Proceedings of the Thirty-first

Western Conference on Linguistics

Volume twenty-five

WECOL 2019

Edited by

Sho Akamine

Hosted by:

California State University, Fresno

November 15 – 17, 2019

Department of Linguistics

California State University, Fresno

Contents

A Morphosyntactic Analysis of Ostensible Lexical Categories in Arabic

Saleem Abdelhady & Phil Branigan
Memorial University of Newfoundland

1 Introduction

Standard Arabic has 30 lexical items that are exceptionally categorized as nouns-verbs-interjections, referred to as *ʾism ʾalfiʿl* '*Names of Verbs'* (NoVs). NoVs show patterns that hinder their categorization; they display nominal features, code epistemic stances of their speakers, and select arguments. These properties, among others, make NoVs look like verbs, nouns, and interjections without entirely behaving like any fixed category.

(1) *hak ʔal-kɪtab-a*
take.OC:2SG DEF-book-ACC
'take the book.'

(2) *ʔaax-ɪn mɪn-ka.*
hurt.OC:1SG-NUN from-you
'I am in pain because of you.'

(3) *hayhaat ʔal-ʕawd-u ʔɪla ʔað-ðul-i.*
impossible.OC.PL DEF-back-NOM to DEF-disgrace-GEN
'Going back to disgrace is impossible.'

(4) *hayyaa ʔɪla ʔal-ʕamal.*
go.OC.2SG to DEF-work
'Let us go to work.'

In this paper, we show how head-movement to a Speech Act Projection (SAP) (Haegeman & Hill, 2013; Hill, 2007, 2013) produces such ostensible categories[1]. NoVs are associated with the type of derivation, which enables such movement. The paper is organized as follows. Section 2 provides a theoretical background. This section aims to present verbs, nouns, and interjections in Arabic. The section also presents a brief review of those categories cross-linguistically. Section 3 analyzes the data and presents allocutivity in Arabic as part of its agreement system. Section 4 concludes the study.

2 Background

Understanding syntactic structural behavior begins with defining syntactic categories (Baker, 2003; Welch, 2016; Wiltschko, 2014). Earlier treatments of categorization have set several principles and features to contrast categories and arrive at their appropriate classification. For instance, according to Chomsky (1970), cited in Baker (2003, p. 3), lexical categories are distinguished by binary features that target two universal lexical items: nouns and verbs. Given the operating binary feature possibilities upon two categories, Chomsky (ibid) illustrates how other categories are classified. If a category is valued as [+ noun] and [- verb], then the lexical item is classified as a noun. Reversing these values results in a verb. Adjectives are valued as [+ nouns] and [+ verbs]. For adpositions, prepositions, and postpositions, the values are set as [- noun] and [- verb].

[1] The term is used following Abdel Hady (2015) and Abdelhady (2013, 2020a).

Proceedings of *WECOL 2019*

Similarly, traditional Arab grammarians have depended on the contrastive classification to distinguish the categories of their language. Nevertheless, their classification system operates upon three major contrastive categories: nouns, verbs, and particles. Elements are categorized through passing a cyclic try-and-error matching of words with predefined features of the three categories. The first cycle begins with testing a word against those features of nouns. If its features do not match those of a noun, then the word is passed to the second cycle. In the second cycle, the word is tested against predefined features for a verb. If the matching is not possible, then the word is defined as a particle (Owens, 1989; Sibawayh, 1970).

In principle, nouns are described as those elements that can have a genitive case, are associated with the definite article, accept nunation, appear in diminutive forms, are used as subjects and objects and can be used with vocative particles (Al-Bataineh, 2019). Verbs are defined as elements that can accept the feminine and energetic suffixes, modal particles (e.g., *qad* and *sawfa*), and negative particles such as *lam* 'not,' show tense, and can have arguments. The following example illustrates predefined features of nouns.

Table 1 Nominal Properties in Standard Arabic

Input	Features	Output
kitāb 'book'	Genitive Case	*fi ʔal-kitāb-i* in DEF-book-GEN
	Definite Article[2]	*ʔal-kitāb* DEF-book
	Nunation	*kitāb-un* book-NUN
	Diminutive Form	*kutajib* booklet.DIM
	Object	*qaraʔ-tu ʔal-kitāb-a* read-I DEF-book-ACC
	Subject	*ʔal-kitāb-u ʕala ʔal-kursi* DEF-book-NOM on DEF-chair

The example shows that the input *kitāb* 'book' fits almost all pre-set features of nouns. It also contrasts with those features that are predefined for verbs; it cannot be used with the perfective particle *qad,* and it cannot interact with tense. Accordingly, the classification system defines the word as a noun. In a similar way, verbs are defined. To illustrate this, consider the following example.

Table 2 Verbal Properties in Standard Arabic

Input	Features	Output
katab 'write'	Feminine suffixes	*kataba-t* wrote-she
	Energetic suffix	*ʔaktuba-na* write-EMPH
	Perfective particle	*qad ktab* PERF wrote
	Negative particle	*lam ja-ktub* not PRES-write
	Agents	*kataba ʔal-walad-u ʔad-dars-a* wrote DEF-boy-NOM DEF-lesson-ACC
	Tense	*kataba - ja-ktub - s-jaktub* wrote.PAST - PRES-write - FUT-write

[2] See Abdelhady (2019) for the exponents of the definite article across different varieties of Arabic

As the input *katab* 'write' fits most pre-set features of verbs and contrasts with those of nouns, it is defined as a verb. Inputs that do not meet the specifications of nouns and verbs (i.e., what is left over) are categorized as particles. For example, *qad* is a particle as it does not take prefixes or suffixes that are associated with nouns or verbs.

Both the binary system (Baker, 2003) and the contrastive cyclic system (Owens, 1989) are problematic; they are not precise and leave a number of categories undefined. The imprecision of such systems appears when exposed to OCs. Therefore, the systems have been subject to revisions to stop at a more accurate and universal description of categories (Baker, 2003; Owens, 1989). According to Carnie (2011, p. 1027), the set of definitions of lexical categories that Baker (2003) proposes is not inclusive. Nevertheless, the proposed revisions have not been able to capture precisely all categories across languages.

Studies have investigated the nature of OLCs across several languages: Celtic languages (Abdelhady, 2020b; Roberts Borsley et al., 2007; Carnie, 2011, 2006; Li, 2004; Willis, 1988), Algonquian languages (Al-Bataineh, 2020a,b; Wiltschko, 2014) and others. This section introduces these categories and highlights their problematic nature for the previously mentioned categorization systems. Traditional attempts at categorization of NoVs accentuate their verbal character (Owens, 1989) or nominal properties(Sibawayh, 1970), while more recent efforts list them as interjections (Eisenstein, 2011). But this literature is inconsistent with recent conclusions about the nature of verbs, nouns (Baker, 2003), and interjections (Hill 2013). For example, in Arabic, it is ungrammatical to mark a verb with addressee (c.f., **xuð**k* 'take:2SG'). But such morphological agreement is acceptable in (1). Additionally, while regular lexical verbs appear in both SVO and VSO patterns, constructions created by NoVs accept only affirmative VSO order (cf., (3) & (6)). NoVs are also incompatible with negation (5). And NoVs impose restrictions on peripheral movements (Rizzi, 1997); their arguments cannot be topicalized or put in focus. Finally, NoVs do not allow adverbial modification.

(5) * *la t-hak ʔal-kitab-a.*
NEG.PRT t-take.OC:2SG DEF-book-ACC
Intended: 'Do not take the book.'

(6) **ʔal-ʕawd-u hayhaat ʔıla ʔað-ðul-i.*
DEF-back-NOM impossible.OC.PL to DEF-disgrace-GEN
'Going back to disgrace is impossible.'

NoVs display some properties of nouns. Like construct state (7), NoVs select only definite nominals as their arguments (cf. (1) with (8)). And NoVs can be suffixed by nunation (2). Their nominal character is also seen in pluralia tantum forms (3), which are reliable indicators of feature valuation for number in nominal heads (Pesetsky and Torrego 2007). But unlike nouns, NoVs do not select nominals with genitive case (cf., (3) with (7)). Instead, their arguments are spelled out with nominative or accusative case, depending on their grammatical role.

(7) *dar ʔal-radʒul-i.*
house DEF-man-GEN
'the house of the man.'

(8) **hak kitab-an.*
take.OC:2SG book-nun
'take a book.'

NoV data also stands in sharp contrast with interjections. Cross-linguistically, interjections do not assign case (e.g., 1-4), nor can they have arguments (1-4). According to Hill (2013), among others, interjections are optional; nevertheless, the bolded items are required for the grammaticality of the examples above.
NoVs display other characteristics features that set them apart from other categories. Conjunction with an NoV phrase creates *unbalanced conjuncts*, analogous to asymmetrical embedded conjuncts in Dutch (Hoekstra, 2009) (c.f.,(10) with (13)); unlike phrasal conjuncts with verbs (11-12), NoVs phrases must always appear as the left-most part of a phrasal conjunct (c.f., (9) and (10)).

(9) * *ʔiftari datfar-an w dunka ʔal-waraq.*
buy notebook-NUN and take.OC:2SG DEF-paper
'buy a notebook and take the paper.'

(10) dunka ʔal-qalam w ʔuktub …
take.OC:2SG DEF-pen and write
'take the pen and write …

(11) *xuð ʔal-qalam w ʔiqraʔ ʔad-dars.*
take DEF-pen and read DEF-lesson
'Take the pen and read the first lesson.'

(12) *ʔiqraʔ ʔad-dars w xuð ʔal- qalam.*
read DEF-lesson and take DEF-pen
'Read the first lesson and take the pen.'

(13) het irriteert ons dat je te laat thuiskomt en
it irritates us that you too late home-come and
je hebt geen sleutel bij je.
you have no key withyou
'It annoys us that you come home late and you do not have a key with you.'
(Dutch)

3 Analysis

The data requires that that NoVs originate in a position appropriate for verbal roots, but that they move to a higher position where a suppletive form is licensed. In their original position, they ensure verb-like case assignment and selection properties. In the derived position, they adjoin to higher expressive discoursal functional heads, specified by speech act participants - speakers and addressees. The high landing site accounts for word order, and their higher participant specifiers account for the "agreement beyond phi" effect [φ](Miyagawa, 2017). Moreover, external specifiers block movement of internal ones, ensuring the impossibility of topicalisation, etc. with NoV structures. There are two possibilities of this prediction. The first possibility is to assume (following the literature (Kaur, 2018, among others) that *-ka* originates in C with unvalued second-person feature and that it enters in an upward probe-goal relation with *sa* to value its unvalued second-person feature. The second possibility is to assume that *-ka* originates in *sa* as a result of spelling out the valued second person feature. Consider again the following data:

(14) *haaka al-kitaab-a*.[3]
PAR:2M.SG DEF-book-ACC.
'Take the book.'

(15) *dūnka al-qalam-a*
PAR:2M.SG def-pen-ACC
'take the pen.'

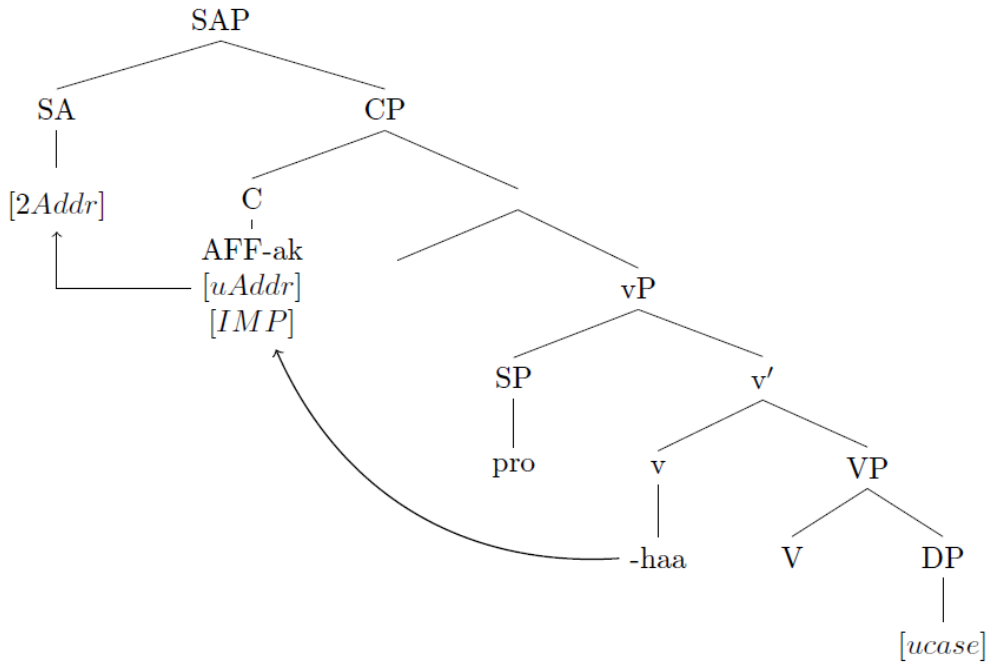
Syntactically speaking, I argue that the particle, *haa*, is verbal in nature; it originates in the V position; it acts as a verb and may end up like verbs. In addition, the particle moves to a position that enables its agreement with the speech act zone, namely agreement with the addressee; at that level, the particle behaves differently. In that position, the particle merges with the allocutive marker that does not have a thematic role;

[3] Note that most imperative NoVs are composed of either prepositions or particles.

that is, the particle attaches to the second person marker. One may ask, then, what is the trigger for the movement to a head that would enable the particle to get that association? The answer to this question is the affixal nature of the allocutive marker and feature valuation. The allocutive marker is a bound morpheme; that is, it cannot stand by itself. Let us assume that the hearer head has a valued second person feature (Hill, 2007a). The allocutive marker attracts the verbal particle to move from its original position to attach to it. Since affixes generally only attach to units of language of specific nature, one application of this process is to assume that the allocutive affix only attaches to prepositions or particles to form a NoV. As proposed by Kaur (2018), We assume that T is defective in imperative constructions[4]; that is, it is not actively participating in the construction of imperatives. After presenting the claim that the motivation for moving the verbal particle to a higher position is the affixal nature of the allocutive marker, we present the role of feature valuation in the process.

We assume that the allocutive marker originates in the C position[5]; the C position has unvalued second person feature and a valued imperative feature (Kaur, 2018, among others); as an active probe the allocutive marker enters in an upward agree mechanism with the addressee head that has a valued second person feature. The marker agrees with the addressee, and because of its affixal nature, it attracts the verbal particle; they form one unit. The outcome is a verbal imperative NoV that is associated with a non-agentive second-person marker.

(16)



The prediction provides evidence for the role of allocutive markers in imperative interpretations. Kaur (2018) shows that Punjabi, an Indo-Aryan language, deploys allocutive markers, such as *-je*, with declarative verbs for endowing such verbs with imperative force.

(17) *(tuu/tussi)* *bacce-nuu* *vekh-Ø/vekh-o*.
(2.SG/2.PL) child-ACC see.IMP-2.SG/see.IMP-2.PL
'Look after/see the child.'

[4] We refrain from presenting T in the following trees; yet, we assume that T has a default interpretation of present.

[5] A fully competitive analysis is to assume that the allocutive marker originates in SA (as an articulation of the interpretable second person feature) and that C has a strong imperative affix; in this scenario, the impetrative affix attracts *haa* to C and then the allocutive marker attracts it again to SA; both views give a similar outcome; nevertheless, we might argue for the latter on the basis that these constructions block topicalization and focus; for now, we follow the literature in establishing the idea that allocutivity is part of imperative NoVs for the clarity of the argument. In the next section, we present effect of the latter on expressive NoVs.

(18) *(tussi)* *bacce-nuu* *vekhyaa-je.*
(2.PL) child-ACC see.PERF.M.SG-ALLOC
'Look after/see the child!' (Kaur, 2018)

Kaur (2018) shows that allocutivity is the primary source for the imperative interpretation in (18). We also conclude that what endows particles with imperative force is their association with allocutive markers. The necessity for syntactic movement is seen directly in the unbalanced conjunctions. Movement of NoV to the SAP head is possible only from within the left conjunct, a pattern seen in Arabic and other languages (Aoun et al., 1994) where an external probe can interact with the leftmost conjunct without engaging with the contents of the one of the right.

The analysis gives straight forward answers to several puzzling questions that have been raised regarding the behavior of NoVs. Why do NoVs reject negation? The answer is that NoVs are higher than the scope of negation. In syntax, NEG c-commands VPs. Since NoVs are only active at SA, then they cannot be a target for negation. The same answer is given to the second question. Why do NoVs reject interrogatives? Since NoVs show an association with non-thematic specifiers, it is not valid to ask questions about who performed them. However, it is entirely legitimate to ask a question that targets their internal VPs. Thus, native speakers can relate and have immediate access to their internal structure. Yet, once they are related to the speech act zone, the zone prevents agentive questions and transfers such questions to their internal structure. A piece of supporting evidence comes from the fact that not only using agentive marking suffixes but also using agentive prefixes along with NoVs is not allowed. A long puzzling question for Arab grammarians has been why a verb can come before NoVs but not vice versa. Given the analysis above, verbs might be able to select expressive phrases. However, spelling out a verb within the same syntactic internal structure of NoVs does not activate the expressive feature of NoVs. Thus, NoVs cannot co-occur with a verb in the same SAPs phrases. Their expressive phrases, nevertheless, can be selected by verbs (cf. verbs of saying such as *ʔaqul* 'say'). NoVs have some expressive power that can be found in interjections.

4 Conclusion

The study concludes with three observations. First, allocutive markers can disguise the interpretation of categories. Second, the imperative interpretation of NoVs is related to allocutivity. Third, NoVs are context-sensitive. The analysis has implications for understanding other cases of ostensible categories whose functions cannot be determined from their surface forms. These range from mixed feature cases (nouns-verbs) to unexpected landing forms associated with unexpected landing sites, such as verbs topicalization structures (Robert Borsley & Roberts, 2006; Carnie, 2006; Li, 2004). The proposed model also advances our understanding of (non)imperative structures; it establishes the significance of speech act specifiers on agreement not only with addressees (1) (e.g., Alcazar & Saltarelli, 2014; Kaur, 2018) but also with speakers (2).

References

Abdel Hady, Saleem. (2015). The Pragmatic Functions of the Ostensible Communicative Act of Invitation in Jordanian Arabic. *SKY Journal of Linguistics*, 1–29.

Abdelhady, Saleem. (2013). *Ostensible Invitations in Jordanian Arabic: A Sociopragmatic Study* [M.A.]. Yarmouk University.

Abdelhady, Saleem. (2019). A Sociolinguistic Study of the Use of the Definite Article (ʔal) in the Jordan Valley. *SKY Journal of Linguistics*, *32*, 7–41.

Abdelhady, Saleem. (2020a). *The Syntax of Ostensible Categories in Arabic* [Ph.D.]. Memorial University of Newfoundland.

Abdelhady, Saleem. (2020b). The Syntax of Sudden Particles in Arabic. *ResearchGate*. https://doi.org/10.13140/RG.2.2.10151.91044

Al-Bataineh, Hussein. (2019). "The Syntax of Vocatives in Arabic." Brill's Journal of Afroasiatic Languages and Linguistics.

Al-Bataineh, Hussein. (2020). "The Nominality of Tłı̨chǫ Classificatory Verb Stems and the Simplicity of Dene Verbal Morphology."*Morphology*.

Al-Bataineh, Hussein. (2020). "Stem Alternation in Tłı̨chǫ Yatıì Classificatory Verbs: A Cognitive Semantic Account."*Folia Linguistica*.

Alcazar, Asier, & Saltarelli, Mario. (2014). *The Syntax of Imperatives*. Cambridge University Press. https://doi.org/10.1017/CBO9780511794391

Aoun, Joseph, Benmamoun, Elabbas., & Sportiche, Dominique. (1994). Agreement, Word Order, and Conjunction in Some Varieties of Arabic. *Linguistic Inquiry*, *25*(2), 195–220. JSTOR.

Baker, Mark (2003). *Lexical Categories: Verbs, Nouns, and Adjectives*. Cambridge University Press.

Borsley, Robert, & Roberts, Ian. (2006). *The Syntax of the Celtic Languages: A Comparative Perspective*.

Borsley, Roberts, Tallerman, Maggie., & Willis, David. (2007). *The Syntax of Welsh*. Cambridge University Press.

Carnie, Andrew. (2011). Mixed categories in Irish. *Lingua*, *121*(7), 1207–1224. https://doi.org/10.1016/j.lingua.2011.01.013

Carnie, Andrew. (2006). Phasing in Mixed Categories: Evidence From Irish. *Syracuse/Cornell Workshop on the Internal Syntax of Nominalized Clauses*, 22.

Eisenstein, Herbert. (2011). Interjections. In Lutz Edzard & Rudolf de Jong (Eds.), *Encyclopedia of Arabic Language and Linguistics* (pp. 378–380). BRILL.

Haegeman, Lilian., & Hill, Virginia. (2013). The syntacticization of discourse. *Syntax and Its Limits*, 370–390.

Hill, Virginia. (2007). Vocatives and the pragmatics–syntax interface. *Lingua*, *117*(12), 2077–2105. https://doi.org/10.1016/j.lingua.2007.01.002

Hill, Virginia. (2013). *Vocatives: How Syntax meets with Pragmatics*. Brill. https://brill.com/view/title/20550

Hoekstra, Eric. (2009). Expletive replacement, verb-second and coordination. *The Linguistic Review*, *11*(3–4), 285–298. https://doi.org/10.1515/tlir.1994.11.3-4.285

Kaur, Gurmet. (2018). *Addresee Agreement as the Locus of Imperative Syntax*. North East Linguistics Society 2018, Cornell University.

Li, Chao. (2004). On Verbal Nouns in Celtic Languages. *Proceedings of the Harvard Celtic Colloquium*, *24/25*, 163–194. JSTOR.

Miyagawa, Shigeru. (2017). *Agreement Beyond Phi*. MIT Press.

Owens, Jonathan. (1989). The Syntactic Basis of Arabic Word Classification. *Arabica*, *36*(2), 211–234. JSTOR.

Pesetsky, David., & Torrego, Esther. (2007). The syntax of valuation and the interpretability of features. In Simin Karimi, Vida Samiian, & Wilkins Wedny (Eds.), *Linguistik Aktuell/Linguistics Today* (Vol. 101, pp. 262–294). John Benjamins Publishing Company. https://doi.org/10.1075/la.101.14pes

Rizzi, Luigi. (1997). The fine structure of the left periphery. In *Elements of grammar* (pp. 281–337). Springer.

Sibawayh, ʔibn ʕutman. (1970). *Al-Kitab* (H. Derenbourg, Ed.). Reprint Hildesheim and New York: Georg Olms.

Welch, Nicholas. (2016). Propping up predicates: Adjectival predication in Tłı̨chǫ Yatìì. *Glossa: A Journal of General Linguistics*, *1*(1), 2. https://doi.org/10.5334/gjgl.7

Willis, Penny. (1988). Is the Welsh verbal noun a verb or a noun? WORD, *39*(3), 201–224. https://doi.org/10.1080/00437956.1988.11435790

Wiltschko, Martina. (2014). *The universal structure of categories: Towards a formal typology* (Vol. 142). Cambridge University Press.

A Comparison of Length Between Geminates and Singletons in Crow

Constantino Damian
California State University, Fresno

1 Introduction

Crow has sets of singletons and geminates represented in the orthography of the language. My aim in this paper is to highlight the different lengths between singleton and geminate obstruents in the language as represented by the orthography. As such, I'll be using the orthographic spelling to represent the sounds in question.

Concerning these sounds, the orthography is similar to IPA, and the geminates are the singletons written twice, with the exceptions being [ʃ] and [ʃː], written 'sh' and 'ssh' respectively, and [ʧ] and [ʧː], written 'ch' and 'tch' respectively (Graczyk 2007). Wallace (1993) also mentions the consonant clusters 'hp,' 'ht' and 'hk' as usually sounding like unaspirated geminates, so I look at those here as well. For ease of distinction between these three geminates and the other geminates, I will simply refer to these three as unaspirated geminates, and the 'pp,' 'tt' and 'kk' stops and aspirated geminates.

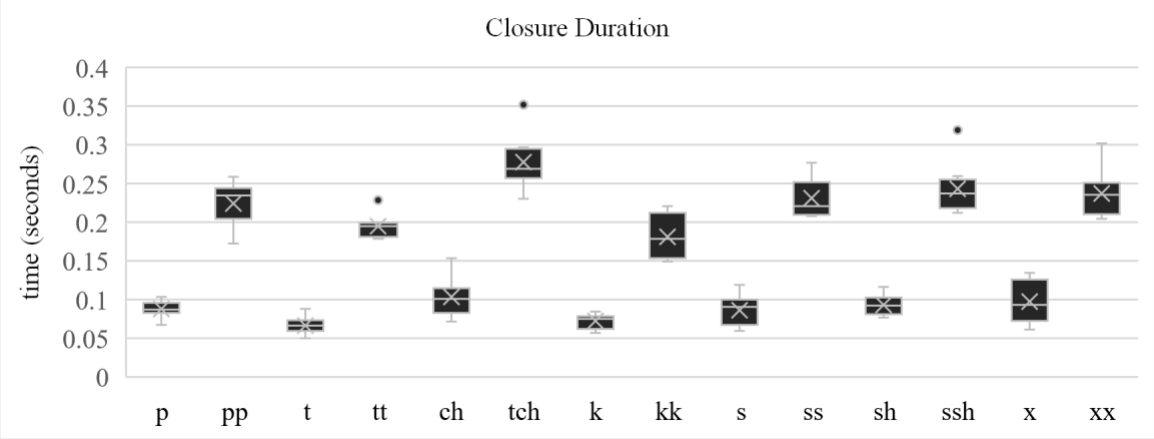
2 Methodology

The recordings I used were of words spoken in isolation by a single Crow speaker. In order to avoid the effects of final lengthening, as well as to minimize any ambiguity as to where an obstruent may begin or end, none of the measurements were taken from word-initial or word-final position. I am not including aspiration or release in the measurements provided, or as a part of the current study, instead comparing only closure duration. Eight samples for each singleton and geminate were measured.

3 Data and Analysis

3.1 *Geminate-singleton comparison* The results for the geminates alongside their singleton counterparts are shown in Figure 1.

Figure 1. Obstruent Closure Duration Ranges



* Thanks to John Boyle, Chris Golston, John Simonian and Zachary Metzler for the material and guidance.

Proceedings of *WECOL 2019*

None of the singletons experienced overlap with their geminate counterparts. The singletons mostly reside below the 150 ms line, and the geminates above it. There is, however, a small overlap between singleton 'ch,' which rises slightly above that line, and geminate 'kk,' which dips slightly below it. The averages and ratios for the two sets of obstruents are shown in Table 1.

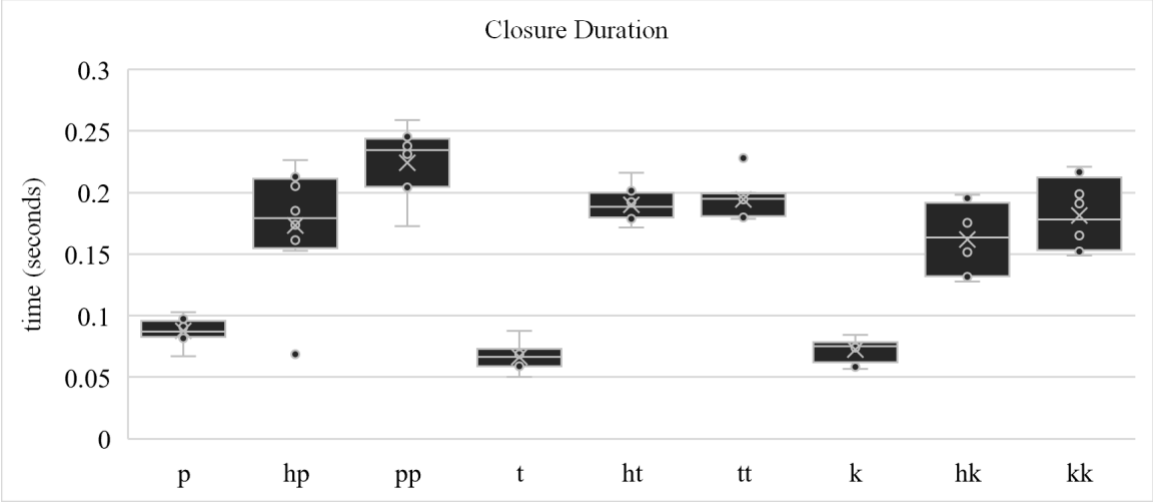
Obstruent	*Long (ms)*	*Short (ms)*	*Ratio (Long:Short)*
pp/p	224	87	2.6:1
tt/t	194	67	2.9:1
kk/k	181	72	2.5:1
tch/ch	278	103	2.7:1
ss/s	231	86	2.7:1
ssh/sh	243	93	2.6:1
xx/x	237	98	2.4:1

Table 1. Geminate-Singleton Average and Ratio

The results in Table 1 show the geminates in Crow to be more than twice the length of their singleton counterparts, with the alveolar stop geminate showing the most extreme difference at almost three times the length. The other geminates show a difference of around 2.5 times the length of their singleton counterparts. The relatively extreme difference shown by the alveolar stops appears to be due mostly to the short length of the singleton. The alveolar stop singleton has the smallest average length among the obstruents, while its geminate counterpart was far from the lengthiest, being the second shortest in duration among the geminates.

3.2 *Unaspirated Geminates* Results for the unaspirated geminate stops can be seen in Figure 2 and Tables 2 and 3 below.

Figure 2. Stop Duration Ranges



Stop	***p value***
hp:pp	.023958
ht:tt	.576898
hk:kk	.198651

Table 2. Geminate Comparison t-test Results

Obstruent	***Long (ms)***	***Short (ms)***	***Ratio (Long:Short)***
hp/p	173	87	2:1
ht/t	190	67	2.9:1
hk/k	162	72	2.2:1

Table 3. Unaspirated Geminate-Singleton Average and Ratio

Two-tailed, equal variance t-tests on the two sets of geminate stops showed an insignificant difference between the two in terms of closure duration in Table 2, with the possible exception of the bilabial stop geminates. Comparing the unaspirated geminate stops with the singletons in Table 3, the results show smaller ratios than in Table 1. The alveolar stop's ratio is about the same in both cases. The velar stops' ratio is slightly smaller in this case than in Table 1. The bilabial stops' ratio is a 2:1 difference here, as opposed to 2.6:1 in Table 1.

4 Conclusion

The results showed there to be a significant difference between the lengths of singletons and their corresponding geminates in Crow. The geminates are around 2.5 times longer than their singleton counterparts, with the unaspirated geminates dropping that number slightly, and the alveolar stop geminates being nearly three times longer than the singleton in both cases.

References

Graczyk, Randolph. 2007. A grammar of Crow. Lincoln, NE: University of Nebraska Press.

Wallace, Karen K. 1993. *Verb incorporation and agreement in Crow.* Ph.D. Diss., University of California, Los Angeles.

Modern Eastern Armenian: SOV or SVO?

Pegah Faghiri and Pollet Samvelian
Université Sorbonne Nouvelle

1 Introduction

The question of whether the 'unmarked' or 'canonical' word order in Modern Eastern Armenian (MEA)[1] is (S)OV or (S)VO is a matter of controversy. MEA can be considered a flexible language with respect to the ordering possibilities of the three major constituents S, O, and, V, at the clausal level. This means that all permutations, which give rise to six possible orders SOV, SVO, VSO, VOS, OVS, and OSV, are grammatical. Among these possible orders, two occur more frequently, SOV and SVO. Given that S occurs in the initial position in both cases, the controversy involves the relative order between O and V: Is MEA an OV or VO language?

With very few exceptions (Dum-Tragut, 2009, Dryer 2013), typological and descriptive studies have generally grouped Armenian with SOV languages (Der-Houssikian, 1978:227-8; Dryer, 1998:286, 310; Dum-Tragut, 2002; Hawkins, 1979:625; Hawkins 1983:286, Kozintseva, 1995: 8; Minassian, 1980: 263; among others[2]). The same holds for theoretical syntax studies of Armenian, mainly within the generative framework, which consider MEA a head-final language (Hodgson, 2013:6; Giorgi and Haroutyunian, 2016: 190; Kahnemuyipour and Megerdoomian, 2011, 2017:81; Tamrazian, 1991: 101; Tamrazian, 1994:7): SOV is thus the 'basic' order, while SVO is 'derived' via movement (extraposition) to the postverbal position. There are also studies that claim that MEA has no dominant word order and is thus nonspecified between OV and VO (see for instance Dum-Tragut, 2009, Dryer 2013 in WALS). More interestingly, however, many Armenian grammars and handbooks on MEA syntax consider SVO to be the canonical or 'recommended' word order in Armenian (Abrahamyan et al., 1975; Arakelyan, 1958; Badikyan, 1976; Papoyan & Badikyan, 2003), indicating that the SOV order is usually used for either narrow focus marking or bare objects.

The claim on Armenian being an SOV language relies on two sets of arguments: 1) the typology of head-direction across various constituents; 2) orderings preferences at the clausal level, especially the position of the focus (section 2). The problem with generalizations on word order preferences at the clausal level is that they are not supported by any quantitative studies. However, it is commonly assumed that identifying the 'unmarked' or 'canonical' word order in a given language is generally a matter of frequency (Greenberg 1966, Lambrecht 1996, among others) and hence needs to be investigated via quantitative methods.

The present study constitutes a first step in this direction. We carried out a preliminary corpus study to investigate the distribution of different possible word orders in MEA transitive sentences, using the Eastern Armenian National Corpus (section 3). The results of this preliminary investigation question the assumption on the markedness of the VO order (section 4).

* We are grateful to Victoria Khurshudyan for her precious help. This work is partly supported by a public grant funded by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program (reference: ANR-10-LABX-0083).

[1] Armenian constitutes an independent branch of the Indo-European language family and has two standards, Modern Eastern Armenian (hitherto MEA) and Modern Western Armenian. This study focuses on MEA, the language spoken in contemporary Republic of Armenia and the Armenian communities of Iran and the ex-Soviet republics. The label Armenian is used when no distinction is made with respect to these two varieties.

[2] Note that some of these studies address the question of word order in Armenian without distinguishing between the Eastern and the Western varieties. In the absence of any large-scale quantitative investigation, we cannot claim that these two varieties diverge (or not) with respect to word order at the clausal level, especially that both are flexible and display both OV and VO orders.

Proceedings of *WECOL 2019*

2 Arguments in favor of MEA being (S)OV

The dominant view of MEA as an (S)OV language is based on two sets of arguments:
1. The direction of branching across constituents;
2. Ordering preferences and constraints, especially the position of the focus, at the clausal and/or VP level.

With respect to word order correlations (or harmonies) highlighted in typological literature (Greenberg 1966, Dryer 1992), Armenian (Eastern and Western) is a fairly consistently left-branching language (Donabedian 2010), which is typical of OV or head-final languages:

- Adpositions are mainly prepositions.
- Modifier adjectives precede the head noun within the NP.
- The genitive complement precedes the head noun within the NP.
- The standard of comparison precedes the parameter of comparison (adjective or adverb) in comparison constructions.
- The lexical verb precedes the auxiliary.
- The predicate precedes the copula.

The only exception with this respect is the position of the relative clause, which follows the head noun. Apart from the consistency of the branching direction across various types of constituents, other arguments can be put forward in favor of basic OV order. Some of these arguments have been mentioned in studies on the syntax of MEA, mainly within the generative framework. Their relevance depends on whether or not one adheres to a specific syntactic representation (configuration). Before going through these arguments, let us clarify that none of the studies cited in this section presents all of them. Actually, some of the arguments have never been evoked in any of the studies on MEA. We have nevertheless included them since they generally appear among arguments which serve to sort out between OV and VO basic orders in various languages (see for instance Asatiani and Skopeteas, 2012). Main (syntactic) arguments in favor of MEA's being OV are the following:

2.1 *The preverbal position of the primary focus* The preverbal position of the primary focus in MEA has generally been a crucial piece of evidence in favor of its classification as an OV language. The correlation between the preverbal focus phenomenon and the OV order is well documented in the literature since the early 70s (cf. Dezső, 1974; Kim, 1988, Herring and & Paolillo, 1995, for a general discussion), where it has been claimed that OV languages are more likely to encode focus preverbally than VO languages.
Comrie (1988: 268) claims that "the basic rule in Modern Eastern Armenian is that the focus must immediately precede the finite verb form"[3]. This claim has generally been maintained in subsequent works (Tamrazian, 1994; Tragut, 2009; Kahnemuyipour and Megerdoomian, 2011, 2017; Giorgi and Haroutyunian, 2016). The main arguments in favor of the preverbal position of the primary focus are:

2.1.1 *The position of (WH-) interrogative pronouns* Interrogative pronouns always occur in the preverbal domain, most commonly immediately before the verb:

(1) ov ēr marina zelyonkina-n
Who COP.PST.3SG Marineh Zelyonkina-DEF
'Who was Marina Zelyonkina?' (EANC)

2.1.2 *The position of the enclitic auxiliary verb* The auxiliary verb in MEA is an enclitic that carries tense and agreement features and attaches to various elements. Tamrazian (1994:12) provides the following principle for the placement of the auxiliary in MEA: "Being a head final language, the auxiliary verb. In this language appears as the rightmost element in declarative sentences", or, in other words, the auxiliary follows the lexical verb. Note however that as mentioned by Tamrazian herself, this is not the unique position for the

[3] Note that Comrie claims at the same time that the obligatory preverbal position involves only interrogative pronouns. Otherwise, the focus can also be placed immediately after the verb.

placement of the auxiliary. The latter is placed in a variety of positions depending on various factors. It is beyond the scope of this paper to thoroughly discuss the rules of the placement of the auxiliary in MEA. Roughly speaking, in focus neutral (or unmarked) affirmative sentences, the auxiliary appears: a) on the lexical verb in intransitive constructions, ex. (2); b) attaches to the preverbal element of the 'compound verb' or the 'complex predicate', ex. (3); c) is placed after the direct object, in case the latter is bare or indefinite, ex. (4); and d) attaches to the verb, in transitive constructions with a definite direct object, ex. (5).

(2) pieŕ-ə k'ayl-um ēr hyurasrah-um
Peter-DEF walk-IPFV be.AUX.PST.3SG sitting.room-LOC
'Peter was walking in the sitting room.'

(3) durs ēm gal-is
out be.AUX.PST.1SG come- IPFV
'I am going out.'

(4) xot ē ut-um kov-i pes
grass be.AUX.PST.3SG eat-IPFV cow-GEN like
'He eats the grass like a cow.'

(5) davit'-ə čaš-n ut-um e
David-DEF meal-DEF eat-IPFV be.AUX.PST.3SG
'David is eating the meal.'

Crucially, in marked (non focus-neutral) sentences, the auxiliary attaches to the focused constituent of the sentence (Hodgson, 2013: 27, among others). This fact is illustrated by the obligatory placement of the auxiliary after interrogative pronouns, regardless of the type of the verbal construction.

(6) bayts' inč'u ē inʒ nayum
but why be.AUX.PST.3SG I.DAT look-IPFV
'But why is [he] looking at me?'

2.2 *The preverbal position of bare objects* Armenian displays bare nouns, that is, nouns used with no determination or quantification. These nouns have either a kind-level or existential reading. When used as DOs, they generally tend to be in a close or cohesive semantic relationship with the verb. Although all studies admit the possibility for the DOs to occur post verbally, bare (non-determined, non-quantified) DOs seem to enjoy less freedom and display a strong preference for the immediately preverbal position (Badikyan 1976:157, Dum-Tragut 2009:562[4], among others).

(7) t'atron-I tomsarkʁ-ic' toms gn-ec'i
theatre-GEN box-office-ABL ticket buy-AOR.3SG
'[I] bought a ticket from the box-office of the theatre.'

2.3 *Word order in VP idioms* The word order in VP idioms can be a cue to the basic word order in flexible languages. It is generally assumed that information structure is one of the main factors in the choice of a given order among various possibilities. Since idiomatic sequences are generally semantically opac (i.e. non compositional), their components cannot bear specific discourse properties (e.g. topicality, givenness, etc.), which implies that contextually triggered word order variations cannot apply to idiomatic sequences. In other words, word order in such sequences is syntactically determined and corresponds to the basic word order in a given language (see Asatiani & Skopeteas 2012, among others). Idiomatic VPs display OV order in MEA:

(8) a. dagaɣ mtnel (lit. coffin enter) 'to die'
b. Layn sirt unenal (lit. large heart have) 'to be lenient'

[4] In reality Dum-Tragut's generalization concerns all indefinite and non-specific DOs, whether they are bare or determined.

2.4 *The preverbal position of low adverbs* Low adverbs are adverbs whose scope is the VP, and therefore are placed "low" in the syntactic structure. They generally occur in the vicinity of the verb, and either immediately precede it or are placed at the left edge of the focus phrase. Their appearance in other positions yields a marked interpretation. In MEA, these adverbs have also been assumed to mark the edge of VP (Tamrazian 1994: 172).

(9) Nrank' arag veradarn-um en u šrǯapatum Misak-in
They quickly return-IPFV be.AUX.3PL and surround-IPFV Misak.DAT-DEF
'They return quickly and surround Misak.'

To sum up, several features of the MEA favor its grouping with OV or head-final languages and consider the VO order to be pragmatically marked and related to information packaging. That is, SVO is viewed as a case of right dislocation resulting from the right movement of the DO to a postverbal position. If indeed SVO is a 'marked' order then it should be identifiable as such by quantitative studies, provided they comply with methodological requirements.

3 Our corpus investigation

We carried out a corpus study to investigate the distribution of different possible word orders in MEA transitive sentences, using the Eastern Armenian National corpus (EANC). EANC is a large-scale open access corpus (http://eanc.net/) with about 110 million tokens consisting of both written discourse and oral discourse, which offers a powerful search engine for making different types of queries. In order to constitute a sample of sentences displaying a transitive construction, we looked for finite verbs with a high potential to appear in transitive constructions, such as *sirel* 'like' or *spanel* 'kill'.

We successively extracted 400 tokens with *sirel* 'like' at the present tense and 150 tokens at the past tense, more precisely the aorist form, as well as 350 tokens with *spanel* 'kill' at the present tense. Note that the aorist is a simple tense, while the present tense is periphrastic, formed by the imperfective participle and the auxiliary *linel* 'be'. Note also that in order to neutralize the issue of the auxiliary placement, we limited our query to occurrences where the participle immediately precedes the auxiliary (i.e. the auxiliary and the lexical verb are contiguous). Consequently, the simple and the periphrastic tenses become comparable, in that the word order variation is limited to the order between the DO and the verb (as a whole).

In sum, we extracted a total of 900 occurrences including 400 simplex and 500 periphrastic forms, that we annotated manually. Filtering out irrelevant occurrences (e.g. passive constructions, interrogative sentences, etc.), we obtained a sample of 570 occurrences containing a subject, a DO and a verb. This sample contains 338 instances of periphrastic verbal forms and 231 instances of simple verbs. We first annotated our sample for word order. Table 1 presents the distribution of word order in this sample and Table 2 presents the distribution of word order by verbal form for SOV and SVO orders (excluding infrequent orders which leaves us with a total of 543 occurrences).

Table 1: Word order distribution

Orders	**Nb. of occ.**	**%**
OSV	7	1.2%
OVS	15	2.6%
SOV	93	16.3%
SVO	450	78.9%
VOS	2	<0.5%
VSO	3	0.5%
Total	570	100%

Table 2: Word order by verbal form

Form \ Order	SOV	SVO	Total
Periphrastic	40	287	327
Simple	53	163	216
Total	93	450	543

We observe that SVO is by far the most frequent order (78.9%) and remains so independent of the verbal form: SVO is by far more frequent than SOV in sentences with simple verbal forms (87.5% vs.12.5%), as well as with the periphrastic form (75.5% vs. 24.5%). Importantly, when we study these occurrences more closely, we see that the SVO order maps into an 'all focus' discourse configuration in a fair share of cases. This is illustrated by the following examples:

(9) yes sir-um ēm ayd kerpar-ə
I.NOM like-IPFV be.AUX.1SG that character-DEF
'I like that character.'

(10) nrank' span-ets'in mez
They kill-AOR.1PL we.DAT
'They killed us.'

We then annotated our sample for other potentially relevant factors, such as the realization of the DO (Table 3), with a particular interest in DO's degree of definiteness. We hence limited our sample to the three possible determinations for DOs in MEA: 1) bare, that is, a noun carrying no determination or quantification, ex. *girk* 'book' 2) indefinite, that is, an NP with an indefinite determination such as *mi girk* 'a book' and 3) definite, that is, an NP with a definite determination such as *girk-ə* 'the book'. Table 4 provides the distribution of word order in this sample (a total of 380 occurrences) with respect to the definiteness DO.

Table 3: Word order by DO type

DO type \ Order	SVO	SOV	Total
Bare	13	4	17
Indefinite	18	5	23
Definite	269	71	340
Infinitive	86	0	86
Clausal	5	0	5
Others	59	13	72
Total	450	93	543

First, we observe that our sample is not well balanced for different types: indefinite and bare DOs are very scarce. For definite DOs, SVO is clearly the most frequent order (79.1% vs. 20.9). Our sample shows a bias towards the SVO order for bare and indefinite DOs also. However, the limited number of occurrences in these categories (respectively, 17 and 23) does allow for a conclusive generalization.

Table 4: Word order by DO definiteness

DO type \ Order	SOV	SVO	Total
Bare	4	13	17
Indefinite	5	18	23
Definite	71	269	340
Total	80	300	380

4 Discussion and conclusion

These findings present a serious challenge for any assumption on the markedness of the SVO order in MEA. Even though there may be a corpus-related bias, the word order distribution in our sample nevertheless undermines such an assumption. Indeed, the rate of the SVO order is too high to qualify as a marked option.

Especially the fact that definite DOs are overwhelmingly postverbal is line with Badikyian's (1976) as well as Dum-Tragut's (2009) claim that the placement of the DO is not random and that it is triggered by the (in)definiteness of the DO, with indefinite (bare included) DOs being preferably preverbal, while definite DOs tend to occur post-verbally.

Badikyan (1976) is one of the first and most detailed studies on the matter and this is also the first attempt to provide corpus-based generalizations. Based on 8 selected corpora of various 'genres' or types (juridical, literature, scientific…) of about 185 words each, Badikyan (1976: 154) claims that "in the ordinary order, the direct object is placed after the predicate in all types of texts." In two cases, however, the preverbal position of the DO is preferred or even obligatory:

1. The speaker wants to draw the attention of their interlocutor on the DO and to emphasize it. In this case, the DO bears what Badikian calls the 'logical accent' of the sentence.
2. The DO is indefinite.

The first case can be considered a marked order, in which the DO realizes the narrow or contrastive focus. Badikyan also mentions that the auxiliary is placed after the DO in this case and thus precedes the lexical verb, which is another cue for considering this order as marked: in other words, the preverbal position of the DO is a deliberate choice on behalf of the speaker to focus it.

The second case, on the other hand, should not be considered to be a 'variation' or a 'choice', as explicitly stated by Badikyan, who claims that the VO order for indefinite DOs is 'unfamiliar' or 'strange' in MEA. It should be mentioned at this point that the only examples Badikyan provides to illustrate the obligatorily preverbal position of indefinite DOs are bare nouns. Consequently, we could safely rephrase his generalization in (2) above as: Bare DOs must occur preverbally. To sum up, according to Badikyian:

1. Bare DOs must occur preverbally;
2. All other DOs occur post-verbally in the basic or unmarked word order. Preverbal placement implies a narrow focus status for definite DOs.

If these generalizations are shown to be empirically sound, then OV and VO orders do not have the same distribution. The relevance of definiteness in the pre- or post-verbal position of the DO has also been noted by Dum-Tragut (2009: 562) 'Both in written and spoken MEA, native speakers, with overwhelming

frequency, prefer the word order SVO for definite direct objects and SOV for indefinite or non-specific direct objects.' Dum-Tragut also notes that even indefinite DOs tend to occur post-verbally provided there are several DOs (i.e. coordination of indefinite DOs or a sequence of multiple indefinite DOs) (p. 563), which could be considered as a case of heavy NP shift.

If future studies confirm the relevance of the determination in the placement of the DO, then taking into account infinitive and clausal DOs, a new picture on word order preferences on the clausal level emerges for MEA:

- Indefinite/Bare DOs tend to precede the verb;
- Definite DOs tend to follow the verb;
- Infinitive DOs generally follow the verb;
- Clausal DOs always follow the verb.

So, from a strictly distributional point of view, OV is in minority (marked) and the default word order is VO. If these preliminary findings were to be confirmed by further quantitative studies, then MEA would represent the case of a typically 'OV' language with respect to word order correlations across constituents, while being VO at the clausal level. This situation is reminiscent to some extent of Georgian, another OV language spoken in the area. As mentioned by Asatiani and Skopeteas (2012: 128) quantitative studies show that both OV and VO orders occur in Georgian at an approximately equal rate. Studies on word order variation in Georgian have mainly focused on information structure, especially the issue of a post-verbal focus (Skopeteas and Fanselow, 2010), but do not look into definiteness. It would be interesting to investigate the same issue in Georgian in order to have a more general view on the role played by definiteness on the position of the DO. Importantly, it has been shown that definiteness plays a role in determining the position of the DO in Persian, another OV language. However, unlike MEA, definite DOs in Persian 'move' leftward, scrambling over PP arguments and are placed at a distance of the verb. While, in MEA they 'move' rightward and are placed after the verb. Despite these opposite directions, both placement tendencies could be the manifestation of the same phenomenon.

Another issue to investigate is the position of indefinite DOs. Like MEA, Persian displays both bare objects (non-determined/quantified) and indefinite objects (marked by an indefinite determiner). Quantitative studies on Persian have shown that these two types of DOs do not behave in the same way with respect to their ordering preferences. Surprisingly (and contrary to common assumptions in many studies), indefinite DOs group with definite DOs. Bare DOs, on the contrary, stay adjacent to the verb. It could well be the case that same tendencies hold for MEA too, which would explain why all examples given by Badikyan (1976) and Dum-Tragut to illustrate the preverbal position of indefinite DOs are in fact bare DOs. If this was to be confirmed, the OV would have an even more limited distribution suggestion that at the clausal level MEA is not really head-final. Further studies are required to evaluate these findings.

References

Asatiani, R. & S. Skopeteas. 2012. The information structure of Georgian. In Manfred Krifka & Renate Musan (eds.), *The expression of information structure*, 127 – 158. Berlin: De Gruyter.

Abrahamyan, S. S., Arakelyan, V. & V. Kosyan. 1975. *Hajoch lezu: šarahjusutjun* [The Armenian Language: Syntax], v. 2. Yerevan.

Arakelyan, V. 1958. *Hajereni šarahjusuthjun* [Armenian Syntax]. v. 1, Yerevan.

Badikyan, Kh. 1976. *Žamanakakich hajereni parz naχadasuthjan šaradasuthjunə* [The Word Order in the Simple Sentences in Modern Armenian]. Yerevan.

Comrie, B. 1988. Topics, Grammaticalized Topics, and Subjects. In *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*. 265-279.

Der-Houssikian, H. 1978. The Semantic Significance of S, O, V Sequences in Armenian. In Jazayery, M.-A., Polomé, E. C. & W. Winter (eds.), *Linguistic and Literary Studies, v. 2, Descriptive Linguistics*, 227-236. Berlin, New York: De Gruyter Mouton.

Dezső, L. 1974. Topics in syntactic typology. In Romportl, M., Skalička, V., Popela J. & B. Palek (eds.), *Linguistica Generalia v. 1: Studies in linguistic typology*, 191- 210. Prague: Charles University.

Dryer, M. S. 1992. The Greenbergian word order correlations. *Language*, 68(1), 81–138.

Dryer, M. S. 1998. Aspects of Word Order in the Languages of Europe. In Siewierska, A. (ed), *Constituent Order in the Languages of Europe*, 283 - 319. Berlin: Mouton de Gruyter.

Dryer, M. S. 2013. Order of Subject, Object and Verb. In Dryer, M. S. & M. Haspelmath (eds.) *The World Atlas of Language Structures Online*.
Dum-Tragut, J. 2009. *Armenian: Modern Eastern Armenian*. Amsterdam: John Benjamins.
Dum-Tragut, J. 2002. *Word order correlations and word order change: an "applied-typological" study on literary Armenian variants (v. 22)*. München: Lincom GmbH.
Giorgi, A. & S. Haroutyunian. 2016. Word Order and Information Structure in Modern Eastern Armenian, *Journal of the Society for Armenian Studies* 25, 185 – 201.
Greenberg, J. H. 1966. *Language universals: With special reference to feature hierarchies*. The Hague: Mouton.
Hodgson, K. 2013. *Discourse Configurationality in Eastern Armenian: The Nominal Projection*. MPhil thesis, University of Cambridge.
Hawkins, J. A. 1979. Implicational universals as predictors of word order change. *Language* 55, 618-648.
Hawkins, J. A. 1983. *Word order universals*. New York: Academic Press.
Herring, Susan C. & J.C. Paolillo. 1995. Focus position in SOV languages. In Downing, P. & M. Noonan (eds.), *Word Order in Discourse*, 163-198. Amsterdam & Philadelphia: John Benjamins.
Kahnemuyipour, A. & K. Megerdoomian. 2011. Second Position Clitics and the vP Phase: The Case of the Armenian Auxiliary. *Linguistic Inquiry* 42, 152-162.
Kahnemuyipour, A. & K. Megerdoomian. 2017. On the Positional Distribution of an Armenian Auxiliary: Second-Position Clisis, Focus, and Phases. Syntax 20(1). 77–97.
Kim, A. H. 1988. Preverbal focusing and type XXIII languages. In Hammond, M., Moravcsik, E. & J. Wirth (eds.), *Studies in syntactic typology*, 147-169. Amsterdam: John Benjamins.
Kozintseva, N. 1995. *Modern Eastern Armenian*. München: Lincom Europa.
Lambrecht, K. 1996. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.
Minassian, M. 1980. *Grammaire d'Arménien Oriental*. New York: Delmar.
Papoyan, A. & Kh. Badikyan. 2003. *Žamanakakich hajoch lezvi šarahjusuthjun* [Modern Armenian Syntax]. Yerevan.
Skopeteas, S. & G. Fanselow. 2010. Focus in Georgian and the expression of contrast. *Lingua* 120. 1370 – 1391.
Tamrazian, A. 1991. Focus and WH-movement in Armenian. *UCL Working Papers in Linguistics* 3, 102-121.
Tamrazian, A. 1994. *The Syntax of Armenian. Chains and the Auxiliary*. PhD Thesis, University College London.

Information Structure and Syntax: Two Positions for Focus in Basque

Matteo Fiorini
University of Utah

1. Introduction

The main goal of this paper is to investigate the role of information structure in the structural and linear organization of Basque. In section 2, I present an overview of some basic descriptive facts of Basque focalization and word order and a discussion about the main syntactic accounts for the topic. Two strategies have been proposed: a syntactic one, baptized 'left-peripheral,' and a purely prosodic one, named 'NSR' (nuclear stress rule) approach. In section 3, I provide examples from the recent literature about focus that show that the investigation of focus typology is crucial in order to obtain a complete picture of the phenomenon. In particular, I point out that different types of *foci* can be analyzed in terms of their requirement or non-requirement of antecedent in the discourse. The requirement is flexible on a continuum that has contrastive focalization and mirative focalization on the two opposite sides. In section 4, I advocate for the presence of two distinct focus projections in Basque: one that hosts discourse-anchored *foci* and one in the T-internal area that hosts *foci* that can be interpreted out of the blue. Moreover, the analysis of mirative *foci* points to the generalization according to which the low-periphery hosts elements that provide epistemic information related to the point of view of the speaker over the propositional content.

2. Focus in Basque

According to É. Kiss (1994), Basque is a 'discourse-configurational language,' i.e., a language whose relatively free word order is strongly affected by information structure. Both experimental studies (Laka and Erdocia, 2012) and typological considerations (Laka, 1994; De Rijk, 2008 a.o.) show a strong tendency for a neutral, canonical SOV order. Whenever this order is not respected, the element that immediately precedes the verbal complex[2] – referred to in prescriptive grammar as *galdegaia* 'the questioned element' – is interpreted as the focus of the sentence. In other words, all the combinations of the three constituents in (1) are grammatical, but (1d) is not pragmatically felicitous because *Jonek* 'Jon' cannot be interpreted as Focus.

(1) a. Jonek Mireni ikusi du
J.erg M.dat see.prt aux.3sg.prs
'Jon saw Miren'
b. [Jonek]$_{FOC}$ ikusi du Mireni
c. Mireni [Jonek]$_{FOC}$ ikusi du
d. Ikusi du [Jonek]$_{FOC}$ Mireni

* I would like to thank the speakers, the audience, and the organizers at WECOL 2019 at the University of Fresno (CA) for the valuable comments and new insight on the claims of this paper. I would also like to thank Aniko Csirmaz for the insightful comments on the drafts of this paper. The material presented here has been extensively discussed in the weekly "Syntax and Semantics meeting" at the department of Linguistics at the University of Utah; I thank my fellow students and faculties for their help with early versions of this work. Finally, this research started as topic of my MA thesis at Ca' Foscari University of Venice; I am thus in debt with many people there but in particular Alessandra Giorgi, Inaki Alfaro Vergarachea, and my informants at *Zornotzako Barnetegi* (Zornotza, Euskal Herria).

[2] For the sake of consistency with previous literature (Laka, 1998; De Rijk, 2008), I use the term 'verbal complex' to identify the analytic verbal forms that represent the vast majority of Basque verbs (e.g., *ikusi du* 's/he saw'), in which the two elements are considered inseparable except for negative constructions where the auxiliary follows the negative particle *ez* which appears in a higher position. The two elements can be separated, however, by a series of particles (see, e.g, De Rijk, 2008). The discussion of this issue is far beyond the scope of this paper.

Proceedings of *WECOL 2019*

Only a handful of particles can intervene between the focus and the verbal complex, and for this reason, the sequence [focus] + [verbal complex] has been standardly assumed to form a cluster (see Elordieta, 2001). I will refer to the group as 'focus cluster' (FC). This basic description can account for the vast majority of structures with some exceptions that I will discuss in section 4.

2.1 *Position of the focus cluster* The Focus Cluster can virtually appear in every position in the sentence; however, FC tends to surface either in sentence-initial (1b) or in sentence-final position (1c). The former has been described as a V2-like configuration in A. Elordieta and Haddican (2013), and I will adopt their terminology[3]. When other phrases precede the FC in V2-like configuration or follow it in FC-final configuration, they are treated as topics (see Ortiz de Urbina, 2002). It is important to remind that immediate adjacency to the verb is the standard way to signal the most informationally-prominent element of a sentence that does not respect the canonical word order (Laka, 1996; Elordieta, 2001 a.o.). I will show that this is not the only available strategy for at least two reasons: (i) there is the tendency to use one or the other configurations to signal additional semantic and pragmatic information; and (ii) there are marked structures in which the adjacency focus-verb is not respected (Ortiz de Urbina, 2002; Etxepare, 1998 a.o.).

2.2 *Previous accounts* For reasons of space, I will not include a detailed review of previous analyses of focus in Basque, but I will provide an overview of the two main strategies proposed in the literature which I organize, in line with Irurtzun (2016), into Left-peripheral approaches and NSR approaches.

2.2.1 *Left peripheral approaches* In some languages, Focus is syntactically marked in that it appears in a left-dislocated position. This position has been treated as a Focus Phrase, part of a layered CP area labeled 'left periphery' of the clause (see Rizzi, 1997 s.w.), in which are hosted discourse-related elements.
Ortiz de Urbina (1995; 1999; 2002) proposes that a focalized element is attracted to [Spec, CP] – [Spec, FocP] in later accounts (e.g., A. Elordieta, 2001; Irurtzun, 2005 a.o.) – and this triggers the movement of the verbal complex to C° so that the [focus] feature of the dislocated constituent can be checked in a Spec-Head configuration ('focus criterion,' Horvath, 1995). A similar analysis is put forward in A. Elordieta (2001) to account for the V2-like configuration described above. Her proposal, however, differs from Ortiz de Urbina's in having the focalized constituent generated in the left periphery, while it is an operator co-indexed with it that triggers movement of the verbal component. This proposal raises a variety of questions related to the way the focalized elements check their features (minimally agreement, case, and theta relations), which are arguably evaluated by the operator that is generated in a theta position. This now reasonably standard analysis of Focus, which is extended to wh-structures, fails to accommodate multiple-*foci* and multiple-*wh* structures, both well attested in the language (see De Rijk, 2008)[4]. Another shortcoming of these accounts, as pointed out in Irurtzun (2005), is that there is nothing in the derivation that explains how a particular constituent comes to bear a [focus] feature[5].

2.2.2 *NSR approach* The NSR (nuclear stress rule)-approach is prosody-driven in that the derivation of focus structures results from conditions on PF. Undoubtedly, prosody plays an essential role in discourse-related structures and, specifically, focus bears main sentence stress, a generalization that is consistent cross-linguistically (see Zubizarreta, 1998). Sentence stress tends to appear in a language-specific canonical position, and while for certain languages (e.g., Italian; Spanish; English), the option of shifting the stress place is available, in others, like Basque, this cannot be done. Departing from these considerations, Arregi (2003) proposes that the adjacency between focus and verbal complex is a by-product of a series of focus evacuation movements. In a nutshell, the stress is assigned to the most embedded branching node, i.e., the preverbal position, according to a modified version of the Nuclear Stress Rule proposed in Cinque (1993). If the target of focalization is not in said position, every unfocalized constituent moves so that the focus can be

[3] Elordieta and Haddican (2013) consider Basque a V2-like language in that – with few exceptions – the verb must be preceded by at least one element (contrary V2 languages that require exactly one constituent in pre-verbal position).

[4] Preliminary analysis of data that I elicited in my informants, point in an interaction between syntax and prosody to mark multiple *foci*. In particular, it appears that one focus is marked by an intonational break that precedes it and the other by the standard pre-verbal position.

[5] Irurtzun (2005) proposes a derivational approach to show how a particular element ended up bearing focus, however, his remark on left-peripheral approaches is still valid since at least one element must enter the derivation already bearing a [focus] feature.

the most embedded element and the main sentence stress is not assigned to the wrong constituent. A. Elordieta (2001) adopts a similar view to account for structures where FC appears in sentence-final position. She proposes that the focalized item reaches the preverbal position via scrambling, in order to receive sentence stress, while the verbal complex remains *in situ*. The central difference between these two proposals is that Arregi's (2003) one is exclusively NSR-driven, and this entails lowering and apparent 'look ahead' operations required by stress-avoiding movements that have been shown to be problematic (see Irurtzun, 2003). On the other hand, according to A. Elordieta (2001, in the spirit of Ortiz de Urbina, 1999 a.o.), the functional portion of the clause is left-headed, that is, the movement of the verbal complex is always to the left.

2.3 *Some theoretical issues with previous analysis* From an interpretational point of view, according to both A. Elordieta's (2001) and Arregi's (2003) analyses, the two configurations of focus structures do not differ in their pragmatic or semantic content. The choice for one or the other strategy, which crucially differs in terms of economy of movement, is due, according to A. Elordieta (2001) to the choice of the speaker to "marks a particular constituent as the single focus of the sentence and ensures that no other potential focus interpretations are available in the sentence" (p. 231). The use of one or the other strategy – which importantly leads to one or the other configuration – consequently, seems to be due to pragmatic factors active at PF. However, it is not clear how this possibility can be implemented in an account in which syntactic operations are responsible for the position of the focalized element. Moreover, it is not clear how the preferences of the speaker are defined, i.e., there seems to be no reason for the speaker to ever use the *in situ* strategy since the interpretation could, as entailed in A. Elordieta's explanation, fail. Both authors explicitly mention the fact that in their works, they only discuss 'information focus' in the sense of É. Kiss (1998), without dealing with other types of *foci*, which are claimed not to play a role in the linear (and structural) position of focus. In fact, to my knowledge, scarce attention has been devoted to focus typology in Basque with the exceptions of Ortiz de Urbina (2002) and Sainz Maza Lecanda (2017). According to the former, *foci* in V2-like configurations are 'focus of correction,' and for the latter, exhaustivity is the driving force for the movement targeting the same position. However, many recent discussions about focus (Rizzi, 1997; É. Kiss, 1998; Dal Farra, 2018) show that the investigation of focus typology is a crucial issue that needs to be addressed when investigating the topic. In the next section, I will sketch a brief description of previous discussions of different types of *foci* before investigating some data for Basque that point to an analysis in which *foci* with different semantic features behave differently from one another in terms of distribution and verbal proximity.

3. Focus typology

Before discussing some of the *foci* identified in the literature, it is important to point out that despite being an extensively investigated topic, the very notion of Focus does not have a standard definition, together with other discourse-related elements. In general, the role of information structure in the syntactic derivation is still a debated issue in the field. On the one hand, one line of research follows Chomsky (1995) in assuming that discourse-related elements such as topic and focus are banned from the syntactic derivation by the *Inclusiveness Condition* (as explicitly mention in Chomsky, Gallego, and Ott, 2019). On the other hand, the role that information structure plays in syntax is central in many works within (at least) the Cartographic tradition (Rizzi, 1997; López, 2009 a.o.). There is, however, no consensus on this point among scholars regardless of the particular framework or line of research they pursue. I cannot offer here a detailed discussion of the debate, but I will adopt the hypothesis that discourse-related items are indeed part of the syntactic component. The notion of focus that I adopt, however, does not strictly correspond to the one adopted within the Cartographic tradition, but I will follow a recent proposal by Kratzer and Selkirk (2020). The authors argue that information structure can indeed be deconstructed into morpho-syntactic features, but that purely newness is not marked grammatically. For the sake of clarity, I will not here adopt their terminology, and I will use the term Information Focus in the sense of É. Kiss (1998), that is, the most prominent element of the structure - typically exemplified by the answer to a *wh*-question - while describing previous researches.

3.1 *Types of foci*

In this section, I will discuss a wide variety of *foci* whose distribution and syntactic behavior have been investigated extensively in the literature. This selection has been made in order to offer a brief overview of some important conclusions of these studies. However, the types of *foci* that are relevant for the discussion about Basque in section 3 are, specifically, corrective and mirative focus.

3.1.1 *Information focus* Information focus (IF) can be defined as the most informationally prominent element in the clause, often exemplified by the answer to a wh-question. In this case, according to É. Kiss (1998), the wh-element activates an open set of alternatives, and in the answer, the appropriate one is chosen (2):

(2) a. What happened?
b. [John has seen Mary]$_{FOC}$
c. Who has John seen?
d. John has seen [Mary]$_{FOC}$

In a language that marks (certain types of) *foci* via dislocation like Italian, the focalized item behaves differently; i.e., it cannot be dislocated (3):

(3) a. Chi ha incontrato Gianni?
Who has meet.prt G.
'who did Gianni meet?'
b. Gianni ha incontrato [Maria]$_{FOC}$ *Focus in situ*
G. has meet.prt M.
'Gianni met Maria'
c. *[Maria]$_{FOC}$, Gianni ha incontrato *Left-dislocated focus*
'It is Mary that Gianni met[6]'

The fact that in many languages two elements, like in Italian, have different distributions, provides evidence for (i) the existence of more than one type of focus, and (ii) the claim that semantic information can affect syntactic operations.
In Basque, as pointed out in the introduction, the most informationally-prominent element of the clause generally occupies the immediately pre-verbal position, the same position occupied by the *wh*-constituent *non* 'where' in (4a). The answer to the whP is the IF of the sentence (4):

(4) a. **non** ikusi duzu umea?
where see.prt aux.3sg.prs child-det.abs
'where have you seen the child?'
b. [**kalean**]$_{FOC}$ ikusi duzu umea
street.loc see.prt aux.3sg.prs.abs
'you have seen the child in the street' [Laka,1996 ex.18 p.15]

Crucially, IF is not only possible as the answer to a *wh*-question, but it can appear in out of the blue contexts, and it does not require an antecedent to be interpreted (see Dal Farra, 2018).

3.1.2 *Corrective Focus* Focus of correction (contrastive focus in, a.o. É. Kiss, 1998 - CF) is responsible for the correction of a previous utterance. In several languages, it can be expressed either by a particular prosody or a particular syntactic distribution. In Italian (5b/c), corrective focus is marked by prosody or left-dislocation. English (6b) can employ cleft-structures to provide a corrective reading:

(5) a. Gianni ha visto Veronica
Gianni aux.3SG.PRS see.prt Veronica
'Gianni saw Veronica'
b. Gianni ha visto [Maria]$_{FOC}$ (, non Veronica)
c. [Maria]$_{FOC}$, Gianni ha visto (, non Veronica)
'Gianni saw Maria (, not Veronica)'

[6] In the English translation I use a cleft because these structures convey in English the same corrective meaning as the dislocation in Italian. However, it is important to notice that (3c) is ungrammatical in unmarked context, hence the judgment.

(6) a. John saw Andrew
b. It is Mary that John saw

In terms of alternative semantic description, CF replaces a previously mentioned item employing another of the alternatives, considered correct by the speaker, activated by the context. This entails a crucial difference between information focus and other types of *foci*: it cannot be uttered out of the blue, but it requires an antecedent in the discourse.
However, Büring (2016) notices that there are cases in which contrastive focalizations do not entail correction, e.g., the pair in (7) in which the contrastive focalization is the elaboration of an element to be contrasted:

(7) a. John bought a big car
b. John bought a RED car

According to Büring, (7b) is an instance of 'unrelated contrastive focus,' which is pragmatically felicitous because the two speakers do not share the same opinion about what is the salient feature of John's new car.

3.1.3 *Contrastive focus* Examples like (7) lead to a closer analysis of those types of *foci* that do not explicitly correct a previous statement, but that involves a contrast between two (or two sets) of the focus-alternative. In particular, Bianchi, Bocci, and Cruschina (2015, 2016) make a distinction between purely corrective focus (contrastive focus in É. Kiss, 1998) and contrastive focus based mainly on Italian data. The former must contain at least one of the alternatives activated by the context, while the latter does not (8):

(8) a. Il mio amico Gianni ha fame, ma so che è a dieta.
'My friend John is hungry, but I know that he is on a diet.'
b. Ti conviene mangiare un'[insalata]$_{FOC}$ (, non la pasta).
To.you be.better eat a salad not the pasta
c. *Un'[insalata]$_{FOC}$ ti conviene mangiare (, non la pasta).
a salad to.you be.better eat not the pasta
'You'd better eat a salad (, not pasta).' Dal Farra, 2018 [4]

Recall that in Italian, Corrective Focus can be fronted, while Contrastive one cannot, as shown by the ungrammaticality of (8c) in which the alternative to the focalized items is a specific set of alternatives – that is, any food that is appropriate for a diet (Bianchi, 2015). The distribution of Contrastive focus, mimics the one of IF, providing evidence for the definition of focus as exclusively contrastive (Kratzer and Selkirk, 2020).

3.1.4 *Mirative Focus* The notion of mirativity as an independent grammatical category has been a debated issue in the literature ever since the seminal work by DeLancey (1997). Mirativity indicates new and surprising information for the speaker who can mark their speech act to convey unexpectedness. A number of scholars argued that mirativity is not an independent category, but it is a feature that can emerge in languages with an evidential system (Hill, 2012). However, DeLancey (2001) points out that although evidentiality and mirativity are conceptually related, "each represents the grammatical indexation of ways in which a proposition can deviate from an ideal of knowledge." (p. 379). In one case, the source of evidence for information (i.e., first-hand perception, hearsay, or inference) while in the other "whether the information represents knowledge […] new to the speaker, or knowledge […] already integrated into the speaker's picture of the world." (p. 380). It seems clear by this definition that the two categories have different roles and, plausibly, they are not in complementary distribution but rather can freely interact – which cannot be true for the possible evidential markers.

Further evidence supporting DeLancey's (1997) proposal comes from the analysis of an additional type of focus identified by Cruschina (2012) in Italian, a language with no evidential system neither syntactic, morphological or phonological. Mirative Focus (MF) expresses surprising and unexpected information, that is, not yet part of the shared knowledge of the interlocutors. It does not allow an antecedent in the discourse; thus, it must be uttered out of the blue, and it is realized using a particular prosodic contour (9):

(9) a. Pensa te! DI VENTI KILI è dimagrito!
think you of twenty kilos he.lost.weight
b. Pensa te! È dimagrito DI VENTI KILI!
think you he.lost.weight of twenty kilos
'Guess what! He lost twenty kilos!' Dal Farra, 2018 [7]

As shown in the example (9), in Italian, MF can appear both in sentence-initial, after an exclamation conveying surprise, or in sentence-final position, much like corrective focus structures like (5). Bianchi (2015) points out that in MF not only the information but also a certain evaluative content is shared by the interlocutors. In the case of (9), neither the speaker nor the hearer expected 'him' to manage to lose so much weight. Both (9a) and (9b) are realized with a particular raising intonation on the focalized portion of the clause with a slight difference in that in (9a) the focus is followed by an intonational break. However, contexts like (10a) and (11a) show that it is not necessary for the surprise over the event to be shared by both the interlocutors. In fact, in order to have a mirative construction, it is only crucial that the information is surprising for one of them. Consider the English example in (10):

(10) a. [*it is the first time that the speaker pays for a medical examination in the US, they share their surprise to an American friend (who reasonably knows the cost of such a visit)*]
b. It's incredible! I had to pay 200 dollars!

The speaker does not think that the interlocutor is not aware of the amount of money they paid, but they want to convey that they pay 'more than expected.'[7]. Moreover, MF is also possible in cases like (11) where a piece of information is not surprising anymore for the speaker, but they expect it is for the hearer.

(11) a. [*there's a new restaurant in town, I had dinner yesterday with a friend and I liked it. I tell my mother about it]*
b. There's a new restaurant downtown and you know? It's really good!

(11b) can probably be interpreted both as 'expected surprise for the hearer' and 'conveying of the surprise of the speaker.' A generalization for mirative focus can be that the focalized item conveys the (assumed) surprise over an event that is the least expected to happen by the speaker, the hearer, or both of them.

Another piece of evidence for the autonomy of mirativity comes from Hindi, which possesses an 'anti-mirative' particle that marks elements that are not surprising, i.e., expected (Benjamin Slade, p.c.).

3.2 *Discourse anchoring and Out of the blue contexts* Different languages employ different strategies to mark focus. Before analyzing the case of Basque, it is important to summarize the crucial difference between the *foci* described above in terms of antecedent requirements. Information Focus and Contrastive Focus are felicitous both with an antecedent in the discourse and out of the blue. I will not discuss them in the Basque data in the next section. On the other hand, Corrective Focus and Mirative Focus must meet a specific requirement in terms of antecedent to be interpreted. While the former requires an antecedent in the discourse, the latter cannot have one. Table 1 shows the characteristics exhibited by the *foci* discussed so far regarding this matter:

	Antecedent in the discourse	**No antecedent in the discourse**
Information Focus	*Possible/not required*	*Possible*
Corrective focus	***Required***	***Not possible***
Contrastive focus	*Possible/not required*	*Possible*
Mirative focus	***Not possible***	***Required***

Table 1

The topic of the remainder of this paper is the discussion of mirative and corrective focalizations in Basque, which, as it can be seen in table 1, represent the two extremes of a scale in which the requirement or the non-requirement of an antecedent in the discourse can block the interpretation for one or the other focus. I will

[7] As far as I know, there is no investigation of mirative focalizations in English.

argue that the interpretation of the two *foci* depends on the presence of two distinct focus projections, one in the left periphery of the clause and one in the T-internal area.

4. Basque focus typology

In this section, we will consider focus-structures in Basque to discuss the presence and distribution of different types of *foci*.
First of all, as pointed out in section 2, only (12e) is infelicitous as a reply to (12a) since the constituent that corresponds to the whP *nori* 'to whom' is not immediately adjacent to the verbal complex.

(12) a. Nori eman zion Jonek kamiseta ?
b. Jonek [Mireni]$_{FOC}$ eman zion kamiseta
c. [Mireni]$_{FOC}$ eman zion Jonek kamiseta
d. Jonek kamiseta [Mireni]$_{FOC}$ eman zion
e. [*Mireni]$_{FOC}$ Jonek eman zion kamiseta
'Jon gave Miren a shirt'

As stated above, this is a generalization that holds for different types of *foci,* and the exceptions discussed in this section represent very marked contexts in terms of pragmatics and semantics.

4.1 *Corrective focus* Consider the patterns highlighted in table 1. It is clear, and cross-linguistically well documented, that pure corrective focus requires an antecedent in the discourse in order to be felicitous. As argued above, a corrective interpretation is active for left-dislocated *foci* in Italian (Rizzi, 1997) and Hungarian (É. Kiss, 1998) and according to Kratzer and Selkirk (2020), contrastivity is the only feature that identifies focus. According to their discussion, purely new information (i.e., information focus) is unmarked, while discourse coherence is signaled by G(iveness) and FoC(ontrast) based on the (mis-)match of an element and its antecedent. My proposal departs from these findings, arguing that it is the presence of an antecedent that gives to a purely contrastive focus a specific corrective reading or a giveness connotation (i.e., topic). I propose that a focalized item moves to the Focus projection in the left periphery of the clause (in line with Rizzi, 1997). In order to collect data to analyze corrective foci in Basque specifically, I have elicited corrective structures from native speakers of Basque from the Bizkaian region. This preliminary, informal data collection consisted of providing a context (13a) to my informants and asking to correct an erroneous statement relative to said context (13b). The partial results show an overwhelming preference for the V2-like configuration of focus structure, as exemplified in (13c):

(13) a. [Jon gave Miren a red shirt]
b. Jonek Eiderri kamiseta gorria eman zion
J.er E.dat shirt.abs red give.prt aux.3sg.pst
c. MIRENI eman zion Jonek kamiseta
M.dat give.prt aux.3sg.pst J.erg shirt.abs
'it was Miren that Jon gave a shirt to'

In the majority of cases, the elements in post-verbal position were ordered following the canonical word-order Subject > Indirect Object > Direct Object, that is S > IO; IO > O; S > O suggesting the presence of movement (while the unfocalized element stays *in-situ*). I follow Ortiz de Urbina (2002) in considering the V2-like configuration as involving a left dislocation of the focalized item, which triggers movement of the verbal complex. The movement is motivated by (at least) a [+corrective] feature that attracts the focalized item and subsequent movement of the verb necessary to check the [+ focus] feature on the head. I argue that movement targets the FocP located in the left periphery of the clause, as identified by Rizzi (1997, s.w.). This hypothesis can straightforwardly account also for those instances of corrective *foci* in which the Focus Cluster is preceded by another element: said element is interpreted as a topic and hosted in the corresponding projection. These elements are realized as de-accented and with a falling intonation, similar to the one that is present for the entire unfocalized portion of the sentence. In (14) I show the derivation of (13c) and its counterpart with topicalization of the unfocalized element:

(14) [$_C$ [$_{FocP}$ Mireni$_j$ [$_{Foc°}$ eman zion$_i$][$_T$ Jonek t$_j$ kamiseta t$_i$]]]] <u>*no topicalization*</u>

$[_C[_{TopP}$ Jonek$_z$ kamiseta$_y$ $[_{Top°}][_{FocP}$ Mireni$_j$ $[_{Foc°}$ eman zion$_i][_T$ t$_z$ t$_j$ t$_y$ t$_i]]]]]$ *topicalization*

The discussion is far from being exhaustive since many factors need to be better investigated, in particular, related to topicalization, which might always be present (in line with Rizzi, 1997) in focus structures. In this case, the topicalized portion of the clause will move to different dedicated projections above or below FocP. I will not discuss any further the matter, but the extensive work on Topic and Topic Typology suggests that, like focus-movements, topicalization can have different results in terms of semantic import.

Crucially, the prosodic contour of corrective focus is different from the one of the 'main sentence stress': as I mentioned in section 2.2.2. discussing the NSR-approach, Basque stress cannot move. However, the extensive data collection in Hualde and al (2002) shows that categorizing Basque following the traditional classification of 'stress languages' and 'tone languages' fails to account for the prosodic properties that it shares with the latter group. I refer to that paper for the detail of the analysis, but it is important to point out that focalized constituents are claimed to be marked prosodically not by main sentence stress, but by a pitch *focal* accent (in line with Hualde, Elordieta, and Elordieta, 1994). Consequently, it seems that the *information focus* of the sentence may indeed be marked by main sentence stress, but this is not the way other *foci* are made prominent by prosody. The consequences of this assumption are consistent with the claims made by Kratzer and Selkirk (2020), according to whom *merely new information* is unmarked. When the information associated with the definition of focus as an element carrying additional semantic values needs to be conveyed, the strategy is a syntactic one.

4.2 *Two positions for focus in Basque* Etxeberria and Etxepare (2008) provide a series of examples of interaction between focus and negation in Basque. Let us first consider the structures in (15) and (16):

(15) [Andoniri]$_{FOC}$ ez diot ardoa ekarri (, eta ez Mikeli)
A.dat neg aux.3sg.prs wine.abs bring.prt and not M.dat
'It is to Andoni (and not to Mikel) that I didn't bring the wine'

(16) Ez diot ardoa [Andoniri]$_{FOC}$ ekarri (, Mikeli baizik)
neg aux.3sg.prs wine.abs A.dat bring.prt M.dat but
'It is not for Andoni (but for Mikel) that I brought the wine'

The focus in (15) lies outside the scope of negation, while (16) is in the scope of negation. The two positions have different interpretations, and only (15) can be uttered in out-of-the-blue contexts. According to Etxeberria and Etxepare (2008), the two positions correspond to specific hosting sites for *foci* in the structure: this leads to an analysis of focalization in terms of movement for at least certain types of *foci*. The landing site for *bound foci* (in the scope of negation) is "in-between NegP and VP" (p.297), that is, in a T-internal position. Let us now turn our attention to example (17), which is claimed to be a corrective focus in Ortiz de Urbina (2002):

(17) Ez diot ardoa ekarri ANDONIRI (, eta ez Mikeli)
neg aux wine brought Andoni-dat, Mikel-dat/and not Mikel-dat
'It is to Andoni (and not to Mikel) that I didn't bring the wine'

It is clear and explicitly mentioned in Etxeberria and Etxepare (2008), that (17) is derived from (15) in that the focalized element is outside the scope of the negation, and it cannot be uttered out of the blue. The adjacency between focus and the verbal element is not respected, and the focalization is marked prosodically by an intonational break before *Andoniri* 'to Andoni.' To derive this structure, Ortiz de Urbina (2002) proposes that after the movement of focus in the left periphery, the rest of the clause moves to a higher position – arguably, a topic position, via remnant movement. In other words, a focus of correction must be left peripheral in order to be interpretable correctly. In (18) it is shown, to avoid confusion, the derivation of the affirmative counterpart of (17):

(18) [TopP[~~Andoniri~~ ardoa ekarri diote]j Top0 [FocP[Andoniri] Foc° [IP ~~ardoa ekarri diote~~]]]

On the other hand, there is no way for the focus in (16) to be at once in the scope of the negation, and in a different position (i.e., in the left periphery). This provides further evidence for the assumption of the presence of a T-internal position for focus in Basque.

The hypothesis of a lower, T-internal, focus position has been proposed by, among others, Jayaseelan (2001) to account for similar facts (focalized item in pre-verbal position) in Malayalam and Belletti (2004) to account for post-verbal subjects in Italian. In particular, Belletti (2004) advocates for the presence of a layered portion of the clause, which she labels 'low periphery,' articulated as the left periphery as proposed in Rizzi (1997). The presence of this periphery is theoretically justified within the theory of Phases (Chomsky, 2001) assuming vP as a phase boundary: the low periphery corresponds to the edge of that phase.

4.3 *Clause internal focus position* So far, the only reason to assume a low focus position in the Basque syntactic structure is related to the interaction between focus and negation. Nevertheless, examples of mirative focalizations can provide additional evidence for this position; consider (19):

(19) Anek ez du bada [milioi bat]$_{FOC}$ irabatzi loterian
Ane.erg neg aux.3sg.prs cond million one.abs win.prf lottery.loc
'Ane won a million at the lottery!

In the structure above, after the subject, the speaker introduces a pragmatic expression (*ez du bada*) to convey their surprise over, crucially, the amount of money won by Ane. Two main considerations need to be spelled out. First of all, in section 3.1.4, I propose that mirativity can be described as the grammatical element that conveys surprise over an event that is the least expected to happen by the speaker. In the case we are discussing, the least expected element is not the winning event, but the amount of the prize which is singled out by the speaker. Secondly, it is important to notice that the exclamative expression cannot be considered to be an 'appendix' to the structure (much like a parenthetical/interjectional phrase) because it makes use of an auxiliary. In Basque negative sentences, the auxiliary moves to the right of the negative particle *ez* – possibly to lexicalize it (Laka, 1994) – leaving the lexical verb stranded. The fact that there is no auxiliary after the lexical verb in (19) implies that the pragmatic marker of surprise is built via the movement operations involved in the derivation of negative structure despite being a semantically empty negation. Similarly, the surprise over the event can be conveyed by a rhetoric question like in (20):

(20) Badakizu Anek [milioi bat]$_{FOC}$ irabazi duela loterian
know.3sg.prs Ane.erg million one.abs win.prf aux-comp.3sg.prs lottery.loc
'Do you know that Ane won a million with the lottery!?'

There is one last instance of mirative focalization identified by Etxepare (1998) and discussed in Irurtzun (2016) in which the focus-verb adjacency is not respected (21), much like the corrective focus in (17):

(21) [Jonek]F ardoa ekarri du
Jon wine bring aux
'[Jon]F brought wine! (he never does so!)' Irurtzun, 2016 [34]

According to Irurtzun (2016), the standard order in which *Jonek* 'Jon' immediately precedes *ekarri du* 'brought' "conversationally implicate the eventuality denoted by the open proposition in the (potential) question they answer," while the one in (21) "conventionally implicate it" (p.259). In his original analysis, Etxepare (1998) argues that this entails a difference in the syntactic position of the two *foci* in that standard focus undergoes movement to [Spec, CP], while mirative moves to a lower position, probably within, in his terms, IP. The crucial difference between (21) and the other mirative structures is that in the former only intonation can mark surprise. The subject *Jonek* 'Jon' is separated from the following element by an intonational break. However, intonation alone cannot convey a focus reading (A. Elordieta, 2001), and the focalized subject has to move to check whatever feature is present in this IP-internal projection.

According to a number of scholars (Cinque, 1999; Alboiu, 2002; Giorgi, 2016 a.o.) epistemic elements, such as adverbs like 'probably' or 'surprisingly,' which represent the point of view of the speaker over the propositional content of the clause, are hosted in a TP (or IP) internal position. Trotzke and Monforte (2019) observe that certain Basque discourse particles (e.g., the interrogative *ote*) occupy a T-internal position,

similarly to the particles found in German *Mittlefeld*. It is reasonable to think that, if these hypotheses are on the right track, the mirative interpretation can be linked to the same portion of the clause, which hosts the pragmatic expressions described in (19) and (20). The nature of these projections varies among the authors. However, the central observation is that, as the CP-layer represents the link between the propositional content and the discourse, the low periphery might have a similar function related to the speaker's presuppositions or opinions over the event.

4.4 *Speaker-oriented head(s)* According to Uriagereka (1995), the presence of a speaker-oriented head that indicates their point of view over the event is not only plausible but necessary to account for the partial freedom in word order found in discourse-configurational languages. As mentioned at the beginning of the paper, these are languages "in which primary sentence articulation is motivated by discourse-semantic, rather than theta role or case" (É. Kiss, 1995). The hypothesis put forward by Uriagereka (1995) is that the category – baptized F short for *Functional* – is the syntactic representation of information-theoretic issues since, according to the author "information theoretic matters do not carry enough conceptual weight for us to posit a separate level of representation […] (and) they affect configurations" (p.155). Another proposal for a speaker-oriented head, which encodes the spatial and temporal coordinates of the speaker, has been made by Giorgi (2010). The proposal has been put forward to account for *double access reading* phenomena, i.e., structures that require two temporal anchoring to be interpreted (**due anni fa mi ha detto che è incinta* 'two years ago she told me she is pregnant'). In Italian, the two anchoring are not simultaneously possible, while in other languages, e.g., Romanian, they are. The projection in question is hosted in the highest portion of the CP area, and this is consistent with the idea that CP connects the context and the clause. I leave the discussion of the advantages and shortcomings of these possibilities for future researches; however, it is crucial to observe how the role of the speaker in syntax is not necessarily a matter of performance, but it can be represented in syntax. The evidence from mirative focalization in Basque presented here seems indeed to point to the direction in which the relationships with the extra-linguistic reality affect the syntactic derivation, as explicitly advocated by Uriagereka (1995). I assume here that Uriagereka's F-head, is indeed the *locus* that encodes the point of view of the speaker over the propositional content. Consequently, it is the landing site of epistemic expressions or, when not present, the focalized item itself in Mirative structures.

4.5 *Mirative focus* The observations made to account for corrective focus can be transferred to the mirative structures discussed in section 2. Recall (19) reported here as (22):

(22) Anek ez du bada [milioi bat]$_{FOC}$ irabatzi loterian
Ane.erg neg aux.3sg.prs cond million one.abs win.prf lottery.loc
'Ane won a million at the lottery!

In the same informal data collection described in 4.1., I asked my informants to produce the appropriate structure to convey a piece of specific information (Ane winning a million at the lottery) to a friend. In the data, there is no instance of V2-configuration or sentence-final one. In fact, in (22), the focalized item is in a T-internal position, lower than the topicalized subject and lower than the negation. The mirative expression *ez du bada* is the element that signals the surprise of the speaker (or the presupposed surprise of the hearer) about the amount won by Ane. From an interpretational point of view, I showed in table 1 that mirative *foci* could not have an antecedent in the discourse, and this provides evidence for analysis as a clause-internal focus position. Let us now discuss the marked case in which the focalized item is not immediately adjacent to the verb in (21), reported here as (23):

(23) [Jonek]$_{FOC}$ ardoa ekarri du
Jon wine bring aux
'[Jon]$_{FOC}$ brought wine! (he never does so!)' Irurtzun, 2016 [34]

The expected word order would be *O* > *S* > *V* so that the subject *Jonek* 'Jon' could be singled out as the most prominent element of the clause. However, in (19) these features are carried on by the exclamative so that the focalized item remains in its first landing site, i.e., the low FocP (24):

(24) [$_{TP}$ Anek$_i$ [$_{F\text{-}P}$ ez du bada [$_{FocP}$ milioi bat$_j$ [$_{Foc°}$ irabatzi] [$_{vP}$ t$_i$ t$_j$ irabatzi loterian]]]]

In (24), since no other element signals mirativity, it is reasonable to think that the correspondent feature – plausibly [mirativity] – needs to be checked by the focus itself that after having checked its [focus] feature, moves further to a higher position. Assuming the presence of a speaker-oriented head as in Uriagereka (1995), this would be the landing site of those elements that need to check the relevant informational feature against the said head (25).

(25) $[_{TPi}\ [_{F\text{-}P}$ Jonek$_j$ $[_{FocP}\ t_j\ [_{Foc^\circ}\ ]\ [_{vP}\ t_j$ ardoa ekarri du]]]]

The F-Phrase has been defined by Uriagereka (1995) as a functional projection encoding a [point of view] feature responsible for the interpretation of any information-theoretical element. The nature of the feature(s) that this head bears, requires further and extensive cross-linguistical investigation. For the time being, let us consider (22) and (23) as evidence for a clause internal position for focus in Basque that hosts (at least) mirative focalizations, which are strictly connected to the point of view of the speaker. The relevant feature that distinguishes the latter to corrective focus is the fact that background information cannot be previously mentioned.

3.6 *Final Remarks* The structures discussed in this section point to a very clear conclusion: despite the apparent freedom in word-order, informationally relevant features are responsible for the different distribution of, at least, two types of *foci*. Their distribution is affected by the syntactic position in which they appear: contrastive *foci* are hosted in the left periphery of the clause, while mirative *foci* in the low-periphery. The relevant feature for which the two types differ is the requirement of an antecedent in the discourse. The prediction is that whenever a focalized item needs to be anchored to the previous discourse, it needs to move to the left-periphery, while when it is anchored to the internal believes of the speaker, it moves to the T-internal focus position. Consequently, whenever an element appears in the left periphery of the clause, it will be interpreted relative to the previous discourse or the external context. It is reasonable to say that correctivity and mirativity are the two extreme positions on a continuum formed by a number of possible ways of anchoring to a discourse-antecedent. For instance, the answer to a *wh*-question does not need an antecedent in the discourse, but it is nevertheless related to the previous utterance. On the other hand, in those cases in which an element is considered to be informationally relevant by the speaker, the same type of focus has different requirements in terms of anchoring. We are not expecting, thus, a clear cut between different types of *foci* but a way to better understand the distribution of *foci* in Basque.

In the spirit of Neeleman et al. (2009), I consider *foci* as formed by feature bundles that further define the information import of a structure. In the cases analyzed in this section, a [corrective] feature needs to be checked in the left periphery, and this is achieved via movement of the focus to the CP area. The low focus position provides the landing site for *foci* that do not require an antecedent in the discourse. If an epistemic feature, in the case discussed here [mirativity], in F-P is not checked, it is the focalized item that moves there. The hypothesis of having a low periphery of the clause that encodes epistemic information can better capture the idea of 'speaker's preferences' as put forward in Elordieta (2001). Moreover, it can shed some light on the way information structure affects syntax, especially regarding word-order preferences in discourse-configurational languages.

5 Conclusion

The main goal of the discussion in this paper was to show that the investigation of focus typology is crucial in the analysis of word-order preferences in Basque. In particular, I have analyzed some corrective and mirative focus-structures that provide evidence for two positions for focus in Basque. The movement of focus to the relevant portion of the clause is driven by features-checking requirements. The main predictions made here are related to the nature of the information encoded in the two peripheries of the clause: discourse-related in the left periphery and epistemic in the low periphery. More data needs to be collected both for Basque and cross-linguistically, in order to confirm or disprove such predictions. However, it appears clear that information structure does affect syntactic operations, and it needs to be taken into account when discussing information-related phenomena.

References

Alboiu, Gabriela (2002). *The Features of Movement in Romanian*. Bucharest: University of Bucharest Press.

Arregi, Karlos (2003). *Focus on Basque movements*. Cambridge, MA: MIT PhD dissertation.

Belletti, Adriana (2004). Aspects of the Low IP Area. Rizzi, L. (ed.), *The Structure of CP and IP*. New York: Oxford University Press, 16-51.

Bianchi, Valentina, & Bocci, Giuliano, Cruschina, Silvio (2016). Focus Fronting, Unexpectedness, and the Evaluative Dimension. *Semantics and Pragmatics*, 9, 1-54.

Bianchi, Valentina (2015). "Focus Fronting and the Syntax-Semantics Interface". Shlonsky, Ur (ed.), *Beyond the Functional Sequence*. Oxford: Oxford University Press, 60-72.

Büring, Daniel (2016). *Intonation and meaning*. Oxford: Oxford University Press.

Chomsky, Noam (1995). *The Minimalist Program*. MIT Press.

Chomsky, Noam (2001). Derivation by phase. In Kenstowicz, M. (ed.) *Ken Hale: A life in language*. Cambridge: MIT Press, 1-52.

Chomsky, Noam, Gallego, Ángel, & Ott, Dennis (2019). Generative Grammar and the Faculty of Language: Insights, Questions, and Challenges. *Catalan Journal Of Linguistics, 0*, 229-261.

Cinque, Guglielmo (1993). A Null Theory of Phrase and Compound Stress. *Linguistic inquiry*. 24, pp. 239-297.

Cinque Guglielmo (1999). *Adverbs and Functional Heads. A Cross-linguistic Perspective*. New York: Oxford University Press.

Cruschina, Silvio (2012). *Discourse-Related Features and Functional Projections*. Oxford: Oxford University Press

Dal Farra, Chiara (2018). Towards a fine-grained theory of Focus. *Annali di Ca' Foscari. Serie occidentale.* 52, Venezia: edizioni Ca' Foscari.

DeLancey, Scott (1997). Mirativity:The grammatical marking of unexpected information. *Linguistic Typology*. 1:33-52.

DeLancey, Scott (2001). The mirative and evidentiality. *Journal of Pragmatics* 33 (3). 369-382.

Elordieta, Arantzatzu (2001). *Verb movement and constituent permutation in Basque*. Leiden: Leiden University dissertation.

Etxeberria, Uribe; Etxepare, Ricardo (2008). Izen eta gertakarien gaineko kuantifikazioa euskaraz. Artiagoitia,X. Lakarra, J. (eds) *Gramatika Jaietan. Patxi Goenagaren omenez*. Bilbao: University of the Basque Country. Supplements of the International Journal of Basque Language and Linguistics 51.259-286.

Giorgi, Alessandra (2016). Epistemic adverbs, the prosody-syntax interface and the theory of phases. *Selected papers from the 43rd Linguistic symposium on Romance languages*. John Bemjamins, 99-117.

Giorgi, Alessandra (2010). *About the Speaker: Towards a Syntax of Indexicality*. Oxford: Oxford University Press.

Grewendorf, Günther (2005). The discourse configurationality of scrambling. In J. Sabel & M. Saito (Eds.) *The free order phenomenon: Its syntactic sources and diversity*, (pp. 75-135). Berlin, New York: Mouton de Gruyter.

Haddican, William, & Elordieta, Arantzatzu (2013). V≥2 in Basque. *University of Pennsylvania Working Papers in Linguistics*. 19, (1).

Hualde, José Ignacio, & Elordieta Gorka, Elordieta Arantzatzu (1994). *The Basque dialect of Lekeitio*. Bilbao: Euskal Herriko Unibertsitatea.

Hill, Nathan W. (2015). Hare lo: the touchstone of mirativity. *SKASE Journal of Theoretical Linguistics*. 13 (2): 24-31.

Irurtzun, Aritz (2016). Strategies for argument and adjunct focalization in Basque. In Fernández, B., & Ortiz De Urbina, J. (eds.), *Microparameters in the grammar of Basque*. Amsterdam/Philadelphia : John Benjamins. 243–263.

Jayaseelan, K. A. (2001). IP-internal topic and focus phrases. *Studia Linguistica*, 55: 39-75.

Kiss, Katalin. É. (1998). Identificational Focus Versus Information Focus. *Language*, 74 (2), 245-73.

Kratzer, Angelika, & Selkirk, Elisabeth (2020). Deconstructing Information Structure. lingbuzz/004201

Laka, Itziar (1996). A brief grammar of Euskara, the Basque language. www.ei.ehu.es/p289content/eu/contenidos/informacion/grammar_euskara/en_doc/index.html

Laka, Itziar, & Erdocia, Kapa (2012). Linearization preferences given "free word order"; subject preferences given ergativity. A look at Basque. In Torrego, E. (ed.), *Of grammar, words, and verses: In honor of Carlos Piera*, 115-142, Oxford: Oxford University Press

López, Luis (2009). *A Derivational Syntax for Information Structure*. Oxford: Oxford University Press.

Ortiz de Urbina, Jon (1994). Verb initial patterns in Basque and Breton. *Lingua*, 94. 125-153

Ortiz de Urbina, Jon. (1999). Focus in Basque. *The grammar of focus,* 24. 311-333.

Ortiz de Urbina, Jon (2002). Focus of correction and remnant movement in Basque. *Erramu Boneta: Festschrift for Rudolf P.G. De Rijk*. Artiagoitia, X.; Goenaga, P.; Lakarra, J. (eds.) Bilbao: Univerdad del Pais Vasco/Euskal Herriko Unibertsitatea, 511-524.

Rizzi, Luigi (1997). The Fine Structure of the Left Periphery. In Haegeman, L. (ed.), Elements of Grammar. *Handbook in Generative Syntax*. Dordrecht: Kluwer, 281-337.

Sainz-Maza Lecanda, Lorena (2017). *Interactions among focus, exhaustivity, and constituent order in Spanish and Basque*. Ph.D. Dissertation: The Ohio State University.

Zubizarreta, Maria Luisa (1998). *Prosody, Focus, and Word Order*. Cambridge, MA: MIT Press.

When Syntax Transforms a Function Word: The Case of Negation

Matteo Greco
Scuola Universitaria Superiore Iuss di Pavia

1 Introduction

Function words are commonly considered to be a small and closed class of words in which each element is associated with a specific and fixed logical meaning (Chierchia 2013). For example, AND expresses coordination and OR expresses disjunction. Undoubtedly, one of the most intriguing functional words is negation, which has been deeply studied since Aristotle (Horn 1989; Speranza & Horn 2012).

From the semantic point of view, negation can be considered a one-place operator reversing the truth-value conditions of the proposition associated to the sentence in which it occurs: if it is true that it is raining, then it is false that it is not raining, and vice versa (more in general, negation can be regarded as a complement-set operator, see Krifka 2010 and Delfitto 2013). However, it is increasingly clear that negation has a rich pragmatic interpretation in natural languages, sometimes even losing the power to deny a proposition, as witnessed by the following Italian sentences:

(1) a. Rimarrò alla festa finché arriva Gianni
stay.1S.FUT to-the party until arrives John
'I will stay at the party until John arrives'

b. Rimarrò alla festa finché *non* arriva Gianni
stay.1S.FUT to-the party until *neg* arrives John
'I will stay at the party until John arrives'

The occurrence of negation in (1) seems to be optional, at least from a genuine semantic point of view, instantiating what has been called *Expletive Negation* (see, among many others, Jespersen 1917; Horn 1989, 2010; Yoon 2011; Makri 2013; Greco 2020), henceforth EN. Starting from data like (1), it seems reasonable to ask whether a word that is established based on its function can be ambiguous or, putting in different words, whether the matching between a (function) word and its function is just employed in some cases. According to (1), the answer to this question seems to be affirmative, however, this contribution aims to challenge it by discussing the role of the syntactic derivation in some instances of Italian EN.

2 The puzzle in the puzzle: a twofold partition of EN

This paper aims to address a very specific question about negation. The sentences in (1) give us the first piece of the puzzle about this particular functional word: the Italian negative marker *non* ('not') is apparently unable to realize its function, i.e. deny a sentence. Crucially, the puzzle is even more complicated since it is possible to detect two subclasses of EN (Greco 2019a) on the ground of a different syntactic and semantic behavior.

* I am very grateful to those who attended the WECOL 2019, particularly the organizers for their amazing job (and their flexibility with me). I also would like to thank Denis Delfitto, Robert Frank, Andrea Moro, Massimo Piattelli Palmarini, Cecilia Poletto, Alessandra Tomaselli, Raffaella Zanuttini and Hedde Zeijlstra for many helpful observations on my work. Finally, my gratitude goes to Luigi Giraffa for his comments on this paper. Of course, all mistakes are mine!

Proceedings of *WECOL 2019*

Let us consider a preliminary discussion on standard and expletive negation. It is well known in the literature (Yoon 2011; Makri 2013; Greco 2020), that languages allowing EN can either display different markers for expletive and standard negation – like the case of Late Middle English in (2) – or display the same marker for both negations – like Italian in (1).

(2) I drede *not* pat *ne* pe curs of God [...] wolde brynge me into a ful yitel
I doubt not_{SN} that neg_{EN} the curse of God would bring me into a very evil
eende if I contynuede pus
end if I continued thus
'I do not doubt that God's curse would bring me to a very evil end if I continued like this' (Testimony of William Thorpe 482, in van der Wurff 1999: 295)

In (2), the semantics of the sentence suggests which is the function of each negative marker: *ne* does not reverse the polarity of the sentence, falling in the EN class, whereas *not* does, falling in the SN one. Unfortunately, it is not always so easy to distinguish the function of a negative marker, particularly when the same morpheme realizes both expletive and standard negation, like in Italian exclamatives:

(3) Che cosa non ha mangiato Gianni!
what *NEG* has eaten John
a. 'What has John eaten!' Expletive negation
b. 'What has not John eaten!' Standard negation

According to Delfitto and Fiorin (2014), one of the possible interpretations of the exclamative sentence in (3) is that the speaker wants to highlight the fact that J. ate everything and this was surprising; however, an alternative reading is the one in which the speaker wants to highlight the fact that J. did not eat something – such as a very expensive lobster that was offered for free – and this was surprising too. A way to discriminate between the two readings is to refer to items that depend on the polarity of a sentence, such as negative polarity items (NPIs) and the so-called "not-also conjunctions". It is well known that NPIs as '*affatto'* ('at all') and not-also conjunctions as "*e ne-anche*" ("and not-also") can occur only within the domain of a negative operator (4a-a') (Espinal (1997); Israel (1997); Giannakidou 2000; Zeijlstra 2004; Delfitto and Fiorin 2014), yielding ungrammaticality when they occur in pure EN structures (Greco 2019a, 2020) – as witnessed by *until* sentences (4b-b') – ruling out the expletive reading when they occur in ambiguous structures such as exclamatives (4c-c'):

(4) a. Luca *(non) ha *affatto* capito
Luke NEG has at all understood
'Luke has not understood at all'
a'. Luca *(non) ha salutato *e ne-anche* Gianni
Luke NEG has greeted and not-also John
'Luke has not greeted and John has not greeted too'
b. Rimarrò alla festa finché non arriva (**affatto*) Gianni
stay.1S.FUT to-the party until NEG arrives at all John
'I will stay at the party until John arrives'
b'. Rimarrò alla festa finché non arriva Gianni (**e ne-anche) /*
stay.1S.FUT to-the party until NEG arrives John and not-also
e anche Luke
and also Luke
'I will stay at the party until John arrives and Luke too'
c. Che cosa non ha *affatto* mangiato Gianni!
what *NEG* has at all eaten John
(i). '*What has John eaten!' ~~Expletive negation~~
(ii). 'What has not John eaten at all!' Standard negation
c'. Che cosa non ha mangiato Gianni *e neanche* Luca!
what *NEG* has eaten John and not-also Luke

(i). '*What has John eaten and Luke too!' ~~Expletive negation~~
(ii). 'What has not John eaten and Luke has not eaten too!' Standard negation

Surprisingly, if we consider some other items depending on the polarity of a sentence, such as weak NPIs and neg-words (cf. among others, Laka 1990; Zeijlstra 2004; and Chierchia 2013), the pattern gets more complicated. The SN sentences allow both weak NPIs as *alzare un dito* ('lifted a finger') and neg-words as *nessuno* ('nobody') (5a.a'), whereas EN sentences show a twofold behavior: *until* sentences allow them (5b-b'), exclamatives do not (5c-c'):

(5) a. Luca *(non) ha *alzato un dito* per aiutarmi
Luke NEG has lifted a finger to help-me
'Luke has not lifted a finger to help me'
a'. Luca *(non) ha aiutato *nessuno*
Luke NEG has helped nobody
'Luke has not helped anybody'
b. Rimarrò alla festa finché Gianni non avrà
stay.1S.FUT to-the party until john NEG have. 3S.FUT
alzato un dito per aiutar-mi
lifted a finger to help-me
'I will stay at the party until John lift a finger to help me'
b'. Rimarrò alla festa finché non arriverà *nessuno* ad aiutar-mi[1]
stay.1S.FUT to-the party until NEG come.3S.FUT nobody to help-me
'I will stay at the party until someone comes to help me'
c. *Chi non ha *alzato un dito* per aiutar-mi!
who NEG has lifted a finger to help-me
c'. Che cosa non ha mangiato *nessuno*!
what NEG has eaten nobody

Crucially, this pattern is consistent with all Italian EN structures, constituting a partition within the class of EN (Greco 2019a). According to this work, EN clauses that allow both weak NPIs and neg-words fall into the class of *weak EN*, EN clauses that do not allow weak NPIs and neg-words fall into the class of *strong EN*, since the former maintain a SN-like behavior, the latter do not.

With regard to our discussion, this twofold partition within the class of EN seems to confirm the ambiguous nature of negation, suggesting an even more complicated scenario: the functional word '*non'* ('not') can instantiate a case of standard negation, as well as a case of expletive negation, and, in this latter situation, it can either be of the *weak* type or the *strong* one. It seems that *non* has three (apparently) different functions even though it realizes the same functional word. Theoretical speaking, it might be the case that three lexical words are homophonous, but this seems, at least, implausible. An alternative is to consider the negative marker *non* as the realization of the only one negation, which displays several non-standard interpretations. Even though this paper adopts a syntactic view, it reaches the same conclusion discussed in Delfitto, Vender, and Melloni (2019) based on the semantics of EN structures: negation is one. Before discussing it, let us consider the case of a particular EN structure, namely the Surprise Negation sentences (SNEGs).

3 Surprise Negation Sentences

So far, we considered two instances of EN structures: the *until* sentences in (1) and the exclamative sentences in (3). We saw that they belong to two subclasses of EN, and further differences can be detected between them. For example, the negative marker is ambiguous in exclamative sentences between the

[1] Some Italian speakers do not accept this sentence because of "*nessuno"*, but they accept if it is changed with "*qualcuno*" (someone). This can be explained by advocating to the regional sub-Italian type to which people belong to. Moreover, differences in the grammaticality judgments are often associated to EN, particularly in languages in which the same negative marker realizes both EN and SN (see Tubau et al. 2017 for Catalan).

standard and the expletive interpretation, whereas it is not in *until* clauses: *non* is just a genuine case of EN. If we look closer to the very rich system of EN structures in Italian, we can find another interesting case: the Surprise Negation sentences (Greco 2019b).

(6) E non mi è scesa dal treno Maria?!
AND EN CL.to me is got off-the train Mary
'Mary got off the train!'

The meaning of (6) could be fully paraphrased as 'That Maria got off the train is a surprise'. The sentence is affirmative regardless of the occurrence of the negative marker "*non*" ("not") and, therefore, it has to be considered an EN clause. Pragmatically, SNEGs are limited to a restricted contest in which speakers are struck by surprising facts (hence, the label "Surprise") and they want to communicate it. Since SNEGs display a marked intonation blending the acoustic features pertaining to both questions and exclamatives, they usually show the combined diacritic "*?!*". SNEGs are often introduced by the expletive coordination marker '*e'* ('and') (Poletto 2005) and they host ethical datives – in this case, '*mi'* (CL.to me), which make SNEGs more natural.

Crucially, SNEGs do not have any affirmative counterpart, unlike *until* clauses: without negation, there is not Sneg at all. Moreover, negation cannot be interpreted as standard in SNEGs, unlike exclamatives. These two facts lead to another puzzle: *non* is both mandatory and vacuous in SNEGs. This puzzle can be clarified by analyzing some grammatical features of SNEGs, which call for a syntactic analysis. Such an analysis will be able to simplify the apparent complexity of the phenomenon by proposing a unified way to treat negation.

3.1 *Some grammatical features of SNEGs* First of all, it might be worth considering the fact that SNEGs belong to the strong class of EN: they do not allow either negative polarity items nor neg-words.

(7) a. *E Maria non mi ha *alzato* *un* *dito* per aiutare Luca?!
AND Mary EN CL.to me has lifted a finger to help Luke
b. *E non mi è sceso dal treno *nessuno*?!
AND EN CL.to me is got off-the train nobody
c. *E non mi è *affatto* scesa dal treno Maria?!
AND EN CL.to me is at all got off-the train Mary
d. *E non mi è scesa dal treno Maria *e* *neanche* Gianni?!
AND EN CL.to me is got off-the train Mary and not-also John

Moreover, SNEGs asymmetrically host discourse-related phenomena, i.e. topicalization and focalization: they allow the former (8a), but not the latter (8b).

(8) a. E *il* *libro* non me lo ha dato a Luca?!
AND the book.TOP EN CL.to me CL.it has given to Luke
'The book, s/he gave it to Luke!'
b. *E LA PENNA non mi ha dato a Luca (non il libro)?!
AND the pen.FOC EN CL.to me has given to Luke not the book

There are no reasons for which SNEGs should show such a behavior, unless we consider topicalization and focalization as the result of two different syntactic transformations, *a là* Rizzi (1997) (see also, among many others, Cruschina 2012 and the references cited there), interacting with the syntax of SNEGs. Assuming Rizzi (1997, 2001), CP is the landing site for topicalized and focalized phrases:

(9) CP…Force°… (TopP°)… Int°… (TopP°)… Foc°… (Top°)… Fin°…TP

Crucially, SNEGs are incompatible with all the structures involving focalization, such as Mirative fronting focus (Cruschina 2012), Quantifier Fronting (Quer 2002; Benincà & Poletto 2004) and Anaphoric Anteposition (Benincà & Poletto 2004):

(10) a. *Ora ricordo. E [una sciarpa rossa]$_i$ non mi ha regalato t_i Gianni
now remember.1ST.S.PRES AND a scarf red EN CL.to me has given John
per Natale?! (*Mirative fronting focus*)
for Christmas

b. *E *qualcosa$_i$* non mi ha concluso t_i Luca stando qui?! (*QP-fronting*)
AND something EN CL.to me has accomplished Luke being here

c. A: Gianni è andato al bar invece di andare a scuola (*Anaphoric Anteposition*)
'John went to a bar instead going to school'
B: *E *questo$_i$* non mi ha fatto t_i Gianni?!
AND this EN CL.to me has done John

Let us consider another fact regarding SNEGs: the interaction with other expletive objects, such as the conjunction "*e*" ("and") (Poletto 2005).

Even though the Italian morpheme *e* is the logic operator corresponding to "*and*", in some cases, it does not seem to connect anything:

(10) E che cosa dovrei fare?
AND what should.1S do
'What should I do then?'

Since its semantic contribution seems to be empty, this kind of coordination has been called *expletive "e"*, henceforth EE. Interestingly, the occurrence of EE in SNEGs is very spontaneous for Italian speakers, even though it is not mandatory. According to Poletto (2005), EE is a Topic marker occupying a functional projection in the CP:

(11) [$_{TopicP}$ (Null) TopP [$_{Topic0}$ e [$_{TopicP}$ [CP]]]

Moreover, Poletto also proposes that EE can only occur in structures sharing the activation of a specific functional projection in the CP-layer, namely the FocP. Again, it seems that SNEGs interact somehow with the informational properties of a sentence, showing a mixing feeling towards focalization: they do not accept focalized phrases, but they are naturally introduced by objects that depend on focalization.

Finally, the last remarkable fact concerning SNEGs is that the whole proposition conveys completely new information. This appears clear when SNEGs are employed as answers to "Propositional Questions" (Greco 2019b), the variable of which refers to events, as witnessed by the following dialog:

(12) A: Che cosa è successo?
'What's happened?'
B: (E) non mi è scesa dal treno Maria?!
EE EN ED.to me is got off-the train Mary
'Mary got off the train!'

On the other hand, SNEGs cannot be employed as answers to "Entity Questions", the variable of which refers to arguments:

(13) A: Chi è sceso dal treno?
'Who got off the train?'
B: *(E) non mi è scesa dal treno Maria?!
EE EN ED.to me is got off-the train Mary

The fact that SNEGs are exclusively allowed in the former class of questions shows that the whole proposition in SNEGs – and not just a single constituent – carries new information. Crucially, according to

Alonso-Ovalle & Guerzoni (2004) and Brunetti (2004), what carries new information in answers represents the focus of the sentence, showing the same syntactic displacement in the CP-layer than the other focalized structures.

To sum up, we saw that SNEGs represent a very peculiar case of EN: negation is mandatory, but it is expletive; SNEGs interact with the informational contribution of a sentence hosting topicalized structures but not the focalized ones; SNEGs are characterized by the association with other expletive items like the coordination, which depends from the activation of focus.

Let us consider a possible way to shed light on these puzzling behaviors.

3.2 *A syntactic analysis for SNEGs* It is well known from the literature (Belletti 1990; Zanuttini 1997; Poletto 2008), that negation in Italian, and many other languages, occupies a syntactic position in the TP domain (see the cited works for a detailed discussion); I would like to change this assumption and propose that in SNEGs negation is externally merged in the CP-domain (*à la* Laka 1990) when the v*P-phase has already been closed and the entire TP has raised to Spec-Foc° (see Greco 2019b, 2020) (phases are underlined):

(14) $[\underline{_{CP}} \ldots[_{TopP}] [_{X^\circ}$ non $] [_{FocP}$ TP [Foc° … $[_{TP} [\underline{_{v^*P}} [\ldots$ v° … $]]]_i]]]$

An immediate consequence of this syntactic structure is that negation is unable to trigger any negative items – falling in the class of strong EN clauses (see 7) – because they belong to different phases: when *not* is introduced in the CP-field, the domain of the vP is impenetrable and, therefore, negation cannot see inside it. Since, according to Giannakidou (1997) and Zeijlstra (2004), a negative operator must bind all free variables in the vP domain in order to allow them, negation in SNEGs loses this possibility since it belongs to another phase, namely the CP-phase. If this is true, we should expect negation allow those elements requiring a positive polarity, such as '*già'* ('already')[2], and this is the case:

(15) [CP E non [v*P mi è *già* scesa dal treno Maria]]?!
EE EN ED.to me is already got off-the train Mary
'Mary already got off the train!'

This particular configuration of negation also suggests another remarkable consequence: since negation takes FocP as its argument, it applies to the presuppositional layer required by the focalization rather than to the propositional one, in contrast with the SN case (even because the v*P-phase is already closed). This hypothesis goes in the same direction of Benincà (1996), Abels (2002, 2005) and Delfitto, Vender & Melloni (2019). This last work is particularly illuminating for the role of EN: according to the authors, "EN is truth-conditionally irrelevant from the fact that the semantics of negation as a truth-value reversal operator is shifted, in the case of EN, to the layer of implicated meaning". If this is true, we should expect SNEGs hosting other presuppositional-like objects, such as *mica* in Italian (Cinque 1976, Frana & Rawlins 2015), and this is the case:

(16) E non mi è *mica* scesa dal treno Maria?!
EE EN ED.to me is *mica* got off-the train Mary
'Mary got off the train!'

From this point of view, it is extremely convenient to consider expletive and standard negation as the result of different syntactic representations involving the same morphological element: when the negative marker *not* is merged in the TP-domain, it gives the standard negation reading; when it is merged in a higher

[2] The fact that '*già'* ('already') is a positive polarity item is witnessed by the following Italian sentence:
(i) Luca (*non) è già arrivato
Luke neg is already arrived
'Luke has already arrived'
The occurrence of negation would yield ungrammaticality in (i), showing the positive nature un '*già'* (see Giannakidou 2011 for a discussion on positive polarity items)

position, i.e. the CP-field, it gives the expletive negation reading (at least, in the SNEGs case). Therefore, the negative marker is always the expression of the unique functional word associated with negation, it just has different interpretations.

Let us consider now the relation between SNEGs and the focalized phrases: according to (14), the whole TP in SNEGs moves to the focalized phrase, leaving no space for other focalized phases. Crucially, this causes the patter discussed above: on the one hand, no space is available for multiple focalized phrases since Italian has only one functional projection for focalization (Rizzi 1997); on the other hand, expletive *and* is so natural in SNEGs since an active focalization is already present, namely the focalization of the TP. Moreover, this also takes into consideration the fact that the whole proposition conveys completely new information in SNEGs.

Let us take a step back to consider Alonso-Ovalle & Guerzoni (2004) and Brunetti (2004)'s proposal mentioned above. In fact, Italian has two different forms to answer to questions: one is "extended" and one is "short": with the former speakers repeat part of the question under discussion (the topic/background) adding the new information (B); with the latter speakers only pronounce the part carrying the new information neglecting all the topic information (B'):

(17) A: Che cosa ha vinto Gianni?
'What did John win?'
B: [Una maglietta]$_{NEW}$ ha vinto Gianni
a shirt has won John
'John won a shirt'
B': [Una maglietta]$_{NEW}$
'A shirt'

The "extended" form is redundant because it displays part of the background knowledge (in the example "*John won*"); the "short" one is more natural and preferred. According to Alonso-Ovalle & Guerzoni (2004) and Brunetti (2004), the "short" form presents the ellipsis of the background information, namely TP, leaving the pronounced part in FocP, which escapes the ellipsis:

(18) [$_{FocP}$ [$_{DP}$ Una maglietta]$_j$ [$_{TP}$ ~~ha vinto Gianni~~ [t$_j$]]

According to this theory, since SNEGs can be used as answers to questions, a part of them has been moved to FocP. Moreover, the fact that SNEGs can only be employed as answers to propositional questions also suggests that the part that has been moved to FocP is the whole TP and this is exactly what (14) proposes.

To sum up, in this section I tried to show how the complexity of SNEGs can be derived from a particular syntactic derivation. The core aspects of my proposal were that negation, which is usually assumed to be merged in the TP domain, can also be merged in the CP-domain and that the whole TP can be focalized. From this particular conformation, several consequences follow: SNEGs interact with the informational contribution of a sentence hosting topicalized structures but not the focalized ones; SNEGs are characterized by the association with other expletive items like the coordination, which depends from the activation of focus, etc. However, the most remarkable consequence of this hypothesis is that EN and SN follow from the same functional word, which always employs the same function, even though it applies to different layers: to the presuppositional layer in SNEGs, to the propositional layer in SN – reaching the same conclusion discussed in semantics by Delfitto, Vender & Melloni (2019). Let us now speculate applying this idea to the case of exclamatives discussed above.

4 Exclamatives: a speculative discussion

We know from (3) (repeated here as 19) that exclamatives has a twofold interpretation: one in which negation is expletive (a) and one in which it is standard (b). In Greco (2019a-b),[3] the former was labeled "Expletive Negation Exclamative" (ENE), and the latter "Negative Exclamative" (NE):

(19) Che cosa non ha mangiato Gianni!
what *NEG* has eaten John
a. 'What has John eaten!' Expletive Negative Exclamative
b. 'What has not John eaten!' Negative Exclamative

I also recall that ENEs belong to the class of strong EN, since they do not allow any negative objects (5c-c'), reproducing the same pattern as SNEGs. Moreover, further similarities can be found between SNEGs and ENEs, such as the fact that both host positive polarity items like *'già'* ('already') (see 15):

(20) Che cosa non ha *già* mangiato Gianni!
what *NEG* has already eaten John
'What has John already eaten!'

According to Zanuttini & Portner (2003) and Delfitto & Fiorin (2014), this kind of exclamatives also brings a surprise evaluation of the content of the sentence. If we consider the example in (20), the surprise was about the fact that John ate, for example, a remarkable amount of food in a limited spam of time. Again, this surprise evaluation resembles the one of SNEGs.

A possible way to take into consideration both the differences between EN and ENE and the similarities between ENE and SNEGs is to assume that negation realizes one of the two options discussed above: when the negative marker *not* is merged in the TP-domain, it gives the standard negation reading in EN; when it is merged in a higher position, i.e. in the CP-domain, it gives the expletive negation reading in ENE and in SNEGs (phases are underlined).

(21) a. [$\underline{_{CP}}$ … [$\underline{_{v*P}}$ [$_{X^\circ}$ non] …] (NE)
b. [$\underline{_{CP}}$ … [$_{X^\circ}$ non] … [$\underline{_{v*P}}$…] (ENE / SNEGs)

However, SNEGs and ENEs cannot be assimilated to each other. In fact, they differ in many ways (see Greco 2019b), such as the fact that ENEs can host focalized phrases, whereas SNEGs cannot, as witnessed by the following dialogs[4]:

(22) A: Che cosa non ha mangiato Gianni!
what *NEG* has eaten John
'What has John eaten!'
B: Che cosa non ha mangiato LUCA (non Gianni)!
what *NEG* has eaten Luke neg John
'What has Luke eaten, not John!'

(23) A: E non mi è scesa dal treno Maria?!
AND EN CL.to me is got off-the train Mary
'Mary got off the train!'
B: *E non mi è scesa dal treno LAURA (non Maria)?!

[3] Many works focused on exclamatives, among many others, see Zanuttini and Portner 2003; Delfitto and Fiorin 2014 and the references cited there.

[4] I will leave the focalized phrase *in situ* in the next sentences. In fact, Italian realizes focalized structures either moving a phrase to a front position at the beginning of the sentence or leaving it *in situ* with a marked intonation. Since ENEs already have a (wh-) phrase that has been moved in the front position, the latter option has to be preferred in this case. See Belletti (2004) for a detailed discussion on the focalization *in situ*.

AND EN CL.to me is got off-the train Laura neg Mary

The pattern in (22-23) suggests that the functional projection dedicated to FocP is still available in the CP-layer of ENE, whereas it is not in SNEGs. To sum up, the two different structures in SNEGs and ENEs can be schematically described as following (phases are underlined):

(24) Surprise Negation Sentences: [**CP** … [X° non] [FocP TP [Foc° … [TP [**v*P** [… v°…]]]i]]]?!
Expletive Negation Exclamative: [**CP** …[*Che cosa*]i [X° non] [FocP] … [**v*P** [… ti …]]]![5]

Negation is merged in CP in both cases, taking the FocP as an argument; however, the TP is focalized only in SNEGs, ruling out other focalizations, whereas is available in ENE.

Many other data should be described in order to fully discuss the case of Exclamatives and further researches will shed light on their full syntactic derivation. At the moment, the crucial points is that the analysis I gave above is able to take into consideration the different interpretations of negation in exclamatives, without giving up a unified approach to this functional word.

5 Concluding remarks

We started this work by asking whether a word that is established based on its function, namely negation, can be ambiguous or, putting in different words, whether the matching between a (function) word and its function is employed just in some cases. This question was raised by the phenomenon of expletive negation, a case in which negation seems to lose its ability to deny the proposition associated with its sentence. However, I proposed that a particular syntactic derivation is able to consider expletive and standard negation in a unified way: when the negative marker *not* is merged in the TP-domain, it gives the standard negation reading; when it is merged in a higher position, i.e. the CP-field, it gives the expletive negation reading. Negative Exclamatives are examples of the former option, whereas Surprise Negation sentences and Expletive Negative Exclamatives are examples of the latter one. Of course, many questions remain untouched, such as the difference between the two classes of expletive negation clauses we discussed in §2. Future studies will shed light on these issues.

References

Abels, Klaus. 2002. Expletive (?) Negation. In Formal Approaches to Slavic Linguistics, ed. Toman Jindrich, vol. 10: 1–21. Bloomington: Michigan Slavic Publications.

Abels, Klaus. 2005. "Expletive Negation" in Russian : A Conspiracy Theory. Journal of Slavic Linguistics 13(1): 5–74.

Alonso-Ovalle, Luis and Elena Guerzoni. 2004. Double negation, negative concord and metalinguistic negation. In Proceedings of CLS 38–1: Main session, eds. Andronis Debenport and Pycha Yoshimura, 15–31. Chicago: CLS Publications.

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

Belletti, Adriana. 2004. Aspects of the low IP area. In The Structure of IP and CP; the Cartography of Syntactic Structures, ed. Luigi Rizzi, Vol. 2: 16–51. New-York & Oxford: Oxford University Press.

Benincà, Paola. 1996. La struttura della frase esclamativa alla luce del dialetto padovano. In Italiano e dialetti nel tempo: saggi di grammatica per Giulio C. Lepschy, eds. Paola Benincà, Guglielmo Cinque, Tullio De Mauro, and Nigel Vincent, 23–43. Roma: Bulzoni.

Benincà, Paola and Cecilia Poletto. 2004. Topic, focus and V2: defining the CP sublayers. in The Structure of CP and IP: the Cartography of Syntactic Structures, ed. Luigi Rizzi, 52–75. New York & Oxford: Oxford University Press.

Brunetti, Lisa. 2004. A unification of Focus. Padova: Unipress.

Chierchia, Gennaro. 2013. Logic in grammar: Polarity, free choice, and intervention. New York & Oxford: Oxford University Press - Oxford Studies in Semantics and Pragmatics.

[5] According to this representation, the subject of ENE should occur between *non* and the T0: "Che cosa non Gianni ha mangiato! (Lit. What neg John has eaten). However, this yields ungrammaticality. Greco (2019) addresses this problem. In a nutshell, according to Zanuttini (1996), the Italian negative marker *non* is a critic-like element and it has to be adjacent to a verb, leaving no space for other elements in between. This requirement also ruled out the interposition of the subject in the Exclamative sentence in (23).

Cinque, Guglielmo. 1976. Mica. Annali della Facolta di Lettere e Filosofia 1: 101–12 (Reprinted in Cinque 1991).
Cruschina, Silvio. 2012. Discourse-Related Features and Functional Projections. Oxford & New York: Oxford University Press.
Delfitto, Denis and Gaetano Fiorin. 2014. Negation in exclamatives. Studia Lingüística 68(3): 284–327.
Delfitto, Denis, Melloni, Chiara and Maria Vender. 2019. The en(rich)ed meaning of expletive negation. Evolutionary Linguistic Theory, 1, 57-89.
Espinal, M. Teresa. 1997. Non-negative negation and Wh-exclamatives. In Negation and polarity; Syntax and Semantics, eds. Danielle Forget, Paul Hirschbühler, France Martineau, and María Luisa Rivero, 75–93. Amsterdam: John Benjamins.
Frana, Ilaria and Kyle Rawlins. 2015. Italian 'mica' in assertions and questions. In Proceedings of Sinn und Bedeutung 20, eds. Nadine Bade, Polina Berezovskaya, and Anthea Scholler, 234–251.
Horn, Laurence R. 1989. A Natural History of Negation. Chicago. Chicago: University of Chicago Press.
Horn, Laurence R. 2010. The expression of negation. Berlin-New York: Mouton de Gruyter.
Giannakidou, Anastasia. 2000. Negative ... Concord? Natural Language and Linguistic Theory 18: 457–523.
Giannakidou, Anastasia. 2011. Negative and positive polarity items: licensing, compositionality and variation. In Semantics: an international handbook of natural language meaning, eds. Claudia Maienborn, Klaus von Heusinger, and Paul Portner, 1660–1712. Berlin: Mouton de Gruyter.
Greco, Matteo. 2019a. Is expletive negation a unitary phenomenon? Lingue e Linguaggio (1): 25–58.
Greco, Matteo. 2019b. On the syntax of surprise negation sentences: A case study on expletive negation. Natural Language and Linguistic Theory, doi: 10.1007/s11049-019-09459-6.
Greco, Matteo. 2020. The Syntax of Surprise: Expletive Negation and the Left Periphery. Newcastle: Cambridge Scholars Publishing.
Horn, Laurence R. 1989. A Natural History of Negation. Chicago. Chicago: University of Chicago Press.
Horn, Laurence R. 2010. The expression of negation. Berlin-New York: Mouton de Gruyter.
Israel, Michael. 1997. The scalar model of polarity sensitivity: the case of the aspectual operators. In Negation and polarity: Syntax and Semantics, eds. Danielle Forget, Paul Hirschbühler, France Martineau, and María Luisa Rivero, 209–229. Amsterdam-Philadelphia: John Benjamins.
Jespersen, Otto. 1917. Negation in English and other languages. Copenhagen: A. F. Høst.
Krifka, Manfred. 2010. How to interpret 'expletive' negation under bevor in German. In Language and Logos: Festschrift for Peter Staudacher on his 70th Birthday, eds. Thomas Hanneforth and Gisbert Fanselow, 214–236. Berlin: Akademie Verlag.
Laka, Itziar. 1990. Negation in Syntax: On the Nature of Functional Categories and Projections. Ph.D. Thesis, MIT.
Makri, Maria-Margarita. 2013. Expletive Negation beyond Romance. Clausal Complementation and Epistemic Modality. MA thesis, University of York.
Poletto, Cecilia. 2005. *Sì* and *E* as CP expletives in Old Italian. In Grammaticalization and Parametric Variation, eds. Montserrat Batllori, Maria-Lluïsa Hernanz, Carme Picallo, and Francesc Roca, 206–235. Oxford: Oxford University Press.
Poletto, Cecilia. 2008. "On negative doubling." Quaderni di lavoro dell'ASIt 8: 57–84.
Quer, Josep. 2002. Edging Quantifiers. On QP Fronting in Western Romance. In Romance Languages and Linguistic Theory 2000, eds. Claire Beyssade, Reineke Bok-Bennema, Frank Drijkoningen, and Paola Monachesi, 253–270. Amsterdam: John Benjamins.
Rizzi, Luigi. 1997. The fine structure of the left periphery. In Elements of Grammar, Handbook of Generative Syntax, ed. Liliane Haegeman, 281–337. Dordrecht: Kluwer.
Rizzi, Luigi. 2001. On the position Int(errogative) in the left periphery of the clause. In Current studies in Italian syntax. Essays offered to Lorenzo Renzi, eds. Gugliemo Cinque and Giampaolo Salvi, 287–296. Amsterdam: Elsevier.
Speranza, J. L. and Laurence Horn L. 2012. A brief history of negation. In Logic: A History of its Central Concepts, eds. Dov M. Gabbay, Francis Pelletier, and John Woods, 11: 127–174. Amsterdam: Elsevier.
Tubau, Susagna, Viviane Déprez, Joan Borràs-Comes, and M. Teresa Espinal. How speakers interpret the negative markers *no* and *no… pas* in Catalan. Probus. doi: 10.1515/probus-2017-0008.
Yoon, Suwon. 2011. 'Not' in the Mood: The Syntax, Semantics and Pragmatics of Evaluative Negation. Ph.D. Thesis, University of Chicago.
van Der Wurff, Wim. 1999. On expletive negation with adversative predicates in the history of English. In Negation in the History of English, eds. Ingrid Tieken-Boon van Ostade, Gunnel Tottie, and Wim van der Wurff, 295–328. Berlin: Mouton de Gruyter.
Zanuttini, Raffaella. 1996. On the Relevance of Tense for Sentential Negation. In Belletti & Rizzi In Parameters and functional heads, eds. Adriana Belletti and Luigi Rizzi, 181–207. Oxford & New York: Oxford University Press.

Zanuttini, Raffaella. 1997. Negation and Clausal Structure. A comparative study of Romance Languages. Oxford – New York: Oxford University Press.
Zanuttini, Raffaella and Paul Portner. 2003. Exclamative clauses at the syntax–semantics interface. Language 79(1): 39–81.
Zeijlstra, Hedde. 2004. Sentential negation and negative concord. Ph.D. Thesis, University of Amsterdam.
Zeijlstra, Hedde. 2006. The Ban on True Negative Imperatives. Empirical Issues in Formal Syntax and Semantics 6: 405–424.

Is Italian Swedish? An uncommon look at Italian laryngeal phonology

Bálint Huszthy
Research Institute for Linguistics (Budapest)

1 Background

Even if Italian and Swedish are distantly related languages, they do not show many common structural features in synchrony, except from laryngeal phonology. Laryngeal contrasts in languages challenge phonological theory, since beyond the voiced-voiceless binary distinction, three-way and four-way laryngeal systems also occur, while contrasts are not only based on [voice] but on other features like [spread glottis] and [constricted glottis], too (Iverson & Salmons 1995; Balogné Bérces & Huber 2010).

Laryngeal Realism (Honeybone 2002, 2005; Petrova et al. 2006; Harris 2009; Beckman et al. 2013; etc.) sorts binary laryngeal systems into two categories, according to the markedness of either the [voice] or the [spread glottis] feature. In the category of the so-called *voice languages* (e.g. Slavic, Romance, Hungarian, etc.) the marked feature is [voice], and thoroughly voiced obstruents are in contrast with voiceless ones; on the other hand, in *aspiration languages* (e.g. most Germanic languages, Mandarin Chinese, etc.) the marked feature is [spread glottis], and the laryngeal contrast stays between voiceless unaspirated and voiceless aspirated stops (e.g. Eng. ***b**ack* [p] ~ ***p**ack* [p^{h}]). In these languages obstruent voicing is usually passive, that is, possible only in intersonorant position (between vowels or sonorants); while in voice languages voiced obstruents have their own voice value, which is considered active and so it can spread, evoking *regressive voice assimilation* (RVA; i.e., *regressive devoicing* /bt/ → [pt] and *regressive voicing* /pd/ → [bd]; e.g. Slavic *vo**dk**a* [tk], *Mc**D**onald's* [gd], etc.).

However, certain languages or varieties with a two-way laryngeal distinction do not properly fit with these generalizations. The classic example is Swedish, whose initial voiced obstruents are fully voiced, but voiceless stops are heavily aspirated (Ringen & Helgason 2004; Helgason & Ringen 2008). Apparently, Swedish presents laryngeal patterns typical for both voice and aspiration languages. Huszthy (2019) claims that the case of Italian is a similar phonological riddle. Italian is usually considered a true voice language in the literature, even if its voiced stops are in contrast with mildly aspirated voiceless stops, and voice spreading cannot be detected in the system. The laryngeal patterns found in Italian permit us to treat the language both as an exceptional voice language or as an exceptional aspiration language. In this paper we argue for the latter proposal, which makes Italian laryngeal phonology similar to that of Swedish.

2 The interface of Element Theory and Laryngeal Relativism

The framework this study is couched in is an extension of Laryngeal Realism proposed by Cyran (2011, 2012, 2014, 2017a, 2017b), which he calls *Laryngeal Relativism* (LR); combined with *Element Theory* (ET), a subtheory of melodic representations in *Government Phonology* (Kaye, Lowenstamm & Vergnaud 1985; Harris 1990, 1994; Scheer 2004; Kaye 2005; etc.).

From an ET-approach, the basic typology of two-way laryngeal systems is identical to that of Laryngeal Realism, expressed through unary phonological primes. Harris (1994) claims that in voice languages the element **L** (used for active voice in obstruents) is active in the marked series of obstruents (so this is the category of *L-systems*); while in aspiration languages the element **H** (used for voicelessness or aspiration in obstruents) is the marked laryngeal prime (so this is the category of *H-systems*).

Cyran's LR is based on the idea that as long as a sufficient phonetic distance is kept between two sets of obstruents (e.g., voiced-voiceless, aspirated-unaspirated) to maintain phonological contrast, both the marked and the unmarked sets may receive any (more or less arbitrary) phonetic interpretation. That is, phonetic interpretation is partly phonological. He also claims that it may even be the case that two

Proceedings of *WECOL 2019*

laryngeal systems which are phonetically identical, stem from two phonological settings in which the marked/unmarked relation is reversed.

Cyran (2011, 2012, 2014) argues for the relevance of LR with the two major dialect groups of his mother tongue: Warsaw Polish and Cracow Polish, which differ phonologically but are phonetically identical in terms of laryngeal features (also cf. Rubach 1996). Warsaw Polish is analyzed as an L-system, while the phonetically identical system of Cracow Polish is analyzed as an H-system, with phonologically active **H** rather than **L**. The fundamental difference between the two systems is that Cracow Polish presents a process called "cross-word pre-sonorant voicing", which must be due to its being an H-system with unmarked obstruents that undergo passive voicing in sonorant contexts. Cyran arrives at a typology in which Warsaw Polish, and generally Slavic and Romance languages (as well as Hungarian) are L-systems with evidence of phonologically active **L**; while Cracow Polish (as well as Germanic languages) are H-systems with phonologically represented **H**.

In our case, Cyran had a crucial innovation, which can also be employed for the case of Italian and Swedish, since Cyran redefined the category of H-systems, "originally" corresponding to Germanic-type aspiration languages. In his system, languages like Cracow Polish have active **H** that spreads. While he is able to elegantly treat laryngeal systems with apparent "cross-word pre-sonorant voicing", for instance, in (standard) English and German, no laryngeal activity in the form of any kind of spreading is attested, which rather suggests the absence of any laryngeal element. However, Balogné Bérces & Huber (2010) claim that aspiration is dominant obstruency in aspiration languages, which is represented by the element **h** in ET; consequently, aspiration has to be represented by **h**, which dominates the phonological expression. If aspiration is the special function of **h**, then actually three types of phonological systems can be identified, rather than two: in addition to Cyran's L- and H-systems, Balogné Bérces (2017) also assumes the existence of the category of what she calls *h-systems* (also cf. Balogné Bérces & Huszthy 2018).

The category of h-languages practically coincides with that of aspiration languages in the classical typology, like most varieties of English and German. In these languages there is no laryngeal activity, like voice spreading (i.e., voice assimilation). Although traditional descriptions report utterance-initial and final (partial) "devoicing" and word-medial "bidirectional devoicing", these phenomena in such languages are not processes but the emergence of the underlying form. The category of h-systems can also include languages whose laryngeal phonology has always been a riddle for phonologists, like Swedish. In fact, if the sufficient discriminability in production and perception is a major driving force in the phonetic implementation of phonological contrasts (Cyran 2012), languages with a voiced unaspirated and a voiceless aspirated series may belong to the same h-system. For instance, Swedish exemplifies the typical Germanic pattern, except for the fact that its voiced obstruents are thoroughly voiced even in utterance-initial position, which has led researchers to classify it as a separate category (cf. Ringen & Helgason 2004; Helgason & Ringen 2008). However, Balogné Bérces (2017) suggests that this is only to enhance discriminability to a degree beyond the minimally required "sufficient"; that is, Swedish simply "overshoots" the phonetic distance required for discriminability, but phonologically it remains an h-system: this is indeed a normal effect of Laryngeal Relativism.

When the marked series of obstruents contains the **H** element, we arrive at Cyran's H-systems, i.e., languages or varieties like Cracow Polish (to which we could add, for instance, Slovak, Catalan, West Flemish, Ecuadorian Spanish, that is, varieties with "cross-word pre-sonorant voicing"). Since the phonetic interpretation is assumed to be arbitrary, the presence of **H** does not in itself guarantee the presence of aspiration; at the same time, **H** is a prime that is able to spread; therefore, H-languages will exhibit voice assimilation, unlike h-languages. If such languages also have final obstruent delaryngealization, they also exhibit cross-word passive voicing manifested in "cross-word pre-sonorant voicing". In addition, the non-existence of an inactive laryngeal prime implies that a system which is binary but does not exhibit voice assimilation, will necessarily be an h-system (cf. Balogné Bérces & Huszthy 2018). This may be the case that we also find in the laryngeal system of Italian.

3 The synchronic laryngeal characterization of Italian

The idea to recategorize Italian laryngeal phonology is based on a recent study which argues that Italian phonology lacks regressive voice assimilation in synchrony; and, in fact, preconsonantal s-voicing found in Italian is not RVA but another phonological process due to lenition (Huszthy 2019). Furthermore,

the study also confirms previous phonetic findings, which report aspiration in Italian voiceless stops. These two arguments are more than enough in order to reanalyze Italian in the combined frameworks of LR and ET as an h-language similar to Swedish.

3.1 *The absence of voice assimilation in Italian* Romance languages are all considered L-languages in the LR-typology, and as such, they are supposed to perform RVA. In Romance languages RVA targets any obstruent cluster with differing input voice values, both word-internally and in sandhi position, e.g. (Port.) *Li**sb**oa* [ʒb] 'Lisbon', *de**z p**atos* [ʃp] 'ten ducks' (Mateus & D'Andrade 2000: 142, 145), (Sp.) *fú**tb**ol* [ðβ] 'football' (Colina 2006: 186), (Rom.) *a**ş v**rea* [ʒv] 'I would like' (Wetzels & Mascaró 2001: 220), (Cat.) *ca**p d**au* [bd] 'no dice', *go**s b**o* [zβ] 'good dog' (Recasens 2014: 165).

In the native vocabulary of Italian /sC/ is the only well-formed obstruent cluster, for phonotactic reasons (cf. Morelli 1999: 162; Krämer 2009: 138). Other obstruent clusters were diachronically eliminated from Italian, mostly by deletion or place assimilation; for instance, in the Italian version of the Latin word *a**bst**ra**ct**us* 'abstract' (→ *a**st**ra**tt**o*) both processes take place, and only /sC/ is preserved (cf. Rohlfs 1966). In Italian /s/ undergoes a voicing process before voiced consonantal segments, which may seem a type of RVA, e.g. [sp]*aro* 'gunshot' vs. [zb]*arra* 'barrier', [zv]*eglia* 'alarm clock', [zm]*ettere* 'to stop', [zl]*itta* 'sled', etc. (cf. Nespor 1993: 74-76; Bertinetto 1999: 271; Bertinetto & Loporcaro 2005: 134; Krämer 2009: 209). The laryngeal behavior of Italian /sC/ does not seem exceptional until this point; still, Huszthy (2019) argues that preconsonantal s-voicing does not equal RVA from a synchronic point of view.

If we compare RVA with Italian preconsonantal s-voicing in phonological terms, we find essential differences between the two phenomena at every level of comparison: in the input, in the trigger, in the domain of application and in the obligatory vs. optional character of the processes, as shown in Chart (1) (for details cf. Huszthy 2019).

Chart 1

	RVA	Preconsonantal s-voicing
a. Input:	Any obstruent	Only /s/
b. Trigger:	Segments with distinctive voice (obstruents)	Obstruents and sonorants (and /w/ in loanwords, e.g. [z]*watch*)
c. Domain:	The utterance (postlexical)	The phonological word (lexical)
d. Occurrence:	Obligatory	Optional in medial position (e.g. *i**ce**berg* [sb]/[zb])

RVA is a typically postlexical process, that is, its domain of application is the phonological utterance (Nespor & Vogel 1986: 229-230). On the other hand, Italian preconsonantal s-voicing does not take place at the word boundary, as it is also reported in the literature, e.g. (It.) *rebu**s d**ifficilissimo* [sd] 'a very hard riddle', (It.) *autobu**s b**ianco* [sb] 'white bus' (Nespor 1993: 74); *lapi**s b**lu* [sb] 'blue pencil' (Bertinetto 1999: 271). Sometimes s-voicing is blocked at morpheme boundaries as well, for instance, at the edge of compound words, e.g. *ga**sd**otto* [sd] 'pipeline' (Bertinetto 1999: 280), *i**ceb**erg* [sb] (Huszthy 2019); while it is optional word-internally in loanwords, e.g. *face**b**ook* [sb]/[zb], *fri**sb**ee* [sb]/[zb], *ba**seb**all* [sb]/[zb], etc. (Huszthy 2016, 2019). According to Nespor & Vogel (1986), Nespor (1993), Bertinetto (1999, 2004), Bertinetto & Loporcaro (2005), we can restrict the domain of application of Italian preconsonantal s-voicing to the phonological word, that is, it is not a postlexical process like RVA.

However, the strongest argument for the phonological separation of RVA and Italian preconsonantal s-voicing lies in the fact that Italians do not use RVA in loanwords and in their foreign accent; that is, neighboring obstruents regularly surface in their pronunciation immediately following each other with differing voice values (Huszthy 2019). The argumentation is based on a loanword test (preceded by several experiments of foreign accent analysis, cf. Huszthy 2013, 2014, 2016). 15 Italian informants have been selected from different regions of Italy and recorded in a soundproof studio. They were asked to read out five times 18 Italian sample texts, involving 108 target words (mostly loanwords and foreign proper nouns) containing laryngeal variables. The recordings clearly show that the perfect adjacency of voiced and voiceless obstruents is entirely possible in the Italian pronunciation of loanwords, e.g. *vo**dk**a* [ˈvɔ·dka],

*rag**t**ime* [regˈtajmə], *Mc**D**onald's* [mekˈdɔˑnald], *foo**tb**all* [ˈfutbalːə], *A**fg**anistan* 'Afghanistan' [afˑˈgaːnistan], *gan**gs**ter* [ˈgaˑŋgster], *so**vk**hoz* [ˈsɔːvkod͡z], *e**cz**ema* [ekˑˈd͡zɛːma], etc. (see examples with acoustic analyses in Huszthy 2019).

The overall statistics of the corpus reveal that the 15 Italian informants retain the underlying voice values in the respective obstruent clusters in 65% of the cases; that is, they avoid RVA in a two-thirds majority, which characterizes the performance of all the informants rather evenly. The speakers also use various repair strategies in order to resolve the clusters; for instance, in 15% of all pronounced clusters schwa insertion appears, that is, an epenthetic schwa separates the obstruents, e.g. *ab*[ə]*side*, *up*[ə]*grade*, etc. A further strategy is *progressive devoicing*, which characterizes 17% of the relevant voiceless+voiced inputs, e.g. *foo**tb**all* [tp], *A**fg**anistan* [fk], etc. RVA also appears among the results, in 15% of all obstruent clusters. However, if we zoom into this phenomenon, it seems rather unbalanced as far as the choice between voicing and devoicing is concerned: of all occurrences of RVA, 28% of the cases exhibit voicing, while in 72% of the potential RVA sites devoicing happens (e.g. *foo**tb**all* [db] and *vo**dk**a* [tk], respectively). On the basis of the claims of LR, only regressive voicing is true evidence of RVA, since what appears to be devoicing may not be the result of a process at all but the realization of an underlying voiceless segment which surfaces unchanged in the absence of passive voicing – stemming from the lack of a source element (cf. Section 2). Among the repair strategies used by the Italian informants of the study, regressive voicing is the least preferred solution which, in total, appears only in 4% of the overall results (see Huszthy 2019: 76). On the other hand, in the vast majority of potential RVA contexts the adjacent obstruents surface in the Italian informants' pronunciation with differing voice values and without any repair strategy.

In conclusion, we claim that Italian phonology does not have classical postlexical RVA in synchrony; instead, it has a morphologically conditioned, heavily restricted lexical voicing process, namely preconsonantal s-voicing. At the same time, we also claim that this process is not an instance of RVA, and the synchronic phonology of Italian lacks voice spreading; that is, the Italian laryngeal system is similar to that of an h-language.

3.2 *Italian VOT values* Our other argument for the recategorization of Italian laryngeal phonology – beyond the absence of voice spreading in the system – is aspiration found in the voiceless series of stops. In LR, the question of *voice onset time* (VOT) has a key role, because it is claimed that the aspiration (i.e., a long-lag VOT) found in voiceless stops is not only a phonetic side-effect of languages, but it has serious phonological consequences. In fact, the LR typology of languages is also based on the VOT mean values of voiceless stops. In Chart (2) below usual VOT averages are shown for initial /p, t, k/, in three L-languages and in three h-languages (data derive from Lisker & Abramson 1964).

Chart 2

	a) L-languages			b) h-languages		
	Spanish	Hungarian	Dutch	English	Cantonese	Korean
/p/	4 ms	2 ms	10 ms	58 ms	77 ms	91 ms
/t/	9 ms	16 ms	15 ms	70 ms	75 ms	94 ms
/k/	29 ms	29 ms	25 ms	80 ms	87 ms	126 ms

As it is obvious from the data shown in Chart (2), VOT duration differs according to the place of articulation of the stops: usually the larger the obstruction, the more aspirated the consonant, since the explosion phase of the articulation takes more time (cf. Johnson 2003). Hence, posterior stops have a longer VOT duration than anterior ones, so the regular lineup is: /p/ → /t/ → /k/. The difference between /p/ and /t/ is not always so obvious, but the VOT values of /k/ are always considerably larger than those of either. VOT duration also differs according to languages, but the researchers of Laryngeal Realism attribute a phonological difference to the distinction among the languages of Group (2a) and Group (2b); at the same time, minor VOT differences among the languages of the same groups are phonologically irrelevant. The VOT averages shown in Group (2b) are of typically aspirated stops, they could also be transcribed as /pʰ, tʰ, kʰ/. Generally speaking, we can talk about unaspirated stops if the VOT value of /p/ is under 10 ms, that of /t/ is under 15 ms, and that of /k/ is under 30 ms.

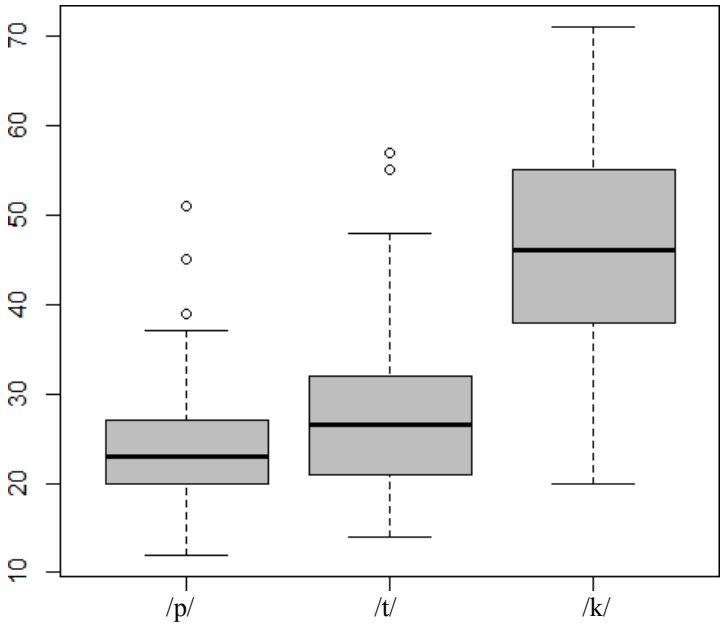
As far as Italian is concerned, the literature does not offer useable VOT values of initial or

intervocalic singleton /p, t, k/ in Standard Italian. The main reason is that voiceless stops tend to lenite (by voicing or by fricativization) in intervocalic position almost in every dialectal variety, which also influences the regional pronunciations of Standard Italian (cf. Marotta 2008: 238). Alternatively, VOT in geminated intervocalic /pː, tː, kː/ has been measured in several Italian varieties (since geminates are not targeted by lenition), and in these long stops preaspiration and postaspiration have equally been identified (Stevens & Hajek 2010a, 2010b). Other studies were dedicated to the aspiration of geminated stops in dialects of Southern Calabria and Southern Apulia (Sorianello 1996; Nodari 2015). They all determine VOT values which fall between the mean values of L-languages and h-languages seen in Chart (2). Some VOT means from Italian varieties from Sorianello (1996): /pː/ in Milan Italian shows 11 ms, in Cosenza Italian (Southern Calabria) 37 ms; /tː/ in Milan Italian shows 22 ms, in Cosenza Italian 51 ms; /kː/ in Milan Italian shows 39 ms, in Cosenza Italian 67 ms. The measurements of Stevens & Hajek (2010a) across 15 Italian cities are as follows: stressed /pː/ shows a VOT mean of 19 ms, stressed /tː/ shows a VOT mean of 26 ms, while stressed /kː/ shows a VOT mean of 60 ms.

Another study has been found which reports VOT values for Standard Italian, even if its purposes are not linguistic but pediatric (Bortolini et al. 1995). They analyze the speech production of "normal and preterm" Italian children, but they also create a control group of 7 adult Italian speakers (their place of origin is not indicated), whose VOT values in word-initial singleton /p, t, k/ are measured. The means are as follows: /p/ 11 ms, /t/ 19 ms, /k/ 34 ms. These values also fall between the general results of L-languages and h-languages seen in Chart (2).

As far as the corpus of this study is concerned, we found that voiceless stops appear to be mildly aspirated in the recordings, too, reaching a VOT average level which is exactly between the two groups of Chart (2). Three target words were selected from the research corpus, containing both a word-initial and a word-internal stop, which were pronounced by 15 informants, a total of 258 occurrences. The target words are ***p**ing**p**ong* [ˈpiŋg(ə)poŋgə], ***t**utt**t**avia* [tutːaˈviːja] 'however', and ***c**hirurgi**c**o* [kiˈrurd͡ʒiko] 'surgical'. VOT values were measured in Praat. The boxplot in (3), prepared with R, summarizes the results of the total occurrences (cf. Huszthy 2019: 139-141).

Chart 3

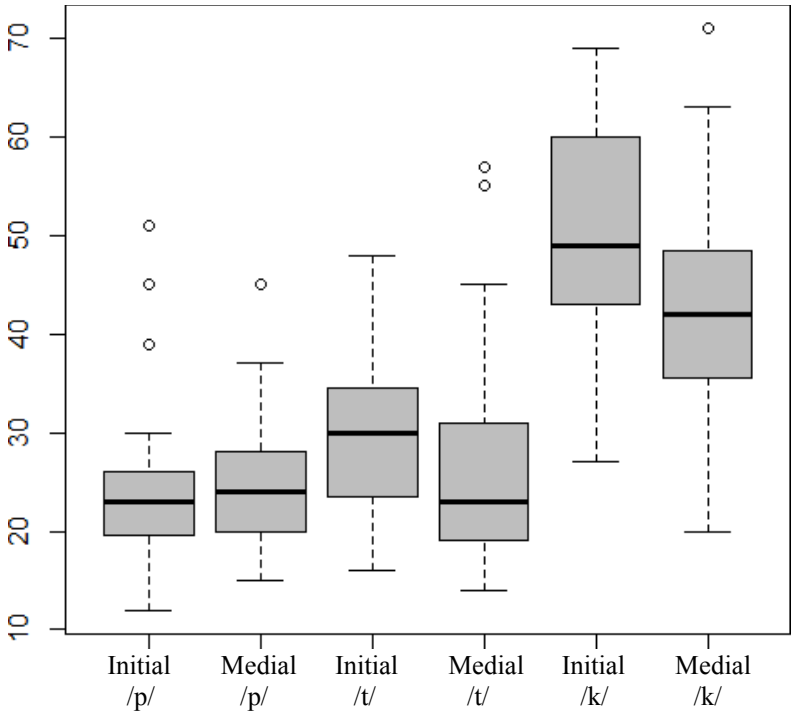


The total mean values are the following: /p/ – 24.04651 ms; /t/ – 27.46512 ms; /k/ – 46.12346 ms. These figures are intermediate between those in classical L-languages and h-languages, seen before in Chart (2), and they broadly coincide with the formerly measured VOT averages of Italian varieties, presented in the previous paragraphs.

However, the VOT results in boxplot (3) show a huge deviation between minimum and maximum

values, which is unusual compared to other languages. The maximum values are the following: /p/ – 37 ms (and an isolated 51 ms); /t/ – 48 ms (and an isolated 57 ms); /k/ – 71 ms: which are clearly aspirated realizations. The minimum values are the following: /p/ – 12 ms; /t/ – 14 ms; /k/ – 20 ms: which, on the other hand, are clearly unaspirated realizations. The boxplot in (4) shows the same VOT results with the division into word-initial and word-medial stops.

Chart 4



In the boxplot in (4) a very slight, but maybe not insignificant difference can be noted between the VOT values of the same stops in initial and medial positions. The case of /p/ appears to be irrelevant: the averages are almost the same, and the only distinction lies in the minimum and the maximum values, which are a little bit higher in word-internal position. But in the case of /t/ and /k/ the word-initial stop appears to be more aspirated than the word-internal one, as expected. In the target words *tuttavìa* and *chirùrgico* both stops are in an unstressed syllable, so the VOT differences may really be due to the position of the stop. This information may lead us to attribute some phonological role to aspiration in Italian.

As we have seen in this section, Italian voiceless stops show a degree of overall aspiration that falls between the standard values of "ordinary" L-languages (like other Romance, Slavic, etc.) and h-languages (like English, German, Chinese etc.). However, from the point of view of LR and ET, the moderate aspiration found in Italian can be crucial. Italian can be shown to exhibit distinctive, phonemic [voice] – in fact, this argumentation is very similar to the one applied to the case of Swedish (cf. Ringen & Helgason 2004; Helgason & Ringen 2008). The main phonetic difference between the two languages is that the voiceless set of stops is heavily aspirated in Swedish, while only mildly aspirated in Italian; however, the voiced set is phonetically prevoiced in both languages.

4 Conclusion: Italian analyzed as Swedish

On the basis of what has been developed in the previous sections, the laryngeal characterization of Italian may be due to a combination of phonological structure and phonetic implementation, that is, Italian may be analyzed as an h-system, similarly to Swedish.

The usual classification of Italian as an L-system is primarily based on two arguments. Firstly, on (impressionistic) evidence of its phonetics, which present fully voiced and voiceless obstruents: these

patterns suggest that it belongs to the same type as, for instance, French, Spanish or (typical) Slavic languages. Secondly, it is supposed to have carried its laryngeal properties throughout its history as part of its genetic inheritance as a Romance language. However, neither of the two arguments is strong enough to be decisive from a synchronic phonological point of view. The first argument cannot be accepted as evidence in phonology according to LR; while the second argument can easily be refuted through diachronic laryngeal modifications found in other languages, e.g., Cracow Polish and Scottish English, which have descended from an ancestor shared by the other varieties, but today diverge from those varieties in their laryngeal settings (Balogné Bérces & Huszthy 2018).

Based on the empirical data shown in Section 3, we can decide where Italian really belongs, or, in fact, whether it can be categorized at all. From a phonetic point of view we find substantial voicing and voicelessness in obstruents; phonologically, however, we fail to identify true laryngeal activity in the data, as no assimilation (i.e., feature spreading) is detected in the vast majority of the informants' outputs. In the present system, this suggests that Italian can be categorized as an h-language, making phonetic use of the sufficient discriminability between fully voiced and voiceless mildly aspirated obstruents. This is the laryngeal characterization of Swedish as well, where the opposition stands between fully voiced and voiceless heavily aspirated stops, without feature spreading.

Further laryngeal strategies used by the Italian informants of the corpus presented in Section 3.1 confirm this hypothesis; in fact, speakers retain underlying voice values in 65% (which is evidence for the absence of voice spreading), while they also apply "bidirectional devoicing processes" in 20% of the clusters, such as regressive devoicing (11%, e.g. *vo*[tk]*a*, *ra*[kt]*ime*) and progressive devoicing (9%, e.g. *foo*[tp]*all*, *A*[fk]*anistan*). This characterization is reminiscent of what is described in Section 2 as the profile of h-languages. If approached on the basis of the spelling, these processes seem "bidirectional devoicing"; phonologically, however, they are better analyzed as the emergence of the underlyingly voiceless forms. That is, from the approach of LR, these apparent "devoicing processes" are not processes at all, since the voiceless forms are not derived but underlying. Italian seems to match this description; in fact, obstruent voicedness in Italian is firmly maintained in sonorant environments only, and it is frequently "lost" next to a voiceless obstruent, manifested in apparent cases of regressive and progressive devoicing. Therefore, obstruent voicing in Italian seems to be passive voicing, which is unexpected from an L-language but a regular feature of h-systems. The only considerable difference between Italian and, for instance, English (or in our case, Swedish), is the phonetic implementation of these obstruents: the absence of heavy aspiration.

At the same time, the slight aspiration found in voiceless Italian obstruents in the corpus (described in Section 3.2), may acquire phonological importance. In fact, the voiceless set shows a degree of overall aspiration that falls between the standard values of "ordinary" L-systems like Slavic or Hungarian and h-systems like most Germanic languages. The contrast among a voiced set and a mildly aspirated voiceless set of obstruents is similar to the case of Swedish: apparently there are languages which "overshoot" the minimal phonetic distance which is required to obtain the laryngeal contrast among obstruents. In conclusion, Italian – which is usually considered an ordinary Romance language (and as such, an L-system) – has several laryngeal peculiarities; nevertheless, it can be analyzed as an h-language, which may explain these peculiarities, and this option is actually predicted by Laryngeal Relativism. A final argument is given by the nature of Element Theory: in this framework we cannot analyze Italian as an L-language, if it does not exhibit L-spreading. Because of the representationally based prerequisites of the model, if we do not find RVA in the system, we are unable to identify the active L-element, which also leads us to the analysis of Italian as an h-language.

References

Balogné Bérces, Katalin (2017). *A* News Feed*em a* Facebook*on*: Fonológiaalapú laringális nyelvtipológia. *Nyelvtudományi Közlemények* 113: 147-166.

Balogné Bérces, Katalin & Dániel Huber (2010). [Voice] and/versus [Spread Glottis] in the Modified Leiden model. *Acta Linguistica Hungarica* 57: 444-457.

Balogné Bérces, Katalin & Bálint Huszthy (2018). Laryngeal Relativism predicts Italian. *Yearbook of Poznań Linguistic Meeting* 47: 153-177.

Beckman, Jill, Michael Jessen & Catherine Ringen (2013). Empirical evidence for laryngeal features: Aspirating vs. true voice languages. *Journal of Linguistics* 49/2: 259-284.

Bertinetto, Pier Marco (1999). Boundary strength and linguistic ecology (mostly exemplified on intervocalic /s/-voicing in Italian). *Folia Linguistica* 33: 267-286.

Bertinetto, Pier Marco (2004). On the undecidable syllabification of /sC/ clusters in Italian: Converging experimental evidence. *Italian Journal of Linguistics* 16: 349-372.

Bertinetto, Pier Marco & Michele Loporcaro (2005). The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome. *Journal of the International Phonetic Association* 35/2: 131-151.

Bortolini, Umberta, Claudio Zmarich, Renato Fior & Serena Bonifacio (1995). Word-initial voicing in the productions of stops in normal and preterm Italian infants. *International Journal of Pediatric Otorhinolaryngology* 31/2-3: 191-206.

Colina, Sonia (2006). Optimality-theoretic advances in our understanding of Spanish syllable structure. In: Fernando Martínez-Gil & Sonia Colina (eds.), *Optimality Theoretic Studies in Spanish Phonology*. Amsterdam & Philadelphia: John Benjamins, pp. 172–204.

Cyran, Eugeniusz (2011). Laryngeal realism and laryngeal relativism: Two voicing systems in Polish. Studies in Polish *Linguistics* 6: 45-80.

Cyran, Eugeniusz (2012). Cracow sandhi voicing is neither phonological nor phonetic: It is both phonological and phonetic. In: Eugeniusz Cyran, Bogdan Szymanek & Henryk Kardela (eds.), *Sound, Structure and Sense: Studies in Memory of Edmund Gussmann*. Lublin: Wydawnictwo KUL, pp. 153-184.

Cyran, Eugeniusz (2014). *Between Phonology and Phonetics: Polish Voicing*. Berlin: Mouton de Gruyter.

Cyran, Eugeniusz (2017a). Hocus bogus? Licensing paths and voicing in Polish. In: Geoff Lindsey & Andrew Nevins (eds.), *Sonic Signatures: Studies Dedicated to John Harris*. Amsterdam: John Benjamins, pp. 34-62.

Cyran, Eugeniusz (2017b). 'Voice' languages with no [voice]? Some consequences of Laryngeal Relativism. *Acta Linguistica Academica* 64/4: 477-511.

Harris, John (1990). Phonological Government. *Phonology* 7/2: 255-300.

Harris, John (1994). *English Sound Structure*. Oxford, Cambridge: Blackwell.

Harris, John (2009). Why final obstruent devoicing is weakening. In: Kuniya Nasukawa & Philip Backley (eds.), *Strength Relations in Phonology*. Berlin & New York: Mouton de Gruyter, pp. 9–45.

Helgason, Pétur & Catherine Ringen (2008). Voicing and aspiration in Swedish stops. *Journal of Phonetics* 36, pp. 607-628.

Honeybone, Patrick (2002). *Germanic Obstruent Lenition: Some Mutual Implications of Theoretical and Historical Phonology*. [PhD dissertation] Newcastle upon Tyne: University of Newcastle.

Honeybone, Patrick (2005). Diachronic evidence in segmental phonology: The case of obstruent laryngeal specifications. In: Marc van Oostendorp & Jeroen van de Weijer (eds.), *The Internal Organization of Phonological Segments*. Berlin & New York: Mouton de Gruyter, pp. 319-354.

Huszthy, Bálint (2013). L'accento straniero degli italiani: Esiste un "accento italiano" comune? *Verbum Analecta Neolatina* 14/1-2: 167-181.

Huszthy, Bálint (2014). Accento italiano: Il ruolo della struttura sillabica. In: Matracki Ivica Peša, Maslina Ljubičić, Filipin Nada Županović, Vinko Kovačić (eds.), *Zbornik Međunarodnoga znanstvenog skupa u spomen na prof. dr. Žarka Muljačića* (1922. – 2009.). Zagreb: FF Press, pp. 299-308.

Huszthy, Bálint (2016). Italian as a voice language without voice assimilation. In: Kate Bellamy, Elena Karvovskaya, George Saad (eds.), *ConSOLE XXIV: Proceedings of the 24th Conference of the Student Organization of Linguistics in Europe*. Leiden: Leiden University Centre for Linguistics, pp. 428-452.

Huszthy, Bálint (2019). *How can Italian Phonology lack Voice Assimilation?* [PhD dissertation] Budapest & Piliscsaba: Pázmány Péter Catholic University.

Iverson, Gregory K. & Joseph C. Salmons (1995). Aspiration and laryngeal representation in Germanic. *Phonology* 12: 369-396.

Johnson, Keith (2003). *Acoustic & Auditory Phonetics*. Malden: Blackwell.

Kaye, Jonathan (2005). GP, I'll have to put your flat feet on the ground. In: Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinhenz & Jan Koster (eds.), *Organizing Grammar: Studies in Honor of Henk van Riemsdijk*. Berlin: Mouton de Gruyter, pp. 283-288.

Kaye, Jonathan, Jean Lowenstamm & Jean-Roger Vergnaud (1985). The internal structure of phonological representations: A theory of charm and government. *Phonology Yearbook* 2: 305-328.

Krämer, Martin (2009). *The Phonology of Italian*. Oxford & New York: Oxford University Press.

Lisker, Leigh & Arthur Abramson (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20/3: 527-565.

Marotta, Giovanna (2008). Lenition in Tuscan Italian (Gorgia Toscana). In: Joaquim Brandão de Carvalho, Tobias Scheer & Philippe Ségéral (eds.), *Lenition and Fortition*. Berlin: Mouton de Gruyter, pp. 235-271.

Mateus, Maria Helena & Ernesto D'Andrade (2000). *The Phonology of Portuguese*. Oxford: Oxford University Press.

Morelli, Frida (1999). *The Phonotactics and Phonology of Obstruent Clusters in Optimality Theory*. [PhD dissertation] College Park: University of Maryland.

Nespor, Marina (1993). *Fonologia*. Bologna: Il Mulino.

Nespor, Marina & Irene Vogel (1986). *Prosodic Phonology*. Dordrecht: Foris.

Nodari, Rosalba (2015). Descrizione acustica delle occlusive sorde aspirate: Analisi sociofonetica dell'italiano regionale di adolescenti calabresi. In: Mario Vayra, Cinzia Avesani & Fabio Tamburini (eds.), *Il farsi e il disfarsi del linguaggio: Acquisizione mutamento e destrutturazionedella struttura sonora del linguaggio*. Milan: Officinaventuno, pp. 139-153.

Petrova, Olga, Rosemary Plapp, Catherine Ringen & Szilárd Szentgyörgyi (2006). Voice and aspiration: Evidence from Russian, Hungarian, German, Swedish, and Turkish. *The Linguistic Review* 23: 1-35.

Recasens, Daniel (2014). *Coarticulation and Sound Change in Romance*. Amsterdam, Philadelphia: John Benjamins.

Ringen, Catherine & Pétur Helgason (2004). Distinctive [voice] does not imply regressive assimilation: Evidence from Swedish. *International Journal of English Studies* 4/2: 53-71.

Rohlfs, Gerhard (1966). *Grammatica storica della lingua italiana e dei suoi dialetti: Fonologia*. Turin: Einaudi.

Rubach, Jerzy (1996). Nonsyllabic analysis of voice assimilation in Polish. *Linguistic Inquiry* 27: 69-110.

Scheer, Tobias (2004). *A Lateral Theory of Phonology: What is CVCV and why should it be?* Berlin: Mouton de Gruyter.

Sorianello, Patrizia (1996). Indici fonetici delle occlusive sorde nel cosentino. *Rivista Italiana di Dialettologia* 20: 123-159.

Stevens, Mary & John Hajek (2010a). Post-aspiration in Standard Italian: Some first cross regional acoustic evidence. In: Takao Kobayashi, Keikichi Hirose & Satoshi Nakamura (eds.), *Proceedings of the 11th Annual Conference of the International Speech Communication Association*. Makuhari: International Speech Communication Association, pp. 1557-1560.

Stevens, Mary & John Hajek (2010b). Preaspirated /pp tt kk/ in standard Italian: A sociophonetic v. phonetic analysis. In: Marija Tabain (ed.), *Proceedings of the 13th Australasian International Conference on Speech Science and Technology*. Melbourne: ASSTA. pp. 1-4.

Wetzels, Leo & Joan Mascaró (2001). The typology of voicing and devoicing. *Language* 77: 207-244.

On the Distribution of the Copula in African American English

Kwang-sup Kim
Hankuk University of Foreign Studies

1 Introduction

The copula is always required in Mainstream English (ME) when the predicate is not a verbal phrase.

(1) a. She *(is) smart.
b. She *(is) a genius.
c. They *(are) in trouble.
d. They *(are) acting silly.

However, it is optionally present in African American English (AAE), as observed by many linguists (Labov et al. 1968, Labov 1969, 1972, 1995, Wolfram 1974, Baugh 1980, Rickford 1998, Walker 2000, Bender 2000, Kautzsch 2002, Green 1998a, 2002, 2011, among many others). Sentences (2)-(5) are all grammatical in AAE even if the copula is absent.

(2) She (is) tall. (Green 2002: 38)
(3) She (is) a woman. (Kautzsch 2002: 90)
(4) She (is) at home.
(5) Bruce (is) running. (Green 2002: 47)

To make matters complicated, there are many constructions in which the copula is obligatory even in AAE (Labov 1969, Bender 2000, Green 2002, 2011). One of them is the $XP_{[+predicate]}$-elided construction. If there is no copula in (6b) and (7), ungrammaticality results.

(6) a. The river wide?
b. Yes, it *(is)
(7) You ain't the best sounder, Eddie! I ain't! He *(is)! (Labov 1969: 720)

This paper shows that this peculiar pattern originates from the fact that T can be merged with a non-verbal predicative phrase in AAE. It is not a universal property of natural language that T must be merged with a bare VP. Many languages, including Arabic, Chinese, Korean, Nahuatl, and sign languages like American Sign Language, do not require a copula when the predicate is not a verbal phrase. This paper claims that AAE has no copula in the narrow syntax. The absence of the copula makes the narrow syntax of AAE simpler than that of ME, but it gives rise to a problem at PF; T must be attached to a verbal host, but it is stranded when its complement is not a bare VP. I propose that the stray affix problem is fixed via either repair-by-insertion or repair-by-ellipsis. More specifically, the problem is repaired by either *be*-insertion or T-deletion, but the option of T-deletion is only available if T is semantically, syntactically, and morphologically unmarked.[1] Therefore, *be* is obligatorily required if T is marked. It will be shown that the repair approach also sheds light on the distribution of dummy *do* and perfective *have*.

*This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (No. 20192213001) and the research fund (2020) of Hankuk University of Foreign Studies. This paper was presented at WECOL, Fresno, 2019. I would like to thank the audience for feedback and helpful suggestions.

[1] The notion 'unmarked' can be replaced with the notion 'recoverable'.

Proceedings of *WECOL 2019*

2 The distribution of the copula in AAE

As mentioned above, the copula is optionally present in AAE, but there are many contexts in which the copula is obligatorily required. This section shows that the contexts requiring the copula can be divided into three types: (i) the contexts in which *be* is required because of a suprasegmental morpheme, (ii) the contexts in which it is required on account of a segmental morpheme, and (iii) the contexts in which T is dislocated.

2.1 *Suprasegmental morphemes and the presence of the copula*

As illustrated in (8)-(10), sentential emphasis, $XP_{[+predicate]}$-ellipsis, and $XP_{[+predicate]}$-dislocation require the copula to be present.

(8) Sentential Emphasis
  a. Bruce IS running!
  b. *Bruce ø running!
(9) XP-ellipsis
  a. They said he wild, and he is.
  b. They said he wild, and *he ø.
(10) XP-dislocation
  a. Couldn't nobody say what color it is.
  b. *Couldn't nobody say what color it ø.

Let us first consider (8a, b). If the whole sentence is to be emphasized, the stress must fall on the first auxiliary, as in (11a, b).

(11) a. John DOES love Mary.
  b. John WILL meet Mary.

This means that if the copula is not present in (8a, b), there is no way to emphasize the whole sentence. Hence, sentence (8b) cannot be emphasized. Ellipsis also involves stress. If *smart* is elided in (12), the constituent preceding the elided site must be stressed, as shown in (13a, b).

(12) John is smart, and Mary is ~~smart~~, too.
(13) a. John is smart, and Mary is, too
  b. *John is smart, and Mary's too.

In the second conjunct of (9b), the ellipsis of the predicate *wild* produces a stress, but there is no auxiliary to bear it. Therefore, (9b) is ill-formed. The ungrammaticality of (10b) can be treated in the same way. In ME, the constituent preceding a dislocated constituent must be stressed.

(14) a. I wonder what color it is.
  b. *I wonder what color it's.

This is expected if movement consists of copy, merge, and deletion, as Chomsky (1995) proposes. Let us generate (10a) under the copy theory of movement. As illustrated in (15a-c), when *what color* is dislocated, ellipsis takes place and it requires that the adjacent constituent must be stressed.

(15) a. Couldn't nobody say what color it is what color: Deletion of *what color*
  b. Couldn't nobody say what color it is ~~what color~~: Assignment of Stress to *be*
  c. Couldn't nobody say what color it IS ~~what color~~

In (10b), however, there is no head that can bear stress. Therefore, it is not well-formed. To sum up, we can generalize that if sentential stress must be assigned, the copula must be present.

2.2 *Segmental bound morphemes and the presence of the copula*

Besides suprasegmental features, some segmental bound morphemes require the presence of the copula. The past tense ending and the non-finite bare inflectional ending require the obligatory presence of the copula. The sentence *Kaya cool* is grammatical, but if *was* is absent in (16), it is interpreted as a present tense construction (Green 2000, 2002). Therefore, *was* must be pronounced if the meaning of the past tense is to be maintained.

(16) Past Tense
Kayla *(was) cool.

Infinitival *to* and imperative C select a non-finite bare form, and the non-finite bare ending must be overtly realized. For example, (17a) and (18a) are ungrammatical.

(17) *To*-Infinitives
a. *You got to ø good, Rednall!
b. You got to be good, Rednall! (Bender 2000: 83)

(18) Imperatives
a. *ø nice to your mother!
b. Be cool, brothers! (Bender 2000: 83)

Tense can be classified into four types, depending on whether or not it is finite and whether or not it is present.[2] Sentences (16-18) show that the finite present tense optionally requires the copula, while the others obligatorily require it.

2.3 *T-movement and the presence of the copula*

T-movement also requires the obligatory presence of the copula. In tag-questions, T-to-C movement takes place, producing a *[T C]* complex. The *[T C]* complex must be pronounced.

(19) Tag-Questions
Bruce not eating, *(is) he? (Green 2002: 43)

The function of $C_{[Q]}$ in interrogatives is to signal that the sentence type is an interrogative, which is expressed by Subject-Aux Inversion.

(20) a. Bruce not eating, [$C_{[Q]}$ [he T]]: T-to-C movement
b. Bruce not eating, [T $C_{[Q]}$ [he ~~T~~]]

Unless *be* is pronounced, the function of $C_{[Q]}$ is erased: that is, there is no way to indicate that T-movement took place. So *be* must be pronounced.

So-Inversion is another construction in which the copula is obligatorily required on account of dislocation. As shown in (21a, b), the copula can be contracted but it cannot be elided in *So*-Inversion.

(21) a. I'm tired and so's my dog.
b. *I'm tired and so my dog. (Bender 2000: 86)

This phenomenon falls under the generalization that the copula is required when T is dislocated. Wood (2008) claims that T-movement takes place in *So*-Inversion as well as in Negative Inversion: *So*-Inversion involves T-movement, as in (22).

(22) [$_{PolarityP}$ so [ø T be]] [$_{TP}$ my dog [~~T be~~] [~~so~~]]]

[2] I assume that the non-finite past tense is realized as *have been*, as illustrated in (i).
(i) John seems to have been here yesterday.

Provided that the analysis in (22) is correct, T occupies a dislocated position in (21a-b), and if the copula is not pronounced, there is no way to indicate that T-movement took place. In this context, the copula must be present.

To recapitulate, the copula is required, (i) when sentential stress requires its host (8-10), (ii) when either a non-finite ending or a non-present ending requires its host (16-18), and when T is dislocated (19 & 21).[3] Put differently, the copula is obligatorily required if T is either stressed, past, non-finite, or dislocated. Given that the unmarked form of T is unstressed, present, finite, and undislocated, we can make a generalization as follows:

(23) The copula is obligatorily required if T is marked, while it is optionally required when T is unmarked.

Section 3 is devoted to providing a principled account of the generalization in (23).

3 Proposal

There are three major functions of the copula: syntactic, morphological, and semantic functions. First, it can be required for syntactic regularity. Suppose that a grammar requires that T must be merged with VP: that is, suppose that every sentence contains a verb and S consists of NP and VP. If so, the copula is required when the predicate is not a verbal phrase. In this case, the presence of the copula is syntactically motivated. Second, in many languages the copula is required to function as a host for affixal T. In these languages the presence of the copula is morphologically required. Third, if the copula has its own inherent semantic content, it is semantically motivated. The major claim of this section is that in AAE the copula is morphologically motivated, while in ME it is syntactically motivated.

3.1 *Deletion?*

Given that (24a, b) give the same interpretation, we can rule out the semantic motivation for the presence of the copula in AAE.

(24) a. Kayla is smart.
b. Kayla smart.

So we are left with syntactic and morphological motivations. Let us first consider the possibility that the presence of the copula is syntactically motivated and it is deleted in some contexts. This is exactly the line of approach pursued in Labov (1969, 1972). According to Labov, the copula is absent because it can be deleted. More precisely, he proposes that AAE and ME are quite analogous in that *be* is base-generated in the underlying form, and the copula is absent in AAE when it is pronounced as a contracted form in ME. As illustrated in (25), if the copula is weakened in ME, it is phonetically realized as a contracted form, and if, on the other hand, it is weakened in AAE, it is pronounced as a null form.

(25) AAE and ME
a. Kayla is smart: weakening of the copula
b. Kalya's smart (ME)
b'. Kayla smart (AAE)

However, as Bender (2000) points out, the variation is not as simple as Labov claims. For instance, *am* can be contracted into *'m* in ME, but it is not realized as a null form in AAE.

(26) a. I am alright.
b. I'm alright.

[3] The aspectual markers like habitual *be* and remote perfective *BIN* must be distinguished from the copula introduced above.
(i) a. Kayla be sick (habitual reading) 'Kayla is always sick'
b. Kayla sick (non-habitual, temporary reading) 'Kayla is sick now'
(ii) I BIN know that guy. 'I have known that guy for a long time' (Labov 1998: 124)

As briefly mentioned at the outset of this paper, I propose that the copula is inserted at PF as a last resort, but I assume that the aspectual markers are base-generated in the narrow syntax.

c. *I alright.

Furthermore, she points out that there are some other constructions in which contraction is not permitted but the copula can be elided. In (27), the copula cannot be contracted but it can be elided.

(27) a. Tha's the man they say is in love
b. *Tha's the man they say 's in love
c. Tha's the man they say ∅ in love (Bender 2000: 91)

The list of the asymmetry between the contracted copula and the elided copula can be further extended. The past tense copula *was* can be contracted in ME, but it cannot be elided in AAE, as we have seen from (16), which is rewritten here as (28c).

(28) a. Kayla was cool.
b. Kayla's cool.
c. Kayla *(was) cool.

In addition, the copula cannot be elided in *So*-Inversion although it can be contracted, as we have seen from (21a, b), repeated here as (29a, b).

(29) a. I'm tired and so's my dog.
b. *I'm tired and so my dog. (Bender 2000: 86)

Sentences in (26-29) show that there is no one-to-one correspondence between the contracted copula in ME and the zero copula in AAE.[4]

Given that the deletion approach based on the syntactic motivation is inadequate, we are only left with the morphological motivation. This section explores the possibility that in AAE the copula is morphologically motivated to fix the stray affix problem.

3.2 *Be-insertion as a last resort*

It is not a universal property of language that T must be merged with VP. There are many languages in which T can take a non-verbal predicate phrase as its complement. Arabic, Chinese, Korean, Nahuatl, and sign languages like American Sign Language, belong to these languages.[5] In Chinese tense is not overtly realized. So there is no morphological motivation for the presence of the copula, and interestingly, the copula is absent in this language.

(30) tā gāoxìng
he happy 'He is happy'

Korean displays a similar pattern. In Korean, if the predicate is a nominal, the copula *-i* is required, but if it is an adjectival, the copula must be absent.[6]

(31) Jay-nun haksayng-i-ess-ta.
J-TOP student-copula-past-ind
'John was a student'
(32) Jay-nun aphu(*i)-ass-ta.
J-TOP sick(*copula)-past-ind
'John was sick'

This is because adjectives can be hosts for affixal T, whereas nominals cannot. The two languages show that the copula

[4] As will be discussed in 3.3 and 4.2, another problem with the ellipsis approach is that it cannot deal with the distribution of perfective *have*.
[5] Jantunen (2007) suggests that in Finnish Sign Language the sign PI is a certainty expressing modal device, but it may be in the process of grammaticalizing into a copula.
[6] Kim (2019) proposes that in (33a) *i* is not a copula but a nominative case marker, although it is traditionally known as a copula.

must be absent if it is not required morphologically. In other words, V is not universally required in sentence formation, and whether or not the predicate phrase must be a verbal phrase is a matter of parameter. I propose that AAE is also on the list of the languages that do not obligatorily require the presence of a verb: that is, T can take a non-verbal predicate as its complement, as in (33).

(33) Joe $T_{present}$ [$_{PP}$ in the kitchen]

In this system *be* is not required in the narrow syntax, but it may be required at PF to meet Lasnik's (1981) Stray Affix Filter. In (33), there is no host for the tense affix. In this case AAE utilizes two options. The first is to save the structure by inserting *be*, as illustrated in (34a, b).

(34) a. Joe $T_{Present}$ in the kitchen: *be*-insertion
b. Joe [be $T_{Present}$] in the kitchen

The second is to make use of the repair-by-ellipsis strategy: the stray affix problem can be fixed when $T_{Present}$ is elided.

(35) a. Joe $T_{Present}$ in the kitchen: Deletion
b. Joe ~~$T_{Present}$~~ in the kitchen

Therefore, the copula is optionally present.[7] Thus far, we have considered the case in which the tense is unmarked. If the tense is marked, the copula is obligatorily present. For instance, if T is past, it must be supported by *be*.

(36) a. Joe T_{Past} in the kitchen: *be*-insertion
b. Joe [be T_{Past}] in the kitchen

Under this repair approach, the generalization in (23) is revised as follows:

(37) When T takes an *ing*-phrase or a non-verbal predicative phrase, it is optionally supported by *be* if T is unmarked, and it must be supported by *be* if it is marked.

3.3 *Three types of insertion: do-insertion, be-insertion, and have-insertion*

It is noteworthy that AAE, just like ME, permits *do*-insertion. In ME, T selects a bare VP as its complement, and *do* is inserted as a last resort when T fails to lower onto a main verb.

(38) a. John does not love Mary.
b. John DOES love Mary.
c. Tom loves Mary, and John does ~~[love Mary]~~, too.
d. John loves Mary, doesn't he?

In fact, AAE displays the same pattern. If T merges with a bare VP, T must be supported by *do* when it is stranded: *do*-support takes place in negative, emphatic, VP-elided, and interrogative sentences. Sentences (39a-d) are examples of AAE, which show a remarkable similarity to ME.

(39) a. (I think) it don't make no sense.
b. he DID eat. (Green 2002: 42)
c. Bruce said he would win the election, and win the election he did. (Green 2002: 44)
d. He ain eat, did he? (Green 2002: 43)

They can be generated if AAE and ME share the insertion rule in (40), as illustrated in (41-44).

[7] It is not a novel idea that T can take a predicate phrase as its complement. Bailey (1965) proposes that 'S → NP PredP' can be a theoretically possible PS rule. Mufwene (1992) argues that it is actually realized in AAE, proposing that the copula is optionally present because AAE makes use of both 'S → NP PredP' and 'S → NP VP'. In this approach *be*-insertion is not required, because *be* is present when 'S → NP VP' is used. There is no T-deletion either; when 'S → NP PredP' is used, tense is not present even in the narrow syntax.

(40) When T takes a verbal phrase as its complement, it is supported by *do* as a last resort if it is stranded.

(41) a. (I think) it T n't make no sense: *Do*-Insertion
b. (I think) it [do [T n't]] make no sense: Phonetic Realization
c. (I think) it don't make no sense

(42) a. he T_{past} Stress eat: *Do*-Insertion
b. he [do [T_{past} Stress]] eat: Phonetic Realization
c. he DID eat

(43) a. Bruce said he would win the election, and win the election he T ~~win the election~~ : *Do*-Insertion
b. Bruce said he would win the election, and win the election he did.

(44) a. He ain eat, C he T ~~eat~~: T-to-C Movement
b. He ain eat, [T C] he ~~T~~ ~~eat~~: *Do*-Insertion
c. He ain eat, did he?

AAE diverges from ME in that T does not impose a strict selectional restriction: it can be merged with a non-verbal predicate phrase or an inflected VP, as well as a bare VP. If it is merged with a non-verbal predicate phrase or an inflected verbal phrase, it is stranded. The generalization in (37) says that *be* is inserted when T takes an *ing*-phrase or a non-verbal predicative phrase. In addition, there is one more insertion rule in AAE: *have*-insertion. As observed by Green (2002), *have* is missing in *BIN*-constructions if T is unstressed, present, and undislocated, but it must be present if T is stressed, past, or dislocated.

(45) She BIN running. 'She has been running for a long time'

(46) a. Have they BIN running? (They BIN running?)
a'. *BIN they running (Green 2002: 68)
b. They BIN left, ain't/haven't they?
b'. *They BIN left, BIN't they? (Green 2002: 69)
c. Bruce BIN running, and Sue have, too.
c'. *Bruce BIN running, and Sue BIN, too. (Green 2002: 69)
d. She had BIN running.

This pattern can be generalized as '*have* is absent when T is unmarked, whereas it is present when it is marked', and the generalization can be captured via the *have*-insertion rule in (47).

(47) When T takes a BIN-phrase as its complement, it is elided as a last resort if it is unmarked, and it is supported by *have* as a last resort if it is marked.

In this approach, (45) is generated via T-deletion because T is unmarked, whereas (46a) is generated via *have*-insertion because T is marked.

(48) a. She T BIN running: T-deletion
b. She ~~T~~ BIN running

(49) a. C they T BIN running: T-to-C movement
b. [T C] they T BIN running: *Have*-Insertion

In summation, AAE has three insertion rules: *do*-insertion, *be*-insertion, *have*-insertion, and they apply in a slightly different way as follows:

(50) a. When T is marked, it must be supported.
b. When T is unmarked, *do*-insertion must take place, *be*-insertion may take place, and *have*-insertion must not take place.

The remainder of this section shows that the three different rules are all governed by the same principles: the Principle of Deletion and the Principle of Insertion.

3.4 *PF-uninterpretable features*

In the Minimalist Program advocated by Chomsky (2000, 2001, 2008, 2013), merger is triggered by an uninterpretable feature; uninterpretable features must be deleted and they can be deleted via merger. For instance, if C has an uninterpretable *wh*-feature—[u(ninterpretable) wh-feature], it must be merged with a *wh*-phrase; otherwise, the derivation crashes.

(51) a. $C_{[uWh]}$ [$_{TP}$ … *Wh*-phrase…]: Merger of the *Wh*-phrase and Deletion of [uWh]
b. [*Wh*-phrase $C_{\sout{[uWh]}}$ [$_{TP}$ … *Wh*-phrase…]]

Wh-movement is overt in English, but it is covert in many languages like Chinese, Japanese, and Korean. Suppose that *wh*-movement takes place at LF in a certain language. If so, the internal merger in (51) must be postponed after Spell-out. In an attempt to capture this phenomenon, Chomsky (1995) proposes the principle of 'procrastinate', according to which operations apply as late as possible if it is triggered by a 'weak' feature.

(52) a. $C_{[Weak\ uWh]}$ [$_{TP}$ … *Wh*-phrase…]: Merger of the *Wh*-phrase at LF and Deletion of [Weak uWh]
b. [*Wh*-phrase $C_{\sout{[Weak\ uWh]}}$ [$_{TP}$ … *Wh*-phrase…]]

This proposal can be reinterpreted as follows: (i) there are three types of uninterpretable features: uninterpretable features with no specification, uninterpretable LF (u-LF) features, and uninterpretable PF (u-PF) features, and (ii) the first type triggers a merger in the narrow syntax in accordance with Pesetsky's (1989) Earliness Principle, the second type requires a merger only at LF, and the third type only at PF.

With the above discussion in mind, let us try to reanalyze the generalizations in (37), (40) and (47) by making use of u-PF feature. First of all, I propose that T has a different u-PF feature, depending on which type of complement it takes.

(53) a. T has an uninterpetable PF-feature [u-$_{PF}$ V] when it is merged with a bare VP.
b. T has an uninterpetable PF-feature [u-$_{PF}$ Be] when it is merged with a non-verbal predicative phrase or an *ing*-phrase.
c. T has an uninterpetable PF-feature [u-$_{PF}$ Have] when it is merged with a *BIN*-phrase.

The uninterpretable PF-feature of T can be deleted if it is merged with a verb that has a corresponding feature. For instance, [u-$_{PF}$ V] can be deleted if T is merged with any type of verb via lowering, but when T is stranded, [u-$_{PF}$ V] is deleted if it is supported by *do*; the unmarked form of a verb is *do*.

(54) a. I $T_{[u\text{-}PF\ V]}$ like it: T-lowering
b. I [like $T_{\sout{[u\text{-}PF\ V]}}$] it
(55) a. I $T_{[u\text{-}PF\ V]}$ like it: VP-ellipsis
b. I $T_{[u\text{-}PF\ V]}$ ~~like it~~: *Do*-Insertion
c. I [do $T_{past\sout{[u\text{-}PF\ V]}}$] ~~like it~~

By contrast, $T_{[u\text{-}PF\ Be]}$ and $T_{[u\text{-}PF\ Have]}$ fail to undergo lowering because their complement cannot be a host for T. Uninterpretable PF features must be deleted at PF, and there are two possible ways to delete them. They can be deleted via either insertion or T-ellipsis: that is, the problem caused by u-PF features can be fixed via either the repair-by-insertion strategy or the repair-by-ellipsis strategy. Sentences (56) and (57) illustrate that the uninterpretable features of $T_{[u\text{-}PF\ Be]}$ and $T_{[u\text{-}PF\ Have]}$ are deleted when they are supported by *be* and *have*, respectively.

(56) a. Kayla $T_{[u\text{-}PF\ Be]}$ smart: *Be*-Insertion and [u-$_{PF}$ Be]-Deletion
b. Kayla [be $T_{\sout{[u\text{-}PF\ Be]}}$] smart
(57) a. She T[u-$_{PF}$ Have] ~~BIN running~~: *Have*-Insertion and [u-$_{PF}$ Have]-deletion
b. She [have $T_{\sout{[u\text{-}PF\ Have]}}$] ~~BIN running~~

They can also be deleted if T is elided. In short, uPF-features on T can be elided either via insertion or deletion. The generalization in (50a, b) stipulates when insertion takes place and when deletion takes place, and now it can be reformulated as follows:

(58) a. When marked, stranded T must be supported.

b. When unmarked, stranded $T_{[u\text{-}PF\ V]}$ must be supported by *do*, stranded $T_{[u\text{-}PF\ Be]}$ may be supported by *be*, stranded $T_{[u\text{-}PF\ Have]}$ must not be supported by *have*.

The rest of this section explores the possibility of explaining the generalization in (58a, b) by making use of the notion '(un)markedness'

3.5 *Markedness and the principle of deletion/insertion*

There are two factors involved in deciding whether T stranding is repaired by deletion or insertion: deletability and insertability. The remainder of this section is devoted to deriving the Principle of Deletion and the Principle of Insertion from the notion '(un)markedness'.

3.5.1 *The principle of deletion*

We have seen that T cannot be deleted unless it is present, unstressed, finite, and undislocated. If T is past, stressed, or non-finite, it is semantically marked,[8] and if it is dislocated, it is syntactically marked. Thus, we are led to the generalization that T cannot be deleted unless it is semantically or syntactically marked. There is an exception to this generalization: the copula must be present when the subject is first person singular, even if T is neither past, nor stressed, nor non-finite, nor dislocated (Labov 1969, Green 2002), as shown by (26a-c), repeated here as (59a-c).

(59) First Person
a. I am alright
b. I'm all right.
c. *I all right.

This suggests that morphologically marked T cannot be deleted either. The verbal ending of AAE is quite simple. A single unmarked bare form is usually used: that is, no distinction is made between first, second and third singular and plural. For instance, the same forms *eat, love,* and *run* are used for first person singular and plural, second person singular and plural, and third person singular and plural in the present tense (Green 1998a, 2002). When regular verbs co-occur with the ending *-s*, they usually carry special emphasis. The morphology of the copula is also simple. In AAE, the distinction between *is* and *are* is not always made: *is* can be used for first plural, second singular and plural, third singular and plural (Green 2002). In other words, *is* can be used for all person-number combinations, except first person singular: if the subject has the features [+first person] & [+singular], the copula must be realized as *am*. In summation, *am* is an exception to the generalization that AAE usually uses an unmarked single verbal form. This means that *am* is a morphologically marked form. $T_{[first\ person\ \&\ singular][uPF\ be]}$ would produce the morphologically marked form *am* if it is phonetically realized. I propose that unmarked T can be deleted, whereas semantically, syntactically, or morphologically marked T cannot be deleted.[9]

(60) The Principle of Deletion
a. Unmarked T can be deleted.
b. Semantically, syntactically, or morphologically marked T cannot be deleted.[10]

3.5.2 *The principle of insertion*

[8] The non-finite ending is semantically marked in that it does not describe a state of affairs in the actual world.

[9] This proposal is reminiscent of the debate about whether or not ellipsis is subject to the morphological identity condition. Sentence (ia) shows that ungrammaticality results if the elided copula is not morphologically identical with its antecedent, as noted by Warner (1985).

(i) a. *Emily was beautiful at the recital and her sister will __, too.
b. Emily was beautiful at the recital and her sister will be beautiful at the recital, too.

Lasnik (1995) makes use of the morphological identity condition while explaining the contrast in (ia-b), and Potsdam (1997) argues that (ia-b) can be explained without recourse to it. This paper remains neutral about this issue. As Chung (2006, 2013) and Merchant (2001, 2013) claim, however, it seems to be descriptively inadequate to propose that ellipsis is only subject to the semantic identity condition.

[10] Semantically and morphologically unmarked $T_{[uPF\ be]}$ is realized as *is* if it is supported by *be*. The Principle of Deletion amounts to saying that the unmarked form *is* can be phonetically unrealized.

Besides 'deletability', we need to consider 'insertability' in deciding whether T stranding is repaired by deletion or insertion. There are three types of V: when unmarked T is stranded, *do* must be inserted, *have* must not be inserted, and *be* can be inserted. This difference can also be derived from the notion 'markedness'. The only syntactic feature of dummy *do* is [+V], which is shared by every verb. In that sense, it is a very general or unmarked verb. By contrast, *have* is only compatible with *BIN*-phrases. It is quite a specific or marked verb. On the other hand, the copula is in between; it can co-occur with a predicative phrase or an *ing*-phrase. I propose that whether or not V can be inserted is dependent on whether or not it is marked.

(61) The Principle of Insertion
  a. When stranded T is unmarked,
    (i) Unmarked V (do) must be inserted.
    (ii) Marked V (have) must not be inserted.
    (iii) Neither unmarked nor marked V (be) can be optionally inserted.
  b. When stranded $T_{[uX]}$ is marked, X is inserted, regardless of the markedness of X.

To sum up, deletion of T and insertion of a verbal head are dependent on whether or not T and V are marked. The patterns in (60) and (61) can be described in an OT-like style as (62a-d), which can be shortened as (63).

(62) a. *Marked T-deletion>> *X-Insertion, where X is either *have*, *be* or *do*.
  b. **Have-Insertion* > *Unmarked T-deletion
  c. **Be-Insertion* = *Unmarked T-deletion
  d. *Unmarked T-deletion> **Do-Insertion*
(63) *Marked T-deletion>> **Have-Insertion* > *Unmarked T-deletion = **Be-Insertion* > **Do-Insertion*

Neither insertion nor deletion can take place in the narrow syntax on account of the Inclusiveness Condition. However, the constraints can be violated at PF as a last resort to save an otherwise ill-formed representation. The generalization in (62) states that (i) the constraint '*Marked T-deletion' must not be violated in any case, (ii) if either '* *Have*-insertion' or '*Unmarked T-deletion' needs to be violated, '*Unmarked T-deletion' is violated, (iii) if either '**Be*-insertion' or '*Unmarked T-deletion' needs to be violated, either of them is violated, and (iv) if either '**Do*-insertion' or '*Unmarked T-deletion' needs to be violated, '**Do*-insertion is violated.

3.6 *Comparison of AAE and ME: PF-based grammar vs. syntax-based grammar*

Let us briefly compare AAE with ME. In ME, T must be merged with a bare VP in the narrow syntax.

(64) ME
  a. $T_{[u\text{-PF V}]\,[\sout{uVP}]}$ [bare VP/vP …]
  b. $T_{[u\text{-PF V}]\,[\sout{uVP}]}$ [Be-P be [PP/AP/NP/ing-P …]]
  c. $T_{[u\text{-PF V}]\,[\sout{uVP}]}$ [Have-P have [–en-P …]]

As a device to turn a non-verbal predicate phrase or an inflected VP into a bare VP, auxiliaries like *be* and *have* are required in the narrow syntax. ME syntactically requires that *be* and *have* co-occur with PP/AP/NP/*ing*-P and *been*-P, respectively.

(65) a. $be_{[u\text{-PP/AP/NP/ing-P}]}$
  b. $have_{[u\text{–en-P}]}$

In AAE, by contrast, T can be merged with either a bare VP, or PP/AP/NP/*ing*-P, or *BIN*-P. T does not have a restriction on the grammatical category of its complement.

(66) AAE
  a. $T_{[u\text{-PF V}]}$ [bare VP/vP …]
  b. $T_{[u\text{-PF Be}]}$ [PP/AP/NP/ing-P …]]
  c. $T_{[u\text{-PF Have}]}$ [BINP …]

In AAE, the syntactic co-occurrence restrictions in (65) are reanalyzed as PF co-occurrence restrictions in the way that T has an uninterpretable PF feature like [uPF *Be*] or [uPF *Have*], depending on the type of its complement. In (64b, c) T is merged with *be* or *have* via head movement, and in (67b, c) T is merged with one of them via repair-by-insertion at PF. A major difference between (64b, c) and (66b, c) lies in the availability of ellipsis. In (64b, c) auxiliaries can move to T, so that the stray affix problem can be fixed without recourse to either insertion or deletion. By contrast, in (66b, c) there is no way to fix the stray affix problem without recourse to the last resort operations like insertion or deletion. The fact that deletion can be an option in AAE gives rise to the difference in output between the two grammars. This approach supports the view that AAE is a reanalysis of ME from the view point of PF.

4 Conclusion

Chomsky (1957) and Hornstein (2001) propose that certain expressions like *do* are not part of the lexicon and are only usable if required for convergence. One of the major claims made in this paper is that the copula in AAE is also not in the lexicon but inserted as a last resort at PF. In ME every matrix sentence contains a verb without exception, which is made possible by the presence of the copula in the narrow syntax.[11] In AAE, however, the copula is not present in the narrow syntax and T is merged with a predicative phrase, regardless of whether or not it is a verbal phrase. As a consequence, the narrow syntax of AAE is simpler than that of ME, but there arises a problem at PF; if T is merged with a non-verbal predicative phrase, it cannot find its host at PF. This paper has shown that this dilemma is resolved via either T-ellipsis or *be*-insertion, and the two operations are subject to the Principle of Deletion and the Principle of Insertion. This paper has extended this approach to the distribution of perfective *have*, providing a uniform account of the distribution of dummy *do*, the copula, and perfective *have*. To conclude, AAE, unlike ME, is a PF-based grammar with regard to the distribution of dummy auxiliaries, which hints at the possibility that AAE results from a reanalysis of ME from the view point of PF.

References

Akmajian, Adrian. 1979. *Aspects of the grammar of focus in English*. New York: Garland.

Bailey, Beryl. 1965. A new perspective on American Negro dialectology. *American Speech* 11, 171-77.

Baugh, John. 1980. A re-examination of the Black English copula. In *Locating Language in Time and Space*, ed. by Labov, 83-106. New York: Academic Press.

Bender, Emily. 2000. Syntactic variation and linguistic competence: the case of AAVE copula absence. Ph.D. dissertation, Stanford University.

Bresnan, Joan. 1972. On Sentence Stress and Syntactic Transformations. In *Contributions to Generative Phonology*, ed. by Brame, 73-107. Austin: University of Texas Press.

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

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on minimalism in honor of Howard Lasnik*, ed. by Martin, Michaels & Uriagereka, 89-155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In In *Ken Hale: A life in language*, ed. by Kenstowicz, 1-52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2008. On phases. In *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, ed. by Freidin, Roger, Carlos P. Otero & Maria L. Zubizarreta, 133–166. Cambridge, MA: MIT Press.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130, 33-49.

Chomsky, Noam. 2015. Problems of Projection. In *Structures, Strategies, and Beyond*: *Studies in honour of Adriana Belletti*, ed. by Domenico, Elisa Di, Cornelia Hamann & Simona Matteini, 1-16. Amsterdam/Philadelphia, Benjamins Publishing Company.

Chung, Sandra. 2006. Sluicing and the lexicon: The point of no return. In *Proceedings of the annual meeting of the Berkeley Linguistics Society* 31, 73–91. Berkeley: University of California, Berkeley Linguistics Society.

Chung, Sandra. 2013. Syntactic identity in sluicing: How much and why. *Linguistic Inquiry* 44, 1–44.

Dikken, Marcel den. 2006. Specificational copular sentences. In *The Blackwell Companion to Syntax*, ed. by Everaert & van Riemsdijk, 292-409. Malden, MA: Blackwell.

Green, J. Lisa. 1998a. Aspect and predicate phrases in African American Vernacular English. In *African American English: Its history, origin, and use*, ed. by Bailey, Baugh, Mufwene & Rickford, 37-68. London: Routledge.

Green, J. Lisa. 1998b. Remote past and states in African-American English. *American Speech* 73, 115–138.

11 It is controversial as to whether or not the copula has a semantic content in ME. There are two opposing views on it: one is that it is polysemous (Huddleston 1971, Akmajian 1979, Safir 1985, Seuren 1985, Rapoport 1987, Higginbotham 1987, Zaring 1996), and the other is that it is semantically vacuous (Stowell 1981, Heggie 1988, Moro 1997, Heycock and Kroch 1999, Den Dikken 2006). I assume that the copula has no semantic content in ME as well as in AAE.

Green, J. Lisa. 2000. Aspectual *be*-type constructions and coercion in African American English. *Natural Language Semantics* 8, 1–25.
Green, J. Lisa. 2002. *African American English: a linguistic introduction*. Cambridge: Cambridge University Press.
Green, J. Lisa. 2011. *Language and African American child*. Cambridge: Cambridge University Press.
Heggie, Lorie A. 1988. The syntax of copular structures. Ph.D. dissertation, University of Southern California.
Heycock, Caroline & Anthony Kroch. 1999. Pseudo-cleft connectedness: implications for the LF interface level. *Linguistic Inquiry* 30, 365-398.
Higginbotham, James. 1985. On semantics. *Linguistic Inquiry* 16, 547-593.
Huddleston, Rodney. 1971. *The sentence in written English*. Cambridge: Cambridge University Press.
Kautzsch, Alexander. 2002. *The historical evolution of earlier African American English: an empirical comparison of early sources*. Berlin: Mouton de Gruyter.
Kim, Kwang-sup. 2019. The copula as a nominative Case marker. In *The grammar of copulas across languages*. ed. by Arche, Fábregas & Marín, 69-89. Oxford: Oxford University Press.
Labov, William. 1969. Contraction, Deletion, and Inherent Variability of the English Copula. *Language* 45, 715-62.
Labov, William. 1972. *Language in the Inner City: Studies in the Black English Vernacular*. University of Pennsylvania Press.
Labov, William. 1995. The case of the missing copula: The interpretation of zeroes in African-American English. In *An Invitation to Cognitive Science*, *vol. 1: Language*, ed. by Gleitman, Liberman & Osherson, 25–54. Cambridge, MA: MIT, 2nd edn.
Labov, William. 1998. Coexistent systems in African-American English. In *African American English: Its history, origin, and use*, ed. by Bailey, Baugh, Mufwene & Rickford, 110-153. London: Routledge.
Labov, William, P. Cohen, C. Robins, & J. Lewis. 1968. A Study of the Non-Standard English of Negro and Puerto Rican Speakers in New York City. Cooperative Research Report 3288, vols. 1 and 2, Philadelphia: U.S. Regional Survey, Linguistics Laboratory, University of Pennsylvania.
Lasnik, Howard. 1981. Restricting the theory of transformations. In *Explanation in linguistics*, ed. by Hornstein & Lightfoot, 152-173. London: Longmans.
Lasnik, Howard. 1995. Verbal morphology: Syntactic Structures meets the minimalist program. In *Evolution and Revolution in Linguistic Theory: Essays in Honor of Carlos Otero*, ed. by Campos & Kempchinsky, 251-275. Washington DC: Georgetown University Press.
Merchant, Jason. 2001. *The syntax of silence*. Oxford: Oxford University Press.
Merchant, Jason. 2013. Voice and ellipsis. *Linguistic Inquiry* 44, 77-108.
Moro, Andrea. 1997. *The raising of predicates: predicative noun phrases and the theory of clause structure*. Cambridge: Cambridge University Press.
Mufwene, Salikoko S. 1992. Why grammars are not monolithic. In *The Joy of Grammar*, *A festschrift in honor of James D. McCawley*, ed. by Diane Brentari, Garry N. Larson & Lynn A. MacLeod, 225-250. Amsterdam/Philadelphia: John Benjamins.
Pesetsky, David. 1989. Language-Particular Rules and the Earliness Principle. Ms., MIT.
Potsdam, Eric. 1997. English verbal morphology and VP ellipsis. *The Proceedings of the 27th Meeting of the North East Linguistic Society,* 353-368. Amherst, MA: GLSA, University of Massachusetts at Amherst.
Rickford, John. 1975. Carrying the new wave to syntax: the case of Black English *been*. In *Analyzing the form and use of language*, ed. by Fasold & Shuy, 162-183. Wasington DC: Georgetown University Press.
Rickford, John. 1978. The question of prior creolization in Black English. In *Pidgin and Creole Linguistics*, ed. by Valdman, 126-146. Bloomington, IN: Indiana University Press
Rickford, John. 1998. A creole origin of African American Vernacular English. In *African American English: Its history, origin, and use*, ed. by Bailey, Baugh, Mufwene & Rickford, 154-200. London: Routledge.
Rapoport, Tova R. 1987. Copula, nominal and small clauses. A study of Israeli Hebrew. Ph.D. dissertation. MIT.
Safir, Kenneth. 1985. *Syntactic chains*. Cambridge: Cambridge University Press.
Sag, Ivan A. & Thomas Wasow. 1999. *Syntactic Theory: A Formal Introduction*. Stanford, CA: CSLI.
Stowell, Tim. 1981. Origin of phrase structure. Ph.D. dissertation, MIT.
Seuren, Pieter. 1985. *Discourse semantics*. Oxford: Blackwell.
Walker, James A. 2000. Present Accounted For: Prosody and Aspect in Early African American English. Ph.D. Dissertation, University of Ottawa.
Wolfram, Walt. 1974. The relationship of Southern White Speech to Vernacular Black English. *Language* 50, 498–527.
Wolfram, Walt. 2004. The grammar of urban African American English. In *Handbook of varieties of English*, ed. by Kortmann & Schneider, 111-132. Berlin: Mouton de Gruyter.
Wood, Jim. 2008. *So*-Inversion as polarity focus. *Proceedings of the 38th Western Conference on Linguistics*, 304-317. Fresno, CA: University of California.
Zaring, Laurie. 1996. Two *be* or not two *be*: Identity, predication and the Welsh copula. *Linguistics and Philosophy* 19, 103-142.

Backness Agreement in Consonant + Glide Onsets in Mandarin

Yang Liu, Lori Repetti
Stony Brook University

1 Introduction

Most languages have restrictions on Consonant + Glide (CG) sequences, and Mandarin has a particularly complex system. An inventory of Mandarin CG sequences can be found in (1).[1]

(1) Mandarin CG sequences[2]

		j Dor [-ba]	ɥ Lab,Dor [-ba]	w Lab,Dor [+ba]
Lab [0ba]	p		■	■
	pʰ		■	■
	m		■	■
	f	■	■	■
Cor [0ba]	t		■	
	tʰ		■	
	n			
	l			
	ts	■	■	
	tsʰ	■	■	
	s	■	■	
	tʂ	■	■	
	tʂʰ	■	■	
	ʂ	■	■	
	ʐ	■	■	
Pal (Cor,Dor) [-ba]	tɕ			■
	tɕʰ			■
	ɕ			■
Dor [+ba]	k	■	■	
	kʰ	■	■	
	x	■	■	

Many linguists study CG restrictions by focusing on articulator features: Labial, Coronal, Dorsal (Duanmu 2000, Hume 1990, Kochetov 2016), and Duanmu (2000: 32) argues that restrictions on Mandarin CG onsets can be accounted for by the Articulator Dissimilation Principle[3]: "Identical articulators cannot

[1] Shaded cells mean that CG sequence is absent.
[2] The consonant [ʐ] has been analyzed as an obstruent or a sonorant (Duanmu 2000). We analyze it as a voiced fricative since it behaves more like an obstruent than a sonorant: (i) all initial sonorants in Mandarin can be followed by [j], yet [ʐ] + [j] is missing; (ii) no non-palatal fricatives + [j] sequences are allowed, and neither is *[ʐj].
[3] See also Yi & Duanmu (2015) and Gong, S. & Zhang, J. (2019) for Articulator Dissimilation.

Proceedings of *WECOL 2019*

occur in succession." However, this proposal does not cover all onset data. Mandarin onsets like [fj] are ungrammatical even though the initial consonant [f] and the following glide [j] have different articulators. What constraints penalize [fj]? Furthermore, Duanmu does not discuss Mandarin palatals [t͡ɕ, t͡ɕʰ, ɕ]. What constraints penalize palatals + [w], but not palatals + [j, ɥ]? In this paper, we discuss restrictions on Mandarin CG onsets and show the need for both articulator features and backness agreement.

2 OCP and Backness Agreement Constraints

We propose that the presence and absence of particular CG sequences in Mandarin can be accounted for using two types of constraints: an OCP constraint (CG: *LabLab) and a group of backness agreement constraints (CG: backness agreement).

(2) CG: *LabLab - labial consonant and a labial glide sequence are not legal
(3) CG: backness agreement - a consonant + glide sequence must have the same backness feature specification

The OCP constraint (2) is in line with Duanmu's (2000) Articulator Dissimilation Principle, so that a labial consonant ([p, pʰ, f, m]) followed by a labial glide ([ɥ, w]) is not a legal sequence. The constraint would rule out sequences such as *[pw] but not [pj] or [tw].[4]

In order to motivate the backness agreement constraints, we have to clarify the backness feature specifications for consonants and glides. Since [back] is a subfeature of the DORSAL node, only velar consonants, palatal consonants, and vowels/glides have a backness feature. Following Riggle (2011), we specify the palatal consonants [ɕ, t͡ɕʰ, t͡ɕ] as [-back], and the velar consonants [kʰ, k, x, ŋ] as [+back]. Furthermore, we follow Duanmu (2000), Hayes (2011), Kenstowicz (1994), Riggle (2011), and others in analyzing all vowels, and therefore glides, as having a DORSAL place node: [j, ɥ] are [-back], [w] is [+back].[5] The labial consonants and non-palatal coronal consonants do not have a back feature specification ([0back]).

(4) [back] specification

palatal consonants	[ɕ, t͡ɕʰ, t͡ɕ]	[-back]
velar consonants	[kʰ, k, x, ŋ]	[+back]
front glides	[j, ɥ]	[-back]
back glide	[w]	[+back]

The backness agreement constraint in (3) does not apply to all CG sequences in Mandarin. It is sub-specified for certain types of consonants and glides only, as defined in (5a-5c).

(5) backness agreement
(a) Agree[back]:DorG - An initial dorsal consonant (palatal or velar) and any following glide ([j ɥ w]) must have the same backness value.
(b) Agree[back]:$C_{[-son]}$ ɥ - An initial obstruent and the following glide [ɥ] must have the same backness value: [-back]
(c) Agree[back]:$C_{[-son, +cont]}$ j - A [-son, +cont] consonant (affricate or fricative) and the following glide [j] must have the same backness value: [-back][6]

[4] OCP-based phonotactic constraints contribute to Mandarin non-word judgment (Gong, Shuxiao. & Zhang, Jie. 2019).
[5] The glides [ɥ] and [w] also have a LABIAL feature, and Duanmu (2000) claims that [w] is LAB but not DOR. Some have argued that the front glides [j, ɥ] can also be analyzed as having a CORONAL articulator (Broselow & Niyondagara 1991, Hume 1994).
[6] We analyze affricates as both [-cont] and [+cont] following Riggle (2011) and others. However, others label affricates as [-cont] (Hayes 2011).

The backness agreement constraints in (5) are always violated if the consonant does not have a backness feature specification. Hence, the relevant labial consonant + glide sequences or non-palatal coronal consonant + glide sequences always violate these constraints.

3 Legal and Illegal Cj, Cɥ and Cw Sequences

In order to account for the attested and unattested Cj sequences, we need two constraints: Agree[back]:C$_{[-son, +cont]}$j and Agree[back]:DorG. The former rules out all fricatives and affricates followed by [j], except the palatal consonants followed by [j] since they share a [-back] feature specification. The latter constraint rules out the velar consonant + [j] sequences since velar consonants are [+back] and [j] is [-back]. Note that we need both constraints because the velar consonants are not all covered by the first constraint since [k, k^h] are [-cont]. All other Cj sequences involving stops and sonorants are legal.

What about Cɥ sequences? Two constraints allow us to account for the attested and unattested forms: *LabLab and Agree[back]:C$_{[-son]}$ɥ. The OCP constraint rules out labial consonants ([p, p^h, m, f]) followed by [ɥ]. The agreement constraint rules out all obstruents followed by [ɥ] that do not share a backness feature, i.e., all of the non-palatal obstruents since only the palatal obstruents share the [-back] feature specification with [ɥ]. Note that another agreement constraint also applies to some of these sequences: Agree[back]:DorG also applies in some of the same contexts as Agree[back]:C$_{[-son]}$ɥ because all dorsals in Mandarin are [-son] (obstruents); however, Agree[back]:DorG would not allow us to account for the absence of the non-palatal coronal obstruent + [ɥ] sequences.

In order to account for Cw sequences, we need two constraints: *LabLab and Agree[back]:DorG. The first one rules out all cases of [w] preceded by a labial consonant. The second one rules out all cases of [w] ([+back]) preceded by a palatal consonant ([-back]). All other sequences are allowed.

4 Conclusion

Duanmu (2000)'s Articulator Dissimilation claim is not explanatory enough for Mandarin onsets. Articulator features and backness values are both needed for the four constraints on Mandarin CG sequences: *CG-LabLab, Agree[back]:DorG, Agree[back]:C$_{[-son]}$ɥ and Agree[back]:C$_{[-son, +cont]}$j. Mandarin requires both OCP and anti-OCP constraints: dissimilation of major articulator features for distinctiveness of phonetic cues, and agreement of dependent features like backness for ease of articulation.

5 References

Broselow, Ellen & Niyondagara, Alice. 1991. "Morphological Structure in Kirundi Palatalization: Implications for Feature Geometry." in F. Katamba (ed.) *Afrikanistische Arbeitspapiere* 25:131-155.

Duanmu, San. 2000. *The Phonology of Standard Chinese*. Oxford: Oxford University Press.

Gong, Shuxiao. & Zhang, Jie. 2019. "The Gradient Acceptability in Mandarin Nonword Judgment." Conference presentation at the Annual Meeting of Phonology. Stony Brook University, US.

Hayes, Bruce. 2011. *Introductory Phonology*. Hoboken, NJ: Wiley-Blackwell.

Hume, Elizabeth. 1990. "Front Vowels, Palatal Consonants and the Rule of Umlaut in Korean" in *Proceedings of NELS 20*, GLSA, Department of Linguistics, University of Massachusetts, Amherst, pp. 230-243.

Hume, Elizabeth. 1994. *Vowels, Coronal Consonants and their Interaction in Nonlinear Phonology.* New York: Garland Publishing.

Kenstowicz, Michael. 1994. *Phonology in Generative Grammar.* Cambridge, Mass. & Oxford: Blackwell.

Kochetov, Alexei. 2016. "Palatalization and glide strengthening as competing repair strategies: Evidence from Kirundi." *Glossa: A Journal of General Linguistics* 1(1): 14. 1–31, DOI: http://dx.doi.org/10.5334/gjgl.32

Riggle, Jason. 2011. Phonological Feature Chart – University of Chicago version 11.02. http://hum.uchicago.edu/~jriggle/

Yi, Li & Duanmu, San. 2015. "Phonemes, Features, and Syllables: Converting Onset and Rime Inventories to Consonants and Vowels." *Language and Linguistics* 16(6), 819–842.

Spanish Language Influence on English Language and Vice Versa

Gia Mosashvili

Ivane Javakhishvili Tbilisi State University

1 Introduction

Spanish is a member of the Indo-European languages, whose earliest ancestor was spoken approximately 5, 000 years ago in the area of the Black Sea. Speakers of Indo-European gradually spread in various directions and varieties of their speech came to be used in enormously broad areas: almost all Europe, the greater part of the Indian subcontinent and many territories in between Armenia, Iran, much of Afghanistan. During the migration process, each group of Indo-European speakers lost contact with other groups, so that innovations and losses originating in one group could not spread to others. Scholars typically recognize nine surviving branches of the Indo-European family: Indo-Iranian, Slavonic, Germanic, Italic, Baltic, Hellenic, Armenian, Albanian and Celtic. As for two other branches, Tocharian and Hittite have left substantial written records but are no longer spoken. The most prominent member of the Italic branch of Indo-European is Latin, which is the ancestor of Spanish language. An alternative way of expressing the relationship between Latin and Spanish is to say that Spanish is Latin and Latin continues to be spoken in parts of Europe, Africa and America. It has been known that Romance languages do not descend from Classical Latin, but from non-literary varieties such as Vulgar Latin and features of Vulgar Latin are revealed as more advanced than the features of Classical Latin (Penny, R. 2002).

English language is related to the dialects spoken by the Germanic tribes, who came to England in the manner described. It is difficult to say how the speech of the Angles differed from the Saxons or from the Jutes. These dialects had been subjected to several centuries of geographical and political separation in England. English language belongs to the Low West Germanic branch of the Indo-European family. This means that it shares certain characteristics common to all the Germanic languages. English belongs with German and certain other languages because of features it has in common with them and that enable us to distinguish a West Germanic group as contrasted with the Scandinavian languages (North Germanic) and Gothic (East Germanic) (Baugh, A.C., Cable, T. 2002). Linguistic forecasters agree that in the 21st century, English and Spanish will number amongst the dominant languages of the world. Both languages have vast numbers of native speakers, falling within the top five languages of the world from this point of view. Spanish language is reported as having already overtaken English in this respect: Mandarin Chinese with speakers of 885,000,000, Spanish 332,000,000, English 322,000,000, Bengali 189,000,000, Hindi 182,000,000 and Portuguese 170,000,000. They are both spoken as native or official languages over vast geographical areas (English comes first, Spanish fourth) which cover a large number of countries. They are both official languages of the United Nations and many other international organizations (Christopher, J. Pountain. 1999). The Bologna declaration for Higher Education has been implemented in all Spanish universities in the academic year 2010-2011. It should be highlighted that this process made many changes in Spanish University system. English gained a more prominent role in Higher education of Spain' (Lujan, C. 2016).

Proceedings of *WECOL 2019*

2 Observations

The University of Texas, with its bilingual studies, suggests some general rules, which are helpful when English speakers learn Spanish or Spanish speakers learn English. For example, if many English words have double consonants (happy, puddle), Spanish words do not, except c and n (*acción*, *innecesario*). Also, words with ph in English (*pharmacy*) are written with f in Spanish (*farmacia*). Cognates are words converted between English and Spanish. English and Spanish cognates share the same Latin root and are very similar. Thousands of English words exist, converted into Spanish with their English meaning (Cobarrubias, R., Green, N. 1996).

- Most words with **tion** ending in English end with **ción** in Spanish.
  invitation/invitación, occupation/ocupación, conversation/conversación

- Nouns with **or** and **our** endings are often identical in both languages.
  doctor/doctor, tractor/tractor, interior/interior

- Many nouns and adjectives in English with **al** endings are often identical of Spanish.
  animal/animal, capital/capital, hospital/hospital

- Nouns with **ist** ending in English end with **ista** in Spanish.
  artist/artista, dentist/dentista, novelist/novelista

- English nouns with **ty** endings end with **dad** in Spanish.
  variety/variedad, society/sociedad, electricity/electricidad

- Nouns with **nce** ending in English end with **ncia** in Spanish.
  distance/distancia, coincidence/coincidencia, independence/independencia

- In English, many adjectives with **ic** ending end with **ico** in Spanish.
  romantic/romantico, metallic/metálico, dramatic/dramatico

- Adjectives with **ive** ending in English end with **ivo** in Spanish.
  negative/negativo, expressive/expresivo, positive/positiv

- Adjectives with **ous** ending in English end with **oso** in Spanish.
  famous/famoso, nervous/nervioso, delicious/delicioso

- Adjectives with **ble** endings in English end with **ble** in Spanish.
  horrible/horrible, impossible/imposible, terrible/terrible

- Many adjectives with **nt** ending in English end with **nte** in Spanish.
  ignorant/ignorante, convenient/conveniente, patient/paciente

- In English, adjectives with **id** ending end with **ido** in Spanish.
  stupid/estúpido, rapid/rápido, splendid/espléndido

- Many English adjectives with **ile** ending end with **il** in Spanish.
  fertile/fértil, hostile/hostil, mobile/móvil

- English adjectives with **ary** endings end with **ario** in Spanish.
  secondary/secundario, dictionary/diccionario, literary/literario

- Almost every English infinitive verb with **ate** (celebrate) ending, can be converted in Spanish infinitive by replacing the final **ate** with **ar** (celebrar).

create/crear, calculate/calcular, concentrate/concentrar

- English verbs (infinitive) ending with **t** (result) can be converted into Spanish infinitive verbs by adding **ar**, **er** or **ir** to the end of the English verb (resultar)
  Insult/insultar, consult/consultar, present/presenter

- Many English infinitive verbs ending with **e** (examine) can be converted into Spanish infinitive verbs by dropping the final **e** and adding **ar** (examinar).
  Accuse/acusar, adore/adorer, authorise/autorizar

- Almost every English infinitive verb ending with **ify** (signify) can be converted into a Spanish verb by replacing the final **ify** with **ificar** (significal).
  unify/unificar, simplify/simplificar, solidify/solidificar

- In Spanish, **mente** combines with (feminine) adjectives to form Spanish adverbs. In English, **ly** combines with many adjectives to form adverbs.
  normally/normalmente, naturally/naturalmente, finally/finalmente

- English nouns with **y** ending often correspond to Spanish nouns with **ia**, **la** or **lo**.
  agency/agencia, day/dia, mercury/mercurio

- Many English nouns with **ism** ending can be converted into Spanish nouns.
  gravatropism/gravatropismo, magnetism/mognetismo, organism/organismo

accident	accidente	bat	bate
activities	actividades	biography	biografía
addition	adición	biology	biología
agriculture	agricultura	biomass	biomasa
air	aire	botany	botánica
air quality	calidad del aire	calcium	calcio
analyze	analice	calculators	calculadoras
analyze	analizar	calories	calorías
animals	animales	camel	camello
appropriate	apropiado	cape	capa
area	área, superficie	carbon cycles	ciclo del carbono
artificial	artificial	career	carrera
asteroids	asteroides	catastrophic	catastrófico
astronaut	astronauta	cells	células
atmosphere	atmósfera	center	centro
atomic	atómica	centimeters	centímeros
atoms	átomos	cereal	cereal
axis	axis	certain types	ciertos tipos
babies	bebés	characteristics	características
balances	balanzas	chemical properties	propiedades químicas
banana	banana	circle	círculo
baseball	beisbol	circular	circular
basketball	básquetbol,baloncesto	classify	clasifican
classify	clasificar	disadvantages	desventajas
color	color	discoveries	descubrimientos
comets	cometas	distance	distancia
communicate	comunique, comunicar	distribute	distribuir

compasses	compases	divide	dividir
complex	complejo	division	división
components	componentes	dominant	dominante
composed	compuestos	east	este
comprehend	comprender	ecological succession	sucesión ecología
computer	computadora	ecosystem	ecosistema
computers	computadoras	electrical	eléctrica
comsumers	consumidores	evidence	evidencia
concept	concepto	examine	examinar
conceptual	conceptuales	exercise	hacer ejercicio
conclusions	conclusiones	exothermic	exotérmico
conservation	conservación	experiments	experimentos
continental	continental	expert	experto
consistency	consistencia	explain	explicar
constantly	constantemente	explain	explique
constancy	constancia	explanations	explicaciones
contact	contacto, contactar	external	externos
contributions	contribuciones	extinction	extinción
correct	correcto	extrapolate	extrapolar
correctly	correctamente	favorites	favoritos
critical	crítico	fever	fiebre
critics	críticos	figure	figura
cube	cubo	finally	finalmente
cycle	ciclo	force	fuerza
data	datos	formalate(-ing	formular
day	día	formula	fórmula
decide	decidir	formulas	formulas
decisions	decisiones	frequency	frecuencia
define	defina(e), definir	fresh	fresco
definition	defínción	fruit	fruta
delicate	delicado	function	función
describe	describa	future	futuro
describe	discribir	galaxy	galaxia
described	descrito	garden	jardín
descriptions	descripciones	gas	gasolina
diagram	diagrama	general	general
different	diferente	generation	generación
direct	directa/directo	generations	generaciones
direction	dirección	genes	genes
electricity	electricidad	genético	genética
electronics	electrónicos	genetic	genetic
elements	elementos	geology	geología
elephants	elefantes	globe	globo terráqueo
emergency	mergencia	gradual changes	cambios graduales
endothermic	endotérmico	graduated cylinders	cilindros graduados
energetic	enérgico	graphically	gråficamente
energy	energía	groups	grupos
estimate	estimar	Guide	guía
ethical practices	prácticas éticas	graphics	gráficas
equilibrium	equilibrio	group	grupo
equipment	equipo	hamburger	hamburguesa
evaluate	evaluar	history	história

events	eventos	human	humano
hurricanes	huracanes	mechanical	mecánica, mecánico
hydroelectric	hidroeléctrica/o	medal	medalla
hypothesis	hipótsis	medicine	medicina
idea	idea	metallic	metálico
ideas	ideas	meteorites	meteoritos
identify	identifican	meteorology	meteorologia
identify	indentificar	meters	metros
impact	impacto	methods	métodos
important	importante	microscopes	microscopios
including	incluyendo	minerals	minerales
indirect	indirecta	minus	menos
individual	individual	minutes	minutos
inexhaustible	inagotable	movement	movimiento
inferences	inferencias	mule	mula
information	información	multiplication	multiplicación
informed	informadas	multiply	multiplicación
inheritance	herencia	multiply	multiplicar
initial	inicial	musica	música
instructions	instrucciones	mustard	mostaza
instrument	instrumento	natural	natural
instruments	instrumentos	nature	naturaleza
interactions	interacciones	nitrogen	nitrógeno
interdependence	interdependencia	non-renewable	no renovable
interesting	interesante	normal	normal
internal	interna	nutrition	nutrición
internal	internos	object	objetos
internet	internet	observations	observaciones
interpret	interpretar	observed	observar
introduction	introduccción	obtain	obtener
jog	hacer jogging	ocean	océano
kinetic energy	energía cinética	oceans	océanos
laboratory	laboratorio	operation	operación
lemonade	limonada	opinion	opinión
lesson	lección	opinions	opiniónes
limit	Limite	orbit	orbita
limitations	limitacioes	orbits	órbita
list	Lista	organ systems	sistemas de órganos
mammals	mamiferos	organizar organize	organismos organisms
manual	manual	organized	organizado
maps	mapas	organs	órganos
marine biology	biología marina	origin	origen
material	material	original	originales
materials	materiales	oxygen	oxígeno
math	matemáticas	paint	pintar
mathematical	matemáticos	paper	papel
matter	materia	parts	partes
pasta	pasta	rock cycle	ciclo de la roca
percent	porcentaje	salt	sal
perimeter	perimetro	sandwich	sándwich

periodic table	tabla periódica	science	ciencia
phases	fases	scientists	científicos
phenomenon	fenómeno	scientist	Cientista, cientifico
physicists	fisicos	sexual reproduction	reproducción
physics	física	simple machines	simple maquinas
pine	pino	society	sociedad
pizza	Pizza	solar system	sistema solar
plan	planear	solid	sólido
plants	plantas	solution	solución
plastics	plásticos	solve	resolver
position	posición	soup	sopa
potatoes	patata (papas)	space	espacio
potential energy	energía potencial	species	especies
presence	presencia	specific	específico
probable	probablemente	statue	estatua
problem	problema	stimuli	estímulos
produce	producir	structure(s)	estrucura
production	producción	substances	substancias
products	productos	suggested	sugeridos
promotion of products and services	proyecto project promoción de productos y servicios	summary	summarize
protect	proteger	surface	superficie
protein	proteína	sustain	Sustentar, sostener
pyramid	pirámide	systems	sistemas
radioactive	radioactivo	tables	Tablas,mesa
radioactivity	radioactividad	tacos	tacos
rat	Rata	technical	técnico
readian energy	energía rediante	telephone	teléfono
reasonable	razonables	telescopes	telescopios
recessive	recesivo(a)	television	televisión
recommendations	recomendaciones	terms	términos
rectangle	rectángulo	text	texto
recyling	reciclaje	theories	teorías
reference	referencia	thermometers	termómetro
references	referencias	title	título
renewable	renovar	tomato	tomate
represent	representar	total	Total
represented	representado	train	tren
reptiles	réptiles	transform	transormar
resource	recurso	transformations	transformaciones
results	resultados	tubes	Tubos
resume	resumír	types	Tipos
universe	universo	volcanic activity	acividad volcánica
uranium	uranio	volcano	volcán
use	Usar	vomiting	vomitar
usually	usualmente	water cycle	ciclo del agua
valid	válidas	water vapor	vapor de agua
varieties	variedades	web site	sitio web
variety	variedad	zebra	Cebra
vet	veterinario/a	zero	Cero
vitamims	vitaminas		

According to the linguistic work 'Spanish Language Influences on Written English' by Betty Rizzo and Santiago Villfane, there are some errors in English language caused by Spanish interference:

*The **first** I would do is to buy myself a beautiful house.*

*But who really is the **responsible** for people's education?*

*I don't understand why if **a young** is fifteen years old **has** to go from junior high school to high school without adequate preparation.*

*The reason a child may see and hear things more differently is because **has** no reason to see things otherwise.*

Is not only the people itself.

*But **is** a matter of fact that I am always trying to get more and more education for myself because I realized that **is** absolutely necessary.*

- The first three errors are related to the fact that in Spanish the combined article and adjective can be normalized. ***The beautiful*** in Spanish means ***the beautiful one***, ***the young*** means ***the young one*** and so on. The remaining five errors are conditioned by the fact that in Spanish the subject pronoun may be omitted because verb forms are differentiated and the person and number of the pronoun are perfectly clear.

- Spanish-speaking students sometimes omit **n** on the indefinite article before a word beginning with a vowel sound, because refinement has no equivalent in Spanish.

- The omission of the **n** from the past tense and from the past participle, as in the example above, is difficult to attribute to a principle connected particularly with Spanish. This omission is common to many non-Spanish-speaking students. But it is the second problem, that the final **t** and **d** are very difficult to pronounce, they tend to be dropped in spoken English. For instance, *It is a place fill with rats and filth lying everywhere, a place where one cannot walk the streets safely at night for fear of being **attack**.*

- Beside the omission caused errors, there are common interference errors of addition: *The problem with the American Students is that they don't worry about **the school**. I hope and I pray that by 1980 **the education** and the **students** will be of a better class.*

- In Spanish the definite article is used before generalized and abstract nouns - *In **public** schools, as far as I can see, seems like nobody cares about the student.* The student added **s** on public after proofreading. *It is because I want to know some **personals** things of that person.*

- The double negative form is permissible in Spanish, so an extra negative may frequently be added to an English sentence by a Spanish-speaking student. For example, *she did not do **nothing** about **nothing**. I feel that now children are **not** learning absolutely **nothing** for schools.*

- Sometimes, Spanish-speaking students are confused, when there is a distinction in English which is not present in Spanish. For instance, in Spanish the relative pronoun ***que*** means ***that***, ***which***, ***what***, ***who*** or ***whom***. The Spanish-speaking student frequently fails to distinguish between ***which***, referring to objects and ***who***, reffering to people. *Probably, comparing a student from another country, **which** has only gotten to eight grade has a better knowledge academically than a High School graduated student from the U.S.*

• An error may occur while two grammatical forms exist in both languages. English verbs may take as a complement either the infinitive or the gerund, while almost all Spanish verbs must take the infinitive. The Spanish speaker tend to use the infinitive, when in English the gerund is the correct form. *I always dreamed **to live**.*

• Error may be caused by the fact that **mucho** in Spanish is the equivalent of both **much** and **many** in English. *The married woman has **much** financial troubles.*

• In Spanish, the most common method of comparison is using of the adverb **más** for both **more** and **most**. Spanish–speaking students have difficulty in handling the distinction between comparative and superlative. **Worse** and **worst** seem most frequently to be confused'. For example: *I had receive a good education better than many people that are around me and **worst** than many around me.*

• As for the problem of infinitives in Spanish and English, **to** infinitive may be omitted, while it is required in the translation. *What we must **to do** is get a better education.* **Must** in English and **tengo que** in Spanish are the same.

• The problem is also literal translation of idiom and prepositions. *If you want your child to grow **in a right way**, give him a better community housing environment.* **In a right way** = **de un modo recto**. ***All the while** the heads of these countries in the Middle East are at war, the people who was not willing to die and want peace are the ones who are fighting*. **All the while = mientras tanto**. *I **made** no notice = no le **hise** caso. I **made** a line = yo **hise** fila.*

• Spelling problems are sometimes apparently connected to pronunciation problems and seem to arise because the student is familiar with different spelling of Spanish word. Also, when student is not very familiar with Spanish spelling, he tries to make English words conform to the Spanish norm. The letter **i** in Spanish and **i** in English seems to cause two different kinds of spelling errors. The student uses **i** to spell the sound **i**. For example: *Nixon has proven himself irresponsible and **deciving** to America and the world. One must **perceive** things as an adult.* An error may occur because the speaker has been mispronouncing the word.

• Consonants are rarely if ever doubled in Spanish and often a doubled consonant in English is the major difference between two closely related words. At this time, the student tends to drop a consonant and when there is no closely equivalent Spanish word, the student may tend to change a doubled to a single consonant. *The needs of writing and reading skills of a high order will be more **necesary (necesario)**. I have received a fragmented education which taught me a lot of knowledge about **diferent (diferente)** thing, without educating me. The English language is teached in a very poor **maner (manera)**.It requires all what this society needs, that is: more interaction, understanding and **comunication (comunicación)** of old and new ideas.*

• After the student becomes aware that a word has a doubled consonant, he/she frequently seems to double the wrong consonant. For example, *I consider for me specially is kind of hard because of the language, but not impossible. The kind of education we receive is bad enough to destroy people from beginning.*

• Spanish-speaking students seem to misspell an English word to make it accord more closely with a Spanish equivalent or with Spanish phonetic principles: *I don't understand why if a young is fifteen years old has to go from junior high school without **adecuate (adecuado)** preparation.*

Cognates are words that share the same Latin or Greek root in Spanish and English. They are very similar in spelling and have the same or similar meaning. However sometimes words seem to be cognates, but are not. They are called false cognates and create one of the biggest challenges for ELL and Spanish language students. False cognates are words that have the same root, sound alike, but different meanings:

Spanish word	**Meaning in English**	**False friend in English or false cognate**	**Translation in Spanish of the false cognate or false friend**
abandonar	to leave, to quit	abandon	desamparar
absoluto	absolute	absolutely	absolutamente
actual; actualmente	present moment	actual	verdadero
advertencia	warning	advertisement	anuncio
agenda	a schedule for an event;	agenda	diario
agonía	dying moments	in agony	dolor fuerte
alcoba	bedroom	alcove	nicho
alterado	angry, upset	altered	modificado, cambiado
ansioso	eager	anxious	nervioso, inquieto
aplicar to	apply (a theory)	apply for (a job)	solicitar
apreciado	esteemed	appreciate	apreciar
arena	sand	arena	anfiteatro
argumento	debate	argument	disputa
asignatura	subject	signature	firma
asistencia	attendance	assistance	ayuda
asistir	to attend	to assist	ayudar
atentado	an attack	attempt	intento
avisar	to warn	advise	aconsejar
bachillerato	four year of college	bachelor, single man	soltero
balde	bucket	bald	calvo

3 Findings

In major research, three works analyzed and researched in which English-Spanish cognates are found in order to prove linguistic influences. The first work is 'Don Quixote' by Miguel De Cervantes, which is written in the XVII century. The second book is 'The Family of Pascual Duarte' by Camilo Jose Cela, written in the XIX century. And the third book is 'Invisible Man' by Ralph Elisson, written in the XIX century. The first chapter of 'Don Quixote' includes 79 true and 4false cognates.

True cognates: *parte/part, famoso/famous, antique/antique, el resto/the rest, fine/fin, diferencia/difference, autores/authors, importa/important, administración/administration, curiosidad/curious, prosa/prosa, razónes/reasons, partes/parts, sinrazón/unreason, perlas/pearls, manera/manner, divinidad/divinity, divinamente/divinely, imaginaba/imagine, grande/grand, promesa/promise, aventura/adventure, muchas/much, mayores/mayors, continuous/continues, competencia/competention, graduado/graduate, barber/barber, acomodada/accommodate, condición/condition, resolución/resolution, celebro/celebrate, fantasia/fantasy, batallas/battles, impossible/impossible, imaginación/imagination, gigantes/giants, affable/affable, historia/history, necesario/necessary, honra/honor, servicio/service, menos/minus, armas/armour, sentia/sense, morrión simple/simple morin, suplió/supplied, industria/industry, cartones/cartoons, entera/enter, facilidad/facility, nuevo/new, barras/bars, satisfecho/satisfied, experiencia/experience, finisima/finest,*

real/real, imaginar/to image, acomodársele/accommodate, formo/form, memoria/memory, significativo/significant, imaginación/imagination, ocasión/occasion, contentado/content, declaraba/declar, confirmándose/confirmation, faltaba/fault, fruto/fruit, ordinario/ordinary, finalmente/finally, presentado/present, talante/talent, discurso/discurse, laboradora/laboratory, titulo/title, natural/natural, músico/music, generación/generation.

False cognates: *nombre/number* - *i*t looks like an English word '*number'*. But while the term 'nombre' means 'the name' in Spanish, English 'number' has not the same meaning. *Campo/camp* – 'campo' means '*countryside'* in Spanish, but English word camp has not the same meaning. *Pluma/plum* – 'pluma' is a pen in Spanish, while 'plum' in English is a kind of fruit. *Disparate/desparate - 'disparate' in Spanish* means '*absurdity', 'folly' and 'blunder'*, while the term '*desperate'* in English means 'different'. Very often, it is difficult to find a difference between these terms.

In the first chapter of 'The Family of Pascual Duarte' includes 101 true and 20 false cognates are found: *motivos/motives, mortales/mortals, destino/destination, variarnos/vary, diferentes/different, ordena/order, marchar/to march, aroma/aroma, fragrance, felicidad/felicity, inocente/innocent, sufren/suffer, violent/violent, defenderse/defend, diferencia/difference, colonia/colony, tatuajes/tattoos, muchos/much, menos/minus, provincial/province, dia/day, larga/large, figurarse/figure, rico/rich, olivos/olives, varios/various, agua/aqua, airosa/airy, elegante/elegant, desnudo/nude, grande/grand, tabaco/tobacco, medio/medium, hostia/host, nueve/new, necesitase/necessary servicio/service, natural/natural, recibidor/receiver, plantas/plants, rara/rare, capital/capital, aparecia/appear, colour/color, odinario/ordinary, portal/portal, antigüedad/antiquity, representar/to represent, parte/part, campanario/companion, manera/manner, contar/to count, memoria/memory, correspondia/corresponded, posición/position, sentia/sense, realidad/reality, único/unique, entrar/enter, moderno/modern, platos/plates, finos/fine, letras/letters, modesto/modest, fotografias/photography, regular/regular, sitio/site, funciono/function, resultaba/result, el resto/the rest, describirlo/describe, vulgaridad/vulgarity, temporadas/temporary, habitaciones/habitation, ocasiones/occasions, principio/principle, ocuparse/occupy, prefirió/prefer, construidas/constructed, Agosto/August, bestia/beast, animales/animals, ocurria/occur, tortuga/tortoise, eléctricas/electric, habitantes/habitants, ignoraban/they ignored, momento/moment, imaginando/imagining, dobla/double, dedicaba/dedicate, modestia/modesty, aparte/apart, sinceridad/sincerity, dio/day, bravia/brave, raya/ray, persona/person, enfrente/in front, dominados/dominated, extendia/extended.*

False cognates: *sendas/send* - the word *'sendas'* means *trails, track in Spanish* and *i*t sounds like English '*send', which has different meaning*. *Fin/fine* - '*fin*'means '*the end'* in Spanish, but the term '*fine*' in English, means beautiful. There is a Spanish word 'fino', which is a true cognate of English word '*fine',* but it is difficult for speakers to find distinction between these words. *Flores/flora* - Spanish word *'flores'*means *'flowers', while 'flora'* in English is a group of plants and so on. *Mirar/mirror – 'mirar'* in Spanish means *'looking at someone'*, but the word *'mirror'* has different meaning in English. *Bañera/banner - 'bañera'* means *'bath'*in Spanish and the word *'banner'* in English, has different meaning. *Son/son* – '*son*' means '*they are'* in Spanish, while the term *'son'* in English is a synonym of Spanish *'hijo'. Tales/tales – 'tales'* in Spanish means *'such', 'like'*, while '*tale'* in English is a short story for kids. *Tallos/tall* – '*tallos*' in Spanish means *'stems',* but in English, the term '*tall'* has different meaning. *Plumas/plum* - these words are described above. *Anterior/interior* - '*anterior'* is *'previous'* in Spanish, while *'interior'* means*'inner part'. Recuerdos/records* – these words are described above. *Flor/floor – 'flor'* in Spanish means '*flower',* while some speakers interpret as *'floor',* which has different meaning in English. *Nombre/number* – these words are described above. *Mobiliario/mobile - 'mobilario'* in Spanish means *'*furniture', while '*mobile'* is *'a gadget'* in English. *Mayor/mayor – 'mayor'* in Spanish means *'higher', 'major'*, while *'mayor'* in English means *'governor of the* city'. *Pero/per - 'pero'* means *'but'* in Spanish, while *'per'* in English means *'the one'. Acostar/cost* -

'*acostar*' means *'to lay down'* in Spanish, but in English, *'the cost'* is means '*the price of a product'*. *Tres/trees* - '*tres'* is a number *'three'* in Spanish, but 'trees' in English has different meaning. *Ruin/ruin – 'ruin'* in Spanish means 'to *mean'*, but English '*ruin*' means the same as other Spanish word '*ruina'*. *Disparar/disappear* - '*disparar'* in Spanish means 'to *shoot'*, but English '*disappear'* means the same as other Spanish word – '*disaparecer'*.

'Invisible Man' by Ralph Elisson involves loads of true cognates, generated from Spanish, German, French, Anglo-French and Latin roots: *contradiction/contradicción* (Middle English, Latin and Greek roots), *discover/discubrir* (Middle English, Anglo French, Old French and Late Latin), *invisible/invisibl*e (Middle English and Latin), *separate/separate* (Late Middle English and Latin roots), *place/platea* (Old English, Middle French, Latin and Greek roots), *destruction/destrucción* (Middle English, Anglo-French and Latin roots), *family/familia* (Middle English and Latin roots), *graduation/graduación* (Late Middle English and Medieval Latin roots), *invite/invite* (Latin root), *occasion/occasion* (Middle English, Old French and Latin), *portable/portable* (Middle English and Latin roots), *suspect/sospechar* (Middle English and Latin roots), *battle/batalla* (Middle English, Old French and Late Latin roots), *important/importante* ('*important'* is Latin origin), *continue/continuar* (Middle English, Anglo French and Latin roots), *complete/completer* (Middle English, Middle French and Latin roots), *terror/terror* (Latin, Middle English and Anglo-French roots), *nigger/negro* (Spanish and French origin), *voice/voz* (Middle English, Anglo-French, Old French and Latin roots), *colour/color* (Middle English, Anglo French, French and Latin roots).

4 Conclusion

English and Spanish are international languages with the same historical roots, which, as the factor, had an influence on the development of both languages. Findings, major research, has shown how deep and historical is mutual influence between English and Spanish. Also, the process of globalization had an important impact on creating cognates in both languages with the same or different meanings. English and Spanish are very popular in students, consequently, difficulties related to the learning of these languages are very common problems between English and Spanish speaking speakers. In order to avoid mistakes on both sides, it is necessary to defend grammatical rules and keep the balance in order not to complicate the way we learn foreign languages. It would be better, if all learning courses and programs involved information about cognates and similarities between the languages. When students learn why Spanish words look like English and vice versa, then this will help them to distinguish words, to guess their meaning and to learn easily, without any complications.

References

Baugh, A. C., Cable, T. (2002). A History of the English Language. 11 New Fetter Lane, London. Routledge.

Cela, C. (2012). La familia De Pascual Duarte. Chapter I.

Cervantes, M. (2004). Don Quijote De La Mancha. El Centro Virtual Cervantes. Chapter I.

Cobarrubias, R., Green, N., Quiroz, S. (1996). Science Cognates English-Spanish for K-6 Bilingual Teacher Candidates. The University of Texas at Brownsville and Texas Southmost College.

Elisson, R. (1952). Invisible Man. Chapter I. 457 Madison Avenue, New York, NY. Random.

Lujan-Garcia, C. (2016). The impact of English on Spanish daily life and so. Universidad de Las Palmas de Gran Canaria.

RBERN. (2015). False Cognates: English/ Spanish.

Penny, R. (2002). A History of the Spanish Language. Trumpington Street, Cambridge, United Kingdom. Cambridge University Press.

Pountain, J. C. (1999). *Spanish and English in the 21st Century*. Queen's College, Cambridge.

Región 13 Education Service Center. (2008). *English-Spanish Cognates.*

Rizzo, B., Villafane, S. (1975). Spanish Language Influences on Written English.

San Diego County Office of Education. (2008). English-Spanish Cognates. San Diego, California.

Nominative-Genitive Conversion in Japanese, Focus, and Improper Movement

Masao Ochi
Osaka University

1 Introduction

This paper investigates the Case alternation phenomenon in Japanese known as Nominative-Genitive Conversion (NGC). NGC is confined to the adnominal clause for the most part. Thus, the the Case alternation for the embedded subject is possible in (1b) but not in (1a).

(1) a. Boku-wa Taro-{ga/*no} warat-ta to omotta.
I-TOP Taro-NOM/GEN laugh-PAST that thought
'I thought that Taro laughed.'

b. Taro-{ga/no} warat-ta riyuu
Taro-NOM/GEN laugh-PAST reason
'the reason that Taro laughed'

NGC has many intriguing properties (see Ochi (2017) for an overview). This paper will examine the interaction of nominative/genitive Case assignment and focus. Building on the insights in the past literature, especially those in Miyagawa (2013, 2017), I will argue that a variant of the traditional ban against improper movement will help us explain why the genitive subject cannot be focused.

This paper is organized as follows. Section 2 discusses core data illustrating the interaction between GNC and focus, proposing that the apparent incompatibility between the genitive subject and focus can be explained in terms of a constraint that echoes the traditional ban on improper movement. In order to further substantiate this proposal, I will examine the distribution of nominative/genitive arguments as negative polarity items in section 3. Section 4 provides concluding remarks.

2 Genitive subject/object and focus

As discussed by Akaso and Haraguchi (2013) and Miyagawa (2013) among others, the genitive subject (as opposed to the nominative subject) cannot be focused, as shown in (2).

(2) Taro-dake-ga/*no yonda ronbun
Taro-only-NOM/GEN read article
'the article that only Taro read'

Miyagawa (2013, 2017), which is an extension of his earlier analyses (Miyagawa 1993, 2011, 2012), argues that this observation follows from his (2010) theory. Adopting Chomsky's (2008) feature inheritance mechanism and allowing some degrees of variation in the features to be transferred across languages, he argues that discourse configurational languages such as Japanese select topic/focus features as the target of the feature transfer operation. To be specific, those formal features originate on the C head (a phase head)

*I would like to thank the audience of WECOL2019 for stimulating discussions. This paper is also an elaboration of sections 4 and 5 of Ochi (in press). This research is financially supported by the Grant-in-Aid for Scientific Research (C) (No. 17K02809), the Ministry of Education, Culture, Sports, Science, and Technology of Japan, awarded to the author.

Proceedings of *WECOL 2019*

and get inherited by T. As a result, focus feature checking requires a CP layer. But the D head cannot probe into CP due to the Phase Impenetrability Condition: CP is a phase and hence the elements inside it cannot be accessed from outside (except for the phase head and its edge). Probing of the subject by D becomes possible when the adnominal clause is TP. Hence the D-licensed genitive and focus are mutually exclusive. On the other hand, the nominative Case value originates on the C head. Naturally, focus and nominative are not mutually exclusive.

Note in passing that I take it that focus under discussion is identificational focus (as opposed to information focus) in the sense of É. Kiss (2002). According to É. Kiss, identificational focus is syntactically realized in the form of movement: for example, an argument modified by *czak* 'only' must undergo focus movement in Hungarian.

(3) a. *János be mutatott csak Pétert Marinak
John VM introduced only Peter Mary-to
'John introduced only Peter to Mary.'
b. János CSAK PÉTERT mutatott be Marinak
John only Peter introduced VM Mary-to
'John introduced only Peter to Mary.' (É. Kiss 2002: 95)

Miyagawa's analysis is consistent with the well-known fact that the *wh*-subject can be genitive (although Miyagawa does not address this point).

(4) Kimi-wa [dare-ga/no kaita] hon-o yonda-no?
You-TOP who-NOM/GEN wrote book-ACC read-Q
'Who is the person x such that you read the book that x wrote'?

Wh-elements are standardly taken to be focus-related, licensed via some focus-related head in the periphery of a clause (see Rizzi 2001). Indeed, *wh*-phrases in Hungarian must undergo focus movement.

(5) a. * János be mutatott kit Marinak?
John VM introduced who Mary-to
'Whom did John introduce to Mary?'
b. János KIT mutatott be Marinak?
John who introduced VM Mary-to (É. Kiss 2002: 90)

In (4), the focus head that licenses the *wh*-subject (i.e., the interrogative C head) is located in the matrix clause. Thus, under Miyagawa's analysis, the adnominal clause can be a bare TP and the *wh*-subject can be D-licensed.

However, one potential problem with Miyagawa's analysis is that the genitive subject and a focus particle are not always mutually exclusive, as the genitive subject construction may have a focus particle on other elements, such as an adverb (6) and a PP argument (7). Further, Akaso and Haraguchi (2013) and Miyagawa (2013) point out that the genitive object can be focused: see (8).[1]

(6) kinoo/sukosi-dake Taro-ga/no nonda kusuri
yesterday/little-only Taro-NOM/GEN took medicine
'the medicine that Taro took only yesterday/only a little'

[1] According to Miyagawa, (8b) is fine because the genitive in this case is licensed via genitive of dependent tense (GDT: see the main text for discussion), a different type of genitive. Thus, nothing prevents the example from projecting the CP zone (including Focus Phrase). However, his analysis would still need to explain why the subject can be genitive when the genitive object is focused, since the genitive Case on the subject must come from D in such a configuration.

(7) Taro-dake-ni Hanako-ga/no okutta ronbun[2]
Taro-only-to Hanako-NOM/GEN sent article
'the article that Hanako sent only to Taro'

(8) a. Taro-dake-ga/*no nihongo-ga wakaru koto
Taro-only-NOM/GEN Japanese-NOM understand fact
'the fact that only Taro understands Japanese'

b. Taro-ga/no nihongo-dake-ga/no wakaru koto
Taro-NOM/GEN Japanese-only-NOM/GEN understand fact
'the fact that Taro understands only Japanese'

In the remainder of this section, I would like to offer a modification of Miyagawa's analysis that maintains the empirical coverage of his analysis and accommodates data such as (6) to (8).

Here are some crucial ingredients of the proposal. First, I assume along with Rizzi (2001) (see also Akaso and Haraguchi (2013) and Miyagawa (2013)) that an informationally focused element is syntactically licensed at the clausal periphery. Second, I assume that adnominal clauses in Japanese are uniformly TPs (Murasugi 1991).[3] Thus, focus is licensed in the C-region when the clause is CP, and in the T-region when the clause is TP.[4] Third, when a nominal argument is focused, it needs to undergo both Case feature checking and focus feature checking. Let me now spell out my proposal:

(9) For any argument α, checking of an 'A-bar type' feature (such as focus and Q) of α cannot precede checking of an 'A-type' feature (such as Case/φ-feature) of α.

Here the phrase "A-type feature" refers to a feature that is checked by a syntactic head whose specifier position is an A-position in the traditional sense. Similarly, the phrase "A-bar type feature" refers to a feature that is checked by a head whose specifier is an A-bar position. The proposal allows derivational scenarios such as those shown in (10a), in which the head Y probes X for φ/Case before another head, Z, probes X for focus, and (10b), where a single probe Agrees with X for both φ/Case and focus. Crucially, it disallows the configuration shown in (10c), where X is probed for focus before it is probed for φ/Case.

(10) a. [[[… X ….] Y] .... Z]
[φ/Case] [Foc]

b. [[[…. X ….] Y] …. Z]
[φ/Case]
[Foc]

c. * [[[…. X ….] Y] …. Z]
[Foc] [φ/Case]

This proposal resurrects an old idea about improper movement: a phrase cannot participate in Case checking after it has undergone A'-movement. An example like the following is barred as it involves movement of *who* into the spec of the embedded CP, an A-bar position, before its Case is checked.

[2] Note that the genitive phrase in this example is not a possessor base-generated in the spec of DP (i.e., outside the relative clause) as it linearly follows a PP argument, which clearly belongs to the relative clause.

[3] Unless adnominal clauses contain an overt complementizer such as *toiuu* (see Nakau 1973), in which case they are CPs.

[4] See Miyagawa (2010) for the proposal that discourse properties such as focus and topic are licensed at the T-region in Japanese. Miyagawa argues that such features originate on C and get inherited by T. In this respect, our suggestion in the main text is different from Miyagawa's view. See also Ochi (2020) for the suggestion that focus may originate on D and get passed down to T.

(11) *I know who$_i$ it was told t_i that Mary would be coming.
cf. I know who$_i$ was told t_i that Mary would be coming.

Note that the English verb *know* may act as an ECM verb as shown in (12a), which means that it can probe a DP located in the edge of the complement clause, as shown in (12b). In the above example, therefore, *know* should be able to probe *who* sitting in the embedded spec of CP, as illustrated in (13), and we would need a way to block such a derivation. Now, if something to the effect of (9) is part of the grammar, it would exclude such a derivation.

(12) a. We know him to be smart.
b. We [$_{vP}$ *v*+know [$_{VP}$ <know> [$_{TP}$ him to be smart]]]

(13) I [$_{vP}$ *v*+know [$_{VP}$ <know> [$_{CP}$ who$_i$ C [$_{TP}$ it was told t_i [that ....]]]
[φ] [~~Foc~~]
*(9) violated

Let us now reconsider (2) in light of (9). Since the adnominal clause is a TP by assumption (see above), it is the adnominal T that bears a focus feature in this case. When the nominative subject of an adnominal clause is focused, both Case and focus are licensed by the adnominal T, as shown in (14a). This derivation does not run afoul of (9): A single probe (T in this case) probes and agrees with the subject, valuing both features simultaneously. When the genitive subject is focused, however, we get a different picture. As shown in (14b), Case checking cannot take place prior to (or in concomitant with) focus feature checking. This would account for the contrast seen in (2).

(14) a. Focused nominative subject
[$_{DP}$ [$_{NP}$ [$_{TP}$ **Taro-only** laugh T] reason] D]
[~~Foc~~]
[~~Case~~] **✓(9) satisfied**

b. Focused genitive subject
*[$_{DP}$ [$_{NP}$ [$_{TP}$ **Taro-only** laugh T] reason] D]
[~~Foc~~] [~~Case~~]
***(9) violated**

Now, the data shown in (6) and (7), which pose potential problems for Miyagawa (2013, 2017), are accommodated straightforwardly under the revised analysis explored here. In (6), for example, *kinoo* 'yesterday' has its focus feature checked by the adnominal T while the subject *Taro* may be assigned genitive by D. The condition in (9) is trivially satisfied.

(15) a. [$_{DP}$ [$_{NP}$ [$_{TP}$ yesterday-only Taro took T] medicine] D]
[+Foc]

b. [$_{DP}$ [$_{NP}$ [$_{TP}$ yesterday-only Taro took T] medicine] D]
[Case]

And the same logic applies to (7). Here, too, focus in on an element other than the genitive DP. Thus, (9) is trivially satisfied.

Now let us return to the contrast in (8). Why is the genitive object compatible with focus? I simply follow

Miyagawa (2012, 2013) and assume that the genitive on the object may be assigned by the weak *v* head in conjunction with what Miyagawa calls dependent Tense. Miyagawa refers to this type of genitive Case as *genitive of dependent tense (GDT)*. Crucially, the GDT-genitive is available for an internal argument but not for an external argument. This is because the weak *v* head is incompatible with the presence of the external argument. See Miyagawa (2012) for more detail. Now once we assume the GDT analysis of Miyagawa (2012), (8) can be explained in a simple manner. Unlike in (2), genitive can be assigned prior to the assignment of the focus feature.

(16) a. [$_{DP}$ [$_{NP}$ [$_{TP}$ [$_{vP}$ [$_{VP}$ Taro French-only understand] *v*] T] fact] D]
[Foc]
[Case]

b. [$_{DP}$ [$_{NP}$ [$_{TP}$ [$_{vP}$ [$_{VP}$ Taro French-only understand] *v*] T] fact] D]
[Case] [Foc]

But is there any support for the supposition that some instances of genitive come from the weak *v* (in conjunction with dependent T) in adnominal clauses in general? The issue is important, since Miyagawa's (2012) GDT analysis was motivated by a restricted availability of genitive subjects in temporal adverbial clauses (see H. Takahashi 2010), but Miyagawa (2013), whose analysis is endorsed here, is an attempt to generalize the availability of the GDT-genitive to all sorts of adnominal clauses, which would mean that, in principle, the genitive Case on the object has two distinct sources, D and weak *v* (in conjunction with dependent T). In the next section, I turn to the distribution of a certain type of Negative Polarity Item (NPI) to provide potential support for Miyagawa's overall approach.

3 Nominative-Genitive Conversion and NPI Licensing

As discussed by Kishimoto (2001), an indeterminate pronoun in Japanese such as *dare* 'who,' *nani* 'what,' and *do-no* 'which-Gen' may function as a negative polarity item (NPI) when it is associated with the Q-particle *-mo* in a negative clause. I will refer to this type of NPI as a *wh*-NPI. As we can see in (17b) and (17c) below, a *wh*-NPI and the particle *-mo* need not be adjacent: *-mo* can 'license' a *wh*-NPI at a distance, as long as the former c-commands the latter. In (17b), for instance, *-mo* is attached to the postposition *to* 'with' and licenses *dare* 'who' as the former c-commands the latter.

(17) a. Taro-wa **nani-mo** kaw-anakat-ta.
Taro-TOP **what-Q** buy-NEG-PAST
'Taro didn't buy anything.'

b. Taro-wa **dare**-to-**mo** aw-anakat-ta.
Taro-TOP **who**-with-**Q** meet-NEG-PAST
'Taro didn't meet with anyone.'

c. Taro-wa **do**-no hito kara-**mo** tegami-o moraw-anakat-ta.
Taro-TOP **which**-GEN person from-**Q** letter-ACC receive-NEG-PAST
'Taro didn't receive a letter from anyone.'

As noted by Kishimoto, an asymmetry arises between the subject and the object when the Q-particle *-mo* appears attached to a verb. While the indeterminate object can function as a *wh*-NPI, the indeterminate subject cannot.

(18) a. ***dare**-ga hon-o kai-**mo** si-nakat-ta koto
who-NOM book-ACC buy-**Q** do-NEG-PAST fact
'the fact that no one bought a book.'

b. Sono hito-ga **nani**-o kai-**mo** si-nakat-ta koto
that person-NOM **what**-ACC buy-**Q** do-NEG-PAST fact
'the fact that that person didn't buy anything'

Essentially following Kishimoto, let us assume that *-mo* takes *v*P as its c-command domain in this type of configuration. (18b) is fine because the *wh*-NPI object, being located inside *v*P, is c-commanded by *-mo* . On the other hand, (18a) is ruled out because the nominative subject is assumed to move to the spec of TP (in overt syntax), which is outside the c-command domain of *-mo*.

Kishimoto further argues that this type of NPI is licensed at LF, not in overt syntax. This part of Kishimoto's proposal comes from the fact that the *wh*-NPI object is not licensed in this type of construction when the object is nominative. To see his point, let us first look at the following pair of examples.

(19) a. Hanako-ga migime-dake-o tumur-e-ru koto
Hanako-NOM right.eye-only-ACC close-can-PRES fact
'the fact that Hanako can close only her right eye'
(can > only; ??only > can)

b. Hanako-ga migime-dake-ga tumur-e-ru koto
Hanako-NOM right.eye-only-NOM close-can-PRES fact
'the fact that Hanako can close only her right eye'
(?can > only; only > can)

Japanese allows nominative marking or accusative marking on the object when the predicate is stative. As noted by many researchers (see Sano (1985) and Tada (1991) among others), the choice between the two case values has an effect on the scope of the object. (19a) has an accusative object, which takes narrow scope with respect to negation: This sentence has the reading in which Hanako can close her right eye with her left eye open (Hanako can wink with her right eye), but it does not have the reading according to which it is only her right eye that Hanako can close (Hanako can close her right eye but she cannot close her left eye). As for (19b), where the object is nominative, the situation is reversed and the wide scope reading of the object is in fact much more salient. A number of researchers, including Koizumi (1998) and M. Takahashi (2010), attribute this scope difference to the different syntactic positions occupied by the nominative object and the accusative object. Simply put, the accusative object remains within *v*P throughout the derivation, with its Case checked by *v*. The nominative object may move to the domain of T that licenses nominative Case. As a result, the accusative object cannot take scope over *(r)are* 'can' whereas the nominative subject can.

(20) a. [$_{TP}$ Hanako [$_{CANP}$ [$_{vP}$ right eye-only-ACC close] can] T]
b. [$_{TP}$ Hanako [$_{TP}$ right eye-only-NOM$_i$ [$_{CANP}$ [$_{vP}$ t_i close] can] T]]

Next, authors such as Yatsushiro (1999), Kishimoto (2001), and Saito (2009) observe that the nominative object may be contained within the fronted *v*P (see (21b) below), which led these authors to conclude that the nominative object (as well as the accusative object) stays inside the *v*P in overt syntax.

(21) a. Hanako-ga migime-ga/-o tumur-e-sae su-ru koto
Hanako-NOM right.eye-NOM/-ACC close-can-even do-PRES fact
'the fact that Hanako can even close her right eye'

b. Migime-ga/-o tumur-e-sae Hanako-ga su-ru koto
right.eye-NOM/-ACC close-can-even Hanako-NOM do-PRES fact
'the fact that even close her right eye, Hanako can'

Combining these two observations, one would be naturally led to the conclusion that the wide scope reading of the nominative object in (19b) is due to the movement of the object in covert syntax (but see below).

Now we are ready to examine the distribution of *wh*-NPI objects. As Kishimoto (2001) notes, the nominative *wh*-NPI object sounds fairly degraded.

(22) a. Sono hito-ga **nani**-o ka-e-**mo** si-nakat-ta koto
that person-NOM **what**-ACC buy-can-**Q** do-NEG-PAST fact
'the fact that that person was not able to buy anything'

b. ??Sono hito-ga **nani**-ga ka-e-**mo** si-nakat-ta koto
that person-NOM **what**-NOM buy-can-**Q** do-NEG-PAST fact
'the fact that that person was not able to buy anything'

On the basis of this observation, Kishimoto concludes that the *wh*-NPI is licensed at LF. Since the nominative object moves out of the *v*P domain in covert syntax (as we saw above), the nominative *wh*-NPI object is not c-commanded by the Q-particle *-mo* at LF.

However, Ochi and Saruwatari (in press) make an observation that raises an interesting problem for Kishimoto's otherwise solid analysis. Ochi and Saruwatari note that when the object remains inside *v*P in overt syntax, including the *v*P-fronting situation that we saw in (21), it does not take wide scope whether it is nominative or accusative.[5]

(23) Migime-dake-ga/-o tumur-e-sae Hanako-ga su-ru koto
right.eye-only-NOM/-ACC close-can-even Hanako-NOM do-PRES fact
'the fact that even close only her right eye, Hanako can' (can > only; ??only > can)

On the basis of an observation like this, Ochi and Saruwatari propose that two derivational paths are available for the nominative object in Japanese. According to their proposal, the nominative object may move to the domain of T in overt syntax, or it may stay in *v*P throughout the derivation. Crucially, they argue that no covert movement is available in Japanese. They also argue that a derivation that involves the movement of the nominative object gives rise to the wide scope reading of the object, and the one in which the nominative object stays inside the *v*P region yields only the narrow scope reading of the object. This line of analysis, if correct, raises an issue for (22b): Why can't the nominative *wh*-NPI object be licensed at LF if the nominative object has the option of remaining inside the *v*P throughout the derivation? In what follows, I will argue that implementing the essence of Kishimoto's analysis in a slightly different manner will resolve this issue.

My proposal (to be revised later) is summed up as follows.

(24) a. Binding of a *wh*-NPI by the Q-particle *-mo* involves feature checking of a formal feature, say, the NPI-feature.
b. For an argument NPI α, checking of the NPI feature of α cannot "precede" Case/φ checking of α.

(24b) is another instantiation of (9): instead of the focus feature checked at the TP level, this time we have the NPI-feature checked at the *v*P level as an A-bar type feature. I assume that the Q-particle *-mo* as a probe participates in A-bar syntax, not A-syntax, given its well-known quantificational and focus-oriented nature. As will be demonstrated in the next paragraph, this line of analysis preserves the essence of Kishimoto's (2001) analysis but implements it in a way that does not resort to covert movement.

Let us now see how this proposal accommodates (18) and, crucially, (22). Since *-mo* is attached to *v*P in both sets of examples, licensing of a *wh*-NPI uniformly occurs at the level of *v*P.[6] Now (18b) is fine because the object is case-licensed at the level of *v*P. On the other hand, (18a) is degraded, not because it moves out of *v*P but because nominative is assigned by T. Given familiar cyclicity considerations, binding of the subject by *-mo*, which is an A-bar head by assumption, takes place before T is introduced into the structure. Consequently, (24b) is violated. The same point holds of (22). When the object is nominative, its Case cannot be checked until T is introduced, which is later in the derivation than the checking of the NPI-feature of the nominative object, which occurs at the level of *v*P. Thus, the contrast in (22) is successfully accommodated.

[5] It is important not to have a pause between the object and the verbal complex in (23b). If there is a pause, the wide scope reading of the nominative object becomes quite salient. I take it that the object and the verbal complex do not form a constituent in such a case, with the object being located outside the fronted *v*P. Interestingly, the accusative object does not seem to yield the wide scope reading even with a pause.

[6] Since the Q-particle *-mo* is c-commanded by the negative head, and since we are dealing here with an NPI, we might instead say that α is licensed as an NPI at the level of NegP, located above *v*P and below TP. Nothing in the discussion hinges on the choice between the two, as far as I can see.

We now turn to adnominal clauses and consider how the nominative and genitive subject behaves in this area. In (25a) *-mo* is directly merged with the subject *wh*-NPI and the example is perfectly grammatical. Keeping this example as the baseline data, let us examine the nominative *wh*-NPI subject (25b) and the genitive *wh*-NPI subject (25c).

(25) a. **dare-mo** hik-anai kyoku
who-Q play-NEG tune
'a tune that nobody plays'
b. ***dare-ga** hiki-**mo** si-nai kyoku
who-NOM play-Q do-NEG tune
'a tune that no child plays'
c. ??**dare-no** hiki-**mo** si-nai kyoku
who-GEN play-Q do-NEG tune
'a tune that no child plays'

(25b) and (25c) are degraded, as expected. Both T and D are located higher than *v*P. (25c) sounds better than (25b), presumably because the genitive subject has the option of staying in a position lower (internal to *v*P) than the nominative subject in overt syntax (see Miyagawa 2011, Ochi 2017), although (24b) is still violated even if the genitive *wh*-NPI subject remains inside the *v*P.

Admittedly, the proposal as stated in (24b) faces a challenge when we examine data in which a *wh*-NPI is contained inside the subject phrase. In the following data, we have the wh-NPI *do* 'which,' which is contained inside the subject DP. The acceptability pattern remains the same as in (25): both the nominative subject and the genitive subject sound degraded.

(26) a. **do**-no ko-**mo** hik-anai kyoku
which-GEN child-Q play-NEG tune
'a tune that no child plays'
b. ***do**-no ko-**ga** hiki-**mo** si-nai kyoku
which-GEN child-NOM play-Q do-NEG tune
'a tune that no child plays'
c. ??**do**-no ko-**no** hiki-**mo** si-nai kyoku
which-GEN child-GEN play-Q do-NEG tune
'a tune that no child plays'

But notice that the Case property of the *wh*-NPI *do* 'which' is satisfied internal to the subject domain if, as seems plausible, this genitive comes from the D head of the subject DP. When the subject DP is merged into the spec of *v*P, to which *-mo* is attached, binding of *do* 'which' by *-mo* should be possible without violating (24b). We may therefore have to modify (24b) in the following way.

(27) For an argument NPI α, checking of the NPI feature of α cannot "precede" Case/φ checking of α *and Case/φ checking of β that (immediately) contains α.*

This modification is of course nothing but a restatement of the problem to be solved. Note that Kishimoto's (2001) original analysis fares better in this respect, for if the subject DP undergoes phrasal movement at LF and if the NPI licensing takes place at that level, (26b) and (26c) would be correctly ruled out: at LF, *do* 'which' would be outside the c-command domain of the Q-particle. Despite this drawback, I will continue to assume (24b), or its variant in (26), since, as discussed earlier, once we assume the availability of covert movement in Japanese, the lack of the wide scope reading of the object in (23) demands an explanation.

Now let us turn to nominative/genitive objects that occur in adnominal clauses containing a stative predicate to which the Q-particle *-mo* is attached. First of all, the following data allow us to confirm that accusative, nominative, and genitive objects are all possible in this type of construction.

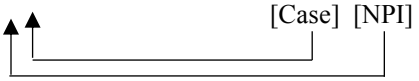
(28) a. Hanako-ga gakufu-**o** yom-e-mo si-nai koto
Hanako-NOM musical.score-ACC read-can-Q do-NEG fact
'the fact that Hanako cannot even read a musical score'
b. Hanako-ga gakufu-**ga** yom-e-mo si-nai koto
Hanako-NOM musical.score-NOM read-can-Q do-NEG fact
'the fact that Hanako cannot even read a musical score'
c. Hanako-ga gakufu-**no** yom-e-mo si-nai koto
Hanako-NOM musical.score-GEN read-can-Q do-NEG fact
'the fact that Hanako cannot even read a musical score'

Now let us see how *wh*-NPI objects behave in the same context. As for the accusative *wh*-NPI object (29a) and the nominative *wh*-NPI object (29b), the results are as expected: the former is good and the latter is degraded, a pattern that we also witnessed in (22). The crucial test case is the genitive *wh*-NPI object in (29c). Unlike the nominative *wh*-NPI subject, the genitive *wh*-NPI object is fine in the V+*mo* clause.[7]

(29) a. Hanako-ga **do**-no gakufu-**o** yom-e-mo si-nai koto
Hanako-NOM **which**-GEN musical.score-ACC read-can-Q do-NEG fact
'the fact that Hanako cannot even read any musical score'
b. ??Hanako-ga **do**-no gakufu-**ga** yom-e-mo si-nai koto
Hanako-NOM **which**-GEN musical.score-NOM read-can-Q do-NEG fact
'the fact that Hanako cannot even read any musical score'
c. Hanako-ga **do**-no gakufu-**no** yom-e-mo si-nai koto
Hanako-NOM **which**-GEN musical.score-GEN read-can-Q do-NEG fact
'the fact that Hanako cannot even read any musical score'

Given (24) (or (27)), the well-formedness of (29c) would be unexpected if D was the only licensor of genitive: D occurs higher than *v* to which *-mo* is attached. But the grammatical status of this example would be accounted for if the weak *v* head (in conjunction with dependent T) could also license genitive in the adnominal clause, as argued by Miyagawa (2012, 2013). As the following schematic representation shows, (27) is observed because both genitive Case and the NPI-feature are checked at the level of *v*P.

(30) [$_{TP}$ [$_{vP}$ ..... **which** musical note-GEN ...] *v*] **-mo** ....
[Case] [NPI]



4 Conclusion

I hope to have demonstrated in this paper that a constraint like (9), which reformulates the traditional ban on improper movement under the Agree-based system and without making reference to movement, enables us to explain why the genitive subject cannot be focused under many circumstances. Let me end this paper by drawing the reader's attention to the ECM construction in English in the context of (9).[8] The

[7] If we have a simplex form of the genitive *wh*-NPI object, the sentence is degraded. But this seems to be for an independent reason. Even when the simplex genitive indeterminate *nan(i)* 'what' appears as a *wh*-interrogative, and not as a *wh*-NPI, as in (ii), the genitive version is still degraded: see (ii).

(i) ??Hanako-ga **nan(i)**-no yom-e-mo si-nai koto
Hanako-NOM **which**-GEN read-can-Q do-NEG fact
'the fact that Hanako cannot even read any musical score'

(ii) Hanako-ga **nan(i)**-{o/ga/??no} yom-e-mo si-nai koto-ga mondai-na no?
Hanako-NOM **which**-ACC/NOM/GEN read-can-Q do-NEG fact-NOM problem-COP Q
'What is [the fact that Hanako cannot even read t] a problem?'

[8] I thank Masako Maeda (p.c.) for raising this issue.

standard view in the literature is that the ECM subject has its Case licensed by the *v* head of the immediately higher clause. In the following *make out* construction (see Lasnik (2001)), we can confirm on the basis of the word order that the focused ECM subject *only Mary* is in the embedded clause (I thank Brian Agbayani (p.c.) for confirming the grammaticality of an example like this).

(31) John made out only Mary to be a liar.

Now if identificational focus is syntactically licensed at the periphery of a clause, as assumed throughout this paper, this focus feature must be licensed by the embedded T. But then we confront a situation that runs counter to (9): checking of the focus feature of *Mary* occurs at the embedded TP level, before *Mary* has its Case checked by the matrix *v* head. Interestingly, by updating the analysis of Davis (1984), Lasnik (2018) proposes that the ECM subject may be Case-licensed by the infinitival T that inherits the Case-assigning property from the matrix *v* head via the operation called Feature Inheritance (see Chomsky (2008)). Lasnik's proposal would allow us to maintain (9).

References

Akaso, Naoyuki and Tomoko Haraguchi. 2013. On the agent/theme asymmetry in Japanese Nominative/Genitive Conversion. In *Proceeding of the 8th workshop on Altaic formal linguistics*: 1-6. MIT Working Papers in Linguistics.

Chomsky, Noam. 2008. On phases. In: Robert Freidin, Carlos P. Otero, and Maria-Luisa Zubizaretta (eds.), *Foundational Issues in Linguistic Theory*. Cambridge, MA: MIT Press.

Davis, Lori. 1984. *Arguments and expletives*. Doctoral dissertation, University of Connecticut, Storrs.

É. Kiss, Katalin. 2002. *The Syntax of Hungarian*. Cambridge: Cambridge University Press.

Kishimoto, Hideki. 2001. Binding of indeterminate pronouns and clause structure in Japanese. *Linguistic Inquiry* 32: 597-633.

Koizumi, Masatoshi. 1998. Remarks on nominative objects. *Journal of Japanese Linguistics*, 16, 39–66.

Lasnik, Howard. 2001. Subjects, objects, and the EPP. In: William D. Davies and Stanley Dubinsky (eds.), *Objects and other subjects: Grammatical functions, functional categories, and configurationality*, 103-121. Dordrecht: Kluwer Academic Publishers.

Lasnik, Howard. 2018. UConn Syntax: Some lasting insights and their interactions UConn 1972-2002, UMD 2002-. A talked given at 50 Years of Linguistics at UCONN, September 7, 2018, University of Connecticut.

Miyagawa, Shigeru. 1993. LF Case-checking and minimal link condition. In *MIT Working Papers in Linguistics* 19, 213-254. MITWPL.

Miyagawa, Shigeru. 2011. Genitive subjects in Altaic and specification of phase. *Lingua* 121: 1265-1282.

Miyagawa, Shigeru. 2012. *Case, argument structure, and word order*. New York: Routledge.

Miyagawa, Shigeru. 2013. Strong Uniformity and Ga/No Conversion. *English Linguistics* 30: 1-24.

Miyagawa, Shigeru. 2017. *Agreement Beyond Phi*, Cambridge, MA: MIT Press.

Murasugi, Keiko. 1991. *Noun Phrases in Japanese and English: A Study in Syntax, Learnability, and Acquisition*, Doctoral dissertation, University of Connecticut, Storrs.

Nakau, Minoru. 1973. *Sentential complementation in Japanese*. Tokyo: Kaitakusha.

Ochi, Masao. 2017. Ga/No conversion. In: Masayoshi Shibatani, Shigeru Miyagawa, and Hisashi Noda (eds.) *Handbook of Japanese Syntax*, 663-700, Mouton de Gruyter, Boston.

Ochi, Masao. 2020. Feature Inheritance and Nominative-Genitive Conversion. Paper presented at WCCFL38, University of British Columbia, March 6, 2020.

Ochi, Masao. In press. Feature Transfer, Left Periphery, and Case Conversion (Review Article of Miyagawa's (2017) monograph, MIT Press), *English Linguistics* 36:2.

Ochi, Masao and Asuka Saruwatari. In press. Shukaku mokutekigo no idoo to sayooiki [On the movement of the nominative object and scope]. In: Keiko Murasugi, Daiko Takahashi, Kensuke Takita, and Masahiko Takahashi (eds.), *Nihongo kenkyuu kara seiseibunpoo riron e [From the study of Japanese language to the theory of generative grammar]*. Tokyo: Kaitakusha.

Rizzi, Luigi. 2001. On the position "int(errogative)" in the left periphery of the clause. In: Guglielmo Cinque and Giampaolo Salvi (eds.), *Current studies in Italian syntax: Essays offered to Lorenzo Renzi*, 267–296. Amsterdam: Elsevier.

Saito, Mamoru. 2009. On the scope properties of nominative phrases in Japanese. Paper presented at the 7th GLOW in Asia Conference.

Sano, Masaki.1985. LF Movement in Japanese. *Descriptive and Applied Linguistics* XVIII, 245–259.
Tada, Hiroaki. 1992. Nominative objects in Japanese. *Journal of Japanese Linguistics* 14: 91-108.
Takahashi, Hisako. 2010. Adverbial clauses and Nominative/Genitive Conversion in Japanese. *Proceedings of the Sixth Workshop on Altaic Formal Linguistics* (WAFL 6), 357-371. MITWPL.
Takahashi, Masahiko. 2010. Case, phases, and nominative/accusative conversion in Japanese. *Journal of East Asian Linguistics* 19: 319-355.
Yatsushiro, Kazuko. 1999. *Case licensing and VP structure*. Doctoral dissertation, University of Connecticut, Storrs.

Another Kind of Negative Concord Items in Japanese

Takeshi Oguro
Chiba University of Commerce

1 Introduction

This paper deals with *mono ka* rhetorical questions in Japanese (*mono ka* RQs for short), which are always interpreted as negative assertions, examined originally by McGloin (1976) and in detail by Oguro (2018) and accounts for the behavior of two types of negative concord items (NCIs) found in *mono ka* RQs in terms of focus.

Section 2 displays some relevant properties of *mono ka* RQs, which are interpreted as negative assertions and license NCIs such as *daremo* 'anyone/no one'.

Section 3 shows that WH-phrases in *mono ka* RQs function as NCIs as well, assuming the work of Miyagawa, Nishioka, and Zeijlstra (2016) on negative sensitive items.

Section 4 shows that WH-phrases employed in *mono ka* RQs can also function as NCIs.

Section 5 observes that the types of NCIs are incompatible with each other and attempts to offer an analysis.

Section 6 is the conclusion.

2 Relevant properties of *mono ka* rhetorical questions

In this section, I show some relevant properties of *mono ka* RQs.

2.1 *Intonation* Whereas ordinary questions (OQs) can end either with rising intonation or falling intonation, *mono ka* RQs only end with falling intonation. This is illustrated in (1).

(1) a. John-ga kuru no ka? ↑ / ↓
John-NOM come C Q
'Will John come?'

b. John-ga kuru mono ka! * ↑ / ↓
John-TOP come C Q
'John will not come!'

Since questions are typically uttered with rising intonation, the effect in (1b) strongly suggests the non-interrogative status of *mono ka* RQs.

2.2 *Answerability* Another aspect that motivates distinguishing *mono ka* RQs from ordinary questions concerns the fact that they fail to be answered. Consider (2) and (3).

(2) a. Dareka kuru no ka? b. Iya, daremo ko-nai yo
someone come C Q no anyone come-NEG SFP
'Will someone come?' 'No, no one will come.'

(3) a. Dare-ga kuru mono ka! b # Iya, daremo ko-nai yo
who-NOM come C Q no anyone come-NEG SFP
'No one will come!' 'No, no one will come.'

(2a), which is an ordinary question, allows (2b) as an answer. On the other hand, (3a), which is a *mono ka*

Proceedings of *WECOL 2019*

RQ, does not invite an answer. Caponigro and Sprouse (2007) claim that rhetorical questions can be answered because they are questions, but their view does not apply to *mono ka* RQs.

2.3 *The addition of* ittai *'the Hell'* Another property of *mono ka* RQs that is not found in ordinary questions is concerned with *ittai*, which is an element for emphasizing the WH-phrase in a question (Pesetsky 1987), as exemplified in (4).

(4) (Ittai) Dare-ga kur-u no?
the.hell who-NOM come-PRES C
'Who (the hell) will come?'

It is important to notice that *ittai* is possible only in WH-questions. It is disallowed in other WH-sentences including concessive clauses like (5) and exclamative clauses like (6).

(5) (*Ittai) Dare-ga kite-mo, boku-wa ureshii.
the.hell who-NOM come-ever I-TOP happy
'Whoever comes, I will be happy.'

(6) Kare-wa (*ittai) nanto kooun na no daroo ka!
he-TOP the.hell how lucky COP C MOD Q
'How lucky he is!'

With this in mind, let us observe the following example.

(7) (*Ittai) Dare-ga kur-u mono ka!
the.hell who-NOM come-PRES C Q
'No one will come!'

This example indicates that *ittai* is incompatible with *mono ka* RQs, which patterns with (5) and (6), suggesting that *mono ka* RQs should not be regarded as ordinary questions.

2.4 *WH*-mo *expressions* One key point which suggests that *mono ka* RQs should be treated differently from ordinary questions has to do with the licensing of WH-*mo* expressions. These expressions are so termed because they are composed of the WH-part such as *dare* 'who' or *nani* 'what' accompanied by the particle *mo* 'also' or 'even'. These expressions are special in that they always have to be in negative sentences, unlike the English item *any*. *Any* can occur in conditional clauses and interrogative clauses as well. Consider the following paradigm.

(8) a. I *(didn't) buy anything.
b. If Mary sees anyone, she will cry.
c. Did she read any book?

(9) a. Daremo konai/*kuru.
anyone come-not/come
'No one will come.'
b. * [daremo kure-ba] boku-wa komaru daroo
anyone come-if I-TOP trouble will
'If John/anyone ever/again comes, I will be in trouble.'
c. * Daremo kuru no ka?
anyone come C Q
'Will anyone come?'

(10) Daremo kuru mono ka!
anyone come C Q
'No one will come!'

Unlike *any*, whose behavior is illustrated in (8), WH-*mo* expressions such as *daremo* 'anyone' are allowed only in negative sentences, as shown in (9). What is important here is that they are not allowed in affirmative questions, as the degraded status of (9c) indicates. As the fine status of (10) clearly shows, however, they are allowed in *mono ka* RQs. Given this distributional property of WH-*mo* expressions, it should be concluded that *mono ka* RQs are negative assertions rather than questions.

2.5 *Compatibility with past tense* An additional point where *mono ka* RQs differ from ordinary questions is that they are subject to a kind of tense restriction. As shown in (11), ordinary questions are compatible with past tense, while *mono ka* RQs are not (Adachi 2005).

(11) a. Dare-ga kinoo ki-masi-ta ka?
who-NOM yesterday come-POLITE-PAST Q
'Who came yesterday?'

b. * Dare-ga kinoo ki-ta mono ka!
who-NOM yesterday come-PAST C Q
'No one came yesterday!'

2.6 *The structure of* mono ka *RQs* I suggest that ordinary questions and *mono ka* RQs have the following structures, based on Rizzi's (1997) view on clause structure.

(12) a. $[_{ForceP}\ [_{FinP}\ TP_{[tensed]}\ [_{Fin}{}^{0}\ no]]\ [_{Force}{}^{0}\ ka_{[+Q]}]]$
b. $[_{ForceP}\ Op_{[NEG]}\ [_{FinP}\ TP_{[subjunctive]}\ [_{Fin}{}^{0}\ mono]]\ [_{Force0}\ ka_{[-Q]}]]$

(12a) depicts the structure of ordinary questions, where *no* heads FinP and selects a tensed clause, and *ka*, with a [+Q] feature, heads ForceP, which is responsible for the interrogative status of the clause, following Hiraiawa and Ishihara (2012), Saito and Haraguchi (2012), and Kuwabara (2013), among others. In (12b), which is the structure for *mono ka* RQs, the head of FinP is *mono*, which is assumed to select a subjunctive clause, based on the observation that *mono ka* RQs are not compatible with episodic tense. The Force head here is *ka* with a [-Q] feature, which makes the clause non-interrogative, and its Spec is occupied by a negative operator, which is responsible for negative interpretation and licenses WH-*mo* expressions.

3 WH-phrases as NCIs

In this section, I show that WH-phrases employed in *mono ka* RQs function as NCIs. Miyagawa et al. (2016) deal with negative sensitive items in Japanese and argue that those that can be in fragment answers are NCIs.

3.1 *WH-*mo *expression as NCIs* Miyagawa et al. examine three types of negative sensitive items: WH-*mo* expressions, XP-*sika*, and *rokuna* Ns. Among them, only WH-*mo* expressions are allowed to be in fragment answers (Nishioka 2000, Watanabe 2004). The other two types, which fail to be in fragment answers, are referred to as negative polarity items (NPIs). (Miyagawa et al. observe that adjunct XP-*sika* can be in fragment answers.)

(13) A: Dare-ga kuru no? B: Daremo/*John-sika/*Rokuna hito.
who-NOM come C anyone/John-SIKA/rokunaperson
'Who will come?' 'No one./No one but John./No decent person.'

Miyagawa et al. propose that fragment answers are a result of the movement of the fragment to the Spec of a functional projection above TP which they call FP, followed by the deletion of the lower projection, as illustrated in (14).

(14) $[_{FP}$ fragment $[_{TP} \ldots t \ldots]]$

They argue that NPIs are semantically dependent on the negative licensor, so they cannot be raised outside

the scope of the licensor. According to them, NCIs, on the other hand, have a syntactic relation with the licensor. That is to say, once the syntactic requirement is met, NCIs are allowed to appear in a position outside the scope of negation. They assume that the licensor is the negative operator in [Spec, NegP], which involves the interpretable negative feature, the [i-NEG] feature and that NCIs contain the uninterpretable negative feature, the [u-NEG] feature. The [u-NEG] features are assumed to be checked off under Upward Agree, which works in the direction that is the opposite of the standard Agree (Chomsky 2000; 2001), as in (15).

(15) $[_{TP} \ldots [_{NegP} \text{Op}_{[i\text{-NEG}]} [_{vP} \ldots \text{NCI}_{[u\text{-NEG}]} \ldots v] \text{na}_{[u\text{-NEG}]}]]$
Upward Agree

That is to say, they are checked off by being c-commanded by the licensor, which is the negative operator. In (15), there are two elements with [u-NEG] features. Both of them have their features checked by being c-commanded by the negative operator.

Let us now see the process of an NCI functioning as a fragment answer.

(16) a. $[_{FP} [_{TP} [_{NegP} \text{Op}_{[i\text{-NEG}]} [_{vP} \ldots \text{NCI}_{\sout{[u\text{-NEG}]}} \ldots]]]]$ checking of the [u-NEG] feature
b. $[_{FP} \text{NCI} [_{TP} [_{NegP} \text{Op}_{[i\text{-NEG}]} [_{vP} \ldots \text{NCI}_{\sout{[u\text{-NEG}]}} \ldots]]]]$ raising of the NCI
c. $[_{FP} \text{NCI} [_{TP} [_{NegP} \text{Op}_{[i\text{-NEG}]} [_{vP} \ldots \text{NCI}_{\sout{[u\text{-NEG}]}} \ldots]]]]$ deleting of TP

In (16a), there is an NCI. It contains a [u-NEG] feature, which must be checked, which motivates the presence of the negative operator. The NCI has its [u-NEG] feature checked by being c-commanded by the negative operator. In (16b), the NCI is raised to the Spec of FP and in (16c) the rest of the clause is deleted, which makes the NCI a fragment.

3.2 *WH-phrases in* mono ka *RQ s as NCIs* In this subsection, I show that WH-phrases employed in *mono ka* RQs behave as NCIs on a par with WH-*mo* expressions.

Let us first consider the paradigm in (17).

(17) A: John to Bill-wa paatii-ni iku kedo, kimi-mo iku no?
John and Bill-TOP party-to go but you-also go C
'John and Bill are going to the party. Are you, too?'
B1: # Daremo iku mono ka!
anyone go C Q
'No one is going!'
B2: (*Ittai) Dare-ga iku mono ka!
the.hell who go C Q
'I am not going!'

(17B1) is not a felicitous answer to the question in (17A), because replying by saying 'no one' contradicts what the questioner has just said, which is that John and Bill are going to the party. (17B2), on the other hand, is fine as an answer. What is remarkable about it is that the subject refers to the speaker, even though it is a WH-phrase.

Let us now observe (18).

(18) A: Kimi-mo paatii-ni iku no? B: (*Ittai) Dare-ga!
you-also party-to go C the.hell who-NOM
'Are you going to the party, too?' 'Not me!'

The short response in (18B) indicates the speaker's refusal to go to the party, though the nominative subject is a WH-phrase. This interpretation suggests that (18B) is a shortened form of (17B2), which is a *mono ka* RQ. The incompatibility of the WH-phrase *dare-ga* with *ittai* in (18B) also corroborates this view.

It is worth mentioning that the use of the WH-phrase here is strongly reminiscent of the use of expressions like *papa* 'daddy' as imposters, which are grammatically not 1st person but nonetheless refer to the speaker. One property of Japanese imposters is that they can bind the bare anaphor *zibun* 'self' but not

other kinds, as shown in (19).

(19) Sensei (=I)-wa kagami-de *watasizisin/*kanozyozisin/zibun-o mita.
Teacher-Top mirror-in myself/herself/self-Acc saw
'Teacher (=I) saw *myself/*herself/self in the mirror.'
(Furuya 2016: 1730)

Similar effects are found with the WH-subject in a *mono ka* RQ, as shown in (20).

(20) A: Kimi-wa sidookyoozyu-o sonkee-siteiru?
you-TOP academic.advisor-ACC respect-ing
'Do you respect your advisor?'
B: Atarimaeda.
Of.course
Dare-ga zibun-no/*soitu-no sidookyoozyu-o mikudasu mono ka!
who-NOM self-gen/the.guy-GEN advisor-ACC look.down C Q
'Of course. I don't look down on my advisor!'

In (20B), where the WH-subject is taken to be the speaker, the pronoun that is bound by the subject must be the bare 'self'-form *zibun* rather than the 'the guy' form *soitu*. If the WH-subject is replaced by a WH-*mo* expression, as in (21), the intended binding reading is readily available.

(21) Daremo zibun-no/soitu-no sidookyoozyu-o mikudasu mono ka!
anyone self-gen/the.guy-GEN advisor-ACC look.down C Q
'No one looks down on his advisor!'

This example, however, lacks the interpretation where the subject refers to the speaker. (21) is a general statement about people who have an advisor, not the speaker's own attitude toward his own advisor. The speaker orientation, or the imposter use, is only found with WH-phrases.

I speculate on how the WH-phrase *dare-ga* can acquire the speaker status. One possibility would be to suggest that this can be achieved by the WH-phrase undergoing movement to Spec, SAP (Speas and Tenny 2003, Miyagawa 2012, 2017, Saito and Haraguchi 2012, and Haegeman and Hill 2013), which dominates ForceP. The relevant derivation would proceed as in (22).

(22) a. [$_{SAP}$ Dare-ga [$_{ForceP}$ t ... mono ka] $_{SA}{}^{0}$] raising of *dare-ga*
b. [$_{SAP}$ Dare-ga [$_{ForceP}$ t ... mono ka] $_{SA}{}^{0}$] deleting of ForceP

This view is motivated by the following contrast.

(23) A: Dareka kuru sooda yo B: Dare-ga ka na?
someonecome I.hear SFP who-NOM Force SFP
'I hear someone is coming' '(I wonder) who?'

(24) A: Kimi-mo iku no? B: Dare-ga (*ka) yo!
you-also go C who-NOM Force SFP
'Are you going, too?' 'Not me!'

In ordinary questions, as shown in (23), reduced questions can involve the complementizer *ka* followed by a sentence final particle. On the other hand, in the case of *mono ka* RQs, *ka* must be deleted but the sentence final particle can survive. Given that *ka* heads ForceP, its obligatory absence in (24B) indicates that ForceP must be deleted in truncated *mono ka* WH-RQs, as suggested in (22b). Following Saito and Haraguchi's (2012) proposal that sentence final particles head Speech Act Phrase, (24B) indicates that the WH-phrase is in Speech Act Phrase, thereby acquiring the speaker status.

Since the WH-phrase *dare-ga* in a *mono ka* RQ can be found in a fragment answer, we can state that it

can function as an NCI, on a par with WH-*mo* expressions. The [u-NEG] feature on *dare-ga* is checked by being c-commanded by the negative operator containing the [i-NEG] feature, so the WH-phrase can be raised out of the scopal domain of the negative operator.

4 Incompatibility between the two types of NCIs

We have seen that the WH-phrase *dare-ga* in a *mono ka* RQ functions as an NCIs on a par with WH-*mo* expressions. Though they are both NCIs, they are not compatible with each other.

(25) a. * Dare-ga nanimo kau mono ka!
who-NOM anything buy C Q
'No one will buy anything!'
b. * Nanimo dare-ga kau mono ka!
anything who-NOM buy C Q
'No one will buy anything!'
c. * Daremo nani-o kau mono ka!
anyone what-ACC buy C Q
'No one will buy anything!'
d. * Nani-o daremo kau mono ka!
what-ACC anyone buy C Q
'No one will buy anything!'

There are cases where these NCIs differ from each other. WH-*mo* expressions like *nanimo* 'anything' can be inside vP (Miyagawa et al.), but the WH-phrase in a *mono ka* RQ must be fronted.

(26) Taroo-wa [$_{vP}$ umaku nanimo tukur]-anakat-ta.
Taro-top skillfully anything make-NEG-PAST
'Taro did not make anything well.'

(27) a. * Mary-ga nani-o kau mono ka!
Mary-NOM what-ACC buy C Q
'Mary will buy nothing!'
b. (?) Nani-o Mary-ga kau mono ka!
what-ACC Mary-NOM buy C Q
'Mary will buy nothing!'

In (26) the WH-*mo* expression *nanimo* appears to the right of the manner adverb *umaku* 'skillfully'. Since manner adverbs generally mark the left edge of vP, *nanimo* in (26) is in vP. As shown in the contrast in (27), the object WH-phrase in a *mono ka* RQ cannot remain in its original position, but it has to be raised to sentence-initial position.

A similar contrast can be found with quirky subject cases. Kishimoto (2009) suggests that subjects marked with *kara* 'from' stay inside vP. Importantly, the WH-subject in a *mono ka* RQ cannot be marked with it, as shown in (28).

(28) a. John-kara ayamaru
John-from apologize
'John will apologize.'
b. Dare-kara ayamaru no?
who-from apologize C
'Who will apologize (first)?'
c. * Dare-kara ayamaru mono ka!
who-from apologize C Q
'No one will apologize!' 'I will not apologize!'

d. Dare-kara-mo ayamaru mono ka!
who-from-mo apologize C Q
'No one will apologize!'

The behavior of the WH-phrase in (28c) contrasts sharply with that of the WH-*mo* expression in (28d). This clearly indicates that WH-phrases in *mono ka* RQs cannot remain in vP.

Another difference concerns the possibility of multiple occurrence. While there can be more than one WH-*mo* expression (Miyagawa et al.), there can be at most one WH-phrase in a *mono ka* RQ.

(29) a. Daremo nanimo kau mono ka!
anyone anything buy C Q
'No one will buy anything!'

b. ?* {Nani-o} Dare-ga {nani-o} kau mono ka!
what-ACC who-NOM what-ACC buy C Q
'No one will buy anything!'

Miayagawa et al. suggest that the behavior of WH-*mo* expressions in (26) and (29a) comes from their unfocused nature. Here, for the sake of exposition, I assume with Rizzi (1997) that there is a unique structural position above TP for a focused phrase and focalization of two elements is excluded (Rizzi 1997: 290). Then it should be that WH-phrases in *mono ka* RQs are focused NCIs.

We can summarize what we have seen as in (30).

(30)

	fragment answer	appearing in vP	multiple occurrence
・WH in MRQs	possible (NCI)	impossible (focused)	impossible (focused)
・WH-*mo*	possible (NCI)	possible (unfocused)	possible (unfocused)

In a nutshell, WH-phrases in *mono ka* RQs are NCIs on a par with WH-*mo* expressions, but the former (WH-phrases) are focused and the latter (WH-*mo* expressions) are not focused.

I suggest that the incompatibility is related to focus. *Mono ka* RQs are uttered differently, depending on which type of NCI is employed.

(31) a. DAre-ga kuru mono ka!
b. Daremo kuru MONO ka!

As shown in (31a), when a WH-phrase is present in a *mono ka* RQ, it must have heavy accent, while as shown in (31b), when a *mono ka* RQ does not contain a WH-phrase, the complementizer *mono* must receive heavy accent. Given the obligatory fronting of the WH-phrase in *mono ka* RQs, I assume that it undergoes focus related movement out of TP, so (31a) could roughly be translated as 'There is no one who will come!' As for (31b), the focus is on the complementizer, which I take to indicate that the whole *mono* clause (FinP) is focused, resulting in the whole clause being raised to a focus related position. It could be translated as 'It is not the case that anyone will come!' Thus, *mono ka* RQs can be viewed as a case of negative preposing (Han and Seigel 1997), whichever material is raised. This state of affairs is schematized in (32).

(32) a. $[_{ForceP}\ Op_{[NEG]}\ [_{FocP}\ \underline{WH}\ [_{Foc'}\ [_{FinP}\ [_{TP}\ t_{WH}\ ...]\ mono]\ Foc^0]]\ ka_{[RQ]}]$ raising of WH

b. $[_{ForceP}\ Op_{[NEG]}\ [_{FocP}\ [\underline{_{FinP}\ [_{TP}\ WH\text{-}mo\ ...]\ mono}]\ [_{Foc'}\ t_{FinP}\ Foc^0]]\ ka_{[RQ]}]$ raising of FinP

We have already seen that multiple occurrence of WH-phrases leads to deviance, due to the ban on multiple occurrence of focused elements. Then the degraded status of the examples in (25) is treated in the same way. In each example in (25), both the WH-phrase and the *mono* clause are focused, violating the ban.

This view is independently supported by the following example.

(33) ?* JOHN-ga nanimo kau MONO ka!
John-NOM anything buy C Q
'It is JOHN who will buy nothing!'

In (33) *John* is assigned heavy stress, so it is focused as well, as the *mono* clause. (33) has two focused elements, leading to the violation of the ban required to exclude cases like (29b).

This idea can capture the behavior of the following strictly negative predicate *hazu-ga nai* 'it is not possible that...', which is allowed only in negative contexts.

(34) John-ga kuru hazu-ga *aru./nai./ aru MONO ka!
John-NOM come possibility-NOM exist/not.exist / exist C Q
'It is *(not) possible that John will come.'

This predicate is not allowed in a conditional clause.

(35) * [John-ga kuru hazu-ga are-ba] boku-wa komaru daroo
John-NOM come possibility-NOM exist.if I-TOP trouble will
'If it is possible that John will come, I will be in trouble.'

This predicate is compatible with *daremo* but not with *dare-ga*.

(36) Daremo/*DAre-ga kuru hazu-ga aru MONO ka!
anyeone/who-NOM come possibility-NOM exist C Q
'It is not possible that anyone will come!'

(36) suggests that this predicate itself is lexically specified as focused, which makes the clause containing it undergo focus movement, rejecting the presence of another focused element.

As is expected, a focalized subject is not compatible with this predicate.

(36) John-ga/*JOHN-ga kuru hazu-ga aru MONO ka!
John-NOM/JOHN-NOM come possibility-NOM exist C Q
'It is not possible that anyone will come!'

The same thing can be said of another negative predicate *tamaranai* 'be unbearable'.

(37) a. Boku-ga aitu-ni make-te tamaru MONO ka!
I-NOM that.guy-to lose-TE bear C Q
'I will not bear losing to that guy!'

b. ? Darenimo make-te tamaru MONO ka!
anyone lose-TE bear C Q
'I will not bear being defeated to anyone!'

c. * Dare-ga aitu-ni make-te tamaru MONO ka!
who-NOM that.guy-to lose-TE bear C Q
'No one/I will not bear losing to that guy!'

d. * BOKU-ga aitu-ni make-te tamaru MONO ka!
I-NOM that.guy-to lose-TE bear C Q
'I will not bear losing to that guy!'

In sum, there are two types of focus-related movement in *mono ka* RQs: one is the movement of WH-phrases such as *dare-ga* and the other is the movement of the *mono*-clause which may contain NCIs such as WH-*mo* expressions and certain negative predicates. They both target [Spec, Foc]. Since FocP is assumed to allow only one specifier, only one instance of movement is allowed, which is the reason for the incompatibility between the two types of NCIs.

5 Conclusion

In this paper, I showed that *mono ka* RQs should be treated as negative assertions rather than questions and WH-phrases found in them are negative concord items like WH-*mo* expressions, the former being focused and the latter being unfocused. These two types of elements are incompatible with each other due to focus related reasons. WH-phrases in *mono ka* RQs are raised to the Focus position but WH-*mo* expressions do not. Instead, the entire *mono* clause that contains them gets raised. Since there can be only one Focus position, WH-phrases and WH-expression fail to occur with other.

6 References

Adachi, Taro. 2005. Rhetorical questions in Japanese. *Memoirs of Kyoto Tachibana Women's University*: 31, 35-50.

Caponigro, Ivan, and Jon Sprouse. 2007. Rhetorical questions as questions. In Puig-Waldmüller, Estella (ed.), *Proceedings of Sinn und Bedeutung 11*, 121-133. Universitat Pompeu Fabra: Barcelona.

Chomsky, Noam (2000) Minimalist inquiries: the framework. In Martin, Roger, David Michaels and Juan Uriagereka (eds.), *Step by step: essays on minimalist syntax in honor of Howard Lasnik*. 89-155. Cambridge, MA: MIT Press.

Chomsky, Noam (2001) Derivation by phase. In Kenstowicz, Michael. (ed.), *Ken Hale: a life in language*. 1-52, Cambridge, MA: MIT Press.

Furuya, Kaori. 2016. PRO and (under)specification of person in imposter constructions. *Theory and Practice in Language Studies* 6, 1723-1731.

Haegeman, Liliane, and Virginia Hill. 2013. The syntacticization of discourse. In Folli, Raffaella, Christina Sevdali and Robert Truswell (eds.), *Syntax and its limits*. 370–390. Oxford: Oxford University Press.

Han, Chung-Hye and Laura Siegel (1997) Syntactic and semantic conditions on NPI licensing in questions. In Agbayani, Brian and Sze-Wing Tang (eds.), *The Proceedings of the Fifteenth West Coast Conference on Formal Linguistics*. 177-192. CSLI Publications: Stanford University.

Hiraiwa, Ken, and Shinichiro Ishihara. 2012. Syntactic metamorphosis: clefts, sluicing, and in-situ focus in Japanese. *Syntax* 15: 142-180.

Kishimoto, Hideki. 2009. Topic prominency in Japanese. *The Linguistic Review* 26(4): 465-513.

Kuwabara, Kazuki. 2013. Peripheral effects in Japanese questions and the fine structure of CP. *Lingua* 126: 92-119.

McGloin, Naomi Hanaoka. 1976. Negation. In Shibatani, Masayoshi (ed.). *Syntax and semantics 5: Japanese generative grammar*. 371-419. New York: Academic Press.

Miyagawa, Shigeru. 2012. Agreements that occur mainly in the main clause. In Aelbrecht, Lobke, Liliane Haegeman and Rachel Dye (eds.), *Main clause phenomena: new horizons*. 79-112. Amsterdam: John Benjamins.

Miayagawa, Shigeru. 2017. *Agreement beyond phi*. Cambridge, MA: MIT Press.

Miyagawa, Shigeru, Nobuaki Nishioka, and Hedde Zeijlstra. 2016. Negative sensitive items and the discourse-configurational nature of Japanese. *Glossa: a Journal of General Linguistics* 1(1): 33. 1–28.

Nishioka, Nobuaki. 2000. Japanese negative polarity items wh-*MO* and XP-*sika* phrases: another overt movement analysis in terms of feature-checking. In Takami, Ken-ichi, Akio Kamio and John Whitman (eds.), *Syntactic and functional explorations in honor of Susumu Kuno*, 159–184. Tokyo: Kurosio Publishers.

Oguro, Takeshi. 2018. Properties of *mono ka* rhetorical questions. *Proceedings of the 13th Workshop on Altaic Formal Linguistics (WAFL13)*. 193-204.

Pesetsky, David. 1987. WH-in-situ: movement and unselective binding. In Reuland, Eric and Alice ter Meulen (eds.), *The representation of (in)definiteness*. 98-129. Cambridge, MA: MIT Press.

Rizzi, Luigi. 1997. The Fine structure of the left periphery. In Haegeman, Liliane. (ed.), *Elements of grammar: handbook of generative syntax*. 281–337. Dordrecht: Kluwer.

Saito, Mamoru and Tomoko Haraguchi. 2012. Deriving the cartography of the Japanese right periphery: the case of sentence-final discourse particles. *Iberia* 4: 104-123.

Speas, Peggy, and Carole Tenny. 2003. Configurational properties of point of view roles. In Di Sciullo, Anna Maria (ed.), *Asymmetry in grammar: vol 1. syntax and semantics*. 315-344. Amsterdam: John Benjamins.

Watanabe, Akira. 2004. The genesis of negative concord: syntax and morphology of negative doubling. *Linguistic Inquiry* 35: 559–612.

Persian: Quantity Sensitive and Iambic

Fatemeh Samavati
California State University, Fresno

1 Introduction

In this paper, I focus on the modern conversational Iranian Persian word-accent. The essential question that I try to answer is how the iambic system works in the Persian language. Therefore, I discuss foot type, directionality (alignment), location of the main prominence, and quantity sensitivity in Persian. To compare the results of this study with the previous studies, I categorize the data into some groups as Ferguson (1957) does. I collected all the data used in this paper from a Persian native speaker, and I analyzed them by Praat. I use the Optimal Theory (Smolensky & Prince 1993) approach to get the constraints ranking for the accent patterns in Persian.

2 Literature Review

Some scholars consider Persian as a stress-accent language, and some others do not. Windfuhr (1997) and Amini (1997) describe the lexicon accent in Persian as a stress-accent. Lazard (1992) expresses the accented syllables in Persian are pronounced with both a high pitch and stress. Bijankhan et al. (2012) have shown that Persian is a pitch-accent language, and the prominence in the word-level falls on the final syllable.

Ferguson (1957) and Eslami (2010) declare that Persian is a final-accented language, and Windfuhr (1997) expresses the word accent is rightmost, and the phrasal accent is leftmost in Persian while Amini (1997) and Tabibzadeh (2010) posit that Persian is an iambic language. Ferguson (1957), Windfuhr (1997), Amini (1997), Eslami (2010), and Tabibzadeh (2010) assert no secondary accent is observed in Persian, and the number of syllables does not affect the accent system in Persian language.

Persian is a foot-binarity based and right-aligned language as Amini (1997) and Tabibzadeh (2010) claim. Amini (1997) assumes that Persian is a syllabic iamb language, and the accent is not affected by the syllable weight in this language, and Persian is a quantity insensitive language. However, all iamb languages are moraic (Hayes 1995). According to Tabibzadeh (2010), Persian is an iambic and quantity sensitive language. Although Tabibzadeh (2010) does not imply to vowel lengthening in Persian, he posits that all light syllables in the final syllable positions are considered as heavy syllables and he tries to get the syllabic foot in Persian through the iambic foot pattern.

Ferguson (1957) claims that if the dictionary entry form of the lexicon is considered, Persian is a final-accented language, and because Persian word accent is dynamic, when a word gets a plural, comparative or superlative markers, or a derivational suffix, the words accent will shift with them. For example, in the word, *xaa.né (house)***,** the accent falls on the final syllable, and in its plural form, *xaa.ne.-háá (houses)*, the plural marker -háá gets the accent. Other examples are the words *maa.híí (fish)* and *máá.h-ii (a moon, a month)* that are minimal pairs. In the second word, *máá.h-ii (a moon, a month)*, the suffix *ii* is a determiner, and it does not get the accent. In his view, the accent falls on the last syllable in Persian infinitive form of the verb. For example, *xor.d-án (to eat)*, is an infinitive from of the verb, and the same as nouns, it is final-accented.

* I'd like to thank Chris Golston that this work would not have been possible without his great help. I also appreciate Zachary Metzler for the helpful discussions and his generous guidance.

Ferguson (1957) and Amini (1997) state there are some exceptional words in Persian such as conjunctions and yes/no expressions that their accent always fall on the first syllable: *vá.lii (but), bál.ke (but), ʔám.maa (but), bá.le (yes), ʔáá.re (yeah), and ná.xeir (no).*

Ferguson (1957) elaborates on how the accent system works in Persian verbs. As he describes, in the is third-person singular simple past verbs, the accent falls on the final syllable because it is the base form of the verbs in Persian, but in the other simple past verbs, the accent falls on the pre-final syllable because they have a dependent pronoun. For example, in the third person, singular, simple past verb *ne.-véft ((he) wrote)*, the last syllable gets the accent while in the first person, singular, simple past verb *ne.-véf.t-am ((I) wrote)*, the accent falls on the syllable before the last as the dependent pronoun *am* does not get the accent. In past perfect and present perfect, the last syllable of the present participle gets the accent. For example, *xor.déé bu.dam ((I) had eaten)* is a past perfect verb and its past participle is *xor.déé*, with the accent on its final syllable. In the past and present progressive verbs, the progression prefix, *mii*, gets the accent. For instance, the progressive verbs, *míí-.ne.vii.s-am ((I) am writing)*, and *míí.ne.veſ.t-am ((I) was writing)*, are initial-accented. In the imperative verbs, the imperative prefix gets the accent: *bé.ne.viis (Write!), ná.ro (Don't go!)*, but if the imperative verb is without the prefix, the accent falls on the final syllable: *ſe.nó (Listen!)*. As a result, according to Ferguson (1957), all verb prefixes such as *míí (progressive marker), bé (imperative marker), and né (negation marker)* get the accent.

Amini (1997) says that Persian is a right-aligned syllabic quantity-insensitive iambic language. She applies the Optimality Theory in Persian accent, and she acclaims that Persian word accent is ranked: **FT-BINARITY, PARSE SYLLABLE, RIGHT-ALIGNED, STRESS-WELL > NONFINALITY**. Tabibzadeh (2010) claims that Persian has a moraic, iambic, and Ft-Bin and right-aligned accent system. To prove this claim, he states the light syllables at the end of the words should be considered as a heavy syllable (the final syllables should be pronounced as a heavy syllable).

3 Vowel Lengthening in Persian

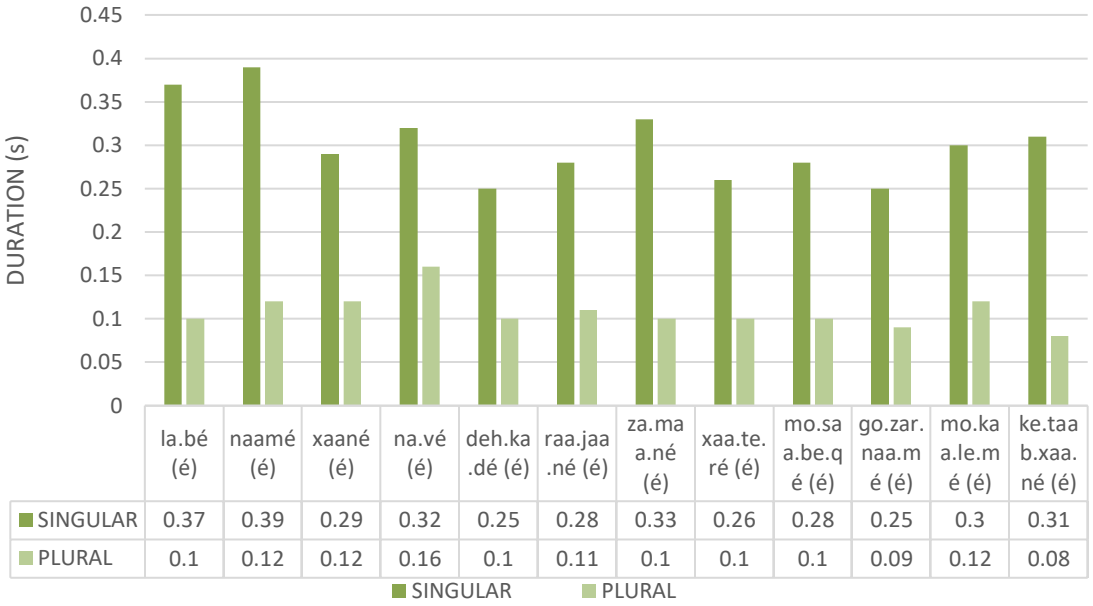
I measure the vowel 'e' duration in the last syllable of the following singular words and their plural forms by Praat. I choose the plural form of the word because Ferguson (1957) and Amini (1997) claim that when a plural marker is added to a noun, the accent falls on the plural marker. Table 1 provides some examples of words with two, three, and four syllables with the lengthened final syllable compared to the plural form of them without those lengthened.

SINGULAR	MEANING	PLURAL
la.béé	edge	la.be.-háá
naa.méé	letter	naa.me.-háá
xaa.néé	home	xaa.ne.-háá
na.véé	grandchild	na.ve.-háá
deh.ka.déé	village	deh.ka.de.-háá
raa.jaa.néé	computer	raa.jaa.ne.-háá
za.maa.néé	era	za.maa.ne.-háá
xaa.te.réé	memory	xaa.te.re.-háá
mo.saa.be.qéé	game	mo.saa.be.qe.-háá
go.zar.naa.méé	passport	go.zar.naa.me.-háá
mo.kaa.le.méé	dialogue	mo.kaa.le.me.-háá
ke.taab.xaa.néé	library	ke.taab.xaa.ne.-háá

Table 1. The Duration of 'e' in Singular and Plural Words in Persian

As you see in Figure 1, the duration of the vowels in the last syllable of the single words are more than when they occur in a plural word. Its duration in single words are about two times of its duration when they

get a plural marker.



	la.bé (é)	naamé (é)	xaané (é)	na.vé (é)	deh.ka .dé (é)	raa.jaa .né (é)	za.ma a.né (é)	xaa.te. ré (é)	mo.sa a.be.q é (é)	go.zar. naa.m é (é)	mo.ka a.le.m é (é)	ke.taa b.xaa. né (é)
SINGULAR	0.37	0.39	0.29	0.32	0.25	0.28	0.33	0.26	0.28	0.25	0.3	0.31
PLURAL	0.1	0.12	0.12	0.16	0.1	0.11	0.1	0.1	0.1	0.09	0.12	0.08

Figure 1. Vowel Duration in Singlular and Plural Words in Persian

Hayes (1985) claims that ITL exposed an underlying durational asymmetry in both trochaic and iambic systems. In trochee feet, the evenness of duration between the foot elements causes it, and in the iambic systems, the final element prominence is strengthened by vowel lengthening, consonant germination, vowel shortening, etc. Kager (1993) argues that there are alternatives to the iambic-trochaic law. His theory extracts such asymmetries from two parameters: first, to avoid clash and lapse in the sequence of moras, and second for moraic prominence due to heavy syllables.

Alber (2005) claims that foot type, directionality of parsing, location of main stress, and quantity sensitivity are the binary parameters to determine the accent types. There is one well-known typological gap in these stress systems generated by these parameters that is right aligned iambic stress system which is considered unattested by Alber (2005). The other three types are left-aligned trochaic, left-aligned iambic, and right-aligned trochaic (see Table 2).

Trochaic system, Left aligning	(ˈσ σ)(ˈσ σ) σ
Trochaic system, Right aligning	σ (ˈσ σ)(ˈσ σ)
Iambic system, Left aligning	(σˈσ)(σˈσ) σ
Iambic system, Right aligning (unattested in Alber (2005), but true in Persian accent pattern)	σ (σˈσ)(σˈσ)

Table 2. Typology of Stress Systems

As Kager (1993) reviews, based on Hayes (1991), trochees contrast in intensity and iamb contrast in duration, so trochaic feet are durational even while iambic feet are durational uneven, so it reflects even syllabic and moraic trochees and uneven iambs. Kager (1993) discusses that quantity-sensitivity occurs in the moraic systems, and one way to get uneven iambs is iambic lengthening. Therefore, the second element of the last foot in a Persian word should be heavy to get accented. Due to iambic lengthening, I make the argument that Persian feet are quantity sensitive iambs, so when a word ends with a light syllable, the short vowel that is 'e' in most cases in the modern Persian language becomes 'e:' to make the syllable heavy, so Persian follows Stress to Weight Principle.

4 Optimal Theory approach to Persian accent

Although Amini (1997) suggests an OT constraints ranking for Persian language, based on my application of Optimality Theory on Persian accent, three different OT constraints rankings can be considered as Ferguson (1957) has divided Persian accent into three main categorizations. The first group includes Persian nouns, adjectives, and adverbs. The second one is Persian verbs, and the third one is yes/no expressions, conjunctions, and adverbial ordinal numbers that Ferguson (1957) claims their accent falls on their initial syllable.

4.1 *OT approach to Persian nouns, adjectives, and adverbs accent* In all nouns (singular and plural), adjectives (including superlative and comparative adjectives), and adverbs, the final syllable gets the accent (Ferguson 1957). Tableau 1 shows how the constrains work in Persian nouns', adjectives', and adverbs' accent. It is important to know that clitics (ezafe constructions, prepositions, contracted forms of the auxiliary verbs, pronouns, determiners, etc) do not get accent in Persian. For example, in the word *máá.h-ii (a moon, a month)*, the suffix *ii* is a determiner, and it does not get the accent. Therefore, the accent falls on the syllable before the determiner *ii*. Again, all plural markers (*háá, áán, úún, íín*, *áát,* and *áá (háá in colloquial Persian))*, superlative and comparative adjective markers, (respectively, *ta.rín* and *tár*) and adverb marker (*aa.néé*) are bound functional suffixes and they get the accent. As the infinitive from of the verb can place as a noun position in the sentences, its accent falls on its last syllable, the same as nouns. An infinitive form of the verb is made up of the 3rd singular form of the past verb + *án*.

(1) singular nouns accent

a. me.dáád (pencil)
b. ɡol.dún (vase)
c. sa.laa.ma.tí (health)
d. daa.neʃ.ʤú (student)

(2) plural nouns accent

a. me.daa.d-áá
pencil-PL
pencils
b. ɡol.du.n-áá
vase-PL
vases

c. de.rax.t-áán
tree-PL
trees

d. ma.qaa.l-áát

article-PL
articles

(3) nouns (with clitics) accent

a. míz-e
table-EZ
table's

b. daf.tá.r-í
notebook-DET
a notebook

c. pi.rá.n-eʃ
shirt- 3SG POSS
her/his shirt

d. kí.f-e.mun
bag-1PL POSS
our bag

(4) adjectives accent

a. meh.ra.bún (kind)
b. zi.báá (beautiful)
c. tu.laa.ní (long)
d. xoʃ.típ (handsome)

(5) superlative adjectives accent

a. meh.ra.bún (kind)
b. zi.báá (beautiful)
c. tu.laa.ní (long)
d. xoʃ.típ (handsome)

(6) comparative adjective accent

a. meh.ra.bun-tár
kind-er
kinder

b. zi.baa.-tár
beautiful-er
more beautiful

c. tu.laa.ni.-tár
long-er
longer

d. xoʃ.tip.-tár
handsome-er
more handsome

(7) adverbs accent

a. xaa.le.s-aa.néé
sincere-ly
sincerely

b. saa.de.q-aa.néé
honest-ly
honestly

c. fo.ru.ta.n-aa.néé
humble-ly
humbly

d. dus.t-aa.néé
friend-ly
friendly

(8) infinitives accent

a. xaa.bi.d-án
sleep-to
to sleep

b. ne.veʃ.t-án
write-to
to write

c. xaan.d-án
read-to
to read

d. ʃe.ni.d-án
hear-to
to hear

IAMBIC, FT-BIN, AL FT R, AL WD R, ALIGN HD R, SWP > NONFINALITY

/mad.re.see/ (school)	IAMBIC	FT-BIN	AL FT R	AL WD R	ALIGN HD R	SWP	NONFINALITY
☞ a. mad.re.séé σ (σˈσ)							
b. mad.re.sé σ (σˈσ)						*!	
c. mad.ré.se σ (ˈσ σ)	*!					*!	
d. mad.réé.se σ (ˈσ σ)	*!		*!	*!	*!		

e. mád.re.se (ˈσ) (σ σ)		*!	*!	*!	*!		

Tableau 1

4.2 *OT approach to Persian verbs accent* In the third-person singular simple past verbs, the final syllable is accented because it does not have any affixes or clitics, but in the other verbs, the accent placement is different. In Persian, the inflectional morphemes get the accent either as a suffix or a prefix. For instance, in progressive, negative, and imperative verbs, the accent falls on the initial syllable which are the progressive marker *(mí)*, the negation marker *(né)*, and the imperative marker *(bé)*. They are inflectional bound prefixes. In the past perfect, present perfect, and passive voice verbs, the accent falls on the last syllable of the past participle. The present perfect, past perfect, and passive voice verbs are made up of a past participle and an auxiliary verb. Therefore, the last syllable of the past participle 000000000000000000000000000gets the accent. The accent constraint in Persian verb with an Optimal Theory approach is ranked as: ALIGN-R (ROOT, H), ALIGN-R (STEM, L), IAMBIC, FT-BIN, PARSE SYLLABLE > NONFINALITY. Tableau 2 shows how it works.

(9) progressive verbs accent

a. mí.-ne.veʃt-Ø
PROG-write.PAST-3SG
(S/he) was writing.

b. mí.-xor.d-im
PROG-eat.PAST-1PL
(We) were eating.

c. mí.-ne.vis-e
PROG-write-3SG
(S/he) is writing.

d. mí.-xor.-im
PROG-eat-1PL
(We) are eating.

(10) negative verbs accent

a. ná.-ne.veʃt-Ø
NEG-write.PAST-3SG
(S/he) didn't write.

b. ná.-raf.t-am
NEG-go.PAST-1SG
(I) didn't go.

c. ná.-xaa.bi.d-im
NEG-sleep.PAST-1PL
(We) didn't sleep.

d. ná.-jaa.mad-Ø
PROG-write.PAST-3SG
(S/he) didn't come.

(11) imperative verbs accent

a. bé.-nuʃ-Ø
IMP-drink-2SG
Drink!

b. bé.-bin-Ø
IMP-look-2SG
Look!

c. bé.-nu.ʃ-in
IMP-drink-2PL
Drink (you all)!

d. bé.-bi.n-in
IMP-look-2PL
Look (you all)!

(12) present perfect verbs accent

a. pu.ʃi.d-éé.-ʔam
wear-PART.PERF.1SG
(I) have worn.

b. re.si.d-éé.-ʔi
arrive-PART-PERF.2SG
(You) have arrived.

c. xaan.d-éé.-ʔim
read-PART-PERF.1PL
(We) have read.

d. ʔaaz.mu.déé.-ʔim
try-PART-PERF.1PL
(We) have tried.

(13) past perfect verbs accent

a. pu.ʃi.d-éé.-bu.d-am
wear-PART-PERF.PAST-1SG
(I) had worn.

b. re.si.d -éé.-bu.d -i
arrive-PART-PERF.PAST-2SG
(You) had arrived.

c. xaan.déé.-bu.d-im
read-PART-PERF.PAST-1PL
(We) had read.

d. ʔaaz.mu.déé.-bu.d-im
try-PART-PERF.PAST-1PL
(We) had tried.

(14) passive verbs accent

a. pu.ʃi.d-éé.-ʃod
wear-PART-become.PAST
It was worn.

b. re.saan.d-éé.-ʃod
arrive-PART-become.PAST
It was arrived.

c. xaan.d-éé.-ʃod
read-PART-become.PAST
It was read.

d. ʔaaz.mu.déé.- ʃod
try-PART-become.PAST
It was tried.

ALIGN-R (ROOT, H), ALIGN-R (STEM, L), IAMBIC, FT-BIN, PARSE SYLLABLE > NONFINALITY

/ mi.-ne.veʃt /((He) was writing)	ALIGN-R (ROOT, H)	ALIGN-R (STEM, L)	IAMBIC	FT-BIN	PARSE SYLLABLE	NONFINALITY
☞ a. / mí.-ne.veʃt /((He) was writing) (σˈ) (σ σ)	*!					
b. / mi.-né.veʃt /((He) was writing) σ (ˈσ σ)	*!	*!	*!			
c. / mi.-ne.véʃt /((He) was writing) σ (σ ˈσ)	*!	*!			*!	

Tableau 2

4.3 *OT approach to initial-accented words in Persian* Ferguson (1957) claims that there are some exceptional words in Persian that their accent falls on the initial syllables. These exceptional words include yes/no expressions such as *bá.le (yes), áá.re (yeah), and ná.xeir (no)*, conjunctions like *vá.lii (but), bál.ke (but), and ʔám.ma (but),* and all ordinal numbers as adverbs that are borrowed from Arabic such as *ʔáv.va.lan (firstly), sáá.ni.jan (secondly), sáá.le.san (thirdly), ráá.be.ʔan (fourthly)*, etc. The data measurements show that in this group, the intensity and pitch have the same pattern, and the accent always falls in the first syllable. As Kager (1993) reviews Hayes (1991), trochees contrast in intensity, and iamb contrast in duration, so trochaic feet are durational even while iambic feet are durational uneven, so it reflects even syllabic and moraic trochees and uneven iambs, and this data shows a syllabic trochee pattern. The constraints of this special category are: TROCHEE, FT-BIN, ALL FT L, AL WD L, ALIGN HD L (Tableau 3).

(15) yes/no expressions accent

a. bá.le (yes)
b. ʔáá.re (yes)
c. bá.li(yes)
d. ʔáá.ri (yes)
e. ná.xeir (no)

(16) conjunctions accent
a. ʔám.ma (but)
b. vá.li (but)
c. lí.ken (but)
d. láá.ken (but)

(17) adverbial ordinal numbers accent
a. ʔáv.va.lan (firstly)
b. sáá.ni.jan (secondly)
c. sáá.le.san (thirdly)
d. ráá.be.ʔan (fourthly)
e. xáá.me.san (fifthly)

/ saa.le.san / (thirdly)	TROCHEE	FT-BIN	ALL FT L	AL WD L	ALIGN HD L
☞ a. sáá.le.san (ˈσ σ) σ					
b. saa.lé.san (σˈσ) σ	*!				*!
c. saa.lé.san σ (ˈσ σ)			*!	*!	*!
d. saa.lé.san σ (σ 'σ)	*!		*!	*!	*!
e. saa.le.san σ σ σ	*!	*!	*!	*!	*!

Tableau 3

5 Conclusion

Persian is a final-syllable accented language if the base form of each word is considered (Ferguson 1957). However, many words in Persian are made from different kinds of affixations, and the OT constraints for the Persian words accent is tricky. Based on my application of Optimality Theory on Persian accent, three different OT constraints rankings can be considered as Ferguson (1957) has categorized Persian accent into three main groups. The first and second group follow an iambic pattern. In the first categorization, the OT constraint ranking is: IAMBIC, FT-BIN, AL FT R, AL WD R, ALIGN HD R, SWP > NONFINALITY. In this paper, the comparison of singular nouns ending with ‘e’ vowel with plural forms of them shows us that in singular words, the accented syllable which is the final syllable is lengthened to get the accent, so it follows SWP. Expressly, the light syllable should be lengthened to get the accent. As a result, Persian is a quantity sensitive, right aligned, and iambic language. Alber (2005) presents four types for accent systems, and she mentions that right-aligned iambic system is unattested. However, Persian language follows this pattern in nouns, adjectives, and adverbs accents.

References

Abolhasanizadeh, V., Bijankhan, M., & Gussenhoven, C. (2012). The Persian pitch accent and its retention after the focus. *Lingua*, *122*(13), 1380-1394.

Alber, B. (2005). Clash, lapse and directionality. *Natural Language & Linguistic Theory*, *23*(3), 485-542.

Altshuler, D. (2009). Quantity-insensitive iambs in Osage. *International Journal of American Linguistics*, *75*(3), 365-398.

Amini, A. (1997). On stress in Persian. *Toronto Working Papers in Linguistics*, *16*(1), 1-20.

Eslami, M. (2010). Stress in Persian language. *Signs and Data Processing, 11 (1),* 1-12.

Ferguson, C. A. (1957). Word stress in Persian. *Language*, *33*(2), 123-135.

Hayes, B. (1985, June). Iambic and trochaic rhythm in stress rules. In *Annual Meeting of the Berkeley Linguistics Society* (Vol. 11, pp. 429-446).

Hayes, B. (1991). Metrical Stress Theory: Principles and Case Studies, unpublished manu-script, UCLA. *HayesMetrical Stress Theory: Principles and Case Studies1991*.

Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. University of Chicago Press.

Kager, R. (1993). Alternatives to the iambic-trochaic law. *Natural Language & Linguistic Theory*, *11*(3), 381-432.

Lazard, G. (1992). *A grammar of contemporary Persian*. Costa Mesa, CA: Mazda Publishers, p.48.

Smolensky, P., & Prince, A. (1993). Optimality Theory: Constraint interaction in generative grammar. *Optimality Theory in phonology*, 3, 1-144.

Verb Types for Deictic Directional Particles

Ronald P. Schaefer and Francis O. Egbokhare
Southern Illinois University Edwardsville and University of Ibadan

1 Introduction

The linguistic expression of motion remains a lively topic of investigation under various guises. Associated motion, for instance, has been re-invigorated of late by typological study of the Amazonian basin in South America (Guilliame 2016) after initial findings from Australia (Koch 1984). Studies of translational motion have drawn inspiration for a number of decades from Talmy (1985, 2000). And deictic motion verbs have been under seemingly continuous investigation in recent times (Anderson and Keenan 1985, Dixon 1991, Fillmore 1966, 1997).

But there is more to deictic directionality (DD) than verbs of motion toward or away from a deictic center (Hooper 2002). Linguistic coding in this overarching domain is not restricted to lexical categories. Coding is also achieved by grammatical forms. Across the language phyla of Africa for example, the primary devices under investigation are suffixes (Creissels 2008), although under an assortment of labels, including andative, itive, translocative, altrilocal, centrifugal for movement away vs. venitive, cislocative, equilocal and centripedal for movement toward. To be sure, we do not intend to expand this terminological maze. For contrasting dimensions of deictic directionality expressed by grammatical forms we will employ "andative" and "venitive."

These terms apply to two particles that occur throughout our corpus of Nigeria's minority language Emai (Edoid, West Benue Congo and Niger Congo, Elugbe 1989, Williamson and Blench 2000), which remains underdescribed (Schaefer and Egbokhare 1999, 2007, 2017), although perhaps less so than most of its Edoid neighbors. Relatively strict SVO, Emai manifests lexical and grammatical tone, remnant segmental inflection, and few prepositions. It has verbs that are intransitive, transitive, and ditransitive, as well as complex predicates articulated by a verb series or a verb plus particle nexus.

Verbs in Emai code translational motion (a figure undergoing directed motion relative to a ground/landmark as reference point) and deictic directional motion (a figure moving to or from a deictic reference point). For the translational type, simple predicates conflate motion+path (non-deictic directed motion) such as *lagaa* 'move round' and *heen* 'move upward, ascend.'

(1) a. ólí ómò ò ó làgáá ùhàì.
the child SC C move.round well
'The child is moving around the well.'

b. ólí ómò héén ólí ókòó.
the child PRP.ascend the hill
'The child has ascended the hill.'

For deictic directional motion, Emai verbs encode motion+deictic-direction. Equivalents for English 'go' and 'come' are *lode* 'go,' *vare* 'perfective come,' and *vade* 'imperfective come.'

* Data incorporated in this paper derive from research sponsored by the National Science Foundation, (BNS #9011338 and SBR #9409552), the Department of State (College and University Affiliations Program grant ASJY 1333), the National Endowment for the Humanities (PD-50004-06), Southern Illinois University Edwardsville and the University of Ibadan, Nigeria. We thank these institutions for their continued support, while not extending to them any responsibility for our data interpretation.

Proceedings of *WECOL 2019*

(2) a. ólì òkpòsò lòdé vbì ìwè.
the woman C.go LOC house
'The woman is going to the house.'
b. ólì òkpòsò váré vbì ìwè.
the woman PRP.come LOC house
'The woman has / will come to the house.'
c. ólì òkpòsò vàdé vbí ímè.
the woman C.come LOC farm
'The woman is coming to the farm.'

Deictic directional particles in Emai are venitive (VN) *re* and andative (AN) *a*. They combine with verbs, as do other post-verbal particles that exhibit functions bearing on valency or aspectual conditions. Particles in Emai cover some of the semantic terrain that verbal suffixes do in Niger-Congo and Bantu, as discussed in Hyman (2007), Schadeberg (2003), among others. In Emai there are three particles with valency effects: applicative *li*, change of location *o*, and projected adherence *e*. None occurs as sole verb of a predication nor shows evidence of ever having done so.

(3) a. ólì òkpòsò shén émà lí ònwìmè.
the woman PRP.sell yam APP farmer
'The woman has sold yam to a farmer.'
b. ólì òkpòsò nwú émà ó vbí úkpódè.
the woman PRP.carry yam CL LOC road
'The woman has put yam onto the road.'
c. ólì òkpòsò óhó úshé'n kú é ólì ònwìmè.
the woman PRP.blow powder disperse PAD the farmer
'The woman has blown powder all over the farmer.'

One post-verbal particle has an aspectual effect: temporal perspective *lee* 'already.'

(4) a. òjè ò ó è ólí émáé lèé.
Oje SC C eat the food TEMP
'Oje is already eating the food.'
b. òjè é ólí émàè léé.
Oje PRP.eat the food TEMP
'Oje has already eaten / finished the food.'

Members of this particle set represent a closed-class. They are largely polysemous, combining distinct functions under single forms. Such is the case with applicative *li*, which signals recipient for change of possession and benefactive otherwise. It is the latter function that exhibits combinatorial potential with other particles, like change of location *o* in (5).

(5) òjè ré óbò ó vbí ébè lí àlèkè.
Oje PRP.take finger CL LOC paper APP Aleke
'Oje has signed the paper for Aleke.'

2 Deictic directionality particles

Deictic directional particles in Emai co-occur with a variety of transitive and intransitive verbs, although not ditransitives and their double object structure. Many of the co-occurring verbs code events of directed motion, change of state or of a non-literal, fictive nature. Verb-particle combinations code events that involve not only change but also deictic center defined by speaker location. Fundamental to DD events is scalar change brought about by a force construed as internal or external to a grammatically expressed event participant. Exertion of force leads to change on a scale from one point to another; it culminates in change that is either directed toward or away from the speaker at deictic center.

TABLE 1. Distribution of DD particles (PT) relative to verb transitivity type (transitive, intransitive), verb directionality type (bidirectional, unidirectional), and force type for entity of change (external, internal).

	PT	TR	INTR	BI-D	UNI-D	EXT-F	INT-F
o 'enter'	re		√		√		√
nwu 'carry'	re	√			√		√
ze 'scoop'	re	√			√		√
shie 'coil'	re	√	√		√		√
filo 'develop'	re	√			√		√
ze 'grow'	re		√		√		√
shoo 'awaken'	re		√		√		√
ranran 'soak'	re		√		√		√
buu 'approach'	re	√			√		√
gbaan 'wind'	a	√	√		√	√	
khuye 'close'	a	√	√		√	√	
gbe 'hit'	a	√	√		√	√	
fu 'decline'	a	√	√		√	√	
fu-no 'decrease'	a	√	√		√	√	
too 'burn'	a	√	√		√	√	
hian 'cut'	a	√	√		√	√	
guogho 'break'	a	√	√		√	√	
anme 'scrape'	a	√			√	√	
kpe 'wash'	a	√			√	√	
hoo 'wash'	a	√			√	√	
aa 'decompose'	a		√		√	√	
me 'move to/fro'	re		√	√			√
me 'move to/fro'	a		√	√		√	
ee 'be anxious'	re	√	√	√			√
ee 'be anxious'	a	√	√	√		√	
koo 'count'	re	√	√		√		√
fi 'project'	a	√			√	√	

Talmy (1991, 2000) refers to events of this type as encoding change in existence from presence to absence or absence to presence. We find merit in this general characterization, distinguishing it from other basic framing events such as simple change of directed motion, change of material state, or change of possession. However, Emai DD particles are constrained in their co-occurrence with verbs. In particular, change in existence by itself does not stipulate which verb selects which particle. Our hypothesis is that verbs select their particle, not the other way round. This being so, the question is why or by what means does a verb select a particle that expresses a venitive or andative meaning.

From our current perspective, a verb selects a given DD particle if its internal semantic structure is impoverished or underspecified with respect to particle meaning while still being compatible with that meaning. Our hypothesis for the DD domain is that an underspecified or covert meaning element stipulates a directional feature whose possible values with respect to deictic center are unidirectional or bidirectional. If unidirectional, possible overt specification would be either toward (venitive) or away (andative) from the deictic center. If bidirectional, specification is not limited to one deictic directional, either would be compatible with a verb.

As we will see, these three options characterize Emai verb selection of a deictic directional particle. Some verbs select only venitive. Others elect only andative. And still a few pick venitive and andative, although not both for a single predication. In addition, change in existence type, i.e. directional specification as andative or venitive, correlates in large measure with force type. Force is another covert feature of verbs that is attributed to one of its arguments. Force construed as internal to the entity of change aligns with venitive marking, while force external to that entity allows an andative mark. An apparent complicating factor with respect to this overall interpretation arises from two events whose coding equivalent is English 'leave.' Noteworthy about such events is that while they are transitive and

unidirectional, they engage not so much force as the intentional non-exertion of force on a potential entity of change. Direction for such events is measured against not a potential entity of change but an entity that does not undergo change. Further discussion of 'leave' verbs occurs in § 2.4.

We direct attention now to an overview of the distributional character of venitive *re* and andative *a*. We then explore their nature relative to verb transitivity, specification of change as bidirectional or unidirectional, and force exertion as internal or external to the entity of change. These elements of event character are summarized in Table 1.

2.1 *Particle with unidirectional feature and internal force* One class of verbs takes only a venitive particle. Members are transitive or intransitive. Verbs in this class are underspecified for a unidirectional feature of change that pertains to movement. The force controlling change is attributed to the entity of change that is encoded as grammatical subject. One of the clearest examples of this class is the verb *o* 'enter.' It appears in its neutral, underspecified form in (1a). Change for this verb is unidirectional, i.e. from the unspecified location of *ólí ómòhè* 'the man,' as entity of change and grammatical subject, to the locative marked reference point *ìwè* 'house.' The force that engenders change is the referent of the grammatical subject. In a non-neutral predication with venitive *re*, (1b), particle usage specifies that the house falls within the speaker's sphere of existence.

(6) a. ólí ómòhè ó vbì ìwè.
the man PRP.enter LOC house
'The man has entered the house.'
b. ólí ómòhè ó vbì ìwè ré.
the man PRP.enter LOC house VN
'The man has entered the house where speaker is.'

Many additional verbs in the venitive-only class are transitive. They include a relatively large number of verbs represented by *nwu* 'carry' and *zẹ* 'scoop.' In their neutral clausal form they refer to an activity. It is the referent of their subject (as well as direct object) that undergoes a change whose end point is underspecified in (7a) and (7c). Under a non-neutral venitive condition (7b) and (7d), change concludes at a deictic point in space within the speaker's sphere of existence. The force controlling change is the entity of change, i.e. the referent of the grammatical subject.

(7) a. ólì òkpòsò nwú émà.
the woman PRP.carry yam
'The woman has carried yam.' / *brought'
b. ólí ókpósó nwú' émà ré.
the woman PAP.carry yam VN
'The woman brought yam.'
c. ólí ókpósó zẹ́' èkẹ̀n.
the woman PAP.scoop sand
'The woman scooped sand.'
d. ólí ókpósó zẹ́' èkẹ̀n ré.
the woman PAP.scoop sand VN
'The woman brought sand.'

Most remaining verbs in the venitive-only class occur as transitive or intransitive verbs. In their neutral form (8a, 8c, 8e, 8g) the feature of directional change for these verbs is underspecified. In their non-neutral form (8b, 8d, 8f, 8h, 8i), a deictic directional feature is explicit in their venitive particle. Each instance of change terminates in a sphere of existence that incorporates the speaker. As well, the force of change feature is attributed to the subject referent, whether expressed grammatically or logically understood (8h-3i).

(8) a. ólí ókà ò ó fì-lò íké'.
the maize SC C develop-DS shoot
'The maize is growing / developing shoots.'

b. ólí ókà fí-ló íké' ré.
the maize PRP.develop-DS shoot VN
'The maize has sprouted shoots.'

c. ólí ókà ò ó zé.
the maize SC C grow
'The maize is growing.'

d. ólí ókà zé ré'.
the maize PRP.grow VN
'The maize has sprouted / grown out / has come into view.'

e. ólí ómòhè shóó ré'.
the man PRP.awaken VN
'The man has arisen / gotten up out of bed.'

f. ólì òkpòsò shóó ólí ómòhè.
the woman PRP.awaken the man
'The woman has awakened the man.'

g. òjè shié ólí úì ó vbí óràn.
Oje coil the rope CL LOC pole
'Oje wound / coiled the rope onto the pole.

h. òjè shié ólí úì ré.
Oje PRP.coil the rope VN
'Oje has wound / coiled up the rope.'

i. ólí úkpúì shié ré'.
the rope PRP.coil VN
'The rope has gotten wound / coiled up.'

The next example of the venitive-only verb class is intransitive. It reveals that particle use is not limited to literal usage. The figurative expression in (9b) expresses unidirectional change that culminates in the deictic sphere of the speaker. As with preceding verbs, the force of change is attributed to the entity of change, which is the referent of the grammatical subject.

(9) a. ólí úkpùn ránrán-ì.
the cloth PRP.soak-F
'The cloth is soaked.'

b. òjè ránrán ré'.
Oje soak VN
'Oje has put on weight.'

The final example in this section reveals the semantic interaction of verb, direct object, and venitive particle. The verb *buu* 'approach' in its neutral form permits as direct object personal pronouns for second and third person as well as lexical nouns (10a). When its direct object articulates as a first person pronoun (10b), venitive particle *re* is obligatory; otherwise venitive is prohibited. Pronoun choice overtly signals that change advances to the speaker's sphere of existence. It is noteworthy that second person pronouns, assumed most often to be another character of the deictic center, do not induce a similar requirement on venitive occurrence.

(10) a. ólí ómò búú ói / é / ólì òkpòsò.
the baby PRP.approach her you the woman
'The baby has approached her / you / the woman.'

b. ólí ómò búú mè ré.
the baby PRP.approach me VN
'The baby has approached me.' (here where I am)

c. *ólí ómò búú é ré.
the baby PRP.approach you VN
'The baby has approached you.'

2.2 *Particle with unidirectional feature and external force* Another class of verbs accepts only an andative particle. Most members are transitive with an intransitive option. Verbs in this class are underspecified for deictic determination of a unidirectional feature of change. The force driving change is attributed to an event participant other than the entity of change, whose grammatical expression is transitive direct object or intransitive subject. One of the clearest examples of this class is the verb *gbaan* 'wind.' It appears in its neutral, underspecified form in (11a). In its non-neutral form with andative *a* (11b-c), *gbaan* and its unidirectional nature articulate change away from deictic center to space that is removed from the speaker's sphere of existence. The force for change resides not in the entity of change, as happened in venitive constructions; rather, it rests with an event participant that does not undergo change, i.e. the subject expressed overtly in transitives or covertly in intransitives.

(11) a. òjè ò ó gbààn òú.
Oje SC C wind thread
'Oje is winding thread.'
b. òjè gbáán òú á.
Oje PRP.wind thread AN
'Oje has unwound the thread.'
c. ólì òú gbáán à.
the thread PAP.wind AN
'The thread got unwound.'

A host of other verbs in our corpus manifest a similar acceptance of the andative in transitive and intransitive constructions. They have an inherent sense that in construction with particle *a* pairs up as do 'close' to 'open,' 'beat' to 'break up,' 'decline' to 'vanish,' 'decrease' to 'extinguish,' 'burn' to 'burn up,' 'cut' to 'cut off,' 'break' to 'break apart,' 'scrape' to 'scrape off,' and 'wash' to 'wash off,' as shown in (12-21). In all examples the force for change is attributed not to the entity of change but to another event participant.

(12) a. òjè khúyé ólí úkhùèdè.
Oje PRP.close the door
'Oje has closed the door.'
b. òjè khúyé ólí úkhùèdè á.
Oje PRP.close the door AN
'Oje has opened up the door.
c. ólí úkhùèdè khúyé-ì.
the door PRP.door-F
'The door is closed.'
d. ólí úkhùèdè khúyé á'.
the door PRP.close AN
'The door has gotten opened up.'

(13) a. ólí ómòhè ò ó gbè ólí óvbèkhàn.
the man SC C hit the youth
'The man is beating /hitting the youth.'
b. ólí ómòhè gbé ólì ìbè.
the man PRP.hit the drum
'The man has beaten / played the drum.'
c. ólí ómòhè gbé ólí ákhè á.
the man PRP.hit the pot AN
'The man has broken the pot.'
d. ólí ákhè gbé á'.
the pot PRP.hit AN
'The pot has gotten broken.'

(14) a. ìbọ̀bọ̀dí fú ọ́ vbí ímè.
cassava PAP.decline CL LOC farm
'Cassava has decayed on the farm.'
b. ọ́lì ìbọ̀bọ̀dí fú à.
the cassava PAP.decline AN
'The cassava vanished.'

(15) a. ọ́lì òkpòsò fú-nọ́ ùrùkpà á.
the woman PRP.decrease-DS lantern AN
'The woman has extinguished / quenched the lantern.'
b. ọ́lì ùrùkpà fú-nọ́ á'.
the lantern PRP.decrease AN
'The lantern has gotten extinguished / quenched.'

(16) a. ọ́lì òkpòsò ò ọ́ tòò ọ́lì ògò.
the woman SC C burn the bush
'The woman is burning the bush.'
b. ọ́lí ọ́mọ̀hè tóó ọ́lì ògò á.
the man PRP.burn the bush AN
'The man has burned up the bush.'
c. ọ́lì ògò ò ọ́ tòó.
the bush SC C burn
'The bush is burning.'
d. ọ́lì ògò tóó á'.
the bush PRP.burn AN
'The bush has burned up.'

(17) a. ọ́lí ọ́mọ̀hè híán ọ́lí úì.
the man PRP.cut the rope
'The man has cut the rope.'
b. ọ́lí ọ́mọ̀hè híán ọ́lí úì á.
the man PRP.cut the rope AN
'The man has cut the rope off.'

(18) a. òjè gúóghó ọ́lí úkpóràn.
Oje PRP.break the stick
'Oje has broken the stick.'
b. òjè gúóghó ọ́lí úkpóràn á.
Oje PRP.break the stick AN
'Oje has broken the stick into disjointed pieces.'
c. ọ́lí úkpóràn gúóghó-ì.
the stick PRP.break-F
'The stick is broken.'
d. ọ́lí úkpóràn gúóghó á'.
the stick PRP.break AN
'The stick has broken into disjointed pieces.'

(19) a. ọ́lì òkpòsò ò ọ́ ànmè ọ́í étò.
the woman SC C scrape his hair
'The woman is scraping his hair.'
b. ọ́lì òkpòsò ánmé ọ́í étò á.
the woman PRP.scrape his hair AN
'The woman has scraped his hair off.'

(20) a. ólí óvbékhán kpé' ólì òkpàn.
the youth PAP.wash the gourd
'The youth washed the gourd.'
b. ólí óvbékhán kpé' ólì òkpàn á.
the youth PAP.wash the gourd AN
'The youth washed off the gourd.'

(21) a. ólì òkpòsò hóó ólí úkpùn.
the woman PRP.wash the cloth
'The woman has washed the cloth.'
b. ólì òkpòsò hóó ólí úkpùn á.
the woman PRP.wash the cloth AN
'The woman has washed off the cloth.'

The final andative-only verb is intransitive *aa* 'decompose.' It conveys unidirectional change that culminates in a sphere of existence that excludes the speaker. As with the preceding verbs, the force for change resides outside the entity of change.

(22) a. ólí éánmí ló àà ó vbí ògò.
the meat PRED decompose CL LOC bush
'The meat will decompose in the bush.
b. ólí éànmì áá á'.
the meat PRP.decompose AN
'The meat has rotted away.'

2.3 *Particles with bidirectional feature and internal/external force* Emai verbs in a third class allow a venitive or andative particle. Some members are exclusively intransitive; others appear as transitive or intransitive. Verbs in this class are underspecified for a bidirectional feature of change. They permit change toward or away from the speaker's sphere of being at the deictic center.

The force for change differentially correlates with particles *re* and *a*. For *re*, force is assigned to the entity of change, which is encoded as grammatical subject. One of the clearest examples of this class is the verb *mee* 'move to and fro.' It appears in its neutral, underspecified form in (23a), where its bidirectional nature is evident in the verb gloss. In construction with venitive *re* (23b-c), it specifies directional change toward the deictic center, the speaker's sphere of reality.

(23) a. òjè ò ó mèè égbè.
Oje SC C move.to.and.fro body
'Oje is swaggering about.'
b. òèèn méé ré'.
sun PRP.move.to.and.fro VN
'The sun has emerged / came out.'
c. ólì ùdúkpù méé ré'.
the coconut PRP.move.to.and.fro VN
'The coconut has emerged / come to the top (of the water).'

For andative *a*, illustration again employs the verb *mee* 'move to and fro.' It combines with particle *a* to specify change directed away from the deictic center and the speaker's sphere of being. Force again is not attributed to the entity of change. Rather, it is expressed by nominal *òvòn* 'sun,' which is marked as oblique by locative *vbi* (24c).

(24) a. òjè ò ó mèè égbè.
Oje SC C move.to.and.fro body
'Oje is swaggering about.'
b. ólí úkpùn méé á'.
the cloth PRP.move.to.and.fro AN
'The cloth has faded.'

c. ólí úkpún ló mè á vbì òvòn.
the cloth PRED move.to.and.fro AN LOC sun
'The cloth will fade in the sun.'

Another verb that allows venitive or andative does so in expressions of fictive, non-physical motion rather than physical motion. The verb *ee* 'be anxious,' as a neutral form in a simple transitive or intransitive construction, conveys simple anxiety as a dynamic mental condition (25a-b). When it combines with *re* or *a*, verb *ee* articulates a mental condition associated with either remembering or forgetting information. In construction with venitive *re* (25c-d), the entity of change is expressed by the grammatical direct object, as is the force of change. The direction of change is toward the sphere of existence that includes the speaker as well as the referent of the grammatical subject. In construction with andative *a* (25e-f), the entity of change is again coded by direct object but the force for change resides outside that entity. The direction of change is away from the sphere of being for the speaker and grammatical subject referent.

(25) a. ólí ómòhè ò ó èè óvbí ói.
the man SC C be.anxious offspring his
'The man is anxious about his child.'

b. ólí ómòhè ò ó èé.
the man SC C be.anxious
'The man is anxious.'

c. ólí ómòhè éé ólí úkpùn ré.
the man PRP.be.anxious the cloth VN
'The man has remembered the cloth.'

d. ólí ómòhè éé ré'.
the man PRP.be.anxious VN
'The man has remembered.'

e. ólí ómòhè éé ólí úkpùn á.
the man PRP.be.anxious the cloth AN
'The man has forgoten the cloth.'

f. ólí ómòhè éé á'.
the man PRP.be.anxious AN
'The man has forgotten.'

2.4 *Particles with unidirectional features and internal/external force* The final two verbs we consider participate in constructions whose equivalents are English 'leave.' Member verbs are *koo* 'count' (26a) and *fì* 'project' (27a-b). They partner respectively with venitive *re* and andative *a*. Initially, one would expect that each verb-particle frame would behave in a manner consistent with previously analyzed venitive and andative constructions. Unfortunately they do not. Neither of these constructions manifests an entity of change with respect to directed motion, *émàè* in (26b), or change of state, *ólì òkpòsò* in (27b). While events coded by these verb-particle pairs are transitive and unidirectional, they do not reflect overt, physical force. Instead, they reveal the exertion of mental force or the intentional non-exertion of physical force on a potential entity of change.

This cultural complication necessitates an analysis with detail drawn from the respective clauses. In the venitive *re* construction, *émàè* 'food' advances into the sphere of existence of the speaker. A presumption of this construction, however, is that *émàè* would have been consumed and thus would not have become available to the speaker. Due to the mental force of the subject referent, *émàè* was not completely consumed. It thus became available to the speaker. In the andative *a* structure, *ólì òkpòsò* 'the woman' withdraws from the speaker's sphere of existence. A presumption of this construction is that *ólì òkpòsò* would have remained in her role as wife and therefore would have retained her position in the speaker's sphere of existence. Mental force of the subject referent leads in this instance to a situation where *ólì òkpòsò* is banished to her home household. She thus no longer remains in the speaker's sphere of existence. In both constructions it is not change in the physical world that determines particle use; rather, it is change as applied to the relation between the world of cultural presumption and the world of physical experience. Deictic direction is assessed against an entity of potential change as it straddles the relation between a presumed world and the physical world.

(26) a. òjè ò ó kòò élí émà.
Oje SC C count the yam
'Oje is counting the yam.'
b. ólí ómòhè kóó émàè ré.
the man PRP.count food VN
'The man has left food behind.'
c. ólí émàè kóó ré'.
the food count VN
'The food has remained behind.'

(27) a. òjè fĩ ètèkùn.
Oje PRP.project arrow
'Oje has shot an arrow.'
b. òjè fĩ úkpóràn.
Oje PRP.project stick
'Oje has thrown a stick.'
c. ólí ómòhè fĩ ólì òkpòsò á.
the man PRP.project the woman AN
'The man has left behind / divorced the woman.'
d. *ólí ómòhè fĩ ólì òkpòsò.
the man PRP.project the woman
'The man has left behind / divorced the woman.'

3.0 Discussion

In the preceding we have explored one clause level relation that is characteristic of Emai, an underdescribed Edoid language of southern Nigeria. The relation pertains to the speaker's sphere of existence. It is manifested by a verb in construction with a grammatical particle that expresses deictic directionality. There are two particles: venitive *re* and andative *a*. They convey change that identified the post-event position of an event participant as either toward or away from the speaker's sphere of being.

A major challenge to our understanding of deictic directional events pertains to verb-particle relations, in particular verb selection of one particle or the other. Verbs in our corpus co-occur with venitive and andative in a constrained fashion. Some select only venitive, others only andative, and a few either venitive or andative, although not both for a single predication. To characterize this relation we assumed that verbs select their particle or particles on the basis of semantic information. Hence we proposed that verbs were underspecified with respect to particle meaning while still being compatible with that meaning. We hypothesized that this meaning relevant to the deictic directional domain incorporated two components. One was an underspecified or covert meaning element in the verb that stipulated a directional feature. Its possible values with respect to a deictic center were unidirectional or bidirectional. If unidirectional, possible overt specification was either venitive or andative. If bidirectional, specification was not limited to one direction or the other, either venitive or andative would be compatible with a verb. Another element of verb meaning pertained to a covert element of force that could be attributed to an event participant. Force was construed as internal to the entity of change for venitive but external to the entity of change for andative.

An apparent complication with respect to this overall interpretation arose from two events whose coding equivalent was English 'leave.' Noteworthy about these events was that while they were transitive and unidirectional, they engaged force relative to presumed rather than actual conditions. For the *re* construction with 'leave,' the presumptive condition invoked change for an event entity that did not change in the actual condition. By not undergoing change, the event entity became available to the speaker's sphere of existence. For the *a* construction with 'leave,' the presumptive condition did not invoke change for an event entity that did change in the actual condition. By undergoing change, this event entity did not become available to the speaker's sphere of being.

While our database for existence change largely fell on predications of directed motion and change of state, other predication types appeared to be excluded. Nowhere to be found were copulas, stative verbs, verbs of possession (e.g. 'sell'), and verbs of incremental consumption (e.g. 'eat,' 'drink'). Clearly, further investigation of the range of predicates susceptible to overt marking of a change in existence state will be

required. There seems to be a definite need to clarify event types and conditions under which venitive and andative markers operate.

Orthographic conventions

Orthographic conventions for Emai are consistent with those in Schaefer and Egbokhare (1999) and Schaefer and Egbokhare (2007, 2017), where **o̲** represents a lax mid back vowel, **e̲** a lax mid front vowel, and **vb** a voiced bilabial approximant. With respect to tone, acute accent marks high, grave accent signals low, and acute accent followed by an apostrophe designates high downstep. Abbreviations for grammatical morphemes used throughout this paper include: AN=andative, APP=applicative, C=continuous, CL=change of location, DS=distributive, F=factative, LOC=locative, PAD=projected adherence, PAP=past perfective, PF=positive focus, PRED=predictive, PRP=present perfective, SC=subject concord, VN=venitive.

References

Anderson, Stephen R. & Edward Keenan. 1985. Deixis. In T. Shopen (ed.), *Language Typology and Syntactic Description III*. 259-308. New York: Cambridge University Press.

Creissels, Denis, G.J. Dimmendaal, Z. Frajzyngier & C. Konig. 2008. Africa as a Morphosyntactic Area. In Bernd Heine & Derek Nurse (eds.), *A Linguistic Geography of Africa*. 86-150. New York: Cambridge University Press.

Dixon, R.M.W. 1991. *A New Approach to English Grammar, on Semantic Principles*. Oxford: Clarendon Press.

Elugbe, Ben. 1989. *Comparative* Edoid*: Phonology and Lexicon*. Port Harcourt: University of Port Harcourt Press.

Fillmore, Charles J. 1997. *Lectures on Deixis*. Stanford: CSLI.

Fillmore, Charles J. 1966. Deictic categories in the semantics of 'come.' *Foundations of Language* 2: 219-227.

Guillaume, Antoine. 2016. Associated motion in South America: Typological and areal perspectives. *Linguistic Typology* 21.1: 81-177.

Hooper, Robin. 2002. Deixis and aspect: The Tokelaun directional particles *mai* and *atu*. *Studies in Language* 26.2: 283-313.

Hyman, Larry H. 2007. Niger-Congo verb extensions: Overview and discussion. In Doris Payne & Jaime Peña (eds.), *Selected Proceedings of the 37th Annual Conference on African Linguistics*. 149-163. Somerville, MA: Cascadilla Proceedings Project.

Koch, Harold. 1984. The category of associated motion in Kaytej. *Language in Central Australia* 1: 23-34.

Schadeberg, Thilo C. 2003. Derivation. In Derek Nurse & Gérard Philippson (eds.), *The Bantu Languages*. 71-89. London: Routledge.

Schaefer, Ronald P. & Francis O. Egbokhare. 1999. *Oral Tradition Narratives of the Emai People* Parts I and II. Hamburg: LIT Verlag.

Schaefer, Ronald P. & Francis O. Egbokhare. 2007. *A Dictionary of Emai: An Edoid Language of Nigeria*. Köln: Rüdiger Köppe Verlag.

Schaefer, Ronald P. & Francis O. Egbokhare. 2017. *A Grammar of Emai.* Mouton Grammar Library Series. Berlin: Mouton de Gruyter.

Talmy, Leonard. 1985. Lexicalization patterns: semantic structure in lexical forms. In T. Shopen (ed.), *Language Typology and Syntactic Description III*. 57-149. New York: Cambridge University Press.

Talmy, Leonard. 1991. Path to realization: A typology of event conflation. *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society* 17: 480-519.

Talmy, Leonard. 2000. *Toward a Cognitive Semantics: Volumes I and II*. Cambridge: MIT Press.

Williamson, Kay & Roger Blench. 2000. Niger Congo. In Bernd Heine & Derek Nurse (eds.), *African Languages: An Introduction*. 11-42. New York: Cambridge University Press.

Emphasis Spread in Saffarini Arabic

Reem Saleh & Chris Golston
California State University, Fresno

1 Introduction

Emphatics are pharyngealized consonants that occur in many semitic languages. Emphasis is a secondary articulation usually involving retraction of the back of the tongue root (RTR), which accompanies a primary articulation at another point in the vocal tract, usually coronal (Younes 1993). Emphasis is contrastive in Saffarini Arabic, the dialect we examine here, as the following show:

(1) plain

tiːn	'fig'	saːħɪl	'coast'
ðahab	'gold'	dam	'blood'

(2) emphatic

tˤiːn	'mud'	sˤɑːber	'patient'
ðˤɑrf	'envelop'	dˤefdˤɑʕ	'frog'

This paper examines the spread of emphasis left- and right-wards in Saffarini, which affects both consonants and vowels, albeit in different ways. Briefly, emphasis spreads left- and rightwards to low vowels but only leftwards to consonants. The details of the spreading are difficult to model with current understandings of featural spread and consonant harmony, whose shortcomings will be dealt with below.

2 The sound inventory of Saffarini Arabic

Saffarini is spoken in a small town in northern Palestine. Data were collected from three native speakers by the first author, herself a native speaker; although the speakers themselves have been exiled, they learned the language from grandparents who were raised there. The consonants and vowels of Saffarini appear in Tables 1 and 2. Note the contrastive pairs of consonants, either plain [t d k θ ð s z] or emphatic [tˤ dˤ kˤ θˤ ðˤ sˤ zˤ]. All of these are coronal obstruents, except for velar [k]. Phonetically, the emphatics probably have a retracted tongue root (RTR) or are uvularized, something we have yet to determine; we use 'emphatic' here as a cover term for whatever the exact phonetic distinction may be.

	LAB	DENT	ALV	PAL	VEL	UVU	PHAR	GLOT
Stop	b		t d tˤ dˤ	ʤ	k kˤ	q		ʔ
Nasal	m		n					
Trill			r					
Fric	f	θ ð θˤ ðˤ	s z sˤ zˤ	ʃ	x ɣ		ʕ ħ	h
Approx	w		l lˤ	j				

Table 1. Consonants of Saffarini Arabic

* We'd like to thank Arto Anttila for helpful discussion, especially for alerting us to McCarthy 1997, and Nariman Saddouq for help with Arabic. Special thanks to Haifa Samara, Yousef Saffarini, and Ali Saffarini for hours spent checking and double checking the data. This work would not have been possible without their tireless and patient help.

Proceeding of *WECOL 2019*

Front	Central	Back
i: i		u: u
e: e		ɔ: ɔ
	a: a	

Table 2. Vowels of Saffarini Arabic

The vowels of Saffarini come in long/short pairs, as shown above. Note that the low vowels are central [a aː]. Instances of back [ɑ ɑː] that we will see below are all derived from harmony with emphatics. [ɑ ɑː] do not occur underlyingly; they only occur in the context of emphatic consonants.

3 Emphasis spread

Emphasis spread in Arabic is not uncommon and has been found in various dialects including Cairene Egyptian (Broselow 1976), Qatari (Bukshaisha 1985), Abha Saudi (Younes 1991), Ṣanānī Yemeni (Watson 1999), and in certain Palestinian dialects (Herzallah 1990, Davis (1995). For an overview, see Watson 2002; an interesting parallel case we will not discuss is Tsilhqot'in, found in Hansson 2010. Here we look at how the emphatic feature of [tˤ dˤ ðˤ sˤ] spreads to vowels and consonants in Saffarini. Briefly, consonant-to-vowel (C-to-V) applies leftwards and isn't blocked by intervening consonants of any kind or vowels of any kind; it also applies rightwards but is blocked in that case by bimoraic palatal feet [iː] and [ijj]. Consonant-to-consonant (C-to-C) spread applies bidirectionally within roots but only leftwards within complex words and is not blocked by anything in either direction. We first explore the data, then the ramifications for theories of spread and harmony.

3.1 *C-to-V spread* Saffarini emphasis spreads from emphatic consonants to low vowels within the same word, changing [a] to [ɑ], regardless of distance from the emphatic. Recall that low vowels are underlyingly all central [a aː] and remain so in words that contain only plain (unemphatic) consonants. For this reason, none of the words below have low back vowels [ɑ ɑː] on the surface:

(3) without emphatic, [a]

talle	'hill'
darrab	'he trained'
kalaːb	'dogs'
θaʕaːleb	'foxes'
ðablaːn	'wilting'
salaːm	'peace'
zaraːfe	'giraffe'

Rightward spread is shown below, where the initial emphatic consonant causes all following low central [a] vowels to back to [ɑ], whether long or short:

(4) with emphatic, [ɑ]

tˤɑlle	'view'
zˤɑref	'plastic bag'
θˤɑlɑtˤ–tˤɑːʃ	'thirteen'
dˤɑːħje	'area, suburb'
ðˤɑrɑb	'he hit'
ðˤɑːlim	'dictator'
sˤɑːħi	'awake'

Leftward spread also occurs, again backing underlying central vowels from [a] to [ɑ]. This is clear if we compare roots that lack emphatics (5) to similar roots whose last consonant is an emphatic (6):

(5) without emphatic, [a]

xabba:z	‘baker’
naħħa:s	‘coppersmith’
sabba:k	‘plumber’
falla:ħ	‘farmer’
merta:ħ	‘relaxed’
mara:teb	‘levels’

(6) with emphatic, [ɑ]

xɑjjɑ:tˤ	‘tailor’
mɑʕɑ:redˤ	‘exhibits’
ħɑ:fiðˤ	‘protector’
fɑ:ħesˤ	‘tester’
tˤɑ:zˤɑ	‘fresh’
xɑllɑ:tˤ	‘mixer’

The result of all this left- and rightward spread is that [ɑ] and [a] are in strict complementary distribution: back [ɑ] only occurs in words with emphatics and central [a] only occurs in words without them. Non-low vowels [e ɔ i u] seem to be affected by emphatics too, but we have not investigated this with enough care to report on it here. We leave the qualities unchanged in our phonetic transcription, but this should not be taken to mean that the vowels surface unchanged.

Leftward spread from C-to-V is not blocked by any class of sounds in Saffarini and is completely productive and exceptionless as far as we can tell. It also occurs in polymorphemic words, where we can see it spread, e.g. from root to prefix. (7) shows words where the underlying form of the discontinuous negative *ma… ef* surfaces unchanged in words without emphatic consonants:

(7) without emphatic, [a]

ma–ʃa:f–at–eʃ	ma–ma:t–at–eʃ	ma–ðabħ–at–eʃ
NEG–see–3F–NEG	NEG–die–3F–NEG	NEG–kill–3F–NEG
‘she didn’t see’	‘she didn’t die’	‘she didn’t kill’
ma–saħb–at–eʃ	ma–ra:ħ–at–eʃ	
NEG–withdraw–3F–NEG	NEG–go–3F–NEG	
‘she didn’t withdraw’	‘she didn’t go’	

(See Mousa 2019:87ff. for syntactic analysis of this discontinuous negative in another rural Palestinian dialect.) The prefixal part of the negative surfaces as [mɑ] when the following verb root contains an emphatic:

(8) with emphatic, [ɑ]

mɑ–tˤɑr–eʃ	mɑ–dˤɑrɑb–eʃ	mɑ–sˤɑ:ħɑb–eʃ
NEG–fly–NEG	NEG–hit–NEG	NEG–make.friends–NEG
‘he didn’t fly’	‘he didn’t hit’	‘he didn’t make friends’
mɑ–ðˤɑll–eʃ	mɑ–zˤebtˤ–ɑt–eʃ	
NEG–remained–NEG	NEG–work.well–3F–NEG	
‘it didn’t remain’	‘it didn’t work well’	

Rightward C-to-V spread is also fully productive within words, as we see below with words that begin with an emphatic and have only [ɑ] never [a]:

(9) with emphatic, [ɑ]

tˤɑːwle	‘table’
ðˤɑːlem	‘dictator’
sˤɑqer	‘falcon’
tˤɑble	‘drum’
tˤɑːher	‘pure’

Rightward C-to-V spread also applies within morphologically complex words, as we can see with suffixes. Below we see FEMININE realized as [at] following roots without emphatics (10) and as [ɑt] following roots with emphatics (11):

(10) without emphatic

teʕb–at	‘got tired–F’
daːr–at	‘turned–F’
ðaːb–at	‘melted–F’
saːħ–at	‘melted–F’
kammal–at	‘completed–F’

(11) with emphatic

tˤɑːr–ɑt	‘flew–F’
rɑkdˤ–ɑt	‘ran–F’
ðˤɑll–ɑt	‘stayed–F’
sˤɑːr–ɑt	‘became–F’
kˤɑtˤtˤɑʕ–ɑt	‘cut–F’

Similarly, for the FEMININE PLURAL, realized as [aːt] after plain roots (12) and as [ɑːt] after roots with emphatic consonants (13).

(12) without emphatic

kanabaj–aːt	‘sofa–F.PL’
daːjr–aːt	‘going out–F.PL’
ðablan–aːt	‘weak–F.PL’
saːjħ–aːt	‘melted–F.PL’

(13) with emphatic

tˤɑːwl–ɑːt	‘table–F.PL’
dˤɑːjʕ–ɑːt	‘lost–F.PL’
ðˤɑːjl–ɑːt	‘stayed– F.PL’
sˤɑːjr–ɑːt	‘became– F.PL’

The last two examples in (13) show that rightward C-to-V spread is not blocked by a single palatal like [j], though it is in some dialects of Arabic: Younes (1993) cites dialects in which emphasis spread is blocked by [ʃ j w] and Davis (1995) discusses a dialect where rightward spread is blocked by /ʃ j i ʤ/. Palatal blocking is complicated in Saffarini, and we explore it next.

Emphasis spread is not blocked by singleton palatals, as we saw above, nor by geminate palatals, as we see below, where emphasis spreads from initial tˤ and sˤ to the final vowel, passing right through the geminate palatal glide:

(14) rightward spread not blocked by geminate [jj]

tˤɑjjɑr–ɑt–hɑ	‘flew–F–it’
sˤɑjjɑd	‘hunter’

Nor is such spread blocked by short [i].

Rightward spread is blocked by what we'll call a palatal foot, a two mora sequence [iː] or [ijj]. The blocking effect of bimoraic [iː] can be seen by comparing (15) and (16). In (15) [iː] lies to the left of the emphatic and of course fails to block its spread to the following low vowels changing /a/ to [ɑ]:

(15) [iː] preceding the emphatic

xiːtˤɑːn=hɑ	'her threads'
ħiːtˤɑːn=hɑ	'her wall'
sˤiːsˤɑːn=hɑ	'her checks'

In (16) on the other hand, long [iː] falls between the emphatic and the suffix low vowel and blocks the spread of emphasis, so that =ha surfaces with [a] rather than [ɑ]:

(16) [iː] following the emphatic

ʔenti ħɑtˤ–iːt=ha hnaak	'did you put it there?'
ʔenti tˤɑːbex–tiː=ha	'did you cook it?'
ʔenti mðˤɑjjeʕ–tiː=ha	'did you lose it?'
ʔenti tˤɑːleb–tiː=ha	'did you order it?'

Rightward C-to-V spread is also blocked across bimoraic [ijj], [i] followed by a geminate (moraic) palatal glide. We see this below where the suffix [–aːt] surfaces unchanged following [ijj]:

(17) rightward spread blocked by a [ijj]

tˤɑblijj–aːt	'round.table–F.PL'
tˤɑlɑbijj–aːt	'order–F.PL'
mɑtˤlijj–aːt	'painted–F.PL'
mɑsˤrijj–aːt	'Egyptian–F.PL'
mɑðˤɑllijj–aːt	'parachute–F.PL'

We attribute this to blocking by a palatal foot, since [ijj] has one μ for the short vowel and one for the geminate (Hayes 1989), and [iː] is bimoraic as well, distinguishing these cases from [j], [i], [ij] and [jj].

3.2 *C-to-C spread* Saffarini also has consonant harmony, but it is morphologically conditioned: it spreads bidirectionally within roots, but only leftwards within words—there is no C-to-C spread from prefix to root or from root to suffix.

Emphasis spreads *leftward* from [tˤ dˤ ðˤ sˤ] to [t d ð s], within roots and words, so that we never find a plain consonant that could be emphatic to the left of an emphatic consonant in the same word (*t....tˤ, *t....sˤ). Thus, we find words without emphatics (18), words with a single emphatic lacking any available target sounds (19), words with multiple emphatics (20):

(18) no emphatics

maːt	'died'
rattab	'arranged'
tesʕa	'nine'
saːb	'left'
sette	'six'
ðaːb–at	'melted–F'

(19) one emphatic

tˤɑːr	'fly'
sˤɑːr	'become'
ðˤɑrɑb	'hit'
sˤɑrɑf	'spend'
ħɑðˤ	'luck'
bɑːsˤ	'bus'

(20) multiple emphatics

sˤɑkˤɑtˤ	‘failed’
zˤɑbɑtˤ	‘worked’
kˤɑsˤ	‘cut’
sˤɑfɑtˤ	‘box’
tˤɑkˤsˤ	‘weather’
sˤɑtˤer	‘line’

But there are no words like those in (21) or (22), all taken from Modern Standard Arabic:

(21) emphatic followed by unemphatic target (MSA)

tˤɑːsatun	‘small bowl’
ʕatˤɑsa	‘sneezed’
ʔatˤ–tˤɑqs	‘the weather’
ɣatˤɑsa	‘dived’
tˤɑmasa	‘deleted’

(22) emphatic preceded by unemphatic target (MSA)

saratˤɑːn	‘cancer’
satˤtˤɑr	‘made a line’
mastˤɑratun	‘ruler’
sa–ju–sajtˤiru	‘he will control’
ta–dˤtˤɑribu	‘she gets disturbed’

Words like those above occur in Saffarini but are pronounced differently. Cognates in Saffarini have C-to-C harmony (leftward within words, bidirectionally within roots):

(23) emphatic followed by emphatic target (Saffarini)

tˤɑːsˤe	‘small bowl’
ʕatˤɑsˤ	‘sneezed’
tˤɑqsˤ	‘the weather’
ɣatˤɑsˤ	‘dived’
tˤɑmasˤ	‘deleted’

(24) emphatic preceded by emphatic target (Saffarini)

sˤɑratˤɑːn	‘cancer’
sˤɑtˤtˤɑr	‘he made a line’
masˤtˤɑrɑ	‘ruler’
raħ–esˤajtˤer	‘he will control’
tˤetˤtˤereb	‘she gets disturbed’

Again, C-to-C spread *within roots* is bidirectional and exceptionless in Saffarini.

C-to-C spread is also found in morphologically complex words, but only leftwards: emphasis spreads from roots to prefixes as we see with NEGATIVE IMPERATIVE *t–*. If *t–* is added to a stem without an emphatic, no change occurs (25) if it is added to a stem with an emphatic, it becomes tˤ– (26):

(25) without emphatic

t–zaʕʕel–haːʃ	‘don’t make her sad’
t–eħkiːl–haːʃ	‘don’t tell her’
t–laːʕeb–haːʃ	‘don’t play with her’
te–staʕmel–haːʃ	‘don’t use it–F’
t–ren–eʃ	‘don’t ring’

(26) with emphatic

tˤ–sˤɑʕʕeb–hɑːʃ	‘don’t make it–F hard’
tˤ–sˤɑɣɣer–hɑːʃ	‘don’t make it–F small’
tˤ–xɑrbetˤ–hɑːʃ	‘don’t scramble it–F’
tˤ–ħutˤ–hɑːʃ	‘don’t put it–F’
tˤ–xɑjjetˤ–hɑːʃ	‘don’t stitch it–F

In the following words we see C-to-C spread across an epenthetic [e]. (27) shows epenthetic [e] in a prefix, added to avoid an unsyllabifiable sequence of three consonants word-initially: tħkiːʃ → teħkiːʃ.

(27) without emphatic

te–lʕab–eʃ	‘don’t play’
te–msaħ–eʃ	‘don’t erase’
te–ħkiːʃ	‘don’t talk’
te–sħab–eʃ	‘don’t pull’
te–kteb–eʃ	‘don’t write’

Below we see the effect of leftward C-to-C spread on the initial [t], across epenthetic [e]:

(28) with emphatic

tˤe–tˤbɑʕ–eʃ	‘don’t print’
tˤe–tˤlɑʕ–eʃ	‘don’t go out’
tˤe–sˤref–eʃ	‘don’t spend’
tˤe–ðˤlem–eʃ	‘don’t be unfair’
tˤe–ɣlɑtˤ–eʃ	‘don’t make mistakes’

Another case with C-to-C leftward harmony involves the numbers in the teens, where –tˤɑːʃ ‘teen’ spreads emphasis to consonants in the root. Compare *five* to *nine* without (29) and with (30) a following emphatic.

(29) without emphatic

xamse	‘five’
sette	‘six’
sabʕa	‘seven’
θamanje	‘eight’
tesʕa	‘nine’

(30) with emphatic

xɑmesˤ–tˤɑːʃ	‘fif–teen’
sˤetˤ–tˤɑːʃ	‘six–teen’
sˤɑbɑʕ–tˤɑːʃ	‘seven–teen’
θˤɑmɑn–tˤɑːʃ	‘eight–teen’
tˤesˤeʕ–tˤɑːʃ	‘nine–teen’

The prefix *sta–* shows the effects of leftward C-to-C spread when there is an emphatic in the following stem, as is clear comparing (31) with (32):

(31) without emphatic

sta–ʕmal	‘he used’
sta–bʃar	‘he became optimistic’
sta–qaːl	‘he resigned’
sta–slam	‘he gave up’
sta–ʕmar	‘he occupied’
sta–fsar	‘he asked’

(32) with emphatic

sˤtˤɑ–ðˤrɑf–ɑt	‘she pretended to be funny’
sˤtˤɑ–kˤsˤɑdˤ–ɑt	‘she meant’
sˤtˤɑ–tˤʕɑm–ɑt	‘she tasted’
sˤtˤɑ–ʕbɑtˤ–ɑt	‘she pretended to be stupid’
sˤtˤɑ–sˤʕɑb–ɑt	‘she found it hard’
sˤtˤɑ–tˤwɑl–ɑt=hɑ	‘she found it–F long’

C-to-C spread is *not* found rightwards within complex words, as we see in below where emphasis fails to spread from emphatics in the root to the [t] in FEMININE [ɑt] (33) or to the [t] in FEMININE PLURAL [ɑ:t] (34):

(33)

tˤɑ:r–ɑt	‘flew–F’
sˤɑrrɑx–ɑt	‘shouted–F’
ðˤɑrb–ɑt	‘hit–F’

(34)

tˤɑbl–ɑ:t	‘drum–F.PL’
sˤɑ:br–ɑ:t	‘patient–F.PL’
ðˤɑ:jʕ–ɑ:t	‘lost–F.PL’

The fact that emphasis spreads rightward from C-to-V here (changing the suffix from –at to –ɑt, but not from C-to-C shows that C-to-V and C-to-C spread are in some sense distinct in Saffarini.

A special case of C-to-C emphasis spread in Saffarini involves spread across words, where it is impossible (as far as we can tell) to tell the directionality of spread. It involves emphasis spread within an intonational phrase, e.g., between a verb and its subject. When the first word has an emphatic but the second does not, a word-final suffix on the first word surfaces unchanged (35); and when the second word has an emphatic but the first word does not, the suffix again surfaces unchanged (36):

(35) with a preceding but without a following emphatic

ðˤɑ:ʕ–ɑt	es–sa:ʕa	
got.lost–F	the–watch	‘The watch is lost.’
sˤɑffɑr–ɑt	l–emʕalme	
whistle–F	the–teacher	‘The teacher whistled.’
ɣɑtˤsˤ–ɑt	l–benet	
dived–F	the–girl	‘The girl dived.’
xɑbtˤ–ɑt	es–sajjara z–zalame	
hit–F	the–car the–man	‘The car hit the man.’
tˤɑħn–ɑt	el–maki:na leħbu:b	
ground–F	the–machine the–seeds	‘The machine ground the seeds.’
tˤɑbx–ɑt	emmi l–ʔakel	
cooked–F	my.mom the–food	‘The mom cooked the food.’

(36) without a preceding but with a following emphatic

ʔasrɑʕ–at	etˤ–tˤɑjjɑrɑ	
got.fast–F	the–airplane	‘The airplane got fast.’
ʔeʃtaɣl–at	etˤ–tˤɑ:ħu:ne	
worked–F	the–grinder	‘The grinder worked.'

wekˢʕ–at etˢ–tˢɑ:ħu:ne	
fell.down–F the–grinder	‘The grinder fell down.’
ʃa:f–at edˢ–dˢefdˢʕɑ l–walad	
saw–F the–frog the–boy	‘The frog saw the boy.’
nazzal–at etˢ–tˢɑbɑ:xɑ l–ʔakel	
served–F the–cook the–food	‘The cook served the food.’
ʃtar–at ʔuxti sa:ʕa	
buy–F sister watch	‘My sister bought a watch.’

But when the suffix has emphatics both before *and* after it, it may surface as emphatic (37). The process is optional, which we indicated as t(ˢ) below: not spreading does not sound terrible. Since it happens across the subject/predicate divide, we assume that this is spread (inter-emphatically) within the intonational phrase, as shown with both transitive and intransitive verbs below.

(37) with a preceding and a following emphatic

ðˢ:ʕ–ɑt(ˢ) edˢ–dˢefdˢʕɑ	
got.lost–F the–frog	‘The frog got lost.’
ɣɑtˢsˢ–t(ˢ) edˢdˢefdˢɑʕɑ	
dived–F the–frog	‘The frog dived.’
xɑbtˢ–t(ˢ) etˢ–tˢɑjjɑrɑ l–be:t	
hit–F the–airplane the–house	‘The airplane hit the house.’
tˢɑħn–t(ˢ) etˢ–tˢɑ:ħu:ne l–ħbu:b	
ground–F the–grinder the–seeds’	‘The grinder ground the seeds.’
tˢɑbx–t(ˢ) etˢ–tˢɑbɑ:xɑ l–ʔakel	
cook–F the–cook the–food	‘The cook cooked the food.’
sˢɑffɑrɑ–t(ˢ) esˢ–sˢɑffɑ:rɑ	
whistle–F the–whistle	‘The whistle whistled.’

The following cases of spread within a DP are more solid and spread is required. This is XP-internal spread, so we assume that the domain is the phonological phrase.

(38) with a preceding but without following emphatic

tˢɑbl–et Ali	‘Ali’s drum’
fursˢ–et leʕeb	‘playing break’
tˢɑri:kˢ–et ʕamal	‘way of making’
kˢetˢʕ–et kna:fe	‘piece of kunafa’

(39) with a preceding and a following emphatic

tˢɑbl–etˢ. tˢɑ:rek	‘Tarek’s drum’
fursˢ–etˢ sˢɑlɑ:	‘prayer break’
kˢetˢʕ–etˢ. sˢelsˢɑ:l	‘piece of clay’
tˢɑri:kˢ–etˢ tˢɑbex	‘way of cooking’

These involve spread within a phonological phrase corresponding to the NP that dominates the noun and its possessor. We leave full exploration of this external sandhi for future research.

4 Analysis

We analyze emphasis spread in Saffarini Arabic as a tug of war between markedness and faithfulness, roughly following McCarthy 1997's analysis of similar data elsewhere in Palestinian Arabic. To avoid the 'sour grapes' effect pointed out by Wilson (2003) and McCarthy (2011), we eschew feature *alignment* in favor of feature *extension*.

The effect of markedness is the tendency for difficult to perceive distinctions, like tongue root retraction, to be lengthened so that they are easier to perceive (Flemming 1995; Kaun 1995, 2004; Walker 2005). The effect of faithfulness is seen in the resistance of consonants to that spread.

Thus, we understand spread as giving the listener more time to hear a difficult phonetic distinction, here the advancement/retraction of the tongue, or uvularization, or whatever the exact dimension is in Saffarini. We assume constraints that *maximize duration* and are directional:

(40) EXTEND(RTR,L): Every RTR feature extends leftwards as far as possible.
EXTEND(RTR,R): Every RTR feature extends rightwards as far as possible.

EXTEND makes featural durations longer, so it is not subject to the sour grapes effect. We hope to pursue this in detail in future work.

EXTEND(RTR,L) is undominated in Saffarini at the level of the prosodic word and spreads RTR from any underlying sound leftwards, affecting all intervening sounds (as far as we can tell), among them [a] and coronal targets like [t s d]. As we have seen, it also extends past a prosodic word edge, if only minimally, within a phonological phrase (39) and optionally within an intonational phrase. For this reason, an analysis based on aligning RTR to the left edge of a word (e.g., Kirchner 1993, McCarthy 1997) will not work for Saffarini. Additionally, this avoids the 'sour grapes' problem raised by McCarthy (2003), which is shown there to be an important defect of spreading in terms of alignment. McCarthy points out for data like this that the failure of emphasis to spread all the way to the right edge of the word (to the final consonant) in the winner does not keep it from spreading as far as it can (to the final vowel). No language has the sour-grapes property of saying *If I can't spread all the way to the edge, I won't spread at all*, though that is exactly what an analysis using ALIGN predicts. EXTEND avoids this problem.

Unlike EXTEND(RTR,L), EXTEND(RTR,R) is dominated by IDENTC, which penalizes change to any consonant:

(41) IDENTC: Every surface consonant is identical to its underlying correspondent.

The constraint ranking is as follows,

(42) EXTEND(RTR,L) > IDENTC, IDENTPALFOOT > EXTEND(RTR,R)

modeling the facts that (i) leftward spread affects both C and V while rightward spread affects only V and (ii) leftward spread is not blocked by an intervening palatal foot while rightward spread is. What this ranking does NOT account for is that IDENTPALFOOT blocks spreading while IDENTC does not, which we return to below.

The tableau in (43) has the first candidate fully faithful, in fatal violation of ALIGN(RTR,L) because the emphasis has not spread leftwards to the prefix. The second and third candidates have emphasis spread all the way to the left and right, except that it short before the final [t] in the second candidate and changes it in the third. The spread to the left in the second candidate sanctions the change of the prefix consonants, but spread to the right in the third does not sanction change to the suffix [t] because ALIGN(RTR,R) is lower ranked than IDENTC-RTR:

(43) Emphasis spread left and right

sta–ðˤraf–at	EXTEND(RTR,L)	IDENTC	EXTEND(RTR,R)
sta–ðˤraf–at	*!		*
☞ sˤtˤa–ðˤraf–at		**	*
sˤtˤa–ðˤraf–atˤ		***!	

IDENTC keeps RTR from spreading to the suffix consonant, but not to the suffix vowel, so that –at surfaces as [at] rather than [atˤ]. This is different than what is generally reported for Palestinian dialects (Herzallah 1990; Davis 1995) where spread to consonants and to vowels is the same. Saffarini is unique in the published literature, as far as we know, in having rightward spread of RTR to vowels but not to consonants.

Rose and Walker (2011) note that 'there appears to be no unified analysis of non-local vowel-consonant harmony' of this sort. In addition to the align/extend kind of analysis proposed here, there are also theories of Agreement By Correspondence that need to be considered. We do this briefly here, because ABC approaches are mainly targeted at consonant harmony (see Hansson 2010). But Saffarini emphasis spread involves both C-to-C and C-to-V spread, so we will consider ABC here.

So we ask, is the emphasis in Arabic *copied*, like an agreement feature in syntax, as is assumed in ABC, or is it *spread*, by phonetically prolonging its duration, as we've assumed above? Current theory allows both, which is not a good thing, and we will cast some doubt on the propriety of ABC, without of course resolving the issue.

C-to-C harmony often seems to require Agreement–By–Correspondence, where a feature is copied from one consonant to a phonetically similar consonant at some distance. Hansson (2010:2) defines consonant harmony as:

> Any assimilatory effect of one consonant on another consonant, or assimilatory cooccurrence restriction holding between two consonants, where:
> a. the two consonants can be separated by a string of segmental material consisting of at the very least a vowel; and
> b. intervening segments, in particular vowels, are not audibly affected by the assimilating property.

which would take Saffarini C-to-C harmony out of the picture. Similarly, Rose & Walker (2011) treat emphasis spread as continuous harmony, where all intervening segments are affected by the assimilating property. While this may be true of much emphasis spread, it is not true in Saffarini, where, as we have seen, rightward spread targets vowels but leaves intervening coronals unaffected. We repeat some relevant cases below:

(44) Emphasis spread left and right
sˤaːħab–at=ha 'she became her friend'
sˤtˤa–tˤwal–at=ha 'she found it–F long'

We can see that emphasis spreads from C-to-V all the way to the final vowel [a], as expected. But note again that the suffix –at has plain [t] rather than emphatic [tˤ], showing that not all of the intervening segments are affected by the assimilating property. Might it then be ABC? We find this unlikely for two reasons, the first specific to Saffarini, the second more theoretical.

Within Saffarini, we could treat C-to-C spread in terms of ABC, since it spreads from coronals to coronals, fulfilling a central tenet of ABC:

> In the ABC model, the main factor that contributes to establishing the correspondence relations which serve as the 'vehicle' for long-distance consonant assimilations is taken to be the relative similarity of the two consonants (as well as their distance from one another). The more similar (and less distant) two consonants are, the stronger the drive towards agreement in any of the feature or features that differentiate them. (Hansson 2010:25)

C-to-C harmony in Saffarini involves very similar consonants (coronal obstruents plus [k]). But treating C-to-V spread in Saffarini as ABC is near impossible as the Cs and Vs share nothing phonetically (coronal obstruents vs. low vowels). The result is that ABC cannot account for (44) even though it would seem to be the type of data that copied features are used for: the RTR-featrue on the final [ɑ] is separated from its trigger (sˤ, tˤ) by a consonant [t] that *could in principle* bear that feature but does not, suggesting (in an ABC model) copy rather than spread.

A more abstract problem arises with what we'll call OCP grapes. Recall the sour-grapes argument that McCarthy used against alignment constraints (above). Something similar should occur, we think, in ABC. Suppose that some dialect of Arabic avoids hetero-morphemic, tier-adjacent sequences of RTR, an OCP restriction of the type we often find with tonal (Leben 1973) and other features. To see the problem, assume an underlying form with one emphatic consonant and leftward C-to-C spread:

t	e	s	–	t	a	–	sˤ	ʕ	e	b	–	e	ʃ
							|						
							RTR						

Without any OCP restrictions we'd get something like the following, which is an attested type of consonant-harmony

tˤ	e	sˤ	–	tˤ	ɑ	–	sˤ	ʕ	e	b	–	e	ʃ
|		|		|		|						
RTR		RTR		RTR			RTR						

But if these features are copied and pasted into the word, they are distinct instances of RTR and should be subject to the OCP. ABC + OCP predicts, as far as we can see, that every other RTR feature will be lost giving us this:

t	e	sˤ	–	t	ɑ	–	sˤ	ʕ	e	b	–	e	ʃ
|		|		|		|						
*		RTR		*			RTR						

Spreading/extending analyses do not make this prediction, since a single RTR feature would be spread, and the OCP would not be an issue.

The problem for ABC + OCP occurs whether the copied *f* is affiliated with the trigger, as we just saw, or with the *targets*. That is, in addition to the misprediction above, ABC + OCP seems to predict the following outcome in some language, where the underlying RTR feature goes unexpressed but every other copy to its left survives:

tˤ	e	s	–	tˤɑ	–	s	ʕ	e	b	–	e	ʃ
|		|		|		|					
RTR		*		RTR		*						

If tauto-morphemic features are exempt from the OCP, we'd get another unattested pattern, where every other morpheme dissimilates in the copied feature:

tˤ	e	sˤ	–	tɑ	–	sˤ	ʕ	e	b	–	e	ʃ
|		|		|		|					
RTR		RTR		*		RTR						

We call this an OCP grapes prediction: that ABC would underapply in cases where the OCP is relevant.

So Saffarini poses two problems for ABC. First, RTR harmony targets coronal obstruents (plus k) and low vowels which do not form a natural class and thus should not form correspondences. Second, ABC + OCP should produce skipping patterns within or across morphemes that don't appear to exist in any language. Again, adherents of ABC will probably not expect it to apply in C-to-V harmony, but Saffarini

has C-to-C harmony as well, which applies in a different way than C-to-V harmony, so it might bear on ABC after all.

Saffarini spread also poses conceptual problems for spreading analyses, including the analysis proposed here. If emphasis spread involves increasing the temporal duration of a feature to make it easier to perceive, it's hard to see how a targetable consonant can be skipped when extending the duration rightwards to a following vowel. And if [t] and [s] can be skipped when extending a feature rightwards, why can't [iː] and [ijj] be skipped? They block spread, as is expected in an autosegmental framework, but it's not clear why [t] and [s] would neither block spread nor be targets of it.

5 Conclusion

We have attempted an admittedly partial description of emphasis spread in Saffarrin Arabic. In some ways the data is tractable with current theory: left- and righward spread affect different kinds of sounds and are differentially blocked by palatals. But in other ways the data does not conform to theoretical expectations, as least as we understand them: the skipping of coronals in rightward C-to-V spread is difficult to model and the blocking of rightward spread by bimoraic palatal feet but not by singleton [i] and [j] is also unexpected.

References

Broselow, E. I. (1976). *The phonology of Egyptian Arabic*. Doctoral dissertation, University of Massachusetts, Amherst.

Bukshaisha, F. A. M. (1985). *An experimental phonetic study of some aspects of Qatari Arabic*. Doctoral dissertation, University of Edinburgh.

Davis, S. (1995). Emphasis spread in Arabic and Grounded Phonology. *Linguistic Inquiry* 26(3), 465-498.

Flemming, E. (1995). *Auditory features in phonology*. Doctoral dissertation, UCLA.

Hansson, G. (2010). *Consonant harmony: Long-distance interaction in phonology*. Number 145 in University of California Publications in Linguistics.

Hayes, B. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry 20*(2), 253-306.

Herzallah, R. (1990). *Aspects of Palestinian Arabic phonology: A non-linear approach.* Doctoral dissertation, Cornell University.

Kaun, A. R. (1995). *The typology of rounding harmony: An optimality theoretic approach*. Doctoral dissertation, UCLA.

Kaun, A. R. (2004). The typology of rounding harmony. In B. Hayes, R. Kirchner & D. Steriade (eds.), *Phonetically based phonology.* Cambridge University Press. 87-116

Kirchner, R. (1993). Turkish vowel harmony and disharmony: an Optimality Theoretic account. Paper presented at Rutgers Optimality Workshop I.

Leben, W. (1973). *Suprasegmental phonology*. Doctoral dissertation, MIT.

McCarthy, J. (1997). Process-specific constraints in Optimality Theory. *Linguistic Inquiry* 28(2), 231-251.

McCarthy, J. J. (2003). OT constraints are categorical. *Phonology 20*(1), 75-138.

McCarthy, J. J. (2011). Autosegmental spreading in optimality theory. *Tones and features (Clements memorial volume)*. 195-222.

Mousa, N. (2019). *Grammatical aspects of rural Palestinian Arabic*. Doctoral dissertation, Arizona State University.

Rose, S., & Walker, R. (2011). Harmony systems. *The handbook of phonological theory 2*, 240-290.

Walker, R. (2005). Weak triggers in vowel harmony. *Natural Language & Linguistic Theory* 23.4, 917–989.

Watson, J. (1999). The directionality of emphasis spread in Arabic. *Linguistic Inquiry* 30(2), 289-300.

Watson, J. (2002). *The phonology and morphology of Arabic.* Oxford University Press.

Wilson, C. (2003) Analyzing unbounded spreading with constraints: marks, targets, and derivations. Ms., UCLA.

Younes, M. (1991). Emphasis spread in three Arabic dialects. Ms., Cornell University.

Younes, M. (1993). Emphasis spread in two Arabic dialects. In M. Eid and C. Holes (eds.), *Perspectives on Arabic linguistics V. Amsterdam: Benjamins.* 119-145.

Superiority & *WH* Scope

Nicholas Sobin
The University of Texas at El Paso

1 Introduction

It is commonly assumed that Superiority (or Attract Closest) defined as a restriction on raising is the basis for the ungrammaticality of sentences such as '**What did who buy?*', and that D–linked *wh* phrases in sentences such as '*Which book did which person buy?*' constitute an exception. If this view were correct, then the obvious problem would be to establish the basis for the exceptional nature of D–linked *wh* phrases. However, this view may not be correct. I argue here that Superiority does not apply to WHQs as a restriction on *wh* raising, and that D–linked phrases are unexceptional. Instead, the grammaticality of the preceding sentences involves a restriction on assigning scope to simple (non–D–linked) *wh* phrases.

In offering this analysis, I will assume that speakers have access to the complete representation of a sentence for carrying out certain grammatical tasks. (See Chomsky 2007 for his 'homunculus' remarks, noting that some phenomena may require access to the whole structure.) Without such access, it is hard to see (i) how a speaker could respond to a question such as '*Who knows where we bought what?*', or (ii) how an EQ such as '*Mary said that Dracula kissed WHO?*' could be formed, or (iii) how to operate Principle C to correctly interpret *she*/*Mary* in a sentence such as '*She$_i$ thinks that Max said that he likes Mary$_j$*'. I will assume that *wh* scope assignment in WHQs also operates under such full access.

2 Wh scope: a first approximation

Consider the following scheme for scope–marking. In English, the interrogative complementizer C_{WH} must have a *wh* phrase in SPEC–C_{WH}, and C_{WH} assigns scope to this *wh* phrase. If the C_{WH} is root, then the *wh* phrase has wide scope—it is intended to elicit an answer. If the C_{WH} is embedded, then the *wh* phrase has narrow scope, and is not intended to elicit a response, though a voluntary response is possible. A *wh* phrase not surface–positioned in SPEC–C_{WH} may in many instances (see below) be construed as having the scope of any C_{WH} which c-commands it, but it must have either wide or narrow scope.

Consider the sentences in (1). (Cf. Lasnik & Saito 1992:118–121)

(1) a. [$_{CP}$ *what*$_{(CWH1)}$ do$_{CWH1}$ you think [$_{CP}$ (that) she likes _ ?]]
b. [$_{CP}$ C_{Decl} I wonder [$_{CP}$ *what*$_{(CWH1)}$ C_{WH1} she likes _]]
c. [$_{CP}$ *Who*$_{(CWH1)}$ C_{WH1} bought *what*$_{(CWH1)}$?]
d. [$_{CP}$ *who*$_{(CWH1)}$ C_{WH1} knows [$_{CP}$ *where*$_{(CWH2)}$ C_{WH2} we bought *what*$_{(CWH1/2)}$?]]
e. [$_{CP}$ *who*$_{(CWH1)}$ C_{WH1} knows [$_{CP}$ *what*$_{(CWH2)}$ C_{WH2} *who*$_{(CWH1)}$ bought _ ?]]
f. [$_{CP}$ *who*$_{(CWH1)}$ C_{WH1} wonders [$_{CP}$ *what*$_{(CWH2)}$ C_{WH2} you told *who*$_{(CWH1)}$ to read _]]

In (1a), *what* has wide scope because it is in SPEC–C of the root C_{WH}. In (1b), *what* has narrow scope because it is in SPEC–C of an embedded C_{WH}. In (1c), *who* and *what* both have wide scope, since root C_{WH} is the only C that can give them a scope. In (1d), since either of the two C_{WH}s can scope mark embedded *what*, we see the scope ambiguity for *what* noted by Baker (1970).

Sentences (1e–f) present a problem: why does embedded *who* in these sentences take only wide scope?

3 Superiority and D-linking

To explore this further, consider previous analyses of Superiority.

Proceedings of *WECOL 2019*

3.1 *Classic Superiority* Classic Superiority refers to a proposed principle that if α c-commands β, a rule cannot apply to α and β if α c-commands γ, and γ is nearer to α than β is (Kuno & Robinson 1972; Chomsky 1973). Since (2a) is possible, the WH Movement of *what* passing over *who* in (2b) violates this version of Superiority.

(2) a. Who C_{WH} _ bought what?
b. *What did_{CWH} who buy _ ?

And in (3b) involving D–linked phrases, the WH movement of *which book* passing over *which person* appears to be exceptional.

(3) a. Which person C_{WH} _ bought which book?
b. Which book did_{CWH} which person buy _ ?

However, sentences (1e–f) pose a problem for this raising view of Superiority because the embedded clauses should not be derivable, but they are. That (1e–f) are viable suggests that the problem with '**What did_{CWH} who buy _ ?*' is not illicit raising (Lasnik & Saito 1992), but scope–marking. In **What did_{CWH} who buy _ ?*, did_{CWH} cannot scope–mark *who*, but a higher C_{WH} could, as it does in (1e). Further, if embedded *who* in (1e–f) is replaced by *which person* as in (4a), the demand for wide scope disappears, as the possible responses in (4b) indicate.

(4) a. [$_{CP}$ $who_{(CWH1)}$ C_{WH1} knows [$_{CP}$ $what_{(CWH2)}$ C_{WH2} $which\ person_{(CWH1/CWH2)}$ bought _ ?]]
b. *Mary does*, OR *Mary knows what which person bought*

3.2 *The Lasnik & Saito (1992) analysis* The present work is not the first to observe that sentences like (1e–f) pose a problem for the raising view of Superiority. Lasnik & Saito (1992) note that sentences like '*Who wonders what who bought*' show that Superiority cannot be a condition on a raising transformation. They propose the following alternative analysis.

(5) **The L&S analysis of Superiority:**
a. A WH–phrase X in SPEC of CP is Op–disjoint (operator disjoint) from a WH–phrase Y if the assignment of the index of X to Y would result in the local A–bar binding of Y by X (in S–structure)
b. If two WH–phrases X and Y are Op–disjoint, then they cannot undergo absorption.

Thus, consider the rough S–structure of L&S's (75b) shown in (6).

(6) [$_{CP}$ who$_1$ [$_{IP}$ t_1 wonders [$_{CP}$ what$_2$ [$_{IP}$ who$_3$ bought t_2]]]] (=L&S's (75b))

If who_3 has LF–moved into the lower SPEC–CP, $what_2$ and who_3 are Op–disjoint since assignment of the index of *what* to *who* (= $what_2$ and who_2) will result in the local A–bar binding of who_3 by $what_2$. Absorption will fail, and the sentence will have no interpretation. However, if who_3 has LF–moved to SPEC of root C_{WH}, who_3 and who_1 will not be marked Op–disjoint. Assignment of a common index to who_1 and who_3 will not result in local A–bar binding of who_3. This is because of the intervening trace t_1. To be 'locally' A–bar bound, the A–bar element must be the most local binder. It would be t_1 and not who_1 that locally binds who_3. Thus absorption of who_1 and who_3 by root C_{WH} is possible, validating the LF–raising of who_3 and resulting in a wide–scope interpretation of who_3.

Interestingly, this analysis in effect says that scope marking considerations are key to Superiority and apply to the structure as a whole, in agreement with the proposal here. However, the L&S analysis raises a number of questions.

First, since it is carried out roughly in the GB framework, syntactic theory has advanced so that a number of the devices utilized in this analysis are unavailable. (LF mov't, D/S-Str, etc.) Second, the analysis involves a hypothetical assignment of shared indices between elements that are otherwise not related to each other; further, such an assignment has nothing to do with the derivation beyond accounting for this wide scope phenomenon. Third, the analysis leaves as a mystery the fact that embedded D–linked phrases are not subject to the Op–disjoint limitation; they don't require wide scope reading. Thus, to a

question like '*Who wonders what which person bought?*' you can respond with '*Mary does*'. Fourth, there is the puzzle of triple WHQs. Triple WHQs as in (7) show no Superiority Effect (Pesetsky 2000).

(7) a. Who C_{WH} _ gave what to who(m)?
b. What did_{CWH} who give _ to who(m)?
c. To whom did_{CWH} who give what _ ?/Who did who give what to _

What and *who* in (7b), and *(to) who(m)* and *what* in (7c) are Op–disjoint, and should be gibberish, but they're not.

In sum, the Superiority Effect fails to appear with (i) D–linked *wh* phrases, (ii) any *wh* phrase in a triple WHQ, and (iii) any *wh* phrase for which there is a higher C_{WH} to assign it scope. All of this suggests a different interpretation of the Superiority Effect.

First, there is no syntactic restriction on positioning a 'lower' *wh* phrase high in SPEC–C_{WH}. Second, by cross–linguistic comparison to languages like German which do not exhibit a strong Superiority Effect (Fanselow et al. 2008; Häussler et al. 2015), D–linked phrases behave normally in not observing Superiority. Third, it is simple (non–D–linked) *wh* phrases within the TP domain of a 'double' WHQ which are 'exceptional' (again by cross-linguistic comparison) in that they cannot be locally scope–marked. The Superiority Effect in WHQs is itself the exception to the norm.

So, let's now consider a scope marking approach to the Superiority Effect involving (i) multiple C_{WH}'s (cf. Pesetsky 2000) and (ii) a Top–Down approach to syntactic derivation.

4 A variety of C_{WH} complementizers

Pesetsky (2000) proposes a range of C_{WH} complementizers differing in their ability to license various quantities and qualities of *wh* phrases ($C_{0\text{-}spec}$, $C_{1\text{-}spec}$, and $C_{m\text{-}spec}$). Taking a cue from that analysis, I propose the variety of C_{WH} complementizers in (8).

(8) a. $C_{WH\text{–}A}$–may scope–mark up to two *wh* phrases
b. $C_{WH\text{–}B}$–may scope–mark more than two *wh* phrases
c. C_{EQ} –unselectively scope–marks any EQ–introduced interrogative phrase
(Sobin 2010)

The Superiority Effect in WHQs is not a restriction on raising, but a restriction on the ability of $C_{WH\text{–}A}$ to scope–mark a second simple (non–D–linked) *wh* phrase. Let's further consider this restriction on $C_{WH\text{–}A}$ in terms of 'Top–Down' syntax.

5 Top–down (T–D) derivation

Based on the work of Phillips (1996, 1997), Richards (1999), Chesi (2007, 2015), and Sobin (2020), T–D derivation proceeds by merging syntactic objects according to the conventions of Phillips (1997) in (9).

(9) MERGE RIGHT
New items must be introduced at the right edge of a structure.

BRANCH RIGHT
Where a terminal can be attached to more than one position in the existing structure with no effect on interpretation, the attachment that results in the more right-branching structure must be chosen.

The conventions in (9) result in the structural configurations in (10).

(10) a. Merge '{X, Y} and Z' => {X, {Y, Z}} (external merge of Z)
b. Merge '{X, {Y, Z}} and Q' => {X, {Y, {Z, Q}}} (external merge of Q)
c. Merge '{X, Y} and X' => {X, {Y, X}} (internal merge of X)

Further, Sobin (2020) proposes the MR–BR derivational cycle in (11).

(11)	a. select an active head	x	Active head	(AH)
	b. merge a selected non-head element	[x, y]	Merge1	(M1)
	c. merge a selected active head	[x [y, z]]	Merge2	(M2)
	d. value features on the merged elements	[x [$y_{(x)}$, $z_{(x, y)}$]	Valuation	(Val)

Following this cycle, '*Mary has been swimming*' is derived (ignoring much detail) as in (12).

(12) a. [C_{Decl} , Mary] AH/M1
b. [C_{Decl} [Mary, T/has]] M2/Val
c. [C_{Decl} [Mary [T/has, ~~Mary~~]]] AH/M1
d. [C_{Decl} [Mary [T/has [~~Mary~~, been]]]] M2/Val
e. [C_{Decl} [Mary [T/has [~~Mary~~ [been, ~~Mary~~]]]]] AH/M1
f. [C_{Decl} [Mary [T/has [~~Mary~~ [been [~~Mary~~, $Voice_{v[EA: Mary]}$]]]]] M2/Val
g. [C_{Decl} [Mary [T/has [~~Mary~~ [been [~~Mary~~ [$Voice_{v[EA: Mary]}$, √*swim*]]]]]] AH/M2/Val

SUBJs simply remerge downward to their positions of interpretation in SPEC–v. For sentence (1e), a partial structure under MR–BR is (13). (See Sobin (2020) for further details about this structure.)

(1e) '*Who knows what who bought?*'

(13) [who_A [$C_{WH–A}$ [~~who~~$_A$ [Tpres [~~who~~$_A$ [Voice*v*–√know [what [$C_{WH–A}$ [who_B [Tpast ...
[D_{UNI}] [D_{UNI}] [D] [D_{UNI}]

Note that embedded *what* is merged before the embedded who_B. Thus, under this T–D approach, Superiority cannot be viewed as a restriction on raising past a closer element. Further, since *wh* phrases are initially merged in their surface positions, neither overt displacement nor LF raising can be essential to or even involved with marking *wh* scope. Only the wholistic approach sketched earlier is operative here.

6 A scope–marking restriction

Though a deeper explanation is yet to be discovered, this T–D approach allows a first approximation of the restriction accounting for the Superiority data considered thus far. This restriction is stated in (14).

(14) **The scope marking restriction on $C_{WH–A}$ (descriptive):**
With a *wh* phrase in its SPEC, $C_{WH–A}$ cannot scope–mark a second simple (non–D–linked) 'SUBJ–related' *wh* phrase which it most immediately c-commands.

Following (14), the scope possibilities for a second *wh* phrase in (15) differ from those in (16).

(15) a. *What did$_{CWH–A}$ <u>who</u> buy _ ?
b. Who $C_{WH–A}$/*Mary knows [what $C_{WH–A}$ <u>who</u> bought?]
c. Who $C_{WH–A}$/*Mary knows [what $C_{WH–A}$ we persuaded <u>who</u> [C_{PRO} <u>PRO</u> to read?]]
d. Who $C_{WH–A}$/*Mary knows [when $C_{WH–A}$ <u>who</u> laughed/arrived?]
e. Who $C_{WH–A}$/*Mary knows [where $C_{WH–A}$ <u>who</u> was taken?]
f. Who $C_{WH–A}$/*Mary knows [where $C_{WH–A}$ <u>what</u> was put?]

(16) a. Who $C_{WH–A}$/Mary knows [where $C_{WH–A}$ we bought what]
b. Who $C_{WH–A}$/Mary knows [what $C_{WH–A}$ we bought where]

In (15a), *who* is blocked from scope marking by (14), rendering the sentence uninterpretable. In (15b), only a higher $C_{WH–A}$ is eligible to scope mark embedded *who*, resulting in mandatory wide scope for embedded *who*. Without the higher $C_{WH–A}$, the case where *Mary* is the root SUBJ, (15b), like (15a), becomes uninterpretable. In (15c), embedded *who* in is a control relation with PRO in SUBJ position, so the scope marking limitation blocks a narrow scope assignment. In (15d–f), embedded *who* is itself in the surface

SUBJ position. In each case, the embedded *wh* phrase is therefore blocked from local scope–marking by embedded $C_{WH–A}$.

In (16), in–situ *what* and *where* are not SUBJ–related and are available for local scope marking either by embedded $C_{WH–A}$, imparting narrow scope or by root $C_{WH–A}$, imparting wide scope – thus we get Baker's (1970) scope ambiguity.

In further support of the claim that it is the scope marking of the second simple *wh* phrase in the sentences of (15) that is restricted, note that substituting a D-linked phrase for the second simple *wh* phrase in the low position and not for the one in the high position results in improved acceptability, as in (17).

(17) a. What did which person buy?
b. *Which book did who buy?
c. Who $C_{WH–A}$/Mary knows [what $C_{WH–A}$ <u>which person</u> bought?]
d. *Who $C_{WH–A}$/Mary knows [which book $C_{WH–A}$ <u>who</u> bought?]

This is only a start. There are many further problems to be addressed, but we can advance some tentative conclusions.

7 Conclusions

In the majority of cases, *wh* phrases of all types appear able to violate 'classic' Superiority. Only simple (non–D–linked) *wh* phrases in 'double' WHQs appear to be subject to Superiority, and those *wh*s are 'salvaged' if there is a higher C_{WH} available to scope–mark them. D–linked *wh* phrases are unexceptional. Thus, the Subjacency Effect is <u>not</u> about an inappropriately raised *wh* phrase, or about raising at all. The Subjacency Effect <u>is</u> about a simple (non–D–linked) *wh* phrase not being able to get scope–marking from its local $C_{WH–A}$. This result is consistent with Pesetsky's general thesis that there is a variety of C_{WH} which differ in the quality and quantity of *wh* phrases which each can license. C_{EQ} contributes further to this thesis in being able to license yet another range of interrogative expressions, including partial *wh* phrases (e.g. *She saw a what?*), and non–*wh* interrogatives (e.g. *Mary/who kissed DRACULA?*). The core principles of scope marking appear to be only that (i) a *wh* phrase must be uniquely scope marked by one C_{WH}, and (ii) each *wh* phrase must have a scope. This is further substantiated by the basic facts of scope in Chinese discussed briefly in Appendix 1.

In linguistic theorizing, we have often assumed that negative untutored judgments are 'unlearnable', and are therefore indicative of underlying constraints/limitations on UG. Superiority in English is very puzzling in this regard, since the Superiority Effect is 'untutored', but nonetheless isn't manifest in a variety of other languages. Among the questions we are now left with are these: First, what is the basis for the existence of a scope–marking limitation on $C_{WH–A}$ (or whatever else the Superiority Effect may be due to), and second, why are simple (non–D–linked) *wh* phrases 'exceptional' in being subject to this limitation?

Appendix 1: WH scope in Chinese

Assume that a C_{WH} must uniquely scope mark at least one *wh* phrase. In English, the *wh* phrase appears in SPEC-C_{WH}, but that's not a universal necessity, as the Chinese data below (Huang 1982) indicate. Neither overt displacement nor LF movement is needed under the wholistic approach to scope marking proposed here. In sentence (18), there can be only one C_{WH} in the sentence, since there's only one *wh* phrase to scope mark.

(18) Zhangsan zhidao shei mai–le shu.
Z. knows who bought books
(i) 'Z. knows who bought books' OR (ii) 'Who does Z. know bought books?'

'Know' can but need not select an interrogative C, so the C_{WH} may be either embedded or root. If the C_{WH} is embedded , then we get meaning (i), and if root, the meaning (ii).

This is further supported by Huang's other data. With only one *wh* phrase present, there can be only one C_{WH}, so consider (19–20). The *wh* phrase in (19) has narrow scope only, because 'ask' requires an embedded interrogative C

(19) Zhangsan wen wo shei mai–le shu
Z. ask me who bought books
'Z. asked me who bought books'

The *wh* phrase in (20) has wide scope only, because 'believe' cannot select an interrogative C

(20) Zhangsan xianxin shei mai–le shu
Z. believe who bought books
'Who does Z. believe bought books?'

So, the same system that accounts for in–situ *wh* phrases in English, also does so for Chinese. Neither overt displacement nor LF raising are required for scope marking. Scope marking operates on the structure as a whole. As stated earlier, the core scope marking principles seem to be that (i) a *wh* phrase must be uniquely scope marked by one C_{WH}, and (ii) each *wh* phrase must have a scope.

References

Baker, C. L. 1970. Notes on the description of English questions: The role of an abstract question morpheme. *Foundations of Language* 6: 197–219.

Cheng, Lisa and L. Downing. 2016. Phasal syntax = cyclic phonology? *Syntax* 19: 156–191.

Chesi, Cristiano. 2007. An introduction to phase–based minimalist grammars: why move is top–down from left–to–right. *Studies in Linguistics* 1: 49–90.

Chesi, Cristiano. 2015. On directionality of phrase structure building. *Journal of Psycholinguistic Research* 44:65–89. (DOI 10.1007/s10936–014–9330–6)

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

Chomsky. Noam. 2007. The Biology of the Language Faculty: Its Perfection, Past and Future.Linguistics and Philosophy, Massachusetts Institute of Technology, MIT (published Feb. 10, 2012).

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

Fanselow, Gisbert & C. Fery. 2008. Missing superiority effects: Long movement in German (and other languages). In Elements of Germanic and Slavic grammars: A comparative view, ed. J. Witkos & G. Fanselow, 67–87. Frankfurt, Germany: Lang.

Häussler, Jana, M. Grant, G. Fanselow, & L Frazier. 2015. Superiority in English and German: Cross–language grammatical differences? *Syntax* 18: 235–265.

Huang, C. T. J. 1982. Move WH in a language without WH movement. *The Linguistic Review,* 1: *369–416.*

Kuno, Susumu & J. Robinson. 1972. Multiple wh-questions. Linguistic Inquiry 3:463–487.

Lasnik, Howard & M. Saito. 1992. Move α. MA: MIT Press.

Pesetsky, David. 1987. Wh-in-situ: Movement and Unselective Binding. E. Reuland and A. ter Meulen, eds., The Representation of (In)definiteness, 98-129. Cambridge, MA: MIT Press.

Pesetsky, David. 2000. Phrasal Movement and Its Kin. Cambridge, MA: MIT Press.

Phillips, Colin. 1996. *Order and Structure*. PhD dissertation, MIT.

Phillips, Colin. 1997. Merge right: An approach to constituency conflicts. In *WCCFL XV: Proceedings of the 15th West Coast Conference on Formal Linguistics*, 381–395. Stanford, CA: CSLI Publications.

Richards, Norvin. 1999. Dependency formation and directionality of tree construction. *MIT Working Papers in Linguistics* (Papers on Syntax and Morphology) 34: 67–105.

Sobin, Nicholas. 2010. Echo questions in the minimalist program. *Linguistic Inquiry* 41: 131–148.

Sobin, Nicholas. 2020. Directing syntactic traffic. *Syntax*. https://doi.org/10.1111/synt.12194