- Hualde, J.I. (2005) *The sounds of Spanish*. New York: Cambridge University Press.
- McCarthy, J. (1979) Formal problems in Semitic morphology and phonology. Doctoral Dissertation, MIT, Cambridge, MA.
- McCarthy, J.J. (1981) A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* 12: 373-418.
- Mihalicek, V., & Wilson, C. (Eds.). (2011) *Language files: materials for an introduction to language & linguistics* (11th edition). Ohio State University Press.
- Robins, R.H. (1959) In defence of WP. *Transactions of the Philological Society* 58(1): 116-144.
- Robins, R.H. (1964) *General linguistics: An introductory survey*. Bloomington: Indiana University Press.
- Stewart, T. (2008) A consumer's guide to contemporary morphological theories. *Working Papers in Linguistics, The Ohio State University*, 58.
- Ussishkin, A. (2005) A fixed prosodic theory of nonconcatenative templatic morphology. *Natural Language & Linguistic Theory* 23(1): 169-218.

Inflected and periphrastic features: issues of comparison and modelling

Gergana Popova
University of London
g.popova@gold.ac.uk

1. Introduction

In recent years there has been renewed discussion in the linguistic literature of periphrases, that is syntactic constructions that express information similar to that realised by inflection and thus seem to straddle the syntax–morphology divide. Structures like the English perfect or progressive are without doubt composed of more than one syntactic element, but at the same time the information associated with them, it has been argued, cannot be distributed amongst the component parts in a compositional manner. What is more, the information associated with the perfect or progressive is similar in nature to that associated with inflected word-forms in the language like the past or present tenses. In morphological studies past and present are often expressed as ‘features’. In some recent approaches to periphrasis, syntactic constructions like the English perfect and progressive have also been associated with ‘features’. In this paper I would like to explore the nature of such ‘periphrastic’ features (in the sense: realised by periphrases). In the next section, I will summarise the arguments that have been put forward in the literature for treating periphrasis as relevant to morphology, as well as syntax. I will then explore some of the relevant properties of inflectional features as a useful starting point for the discussion of periphrastic features. Next, periphrastic features will be the focus of attention. I will first look at the role tensed elements play within periphrases. Then I will look at ‘nested’ periphrases and what can be said about their featural content. In the last sections I look at the consequences of these observations for how periphrastic features can be incorporated in models of periphrasis. It is important to highlight from the outset that the discussion will be underpinned by some theoretical commitments. One of them is the use of features itself: the assumption that morphosyntactic and morphosemantic information is/can be expressed in grammars in this way. Another is the general approach to morphology assumed: an inferential-realizational one, which is based on a particular view of paradigms and a particular view of morphemes.

2. General properties of periphrasis

Periphrases are by definition syntactic. They consist of one or more elements that are syntactic terminals and stand in some syntactic relationship to each other. The syntactic relationship is not necessarily the same across structures and across languages. In fact, Bonami and Webelhuth (2013) make a case that periphrastic constructions are very diverse in this respect.

Periphrases, however, have a morphological side too. The most powerful argument that can be put forward for the morphological nature of periphrases is the fact that they can fill cells in an otherwise inflectional paradigm. This property has been discussed in the literature on morphosyntax at least since Matthews (1991), but is also crucial for Sadler and Spencer (2001) and Ackerman and Stump (2004). Russian can provide an example. In this language verbs distinguish perfective and imperfective aspect and past, present and future tense. There

are no present perfective forms for semantic reasons. In the future, perfective verbs have an inflected form. Imperfective verbs, however, don't have inflected future forms. Instead, they form their future periphrastically: with the help of the auxiliary verb *byt'* 'be' and the infinitive of the lexical verb. This is illustrated with the 3SG forms of the perfective *vypit'* and the imperfective *pit'* 'drink' in (1) below.

(1)

	PRESENT	FUTURE
PERFECTIVE	--	<i>vypit'</i>
IMPERFECTIVE	<i>pjet</i>	<i>budet pit'</i>

Bonami (2015) furnishes a detailed discussion of this property of periphrasis and shows that periphrases can behave on a par with inflection in terms of paradigmatic organization. For some scholars participation in paradigms is not the only property that can properly delimit periphrasis. They extend the definition to cover constructions that are not incontrovertibly part of inflectional paradigms, but express information that is similar and therefore presents itself as an alternative choice to the information expressed by inflection within the same language. The distinction between constructions that fill cells in inflectional paradigms and constructions that express information normally carried by inflection is not necessarily very clear-cut, see the discussion in Spencer and Popova (2015), for example.

Ackerman and Stump (2004) propose as another criterion for distinguishing periphrases from the rest of syntax their tendency to exhibit morphosyntactic non-compositionality. Essentially this means that in periphrasis some of the information carried by the elements of the construction is 'neutralised' and replaced by new information at the level of the construction. For example, the negated future tense in Bulgarian is composed by the present tense form of the special negative form of the verb *imam* 'have' and the inflected present tense form of the lexical verb. But this present tense inflection on the elements that are part of the construction is overridden, as it were, at the level of the construction, which denotes not present, but future time reference.

The fact that periphrastic constructions can be part of inflectional paradigms, the fact that they express grammatical information very similar to the information associated otherwise with inflectional forms and the fact that they can 'neutralise' the information expressed by inflection, i.e. their morphosyntactic non-compositionality, has led to a view of them as syntactic structures that express morphosyntactic or morphosemantic features, features like the perfect or the progressive in English. The purpose of the discussion that follows is to examine the nature of these features further.

In the next section I will look at some of the properties assumed for inflectional features. The following section looks at periphrastic features. The data I am looking at come mostly from the domain of (finite) verbal aspect and tense, so I will look mostly at the non-compositionality of periphrasis with respect to these. I then turn my attention to cases where there is a nesting of periphrastic features, i.e. where in a periphrastic construction the auxiliary element is itself a periphrase. I will make the argument that such 'nested' periphrases are also non-compositional. I then explore briefly what effect adopting non-compositionality in this way might have on existing formalisations, in particular on the most recent and detailed formalisation of periphrasis in Bonami (2015).

3. Inflectional features

To say that a language possesses a feature, say one of number, is to say that there is a regular correspondence between certain forms and certain (grammatical) meanings. The forms in

question are usually word-forms. So, for example, the forms *dogs*, *cats*, and *horses* are systematically associated with the meaning ‘more than one (dog, cat or horse)’ and can be contrasted with *dog*, *cat*, and *horse* which mean ‘one (dog, cat and horse)’. The presence of /s/, /z/ or /iz/ at the end of the first group of words seems to serve as a signal for this interpretation. The lack of this signal is itself meaningful, so the lack of /s/, /z/ or /iz/ in a form becomes as significant as their presence and the forms with and without the ending are in contrast to each other. As a shorthand for this situation and because it helps build economical descriptions, we say that these English nouns are associated with a feature NUMBER, which can have two values: *plural* (the meaning ‘more than one’, associated with the presence of an /s/, /z/ or /iz/ at the end of nouns; some call this ending an exponent of the value) and *singular* (the meaning ‘one’, expressed via the lack of any exponent). The introduction of a feature NUMBER allows the linking of the systematic changes in nouns described above to the differences in function in related forms like *this* and *these*, or *that* and *those* or, in the verbal system, to the difference between a present tense form like *bark* and the present tense form *barks*. In other words, introducing a feature NUMBER in a language like English will allow for an economical statement of grammatical phenomena like agreement. For more on features see Corbett (2012), for example.

There is another sense in which a feature like NUMBER is important for a grammatical description: it is obligatory. A grammar of English generally needs to associate all noun word-forms in the language with some number value: singular or plural. This is independent of the precise way in which exponence works.¹

Given the facts of agreement, NUMBER is a morphosyntactic feature in English. But obligatoriness can apply to morphosemantic features such as TENSE as well. For example, in most cases an English verb is either inflected for the past tense or for the present tense (in 3SG) or is interpreted as a present tense: *barked* vs *bark(s)* or *bark*. Generally, we say that the values of an inflected feature are mutually incompatible, so, for example, a verb can be past or present, but not both.²

4. Periphrasis

If, following recent proposals in the literature, we assume that certain multiword constructions, for example the English *has taken*, are to be treated as forms of a lexeme, on a par with inflected forms like the 3SG present tense *takes* (see, for example, Börjars et al. 1997, Ackerman and Stump 2004, Bonami 2015), then we would associate this string with a feature e.g. PERFECT, and this feature becomes part of the set of features associated with word-forms in the language which might also include, for example, TENSE: *present*.

As we saw above, inflectional features are obligatory and have values that are mutually exclusive. The next question is how periphrastic features fit in with and compare to inflectional features. This seems to be a question that is touched upon, but rarely discussed in detail in the literature. I turn to it next.

¹ So the noun *sheep* is really two forms: *sheep* singular and *sheep* plural, because we can talk about *this sheep* and *these sheep*. One could, of course, also say that this particular form in English is undefined with respect to number and therefore compatible with both singular and plural meanings. And, given that English verbs don’t always give clear information about number, one could imagine a sentence like *The sheep came to graze* where nothing clarifies whether one or more than one sheep are being referred to.

² This glosses over some complexity. On the one hand, it assumes that English doesn’t have future tense, with which some scholars disagree, see for example Salkie (2010) and references therein. On the other, as in the case of *sheep* previously, it omits verbs like *put* which have the same form for past and present (or which, one might say, are unspecified for tense).

4.1 Periphrasis as part of inflectional paradigms: periphrastic values

The most frequently discussed case of the relationship between features realised inflectionally and features realised periphrastically has already been illustrated. This is precisely the case where a multiword construction fills in a cell in an otherwise inflectional paradigm, a situation called ‘feature intersection’ in recent work on periphrasis (Ackerman and Stump 2004, see also Brown et al. 2012). The multiword construction in this case realises not some independent feature, but one or more values of one or more features whose other values are realised via inflection. So in this case, as with the typical inflectional features discussed above, the values are mutually exclusive. A Russian verb is generally either perfective or imperfective in terms of the feature ASPECT, and either past or present in terms of the feature TENSE. When it is both perfective and future, it is inflected (and has a certain exponent) and when it is both imperfective and future it is periphrastic (and its exponent is a multiword construction). This is one way to account for the fact that periphrastic forms in scenarios like this one are limited; for example there are no periphrastic future forms for perfective Russian verbs.

4.2 Periphrastic sub-paradigms: periphrastic features

In many cases where we find multiword syntactic constructions that express grammatical meaning, however, these multiword constructions are not filling cells in otherwise inflectional paradigms in the narrow sense described above, and so it isn’t immediately obvious what the relationship between features expressed periphrastically and those expressed inflectionally should be. Let’s for the moment assume some associations from traditional descriptions, which are often taken over in theoretical discussions as well. As pointed out above, we could say that the English construction *have* + *Ven* is associated with some meaning ‘perfect’, or that the English construction *be* + *Ving* is associated with the meaning ‘progressive’. The question then is, are these features in their own right, or perhaps values of some feature, e.g. ASPECT. If we assume that these are values of aspect, we will end up with a feature whose values are not mutually exclusive, since in English progressive and perfect can combine as in *have been V-ing*. This phenomenon has been called nesting, or stacking, of periphrases (see Popova and Spencer 2012) and is of particular interest to this paper. Such mutual compatibility of values would make English aspect unusual when compared with typical inflectional features like (past and present) tense. Similar arguments would apply to considering perfect and progressive to be values of the feature TENSE.

If we want to maintain that values are always mutually exclusive, we could assume that PERFECT and PROGRESSIVE are independent features. And since they don’t have other values, we will have to make them binary \pm features. This is precisely the assumption that seems to be made in a recent detailed formal model of periphrasis laid out in Bonami (2015).

The next question is how to express the relationship between these features and the feature TENSE: *past* or *present* which we know we need for English verbal inflection. The auxiliary verbs in both the (finite) perfect and the progressive periphrases are tensed, and the tense of the auxiliary can shift between present and past. In English this seems to lead to changes in meaning which are fairly transparent and compositional. For this reason, perhaps, TENSE: *past* and *present* is usually assumed to be in intersection with perfect: \pm and progressive: \pm . If we assume that tense is a feature that is independent of the periphrastic features and that tense is defined over a periphrastic construction and that the tense of the construction is the same as the tense of the auxiliary verb, we will obtain the description in Table 1 (all forms are in the 3SG):

	TENSE	PERFECT	PROGRESSIVE
<i>barks</i>	present	–	–
<i>barked</i>	past	–	–
<i>has barked</i>	present	+	–
<i>had barked</i>	past	+	–
<i>is barked</i>	present	–	+
<i>was barking</i>	past	–	+
<i>has been barking</i>	present	+	+
<i>had been barking</i>	past	+	+

Table 1: Tense, perfect and progressive in English as independent features

This picture suggests that the presence/absence of the auxiliary can be associated with the presence/absence of the feature perfect/progressive, and that distinctions like *perfect*, or *progressive* are additional to existing tense distinctions in the language.

The position that periphrastic features in a language depend on its inflectional potential and that periphrases are in some sense an ‘expansion’ of its existing inflectional features is an attractive one, given the data from languages like Lithuanian, for example. Lithuanian has synthetic past, present, and future forms and also synthetic frequentative past tense forms. The distinctions made in the inflectional system appear to carry over to the system of periphrastic (or compound) tenses via the auxiliary, as the examples below in Table 2 show (the examples are drawn from Ambrazas 1997: 237):

synthetic inflection		periphrasis	
present	<i>dirba</i> ‘works’	perfect	<i>yrà dirbęs (m), dirbusi (f)</i>
past	<i>dirbo</i> ‘worked’	past perfect	<i>bùvo dirbęs, dirbusi</i> ‘had worked’
past frequentative	<i>dirbdavo</i> ‘used to work’	past perfect frequentative	<i>Būdavo dirbęs, dirbusi</i> ‘used to have worked’
future	<i>diršs</i> ‘will work’	future perfect	<i>būs dirbęs, dirbusi</i> ‘will have worked’

Table 2: Synthetic and periphrastic tenses in Lithuanian

The perfect tenses are formed with the help of a finite form of the auxiliary *būti* and the past active participle of the main verb (which distinguishes masculine and feminine gender). The auxiliary is respectively in the following tenses: present for the perfect, past for the past perfect, past frequentative for the past frequentative perfect and future for the future perfect.

The literature on periphrasis, however, suggests that periphrastic constructions can be non-compositional in terms of the morphosyntactic information expressed by the elements in the periphrase and the morphosyntactic information associated with the construction as a whole (see discussions in Ackerman and Stump 2004 or Brown et al. 2012, for example). There are indeed data to suggest that the association of the construction directly with the tense of its elements is not so straightforward. I turn to some such data next.

5. Idiomaticity of tense in periphrasis

Inflectional distinctions are not always simply taken on and incorporated in multiword constructions. Some inflectional distinctions can equally be neutralised in periphrastic forms. A case in point is Bulgarian. This language has two synthetic past tenses: imperfect and aorist. A verb like *rabotja* ‘work’, for example, has two inflected past tense forms (here in the 3SG): the aorist *raboti*, and the imperfect *raboteše*. Bulgarian also has a rich system of periphrastic tenses. Amongst them are the perfect and the pluperfect, formed with the verb *sâm* ‘be’ and a participle of the main verb. The auxiliary in the pluperfect is in the imperfect past tense, but there is no opposition between imperfect and aorist in the pluperfect. So the aorist/imperfect opposition is neutralised when it comes to the periphrastic tenses.

Bulgarian, like other languages with periphrastic tenses, has some forms where an element in the present enters into a construction with an element in the past. An example is the future in the past, formed with an auxiliary meaning ‘want’ inflected in the past tense (imperfect) and the present tense form of the lexical verb (here ‘work’): *šteše da raboti*. If we follow the assumption that the whole multiword expression functions as a word form and then assume that the tense values associated with the elements of such a multiword expression are simply added to the set of feature/values associated with the expression as a whole, then present and past tense values will have to co-exist within the same set, which is in contradiction to them being mutually exclusive.

In addition, periphrastic forms might have their own constraints. Lithuanian provides an example. Amongst the periphrastic forms in Lithuanian are the so-called continuative tenses (Ambrasas 1997: 321f) which are formed by the prefixed (with *be-*) present active participle of the verb and an inflected form of the auxiliary verb *būti* ‘be’. There is no present continuative, only past, past frequentative and future continuative.³ The reason for this gap cannot be formal, as the auxiliary has a full paradigm.

The observations above suggest that in periphrases the tense distinctions made by the inflected verbs within them (most often, but not always, the auxiliary verbs, or ‘ancillary elements’ in the terminology introduced by Bonami) cannot simply be carried over from the inflected verbs. Unless, of course, we manage to find an analysis that will derive the semantic interpretation of the overall construction from the tense values of the elements that comprise it. It is difficult to claim that such an analysis does not exist. But as an indication that it might be difficult, here are some examples of how different languages derive different periphrases from similar inflectional resources.

Basque has a system of periphrastic aspect/tense features not dissimilar to that found in English and Bulgarian. In Basque, most tense/aspect verbal forms are periphrastic. Only a handful of verbs have synthetic inflection for present and past.⁴ The majority of Basque verbs, therefore, have periphrastic tense/aspect forms. According to de Rijk (2008: 143), there are six major periphrastic tenses in the language, illustrated in Table 3 with an intransitive verb ‘fall’ and the auxiliary *izan* ‘be’ used with intransitive verbs.

³ According to Ambrasas (1985: 322), only the past continuative is more widely used in the contemporary language. The other forms are restricted to the Samogitian dialect.

⁴ I ignore for the moment other forms, for example those relating to mood.

The present imperfect: Imperfect participle + present tense auxiliary <i>Ibaira erortzen da.</i> ‘He is falling into the river.’
The past imperfect: Imperfect participle + past tense auxiliary <i>Ibaira erortzen zen.</i> ‘He was falling into the river.’
The present perfect: Perfect participle + present tense auxiliary <i>Ibarira erori da.</i> ‘He has fallen into the river.’
The past perfect: Perfect participle + past tense auxiliary <i>Ibaira erori zen.</i> ‘He fell into the river.’
The (present) future: Future participle + present tense auxiliary <i>Ibaira eroriko da.</i> ‘He will fall into the river.’
The past future: Future participle + past tense auxiliary <i>Ibaira eroriko zen.</i> ‘He was going to fall into the river.’

Table 3: Periphrastic tense/aspect forms in Basque (adapted from de Rijk 2008: 143)

As can be seen from Table 3, the meaning and composition of periphrastic constructions in Basque is analogous to that of other languages. For example, the present and past imperfect in Table 3 bear striking similarities with the progressive in English: an eventuality in progress with present time reference or with past time reference. Both of these tenses are composed with the help of the imperfect participle, and they differ only with respect to the tense of the auxiliary, to which the difference in time reference can be attributed.

The perfect is also analogous to the perfect in other languages, like English or Lithuanian. It is formed with the perfect participle of the verb and a present tense auxiliary. However, the form that is analogous to the past perfect in other languages (perfect template but with a past tense auxiliary) has different semantics. In English and Bulgarian and other languages this form can be associated with locating a moment in time before some other past moment, i.e. with what is most frequently termed the pluperfect in grammatical descriptions. In Basque this form is put to a different use and the meaning is most frequently a simple past. The following example of this non-hodiernal perfective past is given by Oyharçabal (2003: 265):

- (2) *Atzo* *Peru* *ikus* *nuen.*
 Yesterday Peru see.PRF AUX.PST
 ‘Yesterday I saw Peru.’

Very probably, the lack of an inflected past leaves a gap which is filled in by available linguistic resources. Despite the similarity of formal means across languages, the meaning expressed by the forms can vary depending on the needs of the particular language; form-meaning correspondences are decided within the available system of oppositions. One consequence of such observations is the conclusion that periphrastic constructions are, in some sense, idiomatic. Their interpretation cannot be inferred in a straightforward manner from the interpretation of their constituent parts. A further consequence is that constructions that are similar except for the tense of the auxiliary cannot be seen to be ‘derived’ from each other. The perfect and the (formally) past perfect in Basque need to be declared separately in the grammar. Further support for such independence of related constructions comes from Bulgarian. In Bulgarian both the future tense and the future in the past are formed with the help of an ancillary that historically comes from the verb *šta* ‘want’. In the future this ancillary is, historically, a 3SG present tense form which has grammaticalised into an uninflected semi-clitic. In the future in the past, on the other hand, the verb *šta* ‘want’ retains inflection and has a somewhat different syntactic relationship with the main verb. Another

reason not to think of two constructions that differ only in the tense of the ancillary element as being ‘derived from each other’ and being ‘the same with a different tense value’ is that constructions like this can have different meaning, so the formal opposition may not be mirrored by a semantic opposition. For example, in Lithuanian the past continuative seen above can mean an action which was begun as intended, but not finished. The examples below are from Ambrasas (1997: 250-251).

- (3) Jùras jau bũvo beatkeliãs atvỹkstantiems vartũs, bet vỹl juõs privérẽ.
‘Juras was about to open the gate for the visitors, but closed it again.’

The future continuative, by contrast, expresses a supposition: an action which is supposed to have taken place in the future, or sometimes in the present:

- (4) Jis jau trẽãiã pããiã bũs beturĩs.
‘(I think) he has a third wife already’.

The idiomaticity of periphrasis with respect to the inflectional values of the elements that comprise it is recognised in the literature, see references in Brown et al. (2012), or Spencer and Popova (2015), for example. As mentioned before, it is definitional for Ackerman and Stump (2004). The discussion above, if along the right lines, reinforces this view. The more important point here is what consequences idiomaticity would have for the set of morphosyntactic properties associated with a periphrastic construction. If we assume that the present perfect in English is not the present tense of a perfect construction, and past perfect is not the past tense of a perfect construction, but instead both constructions express some feature independent of the present and past values of tense (say, the binary features PRSPERF and PSTPERF), and a similar case is made for the progressive, the situation we saw in Table 1 will resemble the situation in Table 4.

	TENSE	PRSPERF	PSTPERF	PRSPROG	PSTPROG
<i>barks</i>	present	–	–	–	–
<i>barked</i>	past	–	–	–	–
<i>has barked</i>		+	–	–	–
<i>had barked</i>		–	+	–	–
<i>is barking</i>		–	–	+	–
<i>was barking</i>		–	–		+
<i>has been barking</i>		+	–	+	–
<i>had been barking</i>		–	+	–	+

Table 4: Perfect and progressive in English as independent features, but no tense

If we represent the features in this way, we can actually model the present and past tenses, and the present perfect construction, the past perfect construction, the present progressive construction and the past progressive construction as the values of a single feature, because now then are mutually exclusive.

However, *has been Ving* and *had been Ving* still present a problem, because it looks like the perfect and the progressive are both part of these constructions at the same time. This raises the question of how such nested periphrases should be interpreted and represented. Is nesting of periphrasis compositional, or idiomatic? The next section tries to look for an answer.

5.1 Nested periphrases

‘Nesting’ of periphrases can often be understood as resulting from a construction where one of the elements (most frequently the auxiliary) is itself ‘inflected’ for some feature whose realisation is a periphrastic construction. For example, Popova and Spencer (2013) argue that the future perfect in Bulgarian needs to be understood as a perfect (periphrastic) construction in which the auxiliary has been ‘inflected’ for the future, which itself is periphrastic. This is not dissimilar to constructions where an auxiliary is tensed. For example, the nested forms in Bulgarian for the future perfect (see Table 5) are cognate with the forms in Table 6 in Lithuanian. Lithuanian, however, has an inflected future, so there is no nesting.

PERFECT	FUTURE	FUTURE PERFECT
(az) <i>sâm rabotila</i>	(az) <i>šte rabotja</i>	(az) <i>šte sâm rabotila</i>

Table 5: Perfect, Future and Future Perfect of ‘work’ in Bulgarian

PERFECT	FUTURE	FUTURE PERFECT
<i>yrà dirbęs</i>	<i>dirbs</i>	<i>būs dirbęs</i>

Table 6: Perfect, future and future perfect of ‘work’ in Lithuanian
(examples are drawn from Ambrasas 1985: 237f)

The English system of auxiliary constructions can also be understood as a system of nested periphrases. The perfect progressive was mentioned already. It can be understood (and has been modelled in Bonami 2015) as the progressive periphrastic construction in which the auxiliary is ‘inflected’ in the perfect. Thus, the perfect progressive of *work*, *has been working*, can be understood as the progressive template for ‘work’, that is BE + *working*, in which the auxiliary BE appears in the perfect: *has been*.

In the previous section I argued that there is evidence of idiomaticity with respect to the tense values in periphrasis. The question that arises now is whether there is evidence of similar idiomaticity in the case of nested constructions. Current formalisations of the English nested periphrases assume that *has been working* is a form of WORK with feature specification PERF +, PROG + (Bonami 2015). The question is whether this logic could be applied to nested periphrases cross-linguistically.

To establish whether compositionality exists, we could look at the combinations of cognate forms across languages. For example, Udihe, like English, has perfect and progressive. The perfect is inflected, whereas the progressive is periphrastic and is formed by the infinitive of the lexical verb and an inflected copula meaning ‘to be’. The copula can be inflected for the perfect, potentially leading to a combination of perfect and progressive. However, rather than a meaning like the English perfect progressive, this construction is semantically equivalent to the progressive with the auxiliary in the past tense, and both mean past progressive. The examples below are taken from the description of Udihe in the Surrey Morphology Group Periphrasis Project:⁵

- (5) a. *umi-mi bi-si-ni*
 drink-INF be-PAST-3SG
 ‘he was drinking’
 b. *imi-mi bi-s’e*
 drink-INF be-PFV.3SG
 ‘he was drinking’

⁵ For more details please visit <http://www.smg.surrey.ac.uk/periphrasis/>

An even stronger piece of evidence for the lack of predictability in nested periphrasis comes from Basque. Basque has a habitual perfect (past and present) which is used to refer to habits lasting up to a specific point in time (see de Rijk 2008: 149). This periphrase consists of perfect participle of the main verb and a perfect tense form of the auxiliary.

Oyharçabal (2003: 258) adds to the above other uses of this form, for example to express some kind of remoteness, and illustrates it as follows (examples adapted, for sources of examples see original):

- (6) a. *Andiabruak asko aldiz tentatu izan zuan.*
 there devil.ERG many times tempt.PRF be.PRF AUX
 ‘There the devil tempted him many times’.
- b. *Erraztasun onetatik kalte andiak etorri izan dira.*
 Facility these.ABL damage big come. PRF be. PRF AUX
 ‘Great damages have come from these facilities’

It will be difficult to see how the logic suggested above for English, where one possibility for a progressive perfect nested periphrase was to see it as a periphrase which expresses PROG+ and PERF+, can be applied to the Basque habitual, which will then need to be associated with the specification PERF+, PERF+. It is not clear how such a form will be modelled, or how its meaning will be deduced. Instead, the Basque example seems to invite us to set up a new feature for this form, say HABITUAL+, which is then composed of the perfect form of the auxiliary embedded in the perfect construction.

The conclusion to draw from these examples is that we have to treat a nested periphrase as we do the tense specification on an element of a periphrastic construction. To model the construction we need to know that a certain element needs to be in a certain tense (whether inflected or periphrastic), but this tense is not directly part of the feature specification of the overall construction. The feature values presented in Table 4 will now look as in Table 7 below.

	PRSPERF	PSTPERF	PRSPROG	PSTPROG	PRSPERFPROG	PSTPERFPROG
<i>has barked</i>	+	–	–	–	–	–
<i>had barked</i>	–	+	–	–	–	–
<i>is barking</i>	–	–	+	–	–	–
<i>was barking</i>	–	–		+	–	–
<i>has been barking</i>	+	–	–	–	+	–
<i>had been barking</i>	–	–	–	–	–	+

Table 7: Independent inflected and periphrastic feature values in English as independent features

This isn’t the assumption about nesting adopted in one of the most recent formal models of periphrasis, namely Bonami (2015), instead nesting is conceived as compositional, such that the perfect progressive in English, for example, is the result of generating the progressive and then the perfect of the auxiliary. Changing the compositionality assumption in the model has interesting consequences. I next turn to a summary of Bonami (2015) and a discussion of a putative amendment of it.

6. An existing formal model of periphrasis

The most recent formal model of periphrasis is based largely on the assumptions outlined in the initial sections of this paper. It assumes that the morphological system of a language is modelled via the inferential-realisation framework of Paradigm Function Morphology

(PFM), whereas for the syntactic system it employs the mechanisms of HPSG. This paper will focus on the morphological side of the analysis, i.e. on the generation of periphrastic forms via paradigm functions along the lines suggested for inflection in Stump (2001). In PFM the inflected forms of a lexeme are generated by a paradigm function (PF) which takes a root of a lexeme (or its lexemic index in some approaches) and a set of morphosyntactic features that represent a cell in the paradigm of that lexeme, and generates a word-form which has that same set of morphosyntactic features. Formally this is expressed as shown in (7), for more details see Stump (2001: 43f.):

- (7) $PF(\langle X, \sigma \rangle) =_{\text{def}} \langle Y, \sigma \rangle$
 Where X is the root of a lexeme, and Y is a word-form of the same lexeme, and σ is a set of morphosyntactic features.

The analysis in Bonami (2015) is based on the insight that periphrasis can be modelled analogously to (flexible) syntactic idioms like *spill the beans*. The elements of such an idiom can be endowed with some special meaning which when combined gives the meaning of the whole, but the elements need to be constrained to stand in a relationship of mutual selection so they always co-occur. Bonami borrows the mechanism of reverse selection from analyses of idioms in the HPSG literature (for example, Sailer 2000, for further references in Bonami 2015) and allows the paradigm function to introduce reverse selection requirements. More formally, Bonami (2015: 87) expresses this in the following way:

- (8) A reverse selection requirement φ carried by a word w_1 is satisfied if and only if w_1 is syntactically selected by a word w_2 and w_2 verifies property φ .

The paradigm function for the perfect tense can then be expressed in the following way (see Bonami 2015: 101):

- (9) If l is a verb, then for any property set $\sigma \supseteq \{\text{PRF}+\}$,
 $PF(l, \sigma) = \langle \langle \varphi, \sigma \rangle, \{ \langle \textit{have-prf}, \sigma! \{\text{PRF}-\} \rangle \} \rangle$,
 where $PF(l, \sigma! \{\text{VFORM } \textit{pst-ptcp}\}) = \langle \langle \varphi, \sigma! \{\text{VFORM } \textit{pst-ptcp}\} \rangle, \emptyset \rangle$

Essentially, the paradigm function says that the perfect of a verb l is a past participle of the verb that reverse-selects for the *have*-perfect auxiliary. To deliver the past participle form of l and the auxiliary with the right properties, the paradigm function for the perfect of l calls two other paradigm functions. Notable is the stipulation that the auxiliary should be marked as not perfective (PFR-). This stipulation is aimed at preventing the paradigm function for the perfect from applying to the auxiliary in the periphrastic construction, and this preventing forms like **has been been*. Nested periphrasis obtains when a cell in a lexeme's paradigm contains two periphrastic features, for instance both perfect and progressive, that are set to +. To account for such forms, Bonami (2015: 102) suggests the following:

- (10) a. If l is a verb, then for any property set $\sigma \supseteq \{\text{PROG}+\}$,
 $PF(l, \sigma) = \langle \langle \varphi, \sigma \rangle, \{ \langle \textit{be-prog}, \sigma! \{\text{PROG}-\} \rangle \} \rangle$,
 where $PF(l, \sigma! \{\text{VFORM } \textit{prs-ptcp}\}) = \langle \langle \varphi, \sigma! \{\text{VFORM } \textit{prs-ptcp}\} \rangle, \emptyset \rangle$.
 b. If l is a verb, then for any property set $\sigma \supseteq \{\text{PRF}+, \text{PROG}-\}$,
 $PF(l, \sigma) = \langle \langle \varphi, \sigma \rangle, \{ \langle \textit{have-prf}, \sigma! \{\text{PRF}-\} \rangle \} \rangle$,
 where $PF(l, \sigma! \{\text{VFORM } \textit{pst-ptcp}\}) = \langle \langle \varphi, \sigma! \{\text{VFORM } \textit{pst-ptcp}\} \rangle, \emptyset \rangle$.

The paradigm function in (a) above realises the feature progressive. It stipulates that the realisation of the progressive is the present participle of the lexeme which reverse-selects for a *be*-progressive auxiliary. The paradigm function in (b) realises the perfect, but is stipulated to apply only when the progressive is set to $-$. This means that the paradigm function in (b) cannot apply before the progressive paradigm function, ruling out forms like **is having read*. Specifying the PF realising the perfect in this way also means that it can apply to the progressive auxiliary, which has been marked as PROG- , thus producing the necessary nested form.

This approach essentially assumes compositionality of tense in periphrasis, as well as compositionality of nested periphrases, insofar as the set of available features is concerned. Non-compositionality in Bonami (2015) is not precluded, but relies not on how features are set up, but rather on what features are ‘pushed down’ by the periphrastic PFs to the PFs that realise the (inflected) elements of periphrastic constructions. One consequence of approaching non-compositionality this way is that special formal mechanisms are needed to ensure that periphrasis is not recursive and that periphrastic paradigm functions apply in a set order. The question now is what will happen if we assume non-compositionality and encode it in the feature sets, as argued above. This is the subject of the next section.

7. Consequences for the formal modelling of nested periphrases

For reasons of space this section will focus on one example of nested periphrasis, namely the progressive perfect in English. As argued in previous sections, one possible way to translate the non-compositionality assumptions with respect to nested periphrases is to assume that a construction like the perfect progressive is associated not with a set of features containing PRS+ , PERF+ , PROG+ , but rather a single specification, which for convenience was named *prs-perf-prog*. Given the assumptions above, this specification could be treated as one of the values of the feature TENSE . This is now the only feature-value in the set of features associated with a word-form like *has been reading* that can trigger a periphrastic realisation, so the order of progressive and perfect becomes irrelevant. This paradigm function would look something like the following:

- (11) If l is a verb, then for any property set $\sigma \supseteq \{\text{TNS: } \textit{prs-perf-prog}\}$,
 $\text{PF}(l, \sigma) = \langle \langle \varphi, \sigma \rangle, \{ \langle \textit{be}, \sigma! \{\text{TNS: } \textit{prs-perf}\} \rangle \} \rangle$,
 where $\text{PF}(l, \sigma! \{\text{VFORM } \textit{prs-ptcp}\}) = \langle \langle \varphi, \sigma! \{\text{VFORM } \textit{prs-ptcp}\} \rangle, \emptyset \rangle$.

Not only has the order of progressive and perfect become irrelevant, but so has recursion. The paradigm function in (11) above calls in a PF that realises a set of features which contains the specification $\text{TNS: } \textit{prs-perf}$, but as this is an alternative value of tense, the PF realising $\text{TNS: } \textit{prs-perf-prog}$ cannot re-apply.

8. Discussion

The main aim of this paper is to discuss the issue of non-compositionality in constructions that have come to be known as periphrastic, i.e. constructions that are syntactic, but express grammatical meaning usually associated with inflection. The focus is primarily on constructions associated with tense-aspect meanings. Many of the constructions surveyed here consist of a participial form and an inflected ancillary (auxiliary) element. The main point put forward is that periphrastic constructions can be non-compositional, not only when it comes to the tense values associated with their ancillary forms, but also when it comes to the embedding of one periphrase into another, i.e. when their ancillary elements are themselves periphrastic (so called ‘nested’, or ‘stacked’ periphrases). There are various ways to allow for

such non-compositionality in the formal models associated with periphrasis. This paper proposes that one way this can be done is to associate a periphrase with feature-values which are not realised by any of its elements. It then explores what consequences such an approach might have if existing formalisations (such as Bonami 2015) are adapted to express non-compositionality in this way.

The main consequences seem visible when the model is applied to nested periphrases. Nested periphrases seem to pose at least two problems: order of application (should we talk about the progressive template with a perfect auxiliary, or the perfect template with the progressive auxiliary?) and recursion (why not have a progressive template with a progressive auxiliary?). Both of these issues seem to disappear if non-compositionality is approached as suggested here.

This can be argued to be perhaps simply a matter of formal convenience. It remains a valid observation, however, that the order of nested periphrases appears to be largely stipulative, and that periphrasis rarely employs recursion. So stipulative order and non-recursion are another way in which periphrasis resembles morphology rather than syntax. And if so, then non-compositionality could be linked to the mixed (morphosyntactic) nature of periphrasis. In other words, periphrasis exhibits what Dahl (2004) calls ‘featurization’, the association of a form with features. In the case of periphrasis this featurization is also linked to a distribution in a paradigmatic space. The organization of periphrases in a paradigmatic space would explain a form-meaning correspondence which is related not to the meanings of the component parts, but to the oppositions in a system of similar form-meaning pairs.

References

- Ackerman, F. & G. Stump (2004) Paradigms and periphrastic expressions. In Louisa Sadler & Andrew Spencer (Eds.), *Projecting morphology* Stanford Studies in Morphology and the Lexicon. Stanford, California: CSLI Publications, 111–157.
- Ambrazas, V. (Ed.) (1985) *Grammatika litovskogo jazyka*. Vilnius: Mokslas.
- Ambrazas, V. (Ed.) (1997) *Lituanian grammar*. Vilnius: Baltos Lankos.
- Bonami, O. (2015) Periphrasis as collocation. *Morphology* 25(1): 63–110.
- Bonami, O. & G. Webelhuth (2013) The phrase-structural diversity of periphrasis: a lexicalist account. In *Periphrasis: The role of syntax and morphology in paradigms*. Oxford: Oxford University Press, 141–167.
- Börjars, K., N. Vincent & C. Chapman (1997) Paradigms, periphrases and pronominal inflection. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 1996*. Dordrecht: Kluwer Academic Publishers, 155–180.
- Brown, D., M. Chumakina, G. Corbett, G. Popova & A. Spencer (2012) Defining ‘periphrasis’: key notions. *Morphology* 22(2): 233–275.
- Corbett, G. G. (2012) *Features*. Cambridge: Cambridge University Press.
- Dahl, Ö. (2004) *The growth and maintenance of linguistic complexity*. Amsterdam/Philadelphia: John Benjamins.
- de Rijk, R. P. G. (2008) *Standard Basque: A progressive grammar*. Cambridge, MA: The MIT Press.
- Matthews, P. H. (1991) *Morphology*. Cambridge: Cambridge University Press.
- Oyharçabal, B. (2003) Tense, aspect and mood. In José Ignacio Hualde & Jon Ortiz de Urbina (Eds.), *A grammar of Basque*. Berlin and New York: Mouton de Gruyter, 249–284.
- Popova, G. & A. Spencer (2013) Relatedness in periphrasis: a paradigm-based perspective. In M. Chumakina & G. G. Corbett (Eds.), *Periphrasis: The role of syntax and morphology in paradigms*, vol. 180 Proceedings of the British Academy. Oxford University Press/British Academy, 191–225.
- Sadler, L. & A. Spencer (2001) Syntax as an exponent of morphological features. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 2000*. Dordrecht: Kluwer, 71–96.
- Sailer, M. (2000) *Combinatorial semantics and idiomatic expressions in Head-Driven Phrase Structure grammar*. Doctoral dissertation. Tübingen: Eberhard-Karls-Universität.
- Salkie, R. (2010) *Will*: tense or modal or both? *English Language and Linguistics* 14(2), 187–215.
- Spencer, A. & G. Popova (2015) Periphrasis and inflection. In *The Oxford handbook of inflection*. Oxford: Oxford University Press, 197–230.
- Stump, G. (2001) *Paradigm function morphology. a theory of inflectional form*. Cambridge: Cambridge University Press.

The lexicalization of complex constructions: an analysis of adjective-noun combinations

Marcel Schlechtweg
Universität Kassel
 marcelschlechtweg@gmail.com

Holden Härtl
Universität Kassel
 holden.haertl@uni-kassel.de

1. Introduction

Languages can differ with respect to their use of compounds and phrases as naming units that become lexicalized. A comparison of Germanic and Romance languages exemplifies cross-linguistic variation in the choice of the preferred construction type in that the former seem to favor compounds where the latter often use phrases. In the current paper, we aim at discussing the lexicalization of compounds and phrases both from a language-specific and from a cross-linguistic angle. Specifically, we pursue the idea that compounds are naturally more appropriate to become lexicalized than phrases. For this purpose, we will reflect upon some fundamental characteristics of compounds and present psycholinguistic evidence that supports the conception of the lexicalization affinity of compounds.

The paper is structured in the following way. In Section 2, we will discuss the proposal that compounds are well suited to enter the process of lexicalization. In Section 3, we will introduce the psychological process of memorization and show how it is related to the process of lexicalization. In Section 4, we will present empirical evidence suggesting a fundamental difference between the memorization of compounds and phrases. Finally, in Section 5, we will summarize our discussion and conclude the paper.

2. The lexicalization affinity of compounds

While compounds have often been regarded as constructions that typically represent names of specific phenomena, phrases have been argued to fulfill a descriptive function in most cases (Bauer 1988: 102; Hüning 2010: 197). This difference can be easily recognized in German: Whereas the compound *Grünreiher* ‘green_heron’ names the particular kind of heron also called *butorides virescens*¹, the phrase *grüner Reiher* ‘green heron’ can refer to any heron that is green for whatever reason. Nevertheless, we also find examples that show that the functional distinction between compounds and phrases is not a definite rule. The German phrase *gelbes Trikot* ‘yellow jersey’, taken from Schlücker (2014: 148), cannot only function as a descriptive unit but can also name a specific shirt worn by the leader of some sports competitions, e.g. the *Tour de France*.

Although counterexamples exist, the aforementioned functional separation between compounds as typical naming units and phrases, which are usually descriptions, holds in many cases. This has led researchers to claim that compounds are better equipped to fulfill the naming function than phrases. Bücking (2009, 2010) contrasts adjective-noun (AN) compounds to AN phrases in German. Crucially, he concentrates on novel constructions in order to exclude the influence of lexicalization. The author emphasizes that the variable R_{INTEGRAL} , which determines how the adjective and the noun are related, plays a crucial role in compounds. The phenomenon is connected to a fundamental semantic difference between German AN compounds and AN phrases: compounds, as opposed to phrases, are known for their semantically non-compositional character. Therefore, the meaning of a compound goes

¹ <https://de.wikipedia.org/wiki/Grünreiher> (Accessed on January 6, 2016).

beyond the simple sum of the meaning of the adjective and the meaning of the noun and necessitates an additional element, namely R_{INTEGRAL} , that establishes the precise relationship between a compound's constituents. Since names in general, similar to compounds in particular, tend to lack compositional semantics, R_{INTEGRAL} makes compounds more appropriate to fulfill a naming purpose (Herbermann 1981: 334-335; Bücking 2009, 2010).

Härtl (2015) goes in a similar direction and emphasizes the consequences of the structural difference between German AN compounds and AN phrases. Since a compound structurally deviates from a phrase, which represents the default AN combination, it is also likely to differ from the phrase in semantic terms. The creation of a new compound such as *Rotdach* 'red_roof' implies that its meaning does not equal the meaning of the phrase *rotes Dach* 'red roof'; instead, the compound is semantically non-compositional and ready to name a particular complex concept right from the beginning of its existence. The author connects this thought to the idea that compounds are more suited for lexicalization. The special structural and semantic status of compounds, in comparison to phrases, calls for the lexicalization of the complex construction as a name of a particular concept.

The aforementioned contributions by Bücking and Härtl suggest an interesting link between the structure, semantics, function and lexicalization of compounds and phrases. It is of utmost significance that the authors focus on the contrast between AN compounds and AN phrases in German. In this language, the two construction types can be clearly kept apart by means of inflectional agreement. Whereas the two constituents of an AN phrase such as *grüner Reiher* 'green heron' are in agreement with respect to gender, number, case and definiteness, the mere adjectival root of a German AN compound such as *Grünreihler* 'green_heron' does not agree with the noun it precedes (cf. also Booij 2012: 84). Inflectional agreement between the adjective and the noun, in the case of phrases, or the lack of it, in the case of compounds, are reflected on a structural basis through the presence or absence of a suffix. The structural divergence between compounds and phrases seems to be prone to mirror a semantic difference between the two construction types as well. As a consequence of their structural deviation from a normal phrase, compounds tend to carry a meaning that also differs from the default interpretation expressed by a syntactic phrase. The question now arises of how we can find further evidence for the idea that the peculiarities of compounds make them more suitable to become lexicalized than phrases.

3. Lexicalization and memorization

Lexicalization represents a diverse phenomenon, which has been defined in several ways in the literature (Lipka 2002: 111; Bakken 2006: 106-108). For the purpose of the current paper, we consider a construction to be lexicalized if it serves as the linguistic form, or the name, of a specific concept (Lipka 1981: 131; Blank 2001: 1596; Gaeta and Ricca 2009: 38). Moreover, the relation between lexicalization and another crucial term, namely the psychological notion of memorization, turns out to be significant in the context of the present contribution. Several authors have discussed the demarcation between the two terms. Wunderlich (1986: 231) assumes that memorization represents a kind of a pre-stage of lexicalization. Approaches taken by some other authors go well with this idea: While memorization is often considered to be a mental operation attributed to individual language users, lexicalization represents a process that takes place within society if an item enters the shared vocabulary of a language (Pawley and Syder 1983: 208-209; Schwarze and Wunderlich 1985: 16; Lüdi 1986: 226). Specifically, Lüdi (1986: 226) regards lexicalization as "collective memorization". Relying on the proposal that memorization by language users can lead to the lexicalization of a specific expression in a language, we can use the notion of

memorization as a point of departure in order to shed more light on the lexicalization of complex constructions such as compounds and phrases.

4. The memorization of compounds

4.1 General remarks and existing evidence

A recent study by Kotowski, Böer and Härtl (2014: 195-196) compared German AN compounds and AN phrases in terms of memorization. The authors carried out a memorization study by means of an experimental design that consisted of two phases on each of three test days. In the first phase, the task of all subjects was to memorize novel AN compounds as well as AN phrases that were combined with an image showing the entity of interest. In the second phase, a lexical-decision task was conducted where participants had to decide whether a compound or a phrase appeared with the same picture as in the memorization phase. Kotowski et al.'s (2014: 195-196) analysis revealed that subjects gave slower and less accurate responses when reacting to compounds that had not occurred in the memorization phase in comparison to the responses to phrases that had not been part of the memorization phase. However, the reactions to memorized compounds and memorized phrases did not significantly differ. Therefore, the authors argue that memorization can overcome initial difficulties in the processing of compounds. Being exposed to a compound for the first time seems to come with processing difficulties as these constructions are, in contrast to phrases, marked. The process of memorization, however, pushes compounds so that the processing of memorized compounds equals the processing of memorized phrases.

The study presented in the previous paragraph shows that compounds benefit more from memorization than phrases. Although subjects reacted faster and more accurately to non-memorized phrases than to non-memorized compounds, no difference between the responses to memorized compounds and memorized phrases was detected. In other words, compounds were retained better than phrases in the course of memorization over three days. Assuming the above-mentioned relationship between memorization and lexicalization, we can say in a more general sense that compounds and phrases differ in the way they are lexicalized because differences in the process of memorization have an impact on the process of lexicalization. Since memorization represents a crucial step towards lexicalization, we aim at shedding more light on the process of memorization in the context of the demarcation of compounds and phrases in the discussion below. We hypothesize that compounds benefit more from the process of memorization than phrases. At this point, we must specify the exact nature of the expected advantage. Put differently, we have to define a potential memorization advantage and state how it becomes evident. Let us consider two definitions of the term “memorization advantage” for the present contribution (cf. also Schlechtweg and Härtl 2015). First of all, we can speak of a memorization advantage if subjects react faster/more accurately to compounds than to phrases overall, i.e. on all test days taken together. The definition is strictly unidirectional because faster/more accurate responses of phrases cannot be regarded as a memorization advantage. A potential advantage of phrases would be rooted in the fact that phrases represent the normal or standard construction type (ten Hacken 2013: 97). The proposal is connected to the phenomenon of markedness. As stated in Kotowski et al. (2014), compounds are more marked than phrases. Since unmarked items are usually more frequent than marked items and since higher frequency is known to cause faster/more accurate responses in lexical-decision tasks (e.g. Forster and Davis 1984; Bybee 1995: 237 referring to Greenberg 1966), an advantage of phrases on all test days together would not be a memorization advantage but rather an advantage triggered by the higher frequency of the phrasal pattern in general. As a consequence, it is important to take a different definition of

the notion of memorization advantage into account if the responses to phrases are faster/more accurate on all test days together or if the reactions to phrases and those to compounds are similar on all test days together: Although compounds are responded to more slowly/less accurately than phrases on the initial day, they show a memorization advantage in comparison to phrases if responses to compounds do not differ from the responses to phrases on the consecutive day(s). This proposal originates in the fact that the phrasal pattern is more common and frequent (ten Hacken 2013: 97) and, therefore, more likely to cause faster latencies and more accurate responses at an early stage of the memorization process, i.e. on the first test day. If the advantage of phrases disappears through memorization, we can speak of a memorization advantage of compounds because they improved more than phrases in the course of the study.

4.2 Cross-linguistic evidence

Having defined the notion of memorization advantage, we can now turn to the question of how to find further evidence for the idea that compounds show a memorization advantage in comparison to phrases. The study by Kotowski et al. (2014: 195-196) outlined above focused on the analysis of German AN compounds and phrases. In the current paper, we will go a step further and present empirical evidence from three languages, namely German, French and English. Similar to Kotowski et al. (2014: 195-196), we will concentrate on combinations of an adjective and a noun. These constructions represent an interesting group to work on because many potentially confounding variables can be controlled for across the languages under investigation. Remember that we defined an AN phrase as a construction where the adjective agrees with the following or preceding noun and an AN compound as a construction without internal agreement. Applying this definition to the three aforementioned languages and asking what kind of construction (AN compound or AN phrase) each of the languages prefers when naming a new complex lexical concept, we get a clear picture for German and French but not for English. While German tends to use AN compounds as naming units, French favors AN phrases (Van Goethem 2009). Relying on inflection as the decisive factor to differentiate between compounds and phrases, we cannot define AN compounds and AN phrases in English. The only thing we can do is to use our intuitions in the case of English constructions. Specifically, we can assume that English AN constructions with initial stress are compound-like constructions and English AN constructions with non-initial stress are phrase-like constructions (cf. also Schlechtweg and Härtl 2015). The idea is based on Chomsky and Halle's (1968: 17) distinction between compound stress, i.e. initial stress, and nuclear or phrasal stress, i.e. non-initial stress. Despite the criticism of this proposal (Bell and Plag 2012: 487), it represents a good intuition. Although initial stress is not the defining criterion of AN compounds, it usually occurs in AN compounds. At this point, it is helpful to refer to German where the factor of inflection unambiguously identifies AN compounds and phrases. In this language, AN compounds are typically stressed on the adjective (Erben 2000: 43). Since both German and English are languages of Germanic origin, it is plausible to regard initial stress as a common marker of AN compounds (Pereltsvaig 2012: 10) and, thus, to call English AN constructions with initial stress compound-like constructions.

So far, we have assumed that German prefers compounds and French favors phrases to express a complex lexical concept through an adjective-noun combination. Since we have evidence from a single language suggesting a different pattern of memorization for compounds and phrases, we might now ask whether the contrast is also reflected cross-linguistically. Specifically, we can raise the question of whether German AN compounds deviate from French AN/NA phrases in terms of memorization. If languages differ in their preferences for either compounds or phrases and if compounds and phrases differ with regard

to memorization, we can expect that complex constructions from one language, namely compounds, that are composed of two specific constituents are memorized differently than complex constructions from another language, namely phrases, that contain the same constituents. In order to test this hypothesis, a study was conducted that contrasted the memorization of German AN compounds and French AN/NA phrases. In addition to items from these two languages, English complex AN constructions were examined in the same study with respect to how they were memorized. The experiments included not only English AN constructions with initial stress but also AN constructions with non-initial stress.

In the following, we will summarize the experimental study. Note that a very detailed presentation of the study is given in Schlechtweg and Härtl (2015). In this earlier contribution, we describe the participants, the material, the procedure as well as the central hypotheses in detail, present a comprehensive result section and all items under investigation. Contrasting the memorization of German AN compounds, French AN/NA phrases, English AN constructions with initial stress and English AN constructions with non-initial stress, we aimed at investigating whether a memorization advantage of compounds/compound-like constructions can be observed from a cross-linguistic perspective. Note that we regarded the English constructions with initial stress as compound-like constructions and the constructions with non-initial stress as phrase-like constructions. Speakers of the three aforementioned languages participated in the study and were tested on complex constructions of their native language. For this purpose, they were divided into the four groups German, French, EnglishA and EnglishB. Note that two English groups were created in order to examine AN constructions stressed on the adjective in one group (EnglishA) and AN constructions stressed on the noun in the other group (EnglishB). In our study, we included different types of items. While novel AN/NA constructions such as *Jungtourist/jeune touriste/YOUNG tourist/young TOURist* represented the experimental items, existing nouns of the languages under investigation, e.g. *Architekt/architecte/architect/architect*, were used as control items (baseline). Subjects were asked to memorize both the experimental and the control items on three test days, i.e. on day one, day four and day eight. Apart from these items that had to be memorized on three days, we included filler items that did not have to be memorized. Filler items were either other AN/NA constructions or other existing nouns. When creating our set of experimental, control and filler items, we took several potentially confounding variables into account and controlled for them across the three languages examined in the study (e.g. number of syllables, duration in seconds, frequency of the constituents, control and filler nouns, lexicalization status of the complex AN/NA constructions). The experiment was conducted by using the computer program E-Prime (Psychology Software Tools, Inc. 2010). All subjects participated in the experiment on three days and in two phases on each day. In the first phase, the memorization phase, participants heard and memorized the experimental and the control items. Right after that, i.e. in the second or recall phase, subjects heard both the memorized items from the first phase as well as other, non-memorized items, i.e. the filler items. Having heard a memorized item, a participant was expected to press a “Yes”- button in the recall phase. If a subject heard a non-memorized item, however, he or she was supposed to press a “No”-button.

While analyzing the study, we examined both how fast and how accurate subjects responded to the experimental and control items they heard in the recall phase. Note that we will focus on reaction time in the following. Contrasting compounds/compound-like constructions to phrases/phrase-like constructions, we expected differences in the response latencies, i.e. we hypothesized a memorization advantage of the compounds/compound-like constructions. Crucially, however, there should be no difference between the response times of the German, French and English control items. That means, the control items served as a baseline in order to verify that it is possible to compare these languages in a psycholinguistic

study at all. Looking at the response latencies, our statistical analyses confirmed our hypothesis concerning the control items only for the comparisons German-French as well as EnglishA-EnglishB. In other words, the reaction times of the German control items did not significantly differ from the reaction times of the French control items and the response latencies of the control items of the group EnglishA did not significantly differ from the response latencies of the control items of the group EnglishB.² Therefore, when examining the complex AN/NA constructions, we compared the German compounds only to the French phrases. Moreover, we contrasted the English AN constructions with initial stress only to the English AN constructions with non-initial stress. Our analyses revealed that subjects responded significantly more slowly to the French phrases than to the German compounds (Difference of means (henceforth: DM) = 65.0, $t = 3.12$, $p = .046$). We consider this effect to mirror a memorization advantage of compounds in comparison to phrases. Furthermore, the reaction times of the English AN constructions with non-initial stress were shorter than the response latencies of the English AN constructions with initial stress (DM = -65.2, $t = -3.13$, $p = .045$). The results reported so far are presented in Figure 1.

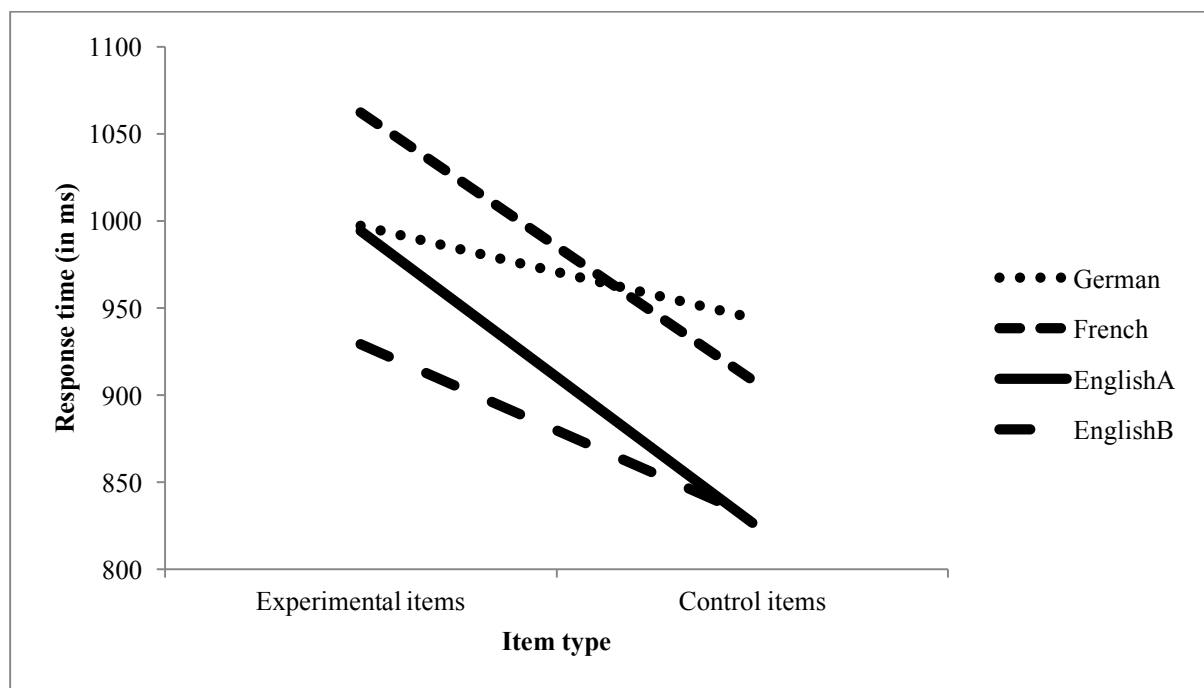


Figure 1: Interaction of LANGUAGE x ITEM TYPE (F_2)

At this point, however, remember that the effect in English does not reflect a memorization advantage but an advantage that originates in the fact that non-initial stress represents the default prosodic structure in English AN constructions (cf. also Fudge 1984: 146; Zwicky 1986: 51; Bauer, Lieber and Plag 2013: 448). Schiller, Levelt and Fikkert's (2004: 237-238) results suggest that responses are faster to the standard stress pattern than to a marked prosodic structure. Therefore, we have to apply the second definition of the term memorization advantage that we introduced above, i.e. the comparison of the response times between the AN constructions with initial stress and those with non-initial stress on the three individual test days. It turned out, however, that the latencies of the two groups did not significantly differ from each other on any of the individual days. In sum, we found a

² All results of this study that are reported in the current paper refer to the item analysis (F_2).

memorization advantage of the German compounds compared to the French phrases but we did not detect a memorization advantage of the English compound-like constructions in comparison to the phrase-like constructions of the same language.

When looking at the set of English AN constructions under investigation in the aforementioned study and reflecting upon the nature of compounds and phrases, we realized that our compound-like constructions were characterized by only one typical feature of compounds, namely initial stress. Therefore, we decided to conduct a second study that investigated not only the factor of stress but also the factor of semantic compositionality. Since compounds are also known to be semantically non-compositional (Downing 1977: 820), we aimed at combining two factors and examining their influence on memorization. Specifically, we redefined our compound-like as well as our phrase-like constructions and investigated whether a memorization advantage can be observed for English compound-like constructions, i.e. AN constructions that carry initial stress and are not semantically compositional (e.g. *HARD shirt*), in comparison to English phrase-like constructions, i.e. AN constructions that bear non-initial stress and are semantically compositional (e.g. *short BRUSH*). Again, as in the case of the first study, the reader is advised to consult Schlechtweg and Härtl (2015) for a very detailed description of this study. In the present work, we can only give a short summary.

The procedure used in this study resembled the procedure of the first investigation described above. However, the examination consisted of two memorization phases and one recall phase on each of the three test days. Our experimental items had to be memorized on three days and could be divided into four groups: Items like *short BRUSH* were stressed on the second constituent and semantically compositional, items like *SHORT brush* were stressed on the first constituent and semantically compositional, items like *hard SHIRT* were stressed on the second constituent and semantically non-compositional and items like *HARD shirt* were stressed on the first constituent and semantically non-compositional. Our filler items were other AN constructions that had not to be memorized and had the function to trigger a “No”-response in the recall phase. When creating the complex AN constructions, we took several potentially confounding variables into consideration (e.g. number of syllables, duration in seconds, frequency of the constituents, lexicalization status of the AN constructions). The decision of whether an item was considered to be semantically compositional or not was based on the results of an online survey (SoSci, Leiner 2014).

In this study, we expected a memorization advantage – in the second sense of the term – of the compound-like constructions, i.e. of the semantically non-compositional constructions with initial stress, compared to the phrase-like constructions, i.e. the semantically compositional constructions with non-initial stress. Put differently, we hypothesized that the reactions to compound-like constructions are slower than the response latencies of phrase-like constructions on the first but not on the following days. Our statistical analyses revealed that the responses to phrase-like constructions were highly significantly faster than the responses to compound-like constructions when considering all three days together ($DM_1 = -68.8$, $t_1 = -5.52$, $p_1 = .000$; $DM_2 = -61.3$, $t_2 = -4.40$, $p_2 = .000$).³ Since phrase-like constructions represent the more common pattern (Lieberman and Sproat 1992: 134; Giegerich 2009: 5-7), this result was expected and simply mirrored an advantage caused by the frequency of the constructions. Looking at the individual days, we found that the phrase-like constructions were responded to (highly) significantly faster than the compound-like constructions on the first but not on the second and third day (Day 1: $DM_1 = -107.1$, $t_1 = -4.96$, $p_1 = .000$; $DM_2 = -81.1$; $t_2 = -3.36$, $p_2 = .048$). These results are presented in Figure 2. We regard this effect as a memorization advantage of compound-like constructions in comparison to phrase-like constructions as the

³ Values with the subscript “1” refer to the subject analysis (F_1) and values with the subscript “2” refer to the item analysis (F_2).

former clearly showed a bigger improvement than the latter from an early stage to a later stage in the memorization process.

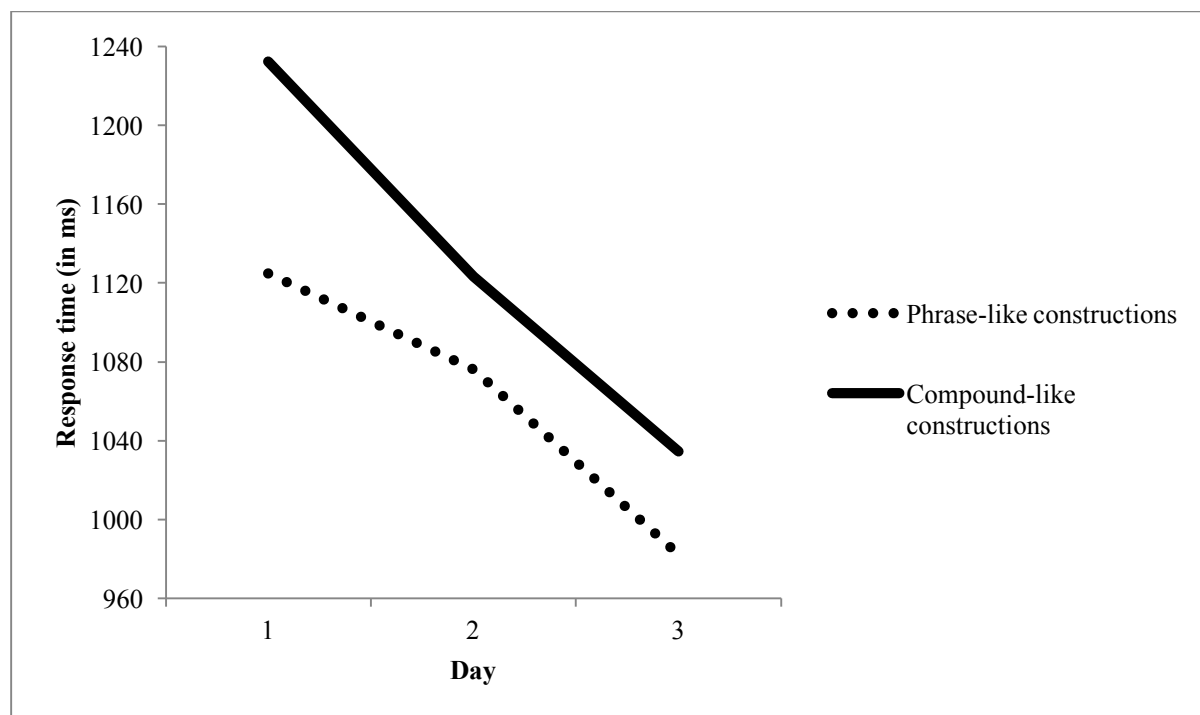


Figure 2: Phrase-like constructions (semantically compositional items with non-initial stress) versus compound-like constructions (semantically non-compositional items with initial stress) on the three test days (F_1)

5. Summary and conclusion

We started our paper with a discussion of the idea that compounds are more likely to become lexicalized than phrases. Then, we analyzed this proposal in more detail with the help of the notion of memorization, which is connected to the process of lexicalization in an important way, and by presenting evidence from studies that investigated the memorization of compounds and phrases.

The results presented in the analysis suggest that compounds/compound-like constructions show a memorization advantage in comparison to phrases/phrase-like constructions. Both the comparison of compounds/compound-like constructions and phrases/phrase-like constructions within a single language and the contrast of these construction types across different languages suggest a fundamental processing difference between compounds/compound-like constructions and phrases/phrase-like constructions. Since memorization represents a crucial step in lexicalization, the evidence reported in the current paper indicates that compounds/compound-like constructions and phrases/phrase-like constructions differ in the way they become a lexicalized complex construction.

References

- Bakken, K. (2006) Lexicalization. In: K. Brown, A.H. Anderson, L. Bauer, M. Berns, G. Hirst & J. Miller (Eds.), *Encyclopedia of language and linguistics*. Oxford: Elsevier, 106-108.
- Bauer, L. (1988) *Introducing linguistic morphology*. Edinburgh: Edinburgh University Press.
- Bauer, L., R. Lieber & I. Plag (2013) *The Oxford reference guide to English morphology*. Oxford: Oxford University Press.

- Bell, M. J. & I. Plag (2012) Informativeness is a determinant of compound stress in English. *Journal of Linguistics* 48(3): 485-520.
- Blank, A. (2001) Pathways of lexicalization. In: M. Haspelmath, E. König, W. Oesterreicher & W. Raible (Eds.), *Language typology and language universals: An international handbook*. Berlin: Walter de Gruyter, 1596-1608.
- Booij, G. (2012) *The grammar of words*. Oxford: Oxford University Press.
- Bücking, S. (2009) How do phrasal and lexical modification differ? Contrasting adjective-noun combinations in German. *Word Structure* 2(2): 184-204.
- Bücking, S. (2010) German nominal compounds as underspecified names for kinds. In: S. Olsen (Ed.), *New impulses in word-formation*. Hamburg: Buske, 253-281.
- Bybee, J. L. (1995) Diachronic and typological properties of morphology and their implications for representation. In: L. B. Feldman (Ed.), *Morphological aspects of language processing*. Hillsdale: Lawrence Erlbaum Associates, 225-246.
- Chomsky, N. & M. Halle (1968) *The sound pattern of English*. New York: Harper & Row.
- Downing, P. (1977) On the creation and use of English compound nouns. *Language* 53(4): 810-842.
- Erben, J. (2000) *Einführung in die deutsche Wortbildungslehre*. Berlin: Erich Schmidt Verlag.
- Forster, K. I. & C. Davis. (1984) Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 10(4): 680-698.
- Fudge, E. (1984) *English word-stress*. London: George Allen & Unwin.
- Gaeta, L. & D. Ricca (2009) Composita solvantur: Compounds as lexical units or morphological objects? *Rivista di Linguistica* 21(1): 35-70.
- Giegerich, H. J. (2009) The English compound stress myth. *Word Structure* 2(1): 1-17.
- Greenberg, J. (1966) *Language universals*. The Hague: Mouton.
- Härtl, H. (2015) Semantic non-transparency in the mental lexicon: On the relation between word-formation and naming. In: C. Brinker-von der Heyde, N. Kalwa, N.-M. Klug & P. Reszke (Eds.), *Eigentlichkeit: Zum Verhältnis von Sprache, Sprechern und Welt*. Berlin: de Gruyter, 395-416.
- Herbermann, C.-P. (1981) *Wort, Basis, Lexem und die Grenze zwischen Lexikon und Grammatik: Eine Untersuchung am Beispiel der Bildung komplexer Substantive*. München: Wilhelm Fink Verlag.
- Hüning, M. (2010) Adjective + Noun constructions between syntax and word formation in Dutch and German. In: A. Onysko & S. Michel (Eds.), *Cognitive perspectives on word formation*. Berlin: De Gruyter Mouton, 195-215.
- Kotowski, S., K. Böer & H. Härtl (2014) Compounds vs. phrases: The cognitive status of morphological products. In: F. Rainer, F. Gardani, H. C. Luschützky & W. U. Dressler (Eds.), *Morphology and meaning: Selected papers from the 15th International Morphology Meeting, Vienna, February 2012*. Amsterdam: John Benjamins Publishing Company, 191-203.
- Leiner, D. (2014) *SoSci Survey* [Computer program]. Version 2.6.00-i. <https://www.sosicisurvey.de> (Access from February 25, 2015 through March 23, 2015).
- Liberman, M. & R. Sproat (1992) The stress and structure of modified noun phrases in English. In: I. A. Sag & A. Szabolcsi (Eds.), *Lexical matters*. Leland Stanford Junior University: Center for the Study of Language and Information, 131-181.
- Lipka, L. (1981) Zur Lexikalisierung im Deutschen und Englischen. In: L. Lipka & H. Günther (Eds.), *Wortbildung*. Darmstadt: Wissenschaftliche Buchgesellschaft, 119-132.
- Lipka, L. (2002) *English lexicology: Lexical structure, word semantics & word-formation*. Tübingen: Gunter Narr Verlag.
- Lüdi, G. (1986) Forms and functions of bilingual speech in pluricultural migrant communities in Switzerland. In: J. A. Fishman, A. Tabouret-Keller, M. Clyne, B. Krishnamurti & M. Abdulaziz (Eds.), *The Fergusonian Impact: In honor of Charles A. Ferguson on the occasion of his 65th birthday: Volume 2: Sociolinguistics and the sociology of language*. Berlin: Mouton de Gruyter, 217-236.
- Pawley, A. & F. H. Syder (1983) Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In: J. C. Richards & R. W. Schmidt (Eds.), *Language and communication*. London: Longman, 191-226.
- Pereltsvaig, A. (2012) *Languages of the world: An introduction*. Cambridge: Cambridge University Press.
- Psychology Software Tools, Inc. (2010) *E-Prime 2 Professional* [Computer program]. Sharpsburg.
- Schiller, N. O., P. Fikkert & C. C. Levelt (2004) Stress priming in picture naming: An SOA study. *Brain and Language* 90: 231-240.
- Schlechtweg, M. & H. Härtl (2015) Memorization and the morphology-syntax divide: A cross-linguistic investigation. Submitted to *SKASE Journal of Theoretical Linguistics*. <http://ling.auf.net/lingbuzz/002770> (Accessed on January 4, 2016), <http://www.marcel schlechtweg.com/articles/> (Accessed on February 11, 2016).
- Schlücker, B. (2014) *Grammatik im Lexikon: Adjektiv-Nomen-Verbindungen im Deutschen und Niederländischen*. Berlin: De Gruyter.

- Schwarze, C. & D. Wunderlich (1985) Einleitung. In: C. Schwarze & D. Wunderlich (Eds.), *Handbuch der Lexikologie*. Königstein: Athenäum, 7-23.
- ten Hacken, P. (2013) Compounds in English, in French, in Polish, and in general. *SKASE Journal of Theoretical Linguistics* 10(1): 97-113.
- Van Goethem, K. (2009) Choosing between A + N compounds and lexicalized A + N phrases: The position of French in comparison to Germanic languages. *Word Structure* 2(2): 241-253.
- Wunderlich, D. (1986) Probleme der Wortstruktur. *Zeitschrift für Sprachwissenschaft* 5(2): 209-252.
- Zwicky, A. M. (1986) Forestress and afterstress. *The Ohio State University (OSU): Working Papers in Linguistics (WPL)* 32: 46-62.

Spatial reduplication in Sicilian: lexicon or grammar?

Giuseppina Todaro
Università di Roma Tre / CLLE-ERSS, CNRS
 & *Université de Toulouse II Jean Jaurès*
 giuseppina.todaro@etu.univ-tlse2.fr

Fabio Montermini
CLLE-ERSS, CNRS
 & *Université de Toulouse II Jean Jaurès*
 fabio.montermini@univ-tlse2.fr

1. Introduction

In this paper we propose an analysis of a particular type of total NN reduplicative structures in Sicilian (and regional Italian spoken in Sicily) with a spatial semantic value. In particular, as we show, these constructions have the function of establishing a trajector / landmark spatial relation between two entities. As we will see, in different subtypes, the reduplicated noun can correspond either to the landmark or to the trajector. In the analysis we present these reduplications are considered as a morphological derivational (lexeme-forming) strategy, creating either adverbs or adjectives. In particular, we show that, in spite of the different input and output categories they specify, the homogeneity in the construction of their semantic interpretation justifies considering that all these strategies are indeed subtypes of the same, larger, construction. Moreover, we show that the semantic interpretation of Sicilian total reduplications crucially depend on the semantic features of the reduplicated noun, but also on the semantics of the related entity (typically expressed by a noun, but which can also be expressed by an explicit or implicit PRO) and of the verb marking the syntactic relation between the two. Consequently, we propose to consider that the morphological construction involving the reduplication is in turn encompassed into a larger construction which also includes an element (typically a noun) corresponding to the entity which is in a spatial relationship with the reduplicated noun. More globally, we consider that the data presented and the analysis proposed provide evidence in favor of considering reduplication as a particular subtype of compounding. Like compounding, it is useful to distinguish between cases in which reduplicated structures are the result of a general, cross-linguistic (and possibly universal) human ability to combine words and cases in which they correspond to grammaticalized lexeme-formation strategies submitted to language-specific constraints.

Our paper is organized as follows: in Section 2 we argue in favor of a joint analysis for compounding and (total) reduplication, and of identifying a dichotomy between two types of compounding / reduplication; Section 3 presents the data our treatment is based on; Section 4 proposes an analysis of Sicilian total reduplications in a Construction Morphology framework; finally, Section 5 contains some concluding remarks.

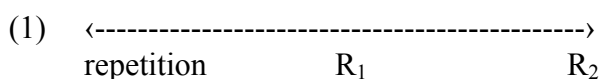
2. The place of reduplication in linguistic competence

Total reduplication (TR) can be considered as a particular type of compounding (cf. Bauer 2003). Generally speaking, compounding consists in creating a lexeme by combining two (rarely more than two) other lexemes. Formally, in most cases the output lexeme results from the juxtaposition of the stems of the two input lexemes (although more complex cases are possible, cf. Montermini 2010). Semantically, the meaning of the output lexeme results from a function applied to the meanings of the two input lexemes (Guevara and Scalise 2009: 104). A rule of TR specifies that the two input lexemes are the same; consequently, the form of the output lexeme consists in the repetition of the stem of the input, and its meaning consists in a

function applied to the meaning of the input (often involving such features as plurality, iterativity, intensification, etc.).

Compounding itself, however, can be seen as a complex and heterogeneous phenomenon. What we normally call “compounding”, in fact, may correspond to at least two linguistic phenomena differing both in their origin and in their properties. On the one side, the presence of compound words in a language can be the outcome of a general cognitive ability to connect semantically two words by simply juxtaposing them. We propose to call this phenomenon Compounding₁ (C₁). In this case, the juxtaposition of lexemes that takes place is closer to what happens in syntax, and, semantically speaking, the compound has a compositional meaning inferable from the sum of the meanings of the lexemes involved. When linguists cite compounding as the “widest-spread morphological technique” (Dressler 2006: 23), it is probably C₁ that they have in mind. On the other side, the rules for forming compound words differ significantly from one language to another and are submitted to various kinds of restrictions and constraints, which include phonological and prosodic constraints on the form of the input stems, constraints on the semantic compatibility between the two lexemes involved, constraints on head placement, constraints on the presence of “linking elements”, which can be obligatory (e.g. *-o-*, or sometimes a different vowel, in Modern Greek, cf. Ralli 2013) or optional (e.g. *-e-*, *-en-*, *-s-* in German and Dutch, cf. Libben et al. 2009 and Don 2009: 380-381, respectively). These facts suggests that, in another sense, compounding, like all other derivational phenomena, consists in a set of morphological patterns that are codified in the grammar of individual languages, and are therefore subject to different language specific parameters (cf. Bauer 2009; Guevara and Scalise 2009; Arcodia and Montermini 2012). We propose to call this phenomenon Compounding₂ (C₂).

Similar considerations can be applied to TR. In particular, a parallel distinction between Reduplication₁ (R₁) and Reduplication₂ (R₂) can be drawn. Cross-linguistically (and sometimes within the same language), we observe a sort of *continuum* going from R₁, the repetition of the same word-form – usually for stylistic purposes –, to clear instances of R₂, i.e. forms that represent the output of a language-specific process. Gil (2005) argues in favor of a *continuum* ranging from stylistically marked repetition to strictly grammaticalized reduplication with a large greyish in-between area for which the classification is not straightforward. We can consider that R₁ stands between the two poles identified by Gil, as we show in (1):



Repetition (or reiteration) is a pragmatic cross-linguistic device generally employed in oral speech (less often in writing) to mark emphasis. Unlike in typical reduplication, the resulting structures are not necessarily binary, as units can be repeated recursively (2a); moreover, the units involved are not necessarily stems or word forms, but may consist in larger syntactic units, such as APs (2b); and finally, the repeated units are not necessarily contiguous, since other units, such as conjunctions, can be intercalated between them (2c):

- (2) a. *He runs, runs, runs and climbs everything like a little monkey.*
 [<http://haberdasheryfun.com/page/15>]
 b. *This is a very beautiful, very beautiful, sturdy table.*
 [<http://www.wayfair.com/Abbyson-Living-Sienna-Dining-Table-AD-DT-018-BYV2437.html>]
 c. *In fact, the tension builds so intensely as Lola runs, runs, and runs that we can only be thankful that the film has a relatively short running time.*
 [<http://www.hccentral.com/cgi-bin/films.cgi?fid=12>]

In its turn, R_1 has often a specific semantic function: it is restricted to the designation of a ‘real’ or prototypical instance of the object designated by the input lexeme. This phenomenon, largely described in the literature, has been observed in the colloquial speech of many languages and is often called Contrastive Focus Reduplication (Ghomeshi et al. 2004, Forza 2011). In oral speech, one of the elements is often pronounced with a special intonation, marking emphasis (a fact we indicate with capital letters in (3)).

- (3) a. *I’ll make the tuna salad, and you make the SALAD-salad*
 b. *Is he French or FRENCH-French?*
 [Ghomeshi et al. 2004: 308]

Like C_1 , R_1 may be considered as a general feature of the human linguistic ability, whereas R_2 is a more specific phenomenon, submitted to language-specific parameters.

Both types of reduplication are attested in Sicilian: R_1 constructions (4a) express one of the semantic values prototypically associated with reduplication with no specific grammatical feature (the same phenomenon can be observed in Italian, among others, cf. Wierzbicka 1986, Dressler and Merlini Barbaresi 1994, Forza 2011); R_2 (4b), on the other hand, corresponds to a morphological lexeme-formation process:

- (4) a. (R_1) *Avi* *l’* *occhi* *nìuri* *nìuri*
 have.3.SG.PRES.IND DET.M.PL eye.M.PL black.M.PL black.M.PL
 ‘He/she has really black eyes’
 b. (R_2) *Camina* *murù* *murù*
 walk.3.SG.PRES.IND wall.M.SG wall.M.SG
 ‘He/she walks along the wall’

(4a) is an example of R_1 : here, no specific effect on the grammatical features of the input lexeme is observed (*nìuri nìuri* functions as an adjective like *nìuri*), and semantically it has the contrastive function referred to above. On the other side, (4b) is an example of R_2 . Here, the effect observed corresponds to what happens with typical derivational, lexeme-forming, phenomena: it has an effect on category (*murù murù* functions as an adverb, while *murù* is a noun), and a specific meaning, non-recoverable from the meaning of the input lexeme (cf. **camina murù*) or from the meanings usually attributed to reduplication. Specifically, the reduplicated form in (4b) expresses the same spatial relationship expressed in English by the preposition *along*. R_2 constructions like the one in (4b) constitute the main focus of this article.

3. The data

The analysis we present is based on a dataset collected from dictionaries, novels and essays, as well as from the scholarly literature on Sicilian. Moreover, some of the data have been collected from occasional exchanges with speakers from the Western area of Sicily (provinces of Trapani and Palermo). Sicilian exhibits the kind of asyndetic total R_2 (cf. Stolz 2009) exemplified above in all its varieties, including in non-standard regional varieties of Italian. The phenomenon involves both nouns and verbs as inputs (Caracausi 1977; Leone 1995; Sgarioto 2005; Amenta 2010; Emmi 2011). In this paper we focus on a particular kind of total NN reduplications that produce adjective- or adverb-like units. The constructions in question do not express any of the meanings generally attributed to reduplications (plurality, emphasis, intensity, iterativity, etc.); in fact, their function is to establish a spatial relationship between a

trajector and a landmark, under specific semantic conditions (cf. Todaro et al. 2014 for a detailed semantic account).

The literature on reduplication generally considers TR as a potentially universal phenomenon (Moravcsik 1978; Stolz 2009). In its turn, reduplication as a proper morphological process (R₂) is often considered to be rare in Romance (and even in Indo-European) languages, although recent studies claimed that TR is not completely absent in these languages. With the help of statistic analyses, it has been shown, for instance, that reduplicated constructions are attested in the Romance languages spoken in the Mediterranean area (see Stolz et al. 2011). These data, like the ones we consider, provide evidence in favor of considering “TR as a potential areal feature in the area of the putative Mediterranean *Sprachbund*” (Stolz et al. 2011: 519). Consequently, we consider R₂ as being part of the basic morphology of Sicilian. In particular, in Sicilian four types of lexeme-forming constructions involving reduplication can be identified:¹

- (5)² a. [VV]_N *curri-curri*
 run-run
 ‘stampede’

Po’ si sintì na gran vociata e ci fu, nella parti opposta della pista, un movimento di curri curri.

‘Then we heard a loud scream and a stampede movement was on the opposite side of the track’

[A. Camilleri, *La pista di sabbia*. Palermo: Sellerio Editore]

- b. [VV]_A *cala-cala*
 go down-go down
 ‘which goes down easily’ / ‘easy-to-drink’

Un vino fatto come lo facevano i nostri nonni, in poche parole un vino cala cala che ti inchiumma a tradimento ma con soddisfazione.

‘A wine made like our ancestors used to make it, in short, an easy-to-drink wine which pins you down unexpectedly but with pleasure’

[<http://www.firriotate.com/2013/12/10/degustazione-vino-veritas/>]

- c. [N_{SG}N_{SG}]_{Adv} *casa-casa*
 house-house
 ‘all over the house’

Noi picciriddri giocavamo a nascondino casa casa.

‘We the children used to play to hide-and-peek all over the house’

[<http://ricerca.repubblica.it/repubblica/archivio/repubblica/2011/02/06/il-bonifico-di-berlusconi-mia-nonna-maria.html>]

- d. [N_{PL}N_{PL}]_A *pirtusa-pirtusa*
 holes-holes
 ‘full of holes’

Genti coi vistiti pirtusa pirtusa che cadivano a pezzi.

‘People with very ripped clothes, falling apart’

[A. Camilleri, *La rivoluzione della luna*. Palermo: Sellerio Editore]

¹ For each type we provide real examples with an indication of the source. All the examples have been translated by the authors.

² Forms like (5a) recall the Italian reduplications that produce nouns from verbs (*fuggifuggi* ‘stampede’, lit. run away-run away, cf. Masini and Thornton 2009), although in Italian the reduplication in question does not seem to be productive anymore.

In this work we deal exclusively with types (5c) and (5d), i.e. with denominal reduplications forming adjectives and adverbs and expressing a spatial relation. In particular, we consider that in Sicilian three types of constructions are available in order to express a spatial relation between two entities: one of the entities is denoted by the input noun of the reduplicated construction, the other is denoted by a syntactically connected unit (in the unmarked case a NP).

Type A (6a')	Type B (6a'')	Type C (6b)
[N _{SG} N _{SG}] _{Adv}	[N _{PL} N _{PL}] _{Adv}	[N _{PL} N _{PL}] _A

Table 1: Three types of NN total reduplications in Sicilian.

- (6) a'. *U picciriddro ioca casa-casa*
 DET.M.SG child.M.SG play.3.SG.PRES.IND house.F.SG-house.F.SG
 ‘The child plays all over the house’
- a''. *U cane camina muntagni-muntagni*
 DET.M.SG dog.M.SG walk.3.SG.PRES.IND mountain.F.PL-mountain.F.PL
 ‘The dog walks in/through the mountains’
- b. *A cammisa è pirtusa-pirtusa*
 DET.F.SG shirt.F.SG be.3.SG.PRES.IND hole.M.PL-hole.M.PL
 ‘The shirt is full of holes’

In two of the types, the reduplicated unit can be either a singular (type A) or a plural (type B) nominal form. The output of the reduplication is an adverb. In type C the reduplicated unit can only be a plural nominal form, and the output functions as an adjective. Semantically, all three constructions denote the relative spatial localization of two entities.

4. The analysis

4.1 Sicilian reduplications as constructions

For the analysis of Sicilian NN R₂ constructions we adopt a Construction Morphology approach (cf. Booij 2010). Therefore, we claim that, like constructions, the reduplications in question instantiate generalizations made by speakers on sets of existing complex words with a systematic correlation between form and meaning. In particular, we propose to represent the TRs exemplified in (6) as constructional schemas ([XX]_α) which, in their turn, are encompassed into larger constructions involving a semantically and syntactically related entity (in most cases expressed by a NP), that we represent as follows: [Y \Re [XX]_α]. Although both levels can be characterized as constructions, the lowest level ([XX]_α) constitutes a typical morphological object, whereas the highest one ([Y \Re [XX]_α]) corresponds to a typical syntactic object. However, as we will see below, both the lowest and the highest level display holistic semantic properties that cannot be accounted for simply on a compositional basis. If the reduplicated structure is an adjective (type C), Y corresponds to its head noun or to the NP it predicates; if the reduplicated structure is an adverb (types A-B), Y corresponds to a syntactically linked NP (see 6a', 6a'').

Sicilian R₂ constructions display some typical properties of morphological derivational processes: i) they have developed a range of specific meanings not recoverable from the combination of the meanings of the base lexemes; ii) their outputs correspond to typical lexical units (no insertion possible between the nouns, single primary stress, no recursivity);

iii) they have an effect on the category of the lexemes they apply to, since they change a noun into an adverb or into an adjective.

From a semantic point of view, Sicilian R_2 constructions mark a spatial localization between a landmark (an anchoring entity) and a trajector (an entity to be located) (cf. Langacker 1987)³. In types A-B, the reduplicated noun designates the landmark and the syntactically connected noun designates the trajector; in type C the situation is reverted: the reduplicated noun designates the trajector and the head noun of the NP in which the reduplication appears designates the landmark. In most cases, the landmark corresponds to a circumscribed entity with discrete boundaries, and the trajector is located within these borders. Virtually any object possessing a physical extension can function as a landmark within which a trajector can be located. Moreover, as we will show in the following sections, the semantic interpretation of the larger construction can be affected by some semantic features of the nouns involved, either as trajectors or as landmarks.

As we observed above, the construction we have globally characterized as $[Y \Re [XX]_\alpha]$ corresponds, in fact, to three different constructions, each one with different specifications for some of their elements, which in turn are encompassed into the larger construction $[Y \Re [XX]_\alpha]$. The general construction and its sub-constructions are given in Figure 1. Spatial relations between the trajector and the landmark are generically indicated as IN, although they can correspond to different spatial configurations and distributions, which depend on the semantic type of each of the nouns involved.

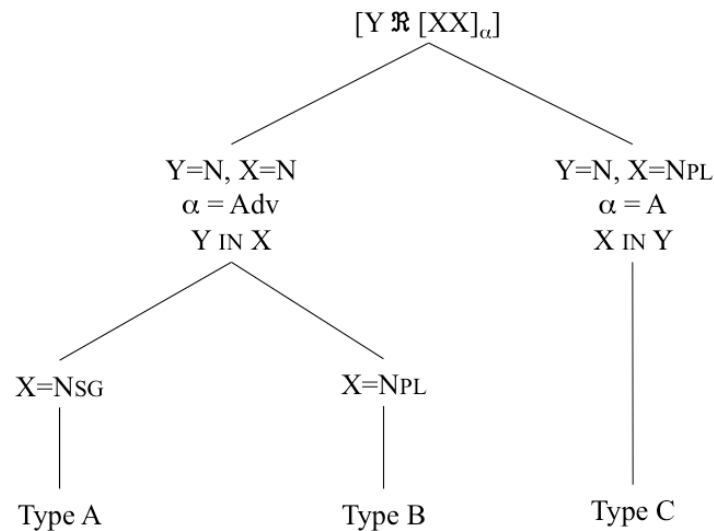


Figure 1: Types and subtypes of reduplicative spatial constructions in Sicilian.

4.2 Semantic constraints on the trajector

The global semantic interpretation of the constructions in question is conditioned by the semantic features of the participants involved. In this section, we focus on the constraints to which the trajector is submitted, whereas in the following sections we will consider the constraints on the landmark and on the verb involved in the construction.

In Types A and B, the trajector can correspond to a mobile entity moving within the boundaries of the landmark. These cases correspond to instances in which the trajector is

³ The terms *landmark* and *trajector* correspond to the notions denoted as *ground* and *figure* respectively, e.g. by Talmy (1983).

animate. Here, the relationship between the landmark and the trajector is dynamic and involves a change of location of the trajector within the circumscribed space of the landmark. The trajector does not move in a linear way, going from one point to another, but rather randomly, occupying different points in an unordered way.

- (7) *U picciottu curria paisi-paisi*
 DET.M.SG boy.M.SG run.3.SG.IMPF.IND village.M.SG-village.M.SG
 ‘The boy was running all over the village’

On the contrary, if the trajector is inanimate the spatial relationship is necessarily static. In this case, it occupies a fixed position within the boundaries of the landmark, and the whole construction can be interpreted in three different ways. The choice of one of the interpretations crucially depends on the semantic properties of the nouns involved, as well as on pragmatic factors.

- (8) a. *Si fici a casa muntagni-muntagni*
 REFL make.3.SG.PST.IND DET.F.SG house.F.SG mountain.F.PL-mountain.F.PL
 ‘He/she built his/her house somewhere in the mountains’
 b. *Chi ci fa sta pianta scala-scala?*
 what there.CL do.3.SG.PRES.IND DEM.F.SG plant.F.SG stair.F.SG-stair.F.SG
 ‘Why is that plant in the middle of the staircase (in the way)?’
 c. *Avemu a rina casa-casa*
 have.1.PL.PRES.IND DET.F.SG sand.F.SG house.F.SG-house.F.SG
 ‘We have sand all over the house’

(8a) is more likely interpreted as expressing a vague localization. In this case, the speaker only knows that the house in question has been built in the mountains, but ignores its exact localization. (8b) is more likely to be interpreted as expressing an incongruous localization. Here, the plant is seen as occupying an unusual, possibly disturbing, place. In these two cases, the choice between the two interpretations pragmatically depends on the referent of the reduplicated noun: it is unlikely that someone cannot precisely locate an object within the limited space of a staircase. On the other side, it is absolutely plausible that someone does not know the exact location of a house in the mountains. Finally, (8c) is more likely interpreted as expressing a homogeneous distribution of the referent of Y within the circumscribed space of X. Note that this interpretation can be activated either when Y corresponds to a mass noun (like *rina*), or when it corresponds to a plural noun:

- (9) *Avemu i ciura casa-casa*
 have.1.PL.PRES.IND DET.M.PL flower.M.PL house.F.SG-house.F.SG
 ‘We have flowers all over the house’

Although pragmatics is clearly the main reason for distinguishing the semantic nuances observed in examples (8a-c), a common, underspecified, reading can be observed. In fact, in each of these cases, the relevant semantic factor is that the whole surface of the landmark is (at least potentially) involved in the spatial relation. The simplest case is represented by (8c) and (9): here Y (*a rina, i ciura*) occupies the totality of X’s (*casa*) surface. On the other hand, in (8a-b) Y (*a casa, sta pianta*) may potentially occupy any point in X’s (*muntagni, scala*) surface, and the exact point it occupies does not change the whole reading: in the case of (8a) it is irrelevant; in the case of (8b), whatever point occupied by Y constitutes an incongruous location.

By contrast with types A and B, in type C the reduplicated noun corresponds to the trajector. In this case, the input of the reduplication is necessarily a plural noun form. The semantic interpretation is similar to that of (8c): here the trajector is interpreted as being homogeneously distributed within the landmark's boundaries. As it can be seen the value [+PL] is directly encoded in the construction of this type of reduplications. In this case, X (the reduplicated noun) can be either animate or inanimate, without any change in meaning, as the examples (10a-b) below show. As we observed above, such structures are not acceptable if X corresponds to a singular noun (10c); they are also unacceptable if Y is [-count] (10d).

- (10) a. *Sugnu* *puci-puci*
 be.1.PL.PRES.IND flea.F.PL-flea.F.PL
 'I am full of fleas'
- b. *Sugnu* *papuli-papuli*
 be.1.PL.PRES.IND blister.F.PL-blister.F.PL
 'I am full of blisters'
- c. **Sugnu* *papula-papula*
 be.1.PL.PRES.IND blister.F.SG-blister.F.SG
- d. **Sugnu* *ogghiu-ogghiu*
 be.1.PL.PRES.IND oil.M.SG-oil.M.SG

In Table 2, we sum up the syntactico-semantic constraints on the trajector we observed for types A, B and C, respectively.

	[animate]	[count]	[±PL]
A-B	+ (dynamic) / – (static)	+ (hom. distribution) / – (incongruity)	+ / –
C	+ / –	+	+

Table 2: Syntactico-semantic constraints on the trajector in Sicilian reduplications.

4.3 Constraints on V

As the representation in Figure 1 shows, the TRs under discussion enter in a larger construction whose function is to relate two entities, one of which is designated by the input noun of the reduplication. The related entity can be designated either by an explicit NP or by a PRO-form. Since Sicilian is a PRO-drop language, the latter may not be explicit (cf. (10)). In types A and B, it usually corresponds to one of the arguments of a verb which is modified by the adverbial reduplicated structure. In type C, it mostly corresponds to a NP modified or predicated by the adjectival reduplicated structure. Types A and B do not display any specific restriction on the types of verbs involved. As we observed above, in fact, in different subtypes the spatial relation between the trajector and the landmark may be either static or dynamic, thus making both dynamic and stative verbs acceptable. Several different readings are available for the same construction and are not necessarily mutually exclusive. Moreover, we claim that, if the trajector is [+animate], a dynamic meaning is intrinsic in the construction itself. This seems to be confirmed by the fact that, in this case, even a neuter verb like *essiri* 'be' entails a dynamic reading, like in (11):

- (11) *U picciriddru è casa-casa*
 DET.M.SG child.M.SG be.3.SG.PRES.IND house.F.SG-house.F.SG
 ‘The child is (moving) all over the house’

On the other side, a genuinely stative verb, in the same context, makes the sentence odd:

- (12)⁴ **U picciriddru rorme casa-casa*
 DET.M.SG child.M.SG be.3.SG.PRES.IND house.F.SG-house.F.SG
 ‘The child is sleeping all over the house’

As far as type C is concerned, when it corresponds to a predicative structure, the verb is necessarily stative or psychological:

- (13) a. *Sta strata è curvi-curvi*
 DEM.F.SG street.F.SG be.3.SG.PRES.IND turn.F.PL-turn.F.PL
 ‘That road is winding’ (lit. full of turns)
 b. *Mi sentu spinguli-spinguli*
 REFL.1.SG feel.1.SG.PRES.IND pin.F.PL-pin.F.PL
 ‘I fell very thrilled’ (lit. I feel pins all over)

In this case, the reduplication has the same function and the same distribution of a qualifying adjective.

4.4 Constraints on the landmark

In type C, there are no specific constraints on the entity that can correspond to the landmark. In fact, as in all the other types of reduplications considered, the only condition is that the landmark corresponds to a physical entity which can delimit a mono- (13a) or a multi-dimensional (15b) space with discrete boundaries. The trajectors (i.e., in this case, several instances of the entity designated by the reduplicated noun) are homogeneously distributed on the surface in question.

- (14) a. *A corda è rruppa-rruppa*
 DET.M.SG rope.F.SG be.3.SG.PRES.IND knot.M.PL-knot.M.PL
 ‘The rope is full of knots’
 b. *A facci è mpuddri-mpuddri*
 DET.M.SG face.F.SG be.3.SG.PRES.IND pimple.F.PL-pimple.F.PL
 ‘The face is pimply’

As far as type A is concerned, on the other hand, the meaning of the entire construction can be slightly modified by some of the features of the landmark, as Figure 1 shows.

If the landmark corresponds to an entity which can be viewed as a mono-dimensional space (along which only a monodirectional movement or distribution is possible), the reduplication is likely to mark a carrier relationship, similar to that expressed by the English prepositions *along* or *on*.

⁴ The sentence in (12) is only acceptable in a context in which (i) the position of the child is considered as incongruous, (something like ‘in the middle of the house’, in contrast to ‘in his bed’), (ii) a child is sleeping somewhere in the house, then he moves and he sleeps somewhere else, then he moves once again and he’s still sleeping, and so on.

- (15) *Fattilla* *ciumi-ciumi* *e* *arrivasti*
 do.2.IMPER=OBJ.F.SG river.M.SG-river.M.SG CONJ arrive.2.SG.PRET.IND
 ‘Follow the river (lit. do it along the river) and you’ll get there’

Note that a structure as the one in example (15) corresponds to the only way available in Sicilian to express a spatial relation which in other Indo-European languages is usually expressed by a preposition (Eng.: *along*, It: *lungo*, etc.). Apart from spaces which are conceived as mono-dimensional, any space is usually interpreted as having different dimensions (at least two, like a sheet of paper or a table) whereon different movements or distributions are possible (along one of the three dimensions). If the landmark is multi-dimensional, then the movement (in the case of an animate trajector) or the distribution (in the case of a mass or a plural trajector) are seen as involving any dimension:

- (16) a. *I* *picciotti* *iocano* *a* *palluni* *strata-strata*
 DET.M.PL boy.M.PL play.3.PL.PRES.IND PREP ball.M.SG street.F.SG-street.F.SG
 ‘Boys are playing football on the street’
 b. *Un* *putiamu* *caminari* *picchi* *c’* *eranu* *assai*
 NEG can.1.PL.IMPF.IND walk.INF because there be.3.PL.IMPF.IND many
petre *strata-strata*
 stone.F.PL street.F.SG-street.F.SG
 ‘We couldn’t walk because there were a lot of stones on the street’

In standard Italian the same semantic nuance, differentiating between a simple localization and a dynamic localization, is expressed, respectively, by the prepositions *a/in* and *per*:

- (17) a. *I* *bambini* *sono* *a/in* *casa*
 DET.M.PL children.M.PL be.3.PL.PRES.IND PREP house.F.SG
 ‘The children are in the house’
 b. *I* *bambini* *sono* *per* *casa*
 DET.M.PL children.M.PL be.3.PL.PRES.IND PREP house.F.SG
 ‘The children are (moving) all over the house’

Concerning type B, as already observed, the landmark can correspond to a plural noun form:

- (18) *U* *nonnu* *passiava* *arvuli-arvuli*
 DET.M.PL grandfather.M.PL stroll.3.SG.IMPF.IND tree.M.PL-tree.M.PL
 ‘Grandfather was strolling among the trees’

In this case, the reading of the construction containing the adverb *arvuli-arvuli* implies a multiple and dynamic localization of the trajector (the grandfather) within the landmark; in its turn, the landmark corresponds to an area containing trees and surrounded by an imaginary outline (cf. Herskovits 1987). Of course, the landmark may also correspond to a space with real boundaries. In this case, the plural indicates a movement taking place successively in several of these spaces, like in (19).

- (19) *U* *dutturi* *firria* *casi-casi*
 DET.M.PL doctor.M.PL go around.3.SG.PRES.IND house.F.PL-house.F.PL
 ‘The doctor goes from an house to another’

5. Concluding remarks

The connection between reduplication and iconicity is often pointed out in the literature, since cross-linguistically these constructions often express plurality, intensity, emphasis, iterativity, etc. Roughly speaking, some of these meanings (e.g. plurality or iterativity) can also be attributed to the Sicilian reduplications we presented in this paper. However, we showed that the range of meanings these reduplications may express are restricted and crucially determined by the syntactic context, the semantics of the elements involved, etc., i.e. by the construction which is responsible for their formation. This observation led us to consider reduplication as a phenomenon proper to the human speech ability which can eventually grammaticalize and give birth to constructions, such as the ones we presented, which are formally and semantically constrained. This grammaticalization process can result in a derivational phenomenon (like the one we investigated here) or an inflectional one (like for instance, in plurals expressed by reduplication in various languages). What we claimed is that TR can be considered a special type of compounding and that, like for compounding (cf. Bauer 2009; Arcodia and Montermini 2012), it is possible to draw a distinction between two kinds of reduplication: R_1 , which is the manifestation of the (potentially universal) ability of humans to reduplicate a linguistic object (e.g. a word) in a fairly iconic way (cf. (2) above) and R_2 , which corresponds to a phenomenon codified in the grammar of individual languages under the form of a construction for which formal, categorial and semantic constraints are specified.

References

- Amenta, L. (2010) La reduplicazione sintattica in siciliano. *Bollettino del CSFLS* 22: 345-358.
- Arcodia, G. F. & F. Montermini (2012) Are reduced compounds compounds? Morphological and prosodic properties of reduced compounds. In: P. Arnaud, F. Maniez & V. Renner (Eds.), *Cross-disciplinary perspectives on lexical blending*. Berlin: Mouton de Gruyter, 93-113.
- Bauer, L. (2003) *Introducing linguistic morphology*. 2nd ed. Washington D.C.: Georgetown University Press.
- Bauer, L. (2009) Typology in compounds. In: R. Lieber & P. Štekauer (Eds.), *The Oxford handbook of compounding*, Oxford: Oxford University Press, 343-356.
- Booij, G. (1992) Compounding in Dutch. *Rivista di Linguistica* 4: 37-59.
- Booij, G. (2010) *Construction Morphology*. Oxford: Oxford University Press.
- Caracausi, G. (1977) Ancora sul tipo 'camminare riva riva'. *Bollettino del CSFLS* 13: 383-396.
- Don J. (2009) IE, Germanic: Dutch. In: R. Lieber & P. Štekauer (Eds.), *The Oxford handbook of compounding*, Oxford: Oxford University Press, 370-385.
- Dressler W. U. (2006) Compound types. In: G. Libben & G. Jarema (Eds.), *The representation and processing of compound words*. Oxford: Oxford University Press, 23-44.
- Dressler, Wolfgang U. & Lavinia Merlini Barbaresi. 1994. Morphopragmatics: Diminutives and intensifiers in Italian, German, and other languages. Berlin & New York: Mouton de Gruyter.
- Emmi, T. (2011) *La formazione delle parole nel siciliano*. Palermo: CSFLS.
- Forza, F. (2011) Doubling as a sign of morphology: A typological perspective. *Journal of Universal Language* 12(2): 7-44.
- Ghomeshi, J., R. Jackendoff, N. Rosen & K. Russell (2004) Contrastive focus reduplication in English (the salad-salad paper). *Natural Language & Linguistic Theory* 22: 307-357.
- Gil, D. (2005) From repetition to reduplication in Riau Indonesian. In: B. Hurch (Ed.), *Studies on reduplication*. Oxford: Oxford University Press, 31-64.
- Guevara, E. & S. Scalise (2009) Searching for universals in compounding. In: S. Scalise, A. Bisetto & E. Magni (Eds.), *Universals of language today*. Amsterdam: Springer, 101-128.
- Herskovits, A. (1987) *Language and spatial cognition*. Cambridge: Cambridge University Press.
- Langacker, R. W. (1987) *Foundations of cognitive grammar: Theoretical prerequisites*. Vol. 1. Stanford: Stanford University Press.
- Leone, A. (1995) *Profilo di sintassi siciliana*. Palermo: CSFLS.

- Libben G., M. Boniecki, M. Martha, K. Mittermann, K. Korecky-Kroll & W. U. Dressler (2009) Interfixation in German compounds: What factors govern acceptability judgements? *Italian Journal of Linguistics / Rivista di linguistica* 21(1): 149-180.
- Masini, F. & A. M. Thornton (2009) Italian VeV lexical constructions. In: A. Ralli, G. Booij, S. Scalise & A. Karasimos (Eds.), *On-line Proceedings of the 6th Mediterranean Morphology Meeting*, 146-186.
- Montermini F. (2010) Units in compounding. In: S. Scalise & I. Vogel (Eds.), *Cross-disciplinary issues in compounding*. Amsterdam: John Benjamins, 79-92.
- Moravcsik, E. A. (1978) Reduplicative constructions. In: J. H. Greenberg (Ed.) *Universals of human language*. Vol. 3: *Word structure*. Stanford: Stanford University Press, 297-334.
- Ralli A. (2013) *Compounding in Modern Greek*. Dordrecht: Springer.
- Sgarioto, L. (2005) 'Caminari riva riva': su un fenomeno di reduplicazione nominale in siciliano. *Quaderni di Lavoro dell'ASIS* 5: 36-49.
- Stolz, T. (2009) Total reduplication: syndetic vs. asyndetic patterns in Europe. *Grazer Linguistische Studien* 17: 99-113.
- Stolz, T., C. Stroh & A. Urdze (2011) *Total reduplication: the areal linguistics of a potential universal*. Berlin: Akademie Verlag.
- Talmy, L. (1983) How language structures space. In: H. Pick & L. Acredolo (Eds.), *Spatial orientation: Theory, research, and application*. New York: Plenum Press, 225-282.
- Todaro, G., F. Villoing & P. Gréa (2014) Internal localisation NN_{adv} reduplication in Sicilian. In: S. Augendre, G. Couasnon-Torlois, D. Lebon, C. Michard, G. Boyé & F. Montermini (Eds.), *Actes des Décembrettes 8*, [Carnets de Grammaire 22]. Toulouse: CLLE-ERSS, 168-188.
- Wierzbicka, A. 1986. Italian reduplication: Cross cultural pragmatics and illocutionary semantics. *Linguistics* 24: 287-315.