Proceedings of the Thirty-fourth
Western Conference on Linguistics

Volume twenty-eight
WECOL 2022

Edited by
Sho Akamine

Hosted by:
California State University, Fresno
November 12 – 13, 2022

Department of Linguistics
California State University, Fresno

Contents

Modeling Conversational Implicature: A Study of Bengali Verb *khaoa*

Soumya Sankar Ghosh
VIT Bhopal University

Due to the conversation's dynamic, emergent nature is frequently thought of as the most challenging topic of formal linguistics. Comparatively to classical sentential analysis, the level of complexity is likewise considerable. The difficulty in creating a formal account for conversational analysis is two - dimensional: Since the utterance is the smallest structural unit at the level of conversational analysis, it is necessary to develop the current theoretical framework in a way that it can account for the utterance. The linkages between the utterances in a discussion should also be explained using a system that should be established. Keeping this thing in mind, the current research work tries to develop a theoretical account of conversational implicature with a special focus on the Bengali Eat predicate, *khaoa*.

Keywords: Conversation Implicature, Intending Function, Contending Function, Volitionality, Non-volitionality.

1 Introduction

The context-based interpretation of any statement is more effective than the technique of feature-based modelling of the lexical expression, which fails in a number of ways. However, a theory that derives predicates' meaning entirely from contextual information is frequently regarded as an unattainable ideal. The solution to this conundrum is suggested in various different ways. Despite their shortcomings, these methods are always being altered in an effort to provide a useful framework to describe how meaning is understood. The focus of this article, which is a part of one of these programmatic traditions, is to develop a theoretical account of conversational implicature with a particular emphasis on the Bengali (henceforth Bangla) verb *khaoa*.

Eat and drink both refer to taking food and liquids into the mouth and digestive system. In both scenarios, an object moves into the mouth where it is tasted and felt. But the mechanisms that take place in the mouth for the two verbs are very different. Following Newman (2009), we distinguish the following descriptive categories that are a part of the "complex matrix" of English eat and drink, as illustrated below.

Proceedings of *WECOL 2022*

INTERNAL COMPLEXITY	*eat* and *drink* are dynamic, involving actions by a person which affect some other entity, but where the person also typically experiences a range of sensations	
SPATIAL-TEMPORAL PROFILE	*eat*	food is taken into mouth and moves through the digestive tract
	drink	liquid is taken into mouth and moves through the digestive tract
ACTIVE ZONE	*eat*	mouth, teeth, tongue, palate
	drink	mouth, tongue, palate
FORCE DYNAMICS	*eat*	forceful crushing and biting of food, controlled by person
	drink	no change to liquid in the mouth, controlled by person
TYPICAL SOCIAL/CULTURAL SIGNIFICANCE	eating and drinking are vital to humans, are usually enjoyed, and are the basis for many social occasions; both can be the means to inducing altered physical and psychological states (e.g., drunkenness)	

Fig 1: Components of the Central Meaning of *Eat* and *Drink* (Newman 2009)

As was said above eating and drinking have several characteristics. But, the verb translated as "eat" can signify a very wide range of things in many different languages (Wierzbicka 2009). Bangla is among them. A thorough review of the Bangla language will indicate that even though there are words like *khaoa, pan kOra, choSa* to indicate the concept of 'eating', 'drinking', and 'sucking' respectively but the language has conceptualized all three activities together with the verb *khaoa*, roughly 'ingest'. Consider the following:

(1)	Onjon	roj	bhat	kha-y
	Anjan.NOM	everyday	rice	eat.PRS.3
	'Anjan eats rice every day.'			
(2)	Onjon	prochur	jOl	kha-y
	Anjan.NOM	a lot	rice	eat.PRS.3
	'Anjan drinks water a lot.'			

For the current task, it is necessary to redefine the Bangla predicate *khaoa* in the manner described below, in accordance with Wierzbicka (2009).

(3) *Lexico-syntactic frame*: The verb specifies the action is executed by the doer within a time frame that affects some things or substances at the same time.
(4) *Prototypical motivational scenario*: "doing something to something with one's mouth"
(5) *Manner*: It specifies the action is iterative and involves the stages of taking the object (x) in mouth, keeping it for a while and then doing something to it with someone's mouth.

This framework appropriately explains other constructions like *sigareT khaoa* 'to smoke', *cumu khaoa* 'to kiss' etc. Furthermore, following Jackendoff (1995), meaning construing capacity of *khaoa* in conversational discourse will indicate the homomorphic relationship between syntax and semantics, as it is evident here:

(6) **Syntactic Template**: $[_S\ [_{NP}[_N\ ]]\ [_{VP}[_{NP}[_N\ ]]][_V]]]$
(7) **Semantic Template:** $[_{Event}\text{CAUSE}([_{Thing}\ ],[_{Event}\ \text{GO}([_{Thing}\ ],\ [_{Path}\ \text{TO}([_{Place}\ \text{IN}([_{Thing}\ ])])])])]$

The syntactic and semantic template of (1) as evident in (6) and (7) will be the same for (2), as the current

paper already argues that the verb *khaoa* in Bangla merged the *selectional restriction* that drink predicate normally calls for. As per the thesis of homomorphism, the algebra explaining semantic composition is similar to that of the algebra explaining syntactic composition; though, the schematic representations in (6) and (7) seem to be a mismatch primarily because of the reason that the contextual information encoded in a semantic template is more in comparison to its syntactic counterpart. This happens primarily because of the reason that the phenomenal world is encoded differently in syntax and semantics. With the incorporation of conversational implicature, the problem becomes more acute.

(8) Onjon roj ghuS kha-y
Anjan.NOM everyday bribe eat.PRS.3
'Anjan takes bribes every day.'

(9) Onjon roj bOka kha-y
Anjan.NOM everyday scold eat.PRS.3
'Anjan gets scolded every day.'

Like conventional implicature, conversational implicature is not fossilized. Being an implicature, (58-9) are amenable to the calculable feature. In Levinson's (1983) diction, "for every putative implicature it should be possible to construct an argument … showing how from the literal meaning or the sense of the utterance on the one hand, and the co-operative principle and the maxims on the other, it follows that an addressee would make the inference in question to preserve the assumption of co-operation". Under this situation, it is now the objective of this paper to explain the meaning construction process involved in the interpretation of the utterances like (8-9).

To accomplish the aforementioned objective, Section 2 aims to explore a theoretical framework that will allow language expressions to more effectively govern the influx of common sense knowledge. This explanation will be supplemented in Section 3 with a discussion of some pragmatic findings about the linguistic behaviour of *khaoa* to clarify how existing knowledge of pragmatics might offer some crucial hints about the formalization of the aforementioned issue. A conceptual foundation that is essential for offering a holistic formal explanation of conversation has been proposed in Section 4 as a final step.

2 Theoretical Background

A language expression, according to Karmakar and Kasturirangan (2011), is thought of as a cognitive fortitude made up of a contending function (= C_f) and an intending function (= I_f). Essentially, the intended function calls on the appropriate conceptual category. A conceptual category denotes the orderly description of connected knowledge systems (Laurence & Margolis, 1999). In the work of Ghosh and Karmakar (2014), a conceptual category is instead envisioned as a cognitive capacitor that stores all potential viewpoints on a reality (Merleau-Ponty, 1995; Millikan, 2004). Since linguistic communication is always context-dependent, they contend that desiring is insufficient on its own to form a discourse. Another cognitive function is required to place conceptual categories in that framework (Zilberman, 1988; Langacker, 2008). This act of relativization is referred to as contending. While arguing, a language expression's purpose is to select a specific point of view within a discourse environment. As an illustration, the act of intending "rose" invokes the category relating to it, which contains details on its structural aspects (such as shape, size, constituencies, etc.), while also indicating the functional aspects (like the symbol of love, friendship, peace etc.). These two instances show that the choice of these structural and functional elements will depend on the communicative circumstance.

As seen in the illustration below with a provision for an intermediary domain necessary for meaning transference, it has also been claimed that meaning building can best be thought of as the combination of intending and competing functions.

(10) $C_f \circ I_f = def$ {(x, z): for some y, (x,y) ∈ I_f & (y,z) ∈ C_f; where x ∈ Domain(I_f) & z ϵ Range(C_f)}

As per the definition given in (10), transference of meaning presupposes a hidden link y between I_f and C_f to compose the domain of intending to the range of contending. Establishing a link between the previously disconnected intending domain and contending range – through an intermediating hidden layer of meaning component (here, in this case, y) – is called convolution following a metaphor from mathematics. The tricks of exploring the way convoluted sense (= $C_f \circ I_f$) is formed lie with the discovery of intermediating sense (= y) which remains crucial in carrying out the task of meaning transference.

The most intriguing aspect of this seems to be the requirement for an intermediary value that both functions must provide for successful meaning transference when two functions combine to create a third meaning. In light of this comparable phenomenon, Goguen (2006) contends that shared underlying substrates of this kind are important because they enable the cognizer to foresee further traits that will be inherited as the composition takes shape. The frame of reference with which the composed-meaning space is interpreted is the provision for intermediating value in construing the underlying substrate.

After discussing these features, we can conclude that expression meanings are frequently products we build and then rebuild upon. They are not necessarily preset elements held in a context. According to Karmakar and Kasturirangan (2010), the ability of linguistic entities to construct meaning in a discourse is influenced by the way that the respective functions connected to an expression—i.e., intending and contending—are rooted in and placed in our subordinate cognitive domains. The conceptual path a cognizer takes, albeit instinctively, in order to access the communicative aim, is what we refer to as grounding and locating. In actuality, research on the conceptual path is an investigation into the organizational principles underlying conceptualization processes.

3 Text Interpretation

Through the analysis of the examples stated in (8) and (9) the statements we have stated so far will be further supported. Please reconsider the examples stated in (8-9), in (11-12):

(11)	Onjon	roj	ghuS	kha-y
	Anjan.NOM	everyday	bribe	eat.PRS.3
	‘Anjan takes bribes every day.’			

(12)	Onjon	roj	bOka	kha-y
	Anjan.NOM	everyday	scold	eat.PRS.3
	‘Anjan gets scolded every day.’			

As per the above discussion, the sense conveyed by *khaoa* in (11) and (12) is ‘to take’ and ‘to get’ respectively whereas in its literal sense, it carries the sense of ingest. Therefore, the task of this paper, in one sense, is to bridge the gap between *what is said* and *what is meant*. If we compare (11) with (1) we will notice that (1) will not implicate the sense of ‘Anjan takes rice’ in Bangla. As a result, it could be argued that the default interpretation that we often assign to the Bangla verb *kha* as is represented, often undergoes the process of meaning transference under the direct influence of the other lexical constituents of a sentence. More explicitly *what is meant* by (11) is due to the appearance of *ghuS* ‘bribe’ in place of *bhat* ‘rice’. The appearance of *bhat* in (1) will always license the sense of eating in contrast to the appearance of *ghuS* in (11) which licenses the sense of *taking*. This is not the only instance to show how the semantics of accompanying nominal forms in a sentence influence the inferential role of the verb. The interplay of the nominal and verbal component in (6) further suggests that the verb *khaoa* is forming the sense of *getting* rather than *taking* because of the nominal component *bOka* ‘scold’. Consider the following figure for a better understanding.

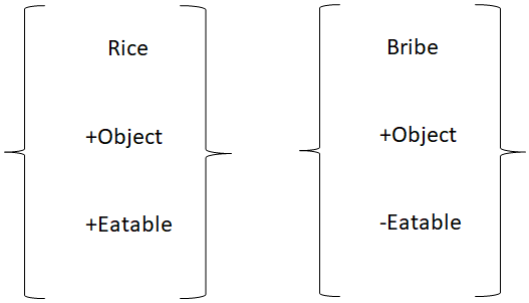


Fig 2: Categorization of the nominal counterpart of *khaoa* predicate

The question that arises from here is, the aberrant behaviour of *khaoa* depends only on its nominal counterpart or there is another factor that is working as a central force behind this? Further reconsideration of examples (11) and (12) triggers the fact that in (11) *khaoa* assigns an agentive-subject theta role and an experiencer-subject theta role in (12). This allows the mentioned verb to take the sense of either TAKE or GET. Thus the current work proposes that *khaoa* will take the sense of TAKE if the agent is volitional/intentional, and it will carry the sense of GET if the agent is non-volitional. With this understanding, the current work will focus on the conceptual mapping of *khaoa* in the next section.

4 Conceptual Framework of *khaoa*

While investing the interrelation between the senses of *eating* and *taking*, as discussed in the previous section, the paper found though *eating* necessarily contains a sense of *taking* as well, the opposite is not always necessarily true. Consider (13-14):

(13) (Anjan eats rice) implies (it is necessarily the case that (Anjan takes rice)).
(14) (Anjan takes rice) does not imply (it is necessarily the case that (Anjan eats rice)).

This results in an understanding that the senses of eat and take constitute a scale <eat, take> where the first member ensures the sense of the second member but the opposite does not hold true. Under this situation, the relation between 'said' and 'meant' could be explained following the theoretical framework outlined in the previous section if we can establish an intermediating meaning function like ***y*** (as defined in 10). This intermediating function will explain how a meant aspect of articulation is constructed and the logic of the scale we have discussed above. A little introspection will reveal the fact that both the verb forms of Bengali namely *khaoa* and *neoa* 'to take' are sharing a sense of consuming. This is also apparent in the following mental map drawn on the basis of the *Think Map visual thesaurus*:

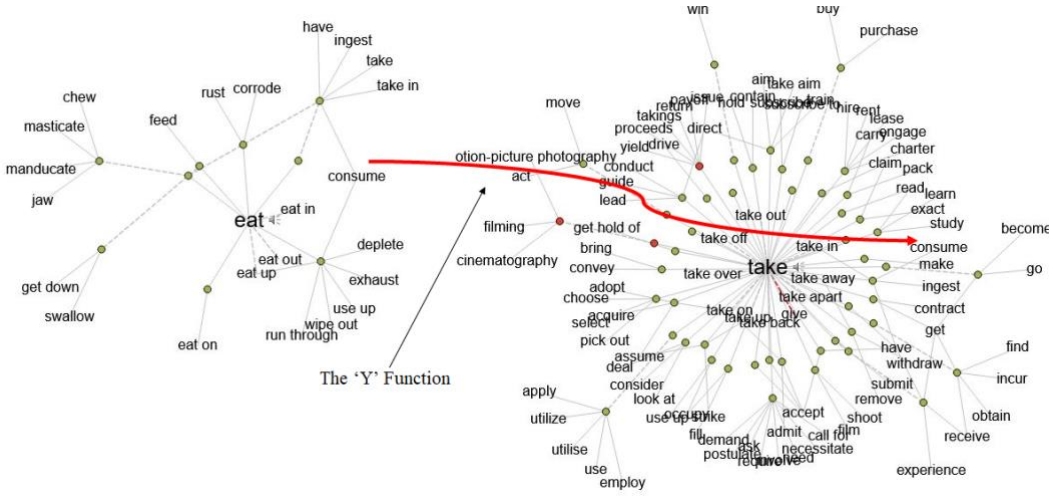


Fig 3: Visual representation of *eat* and *take*

Similar situation could be observed for the scale <eat, get> also. Consider the following:

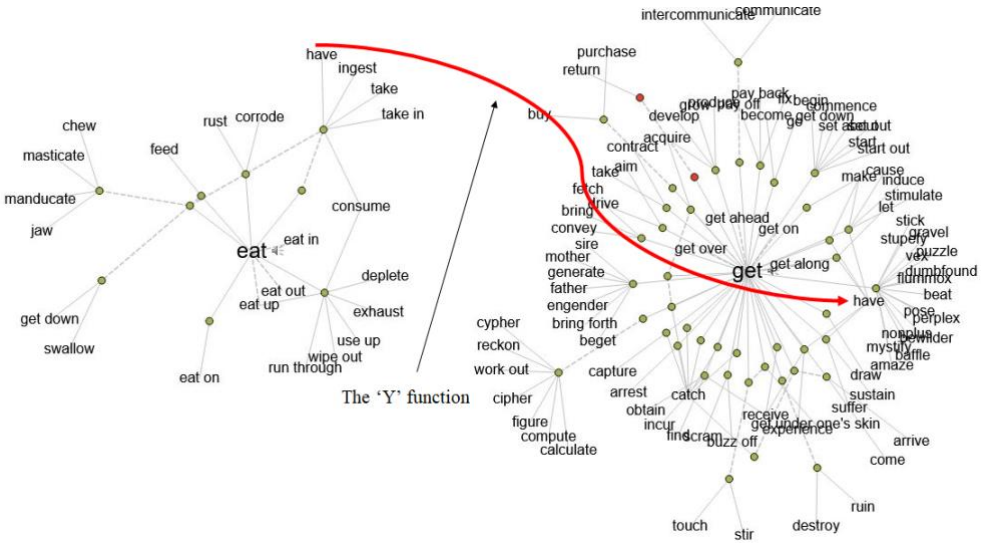


Fig 4: Visual representation of *eat* and *get*

However, as was noted, superimposing these two realms of knowledge activates a number of linkages that are still in their infancy in this articulation. If allowed, one of these networks of connections leads to a mathematical convolution that produces a third sensation of CONSUMPTION that may or may not be connected to either of the knowledge domains previously stated. In this regard, the current article takes into consideration the concept of convolution proposed by Ghosh and Karmakar (2014), where it is viewed as an integral that expresses the degree of overlap between one function I_f and another function C_f. Consequently, it "blends" one function with another. The formal representation, from Ghosh and Karmakar (2014), is given below:

(15) $(C_f \circ I_f)(D) =def \int C_f(\tau) I_f(D-\tau) d\tau$, where D is the domain that arises as a result of convolution and τ is the time-dependent variables

Evidently, the semantics of the term, which is time-independent, takes care of some parts of the background knowledge involved in expanding the richness of the information, while the pragmatics, which is time-dependent, takes care of some other aspects. Together, these two factors provide an explanation for how (11–12) came to be understood, or, to put it another way, they establish how D emerged.

5 Conclusion

The current work tries to provide a framework for Bangla ingest predicate *khaoa* while considering its default interpretation and its conceptual reading. This study analyzes how ideas are incorporated into texts. In addressing this matter, we have mostly focused on the *khaoa*. As we continued to extend our prior viewpoint on this matter, we suggested that certain mathematical concepts, such as convolution and the integral domain, might be used to successfully explain the problems with conceptual meaning. It is to note here that the present paper has not focused on the conventional metaphorical usages of *khaoa*, as represented in (16-17):

(16)	tumi	ki	chokh-er	matha	khe-echh-o
	you.NOM	Q PRT	eyes-GEN	head	eat.PRS-PRF-2

'Have you gone blind!'

(17) Onjon prem-e daga khe-echh-e
Anjan.NOM love-LOC rejection eat.PRS-PRF.3
'Anjan has been ditched in love.'

It is clear from these two examples that *khaoa* behaves very differently in *chokher matha khaoa* and *daga khaoa*. The fundamental reason for this is that, contrary to what we previously discussed, the linguistic items mentioned above do not create any blended space and are instead considered to be one linguistic item in Bangla, denoting, respectively, "act blindly" and "ditched in love."

References

Geach, P. 1962. *Reference and Generality.* New York: Cornell University Press.

Gerber, M, and J Chai. 2010. "Beyond nombank: A study of implicit arguments for nominal predicates." *In Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics.* Uppsala, Sweden: Association for Computational Linguistics. 1583-1592.

Ghosh, S, and S Karmakar. 2014. "Roles of Nominals in Construing Meaning at the Level of Discourse." *Proceedings of the Eleventh International Conference on Natural Language Processing.* Goa University, India: ACL Anthology. 76-81.

Givón, T. 2005. *Context as Other Minds: The Pragmatics of Sociality, Cognition, and Communication.* Amsterdam: John Benjamins.

Goguen, J. 2006. "Mathematical models of cognitive space and time." In *Reasoning and Cognition: Proceedings of the interdisciplinary conference on reasoning and cognition*, by D Andler, Y Ogawa, M Okada and S Watanabe, 125-128. Japan: Keio University Press.

Jackendoff, R. 1996. *Semantic Structure.* Massachusetts: MIT Press.

Karmakar, S, and R Kasturirangan. 2010. "Perspectivizing Space in Bāŋlā Discourse." *Proceedings of the 32nd Annual Conference of the Cognitive Science Society.* Austin, TX: Cognitive Science Society (CSS). 826-830.

Karmakar, S, and R Kasturirangan. 2011. "Structural and Structuring: Issues in Designing an Expression." *Jadavpur Journal of Philosophy* 21 (1): 87-102.

Langacker, R. 2008. *Cognitive Grammar: A Basic Introduction.* UK: Oxford University Press.

Laurence, S, and E Margolis. 1999. "Concepts and Cognitive Science." In *Concepts: Core Readings*, by S Laurence and E Margolis, 3-81. Cambridge: MA: MIT Press.

Levinson, S. 1983. *Pragmatics.* Cambridge: Cambridge University Press.

Merleau-Ponty, M. 1995. *Phenomenology of Perception.* New York: Routledge.

Millikan, R. 2004. *Varieties of Meaning.* Massachussetts: MIT Press.

Newman, J. 2009. "A cross-linguistic overview of 'eat' and 'drink'." In *The Linguistics of Eating and Drinking*, by J Newman, 1-26. Amsterdam: John Benjamin Publishing Company.

Pustejovsky, J. 1995. *The Generative Lexicon.* Cambridge: MA: MIT Press.

Wierzbicka, A. 2009. "All people eat and drink: does this mean that 'eat' and 'drink' are universal human concepts?" In *The Linguistics of Eating and Drinking*, by J Newman, 65-89. Amsterdam: John Benjamins Publishing Company.

Ziberman, D. 1988. *The Birth of Meaning in Hindu Thought.* Dordrecht, Boston: Reidel Publishing Company.

The Cognitive Phonetics of Velar Palatalization

Sayantan Mandal
Concordia Center for Cognitive Science & Linguistics

This paper investigates the nature of the phonetics-phonology interface – the relationship between phonetic substance and phonological primes – in light of a pattern of velar palatalization in Malayalam. Data from psycholinguistic investigations of the process are interpreted in the context of a modular framework with a transductive interface between phonological symbols and phonetic substance. It is argued that the modular, interface-driven account provides for a more economical, and biolinguistically coherent, account than substantive and grounded approaches.

Keywords: *phonology-phonetics interface, substance-free phonology, cognitive phonetics, universal grammar, phonological representations, form, substance.*

1 Introduction

My purpose here is to present a rather well-known paradigm of alternation in theoretical phonology, *velar palatalization*, and review what a case study from Malayalam can tell us about the nature of the phonology-phonetics interface. I discuss the results of two psycholinguistics studies (Mandal et al. forthcoming) of velar palatalization in Malayalam (K. P. Mohanan & Mohanan, 1984), and evaluate the possible interpretations the data receives, with particular focus on two contemporary frameworks – grounded/substantive phonology (D. B. Archangeli & Pulleyblank, 1994; Wilson, 2006) and substance-free phonology (Hale & Reiss, 2008; Volenec & Reiss, 2020). I argue that a modular framework where phonetic-substance plays a role, but without being directly accessible to phonology, provides for a more economic and biolinguistically plausible account over substantive/grounded varieties of phonology.

Palatalization processes encompass a few somewhat related types of alternations (Halle, 2005). Consider the alternations below.

(1) a.	electri[k] ~ electri[s]ity		Dorsal ~ Coronal Alternations [Eng]
	ry[k] "scream" ~ ry[ʧ]ɛć	"to scream"	Stop~Affricat Alternations. [Polish]
b.	kutti[k:]e ~ kutti[kʲ:]e	"child" (DAT.)	Secondary Articulation. [Malayalam]

Of the forms of alternation in (1) above, type-a illustrates a softening process resulting in a complete segmental change, while in type-b the dorsal consonant takes on a secondary palatal articulation. Palatalization has received significant amount of attention in the theoretical literature (Halle, 2005; Wilson, 2006) because it raises important questions regarding the nature of form and substance in phonological representations. For instance, the forms of alternations shown in (1) appear to share certain properties. For one, each change from the underlying representation (UR) to the surface (SR) involves a frontward shift in the place of articulation. Also, the forms above all appear to illustrate a change triggered in the context of a following non-back vowel. A cursory review of existing cross-linguistic studies (Halle, 2005; Keating & Lahiri, 1993; Maddieson & Disner, 1984) would, indeed, show a preponderance of phenomena where palatalization/softening is triggered in the context of non-back vowels. This typological asymmetry is typically treated in theory as reflecting some underlying grammatical principle of a non-trivial nature (Wilson, 2006). For instance, Halle (2005; 36) writes, *"What all types of ''velar softening'' have in common is that they are triggered by front vowels and that the preceding consonant changes from [DA*

Proceedings of *WECOL 2022*

Dorsal] to [DA Coronal][1]. Wilson (2006) proposes a set of *implicational laws of palatalization* – palatalization before a less front vowel (e.g. /e/) in a given language asymmetrically implies palatalization before more fronted vowels (e.g. /i/). Halle's (2005) proposal is driven by articulatory concerns: *"I now believe that for phonology, articulatory considerations are paramount and that acoustic aspects of speech play at best a subsidiary functional role*" (Halle, 2005; 31). Wilson's (2006: 954) proposal, on the other hand, makes extensive use of context-sensitive perceptual similarity and confusion matrices, and their effects on phonological processing and acquisition. Though Halle and Wilson sketch competing frameworks that contrast in significant ways, what they share in common are the assumptions that (a) velar palatalization processes tell us something non-trivial about the nature of phonological cognition (Halle, 2005), and (b) that said cognition is rich in phonetic substance (Wilson, 2006). My argument here will centre around the claim that both these frameworks suffer from overt symptoms of *substance-abuse* (Hale & Reiss, 2000) . Specifically, I argue that the cross-linguisic "generalization" that palatalization processes are always triggered in the context of non-back vowels is not objectively born out by the typology, but is rather a predictable consequence of the functionalist presumption that phonological alternations *must have* phonetic motivations. I discuss a well-known pattern of velar palatalization in Malayalam (Mohanan & Mohanan, 1984) to show that such functionalist presumptions have troubling consequences for how we rationalize about phonological representations, often resulting in ad hoc representational postulations that violate economy concerns. I then propose an alternative, modular framework, where phonetic substance and phonological primes occupy distinct, encapsulated modules, with a transductive phonology-phonetics interface (Volenec & Reiss, 2018) in between. I discuss the results of two behavioral psycholinguistic studies of the Malayayalam process (Mandal et al. Forthcoming) in support of the modular approach presented here.

2 Palatalization

Palatalization, as a cross-linguistic phenomena, is rather well-studied from both phonological (Halle, 2005; Wilson, 2006) and speech processing (Jacobs & Berns, 2013; Kochetov & Alderete, 2011; Rubach, 2003) perspectives. In particular, the acoustic, articulatory and perceptual properties of velars and palato-alveolars have been extensively documented. As a general rule, velar consonants are articulated further forward when they appear before front vowels. Often discussed under the rubric of *anticipatory co-articulation*, this phenomena is stable enough that Keating and Lahiri (1993; 89) conclude that the *"more front the vowel, the more front the velar"*. One would, of course, expect this pattern of co-articulatory effects to reflect in the acoustic properties of the externalized forms, and indeed Keating and Lahiri (1993) confirm this. The more fronted a vowel's articulation is, the smaller the resonant cavity gets, resulting in higher frequency peaks. The peak spectral frequencies of velar releases are directly proportional to the relative frontedness of the contextual vowel (Guion, 1996, 1998). In comparison, the palato-alveolar affricates display spectral frequencies that are (a) on average higher than those of velars, and (b) remain roughly constant across the frontedness of the contextual vowels (Guion, 1996). Guion's (1996) book-length treatment extends the articulatory and acoustic findings to predictable behavioral evidence from speech perception studies. Under experimental conditions native listeners tend to confuse velar stops with palato-alveolar affricates more as the frontedness of the following vowel increases. In other words subjects are more likely to confuse [ki] with [ʧi], than [ka] with [ʧa], and least likely to confuse [ku] with [ʧu].

Perception of speech sounds makes extensive use of the acoustic cues available in signals (Cutler & Butterfield, 1992; Kuhl & Miller, 1975a, 1975b). The acoustic signature of a speech, in turn, are shaped by articulator configurations (Best, 1995; Guion, 1996; Lenneberg, 1967), while also evoking articulatory response from the brain (Berent et al., 2020). Traditionally it has been either implicitly or explicitly assumed by phonologists that these phonetic and perceptual factors figure directly as causal forces shaping phonological grammar (D. Archangeli & Pulleyblank, 2015; Wilson, 2006). This view of a substantive grammar is partly derived from the observation that an overwhelming number of phonological processes

[1] DA = Designated Articulator. Halle (2005) argues that certain features are *articulator-free*, in that they are not articulated by a fixed articulator. Rather, a *designated articulator* is assigned to these features in a context-sensitive and language-specific manner.

appear to be phonetically *sensible.* Investigating the nature of the articulatory and perceptual channels have led to significant findings regarding the shape and distribution of sounds across languages (Best, 1995), the nature of evolution of linguistic speech (Fitch, 2018; Liberman et al., 1967; Tecumseh Fitch & Reby, 2001), and the role of perception in sound change (Guion, 1998; Jacobs & Berns, 2013). According to substantive phonology, speakers possess detailed substantive knowledge of their articulatory and perceptual channels, and phonological computations that guide speech make reference to this knowledge. This is most clearly visible in Wilson's (2006) treatment of palatalization. Wilson argues that across languages palatalization rules reflect two assymetric implicational relationships. First, if a language palatalizes before a less front vowel (e.g. /e/) then it will also palatalize before more front vowels (e.g. /i/). Second, palatalization of voiced velars asymmetrically implies palatalization of voiceless velars (Wilson, 2006; 950). The assumption implicit to this view is that typological trends and actual speech forms constitute direct explicandum for phonological theory. For instance, velar stops co-occur with front vowels with lower frequencies across languages than would be expected from language change (Maddieson & Precoda, 1992). This, Wilson (2006) argues, illustrates the principle that the same factors that lead to sound change diachronically are also visible in the static distribution of sounds in other languages. Velar stops and palatalized affricates exhibit greater substantive similarity in the context of more fronted vowels, leading to fronted vowels conditioning palatalization with greater efficiency. This similarity is, overall, greater for voiced segments than voiceless segments. For substantive/grounded frameworks such substantive similarities and correlations are presumed to reflect underlying cognitive principles that privilege alternations between sounds that are perceptually similar. Thus, Wilson (2006; 951) writes, *[...] the observed laws on velar palatalization derive from mental structures (such as rules or rankings of violable constraints) that are in turn shaped by phonetic substance".*

The substantive/grounded approaches make two particular assumptions which, while they work fine as tentative hypotheses, are by no means objective null hypotheses – (a) the claim that typology and verbal behavior (sounds *qua* sounds) constitute explicandum for phonological theory (A. M. Chabot, 2021; Hale & Reiss, 2008; Reiss, 2007; B. D. Samuels, 2011), and (b) substantive correlates imply substantive representations and computations (A. Chabot, 2022; Odden, 2022; Reiss & Volenec, 2022; B. Samuels et al., 2020). In discussing his *implicational laws* Wilson (2006) points out that in experimental conditions learners are more likely to generalize palatalization in context of /e/ to those of /i/, than they are in the other direction. For Wilson (2006; 968) this behavioral output is accounted for directly by the grammar by incorporating perceptual similarity into constraint weights, and this *"model yields a detailed qualitative and quantitative fits to the pattern of behavioral data"* that is observably better than a model without such substantive bias. Whether such fitness is desirable, however, is debatable. Already, alternative frameworks such as evolutionary phonology (Blevins, 2004) point out that typological asymmetries are derivable from diachronic sound change, and thus need not be attributed to mental grammars. Indeed, as Wilson (2006; 951) himself notes the "observed laws" are derived from "mental structures" that are in turn shaped by "phonetic substance". Presumably, Wilson has in mind diachronic influence of phonetic substance shaping "mental structures", which are then reflected in the observed laws. However, if diachronic influence of phonetic substance is directly capable of accounting for typological asymmetries, there remains little reason to import this effect into the domain of mental structures. To put it in another way, if diachronic factors can shape typological asymmetries directly then there is no need to assume that such factors also make changes to "mental structures" for the very same reason (cf. Chomsky, 2005). Likewise, the same substantive factors that act as diachronic forces of change are also liable to introduce synchronic co-articulatory palatalization in the context of fronted vowels (Guion, 1996; Maddieson & Precoda, 1992). The question, then, arises as to whether any given pattern of palatalization of velar consonants ought to be given a phonological account at all? If so, then how do we decide when a phonological (mental) account is required?

A foundational premise of generative phonology concerns equivalence class driven logic in rule syntax (Chomsky & Halle, 1968; Volenec & Reiss, 2020). Rules are defined in terms of three equivalence classes – the class of segments that are targets of a rule, the class that functions as the trigger/context of rule application, and finally the output. One way to respond to the question raised at the end of the previous paragraph, then, might be to assume that any systematic alternation that can be given class-driven account should receive one. Assuming for the moment that this is a serviceable hypothesis, one is still left with the task of deciding on the appropriate classes. Recall here that Halle (2005) argues that what all palatalization

phenomena have in common is that they are triggered by front vowels. If we interpret this statement as phonological/featural one, then this is technically not true. Wilson (2006) presents evidence that in empirical conditions the low, back vowel /a/ routinely triggers palatalization in naive participants. Moreover, palatalization driven by /i, e, a/ are routinely attested across-languages (Guion, 1996), with a specific paradigm from Dravidian (K. P. Mohanan & Mohanan, 1984) discussed in detail in the following section. In brief, if we assume an algebraic framework for rules, as is usual in rule-based phonology, then rules formulated around the class of [- Back] vowels will fall short of full empirical coverage. In Wilson's (2006) framework this is overcome by incorporating gradient perceptual similarity as constraint weights in the grammar. Rules, however, are algebraic and any rule triggered by [-Back] will skip the low, back vowel. In order to overcome this apparent contradiction – the rule needs to both be triggered by [- Back] segments and include /a/ as a trigger – ad hoc solutions are sometimes presented. In the next section I provide an example of this from Malayalam, a Dravidian language with a rich system of phonological processes (K. P. Mohanan, 1982; K. P. Mohanan & Mohanan, 1984; T. Mohanan, 1989), and discuss the implications for form and substance in phonology.

3 Form and substance in phonology: a Malayalam case-study

Malayalam is a Dravidian language with a basic five vowel system, and systematic length contrasts for all (Asher & Kumari, 1997; T. Mohanan, 1989). Mohanan and Mohanan (1984) report that the palato-alveolar consonants /k^{j}, k^{hj}, g^{j}, g^{hj}, ŋ^{j}/ should be derived from the respective velars preceded by front vowels. The process applies across morpheme boundaries (2a), and within words (2b), to singletons (2c) and geminates (2a, 2b), but not to non-geminate clusters (2d).

(2)

a. kut:i (NOM.) ~ kut:ik^{j}:ə (DAT.)	'child'
b. wek:^{j}al	'cooking'
c. kanak^{j}am	'gold'
d. wikramam	'brave deed'

Based on the general pattern captured in (2) above, Mohanan and Mohanan (1984) propose that the rule for velar palatalization in Malayalam applies only to single melody velars in the conditions given below:

(3) [- Continuant, + High] → [- Back] / [- Consonantal, - Back] __ [- Consonantal]

The single-melody condition on the structural description of the rule means that the rule successfully targets singletons and geminates to the exclusion of clusters. It is important to note here that Malayalam palatalizes velars following all three of the vowels discussed in Wilson's (2006) *implicational laws* – after /i/ (2a), after /e/ (2b) and after /a/ (2c). However, notice that given the Malayalam vowel system in (4) below, the rule in (3) above will not be triggered by /a/.

(4)

	i	**e**	**a**	**o**	**u**
Back	-	-	+	+	+
Round	-	-	-	+	+
High	+	-	-	-	+
Low	-	-	+	-	-

Mohanan and Mohanan (1984) are aware of this apparent contradiction, and write (p. 586, footnote 24) *"Phonetically, a is in fact a back vowel* (Warrier, 1976), *but phonologically it patterns with front vowels"*. In this view, /a/ in the table above would be specified [-Back]. Traditionally generative phonology has assumed that phonological features are grounded in phonetic substance (Kenstowicz, 1994). As Warrier (1976) points out, however, the phonetic qualities of /a/ reflect those typical of back vowels. Thus, for a purely substantive view of features there is no straightforward way to account for why a segment with the

substantive qualities of a back vowel will bear a [- Back] specification. The same substantive qualities that lead the learner to posit [+Back] for /u/ should also classify /a/ as [+Back]. One possible way to reconcile this contradiction is to adopt a version of *language-specific phonetics* – the view that the substantive correlates of features are language-specific – familiar from some versions of substance-free phonology (Dresher, 2018; B. Samuels et al., 2020). In this view, there is no specific degree of backness that qualifies a segment as either [+ Back] or [-Back]. Rather, the specification is derived from phonological activities that engage specific features, while the substantive values associated with its externalization is learned from experience with actual data. Thus, in this view, the otherwise back-like substantive correlates of /a/ are mapped to a [-Back] specification *because* it triggers a fronting process (i.e. velar palatalization). Language-specific phonetics is not without its merits. It is, indeed, a fact that the externalization of any segment exhibits significant substantive variations even when spoken by the same speaker at different times. Factor in other sources of variation like difference in sizes of oral cavity, rate of speaking, and other contextual variations, and it becomes clear that *invariant* substantive correlates of phonological units do not exist. If invariance does not exist, then the specific correlates must be learned on a case by case basis. Indeed, Samuels et al. (2020) use this observation to argue that (a) features are, in fact, substance-free algebraic symbols that merely help create equivalence classes for the purpose of rules, and (b) the specific substantive reading that a feature receives is acquired from ambient data on a language-specific basis. The features are substance-free, in this view, because they are merely symbolic primes used to capture equivalence classes as generalized from UR~SR alternations. Consequently, phonological grammar also manipulates these primes as variables, without taking into consideration their phonetic readings. Thus, phonological computation itself is also substance-free. The mapping from phonological SR to phonetic substance is, both, learned and arbitrary (Dresher, 2018; B. Samuels et al., 2020; Scheer & Romance, 2009).

While this approach, often referred to as *radically substance-free phonology* (rSFP), does resolve the apparent paradox with Malayalam /a/, it suffers from two inherent problems. First, rSFP appears to assume that learning of elementary primes (such as distinctive features) is possible in the complete absence of any priors[2]. As pointed out by Reiss and colleagues (Hale & Reiss, 2008; Reiss, 2007; Reiss & Volenec, 2022; Volenec & Reiss, 2018, 2020), to date there has been no formal refutation of Fodor's (Fodor, 1983, 1998, 2001) argument that any mathematically plausible theory of learning only consists of novel combinations of a priori available primes. In other words, unless the learner's acquisition device was already equipped with a "sense" of, say [± Back], the learner could not possibly parse any stimuli in terms of [± Back]. There is no learning primes from a completely blank slate (Piattelli-Palmarini, 1980). Second, rSFP typically provides a number of arguments for why features must be learned/emergent (Dresher, 2014, 2018; B. Samuels et al., 2020). The most relevant of these, for the current purposes, is the *lack of invariance* problem discussed before. Crucial to this belief is the conviction that innate features imply invariant correlates (Samuels et al., 2020). For proponents of rSFP the fact that the apparent problem of lack of invariance must be overcome automatically negates the possibility of innate universal feature sets. This objection[3] is, however, untenable in light of what is already known about the human performance systems (Hickok, 2022; Hickok & Poeppel, 2016; Lenneberg, 1967). The next paragraphs first provide an explication of why *lack of invariance* is not a problem for innate universal features, and then proceeds to demonstrate that neither language-specific phonetics nor arbitrary mappings between substance-free features is a logical necessity to account for velar palatalization.

The framework that I will sketch out here is often called the *Concordia School of Substance Free Phonology (Hale & Reiss, 2000, 2008; Reiss & Volenec, 2022; Volenec & Reiss, 2020)*, or just substance-free phonology (SFP) for short. It shares with rSFP its two central tenets, namely (a) phonological features are substance-free algebraic primes, and (b) phonological computations are blind to phonetic substance. It

[2]Though, see recent discussions by Dresher and colleagues (Dresher, 2014; B. Samuels et al., 2020; Scheer, 2022) conceding that such learning may require some notion of prior innate features after all. This significantly reduces the differences in representational assumptions made by rSFP and SFP, at least as concerns the phonological primes themselves.

[3]A complete review of either of these two claims of rSFP is beyond the scope of this article. For a greater discussion the reader is referred to the discussions by Reiss and colleagues (Reiss, 2017; Reiss & Volenec, 2022; Volenec & Reiss, 2018, 2020), contra Dresher and colleagues (Dresher, 2014, 2018; B. Samuels et al., 2020).

differs from rSFP, also, in two irreconcilable aspects. First, SFP assumes (contra rSFP) that the set of features provided for by UG is fixed, and invariant across languages. Second, SFP firmly posits (contra rSFP) that the phonetic readings of features at the phonetics-phonology interface (PPI) is completely deterministic and universal. This is discussed in greater detail in the following section, while the discussion here is focused on rules and phonological representations.

As mentioned previously, SFP assumes that there is a fixed, invariant, universal set of distinctive features that is provided for by UG (Hale & Reiss, 2008), while admitting the exact contents of the set are still up for debate (Volenec & Reiss, 2020). The innateness and universality of features are more conceptual necessities, than empirical facts (Reiss & Volenec, 2022). These features are substance-free binary variables, with specifications ranging over the set {+, -}. Thus, segments are sets of valued features, while classes are sets of sets of valued features. Phonological processes are conceptualized as set-theoretic operations. Thus, equivalence classes are formed by generalized intersection of segments. Rules are defined in terms of these classes, and for an alternation to be a phonological rule it *must be expressable* in terms of classes. Phonological processes make changes from the UR to SR via set-theoretic operations such as set-subtraction (for feature deletion), set-unification or priority union (for feature insertion) etc. (Bale et al., 2014; Bale & Reiss, 2018). This can be made clear by applying this logic to Malayalam. We know from (2) above that the process is triggered by the vowels /i, e, a/. Consulting the table in (4) above it is easy to see that these vowels are represented by the following sets of valued features.

(5) i = {+ High, -Low, -Back, -Round}
e = {-High, -Low, -Back, -Round}
a = {-High, +Low, +Back, -Round}

Intersection of these sets yields the set {-Round}. In SFP classes derived from such intersection of segments are represented within square brackets (here, the class of [-Round] vowels). The contrast here should be immediately relevant. Unlike Mohanan and Mohanan (1984), and indeed also Halle (2005), in SFP the trigger of the palatalization process is assumed to be [-Round] vowels. Given the axiom that phonological rules are blind to substantive correlates of features, in substance-free approaches there is no provision to privilege a fronting change like velar palatalization in the context of [-Back] vowels over those in the context of [-Round] vowels (or, indeed, any other feature).

I turn, now, to the structural description of the Malayalam palatalization rule. Looking back at the rule proposed by Mohanan and Mohanan (1984) in (3) above, we can retain the condition that the rule applies to single melody velars, but change the trigger to the [-Round] class of vowels. Next, recall that rules make changes to feature specification through logical set-theoretic operations. Here, we decompose the traditional right-arrow '→' familiar from traditional rule-based phonology into distinct sub-routines. In the case of Malayalam, the velar palatalization rule makes use of two consecutive operations. First, set-subtraction (-) removes the {+Back} specification of the target velars. Following this, the underspecified velars undergoe set-unification (∪) with the valued feature-set {-Back}.

(6) a. [- Continuant, + High] - {+Back} / [- Round] __ [- Consonantal] Set-Subtraction (SS)
b. [- Continuant, + High] ∪{-Back} / [- Round] __ [- Consonantal] Set-Unification (SU)

Now, consider the derivation in (7) below. Here I use the capital font to denote the underspecified velars that are lacking {±Back}.

(7)	**Input UR:**	/kut:ik:ə/	/wek:al /	/kanakam/	/wikramam/
	SS:	/kut:iK:ə/	/weK:al/	/kanaKam/	-
	SA:	/kut:ikʲ:ə/	/wekʲ:al/ /	kanakʲam/	-
	Output SR:	**[kut:ikʲ:ə]**	**[wekʲ:al]**	**[kanakʲam]**	**[wikramam]**

The output SRs in (7) above match the ones from (2), with palatalized forms being systematically derived from their corresponding velars in appropriate contexts. The crucial difference is that, unlike Mohanan and Mohanan's (1984) formulation the rule(s) in (6) are framed in terms of a triggering class that does not require ad hoc assumptions regarding the phonetic correlates of otherwise familiar features. Further, unlike Wilson's (2006) analyses the approach presented here does not incorporate performative psycho-phonetic factors (perceptual similarities and confusion matrices) as substantive constraints, thus avoiding the duplication of phonetic information in phonology (Dresher, 2014; Hale & Reiss, 2000; Scheer & Romance, 2009).

Substance-free approaches to phonology view phonological grammars, of the sort discussed so far, to be the matter of linguistic *competence* (Chomsky, 2007; Scheer, 2022; Volenec & Reiss, 2020). Grammar is a set of primes, and rules to combine and recombine said primes, that underlie the neuromuscular execution of behavior (Volenec & Reiss, 2018). It is important to note here that technically there is nothing inherent in the assumptions of rSFP itself that syntactically excludes the possibility of a rule of the sort presented in (6). rSFP, like the current framework, is explicit about phonological computations being agnostic to phonetic substance. As such, a palatalization rule triggered by [-Round] is not a *crazy rule* (Scheer & Romance, 2009), even in rSFP, because in substance-free approaches (in general) there is no notion of crazy or "unnatural" rules. Some rule-effects are attested less often than others because of diachronic factors (Blevins, 2004; Hale, 2000), the nature of perceptual assimilation (Best, 1995; Cutler, 2012) etc., but they are not less privileged by the grammar in any sense (Scheer & Romance, 2009). However, in practice proponents of rSFP tend to derive feature-taxonomy on a language-specific basis using phonological alternations to create contrastive feature-hierarchies (Dresher, 2018). Thus, vowels that pattern together are grouped together, under relevant features. In Malayalam, two of the three vowels that trigger palatalization are, in fact, front vowels (/i, e/) in the usual sense. This, alongside the fact that rSFP also allows arbitrary mappings between phonological primes and substantive correlates, means that ad hoc assumptions of the sort made in Mohanan and Mohanan (1984) regarding the phonological and phonetic nature of /a/ tend to appear less unwarranted. SFP and rSFP make radically different assumptions regarding the nature of the phonology-phonetic interface and acquisition, while sharing broad axioms regarding the nature of mental faculties and competence (Dresher, 2014; contra, Reiss & Volenec, 2022).

Empirical probing of competence, however, usually rely on indirect evidence gleamed from close examination of *performance*. The next section discusses the issues surrounding the externalization of the output of phonology, variations, universals, and the nature of the phonetics-phonology interface, in the context of a behavioral psycholinguistic study of Malayalam palatalization. Particular attention is paid to accounting for the implicational distribution reported by Wilson (2006), and clearly visible in the data reported (Mandal et al., forthcoming), without importing substance into the phonological module.

4 Perception, production and representations

The previous section discussed general articulatory and perceptual properties exhibited by palatalized segments across languages. Specifically, it was noted that cross-linguistically the distribution of palatalization processes show implicational asymmetry with regard to the relative frontedness of the adjacent vowel (Wilson, 2006). In brief, the perceptual and articulatory overlaps between velars and palatals increase as the vowel gets progressively more fronted. This implies that more front vowels such as /i/ are more likely to trigger palatalization than less fronted ones like /e/, or /a/. Mandal et al. (forthcoming) report two experimental investigations of the Malayalam palatalization process with early Malayalam-English (Aus.) bilinguals, testing the productivity of each of the triggering vowels (/i, e, a/) in both perception and production modalities. The study focuses exclusively on palatalization across suffix-boundaries, and takes advantage of the fact that certain Malayalam suffixes are known to block palatalization. Mohanan and Mohanan (1984) report that among other idiosyncrasies Malayalam suffixes appear to be selectively licensed for palatalization. For instance, while the suffixes -kːuka] "Inf." and -kːə] "Dat." regularly undergo palatalization, the plural suffix -kal] never does.

(8) **UR** **SR**

kut:ik:ə	kut:ikʲ:ə	"child" (DAT.)
marak:uka	marakʲ:uka	"cover" (INF.)
kut:ikal	kut:ikal	"children"

Taking advantage of this pattern, the study probed the rate of palatalization of the aforementioned suffixes with nonce stems of the form CVCV. Stems were designed to ensure equal number of exemplars ending in each of the three triggering vowels. This yields a 3 x 3 design, with three suffixes (-k:uka], -k:ə], and -kal]) and three stem-final vowels (/i, e, a/). Participants were tested in two distinct conditions. The first condition utilized a perceptual forced-choice paradigm, where participants were played conjugated forms in pairs, with one member of each pair containing palatalized velar and the other one containing non-palatalized velar. The participants indicated their preferred token by pressing either 1 (for the first token of a pair) or 2 (for the second token). The second test condition made use of a modified-WUG task. Participants were shown orthographic representations of "stem + affix", and were instructed to produce the conjugated form out loud. The stimuli were once again presented in the same 3 x 3 design as the first condition, utilizing the same stem-final vowels and suffixes. The utterances spoken by the participants were recorded, and subjected to phonetic analyses as well as classification (palatalized vs. non-palatalized) by a phonetically trained L1-consultant. In both conditions the proportions of palatalized to non-palatalized responses constitute the main data points.

In the first condition, the authors report a main effect of vowel, $F(1,14) = 24.13$, $p < 0.01$, and suffix-type, $F(1,14)=917.72$, $p < 0.01$, alongside a vowel*suffix interaction $F(1,14) = 7.09$ and $p = 0.01$. Planned contrasts showed that while the low-back vowel /a/ was significantly different from the other two vowels, there was no noticeable difference between /i/ and /e/. The statistical trend is also reflected in the second condition, where the authors report a main effect of vowels, $F(1,14) = 10.38$, $p < 0.01$, suffixes, $F(1,14) = 1747.79$, $p > 0.01$, but no interactions. Planned contrasts, once again, grouped /i, e/ together, with /a/ showing statistically significant difference from the former two. No difference is reported between the two licensed suffixes, while the unlicensed one systematically fails to trigger palatalization in both conditions. The figure below provides an over-view of the means of proportions of palatalized responses in each condition.

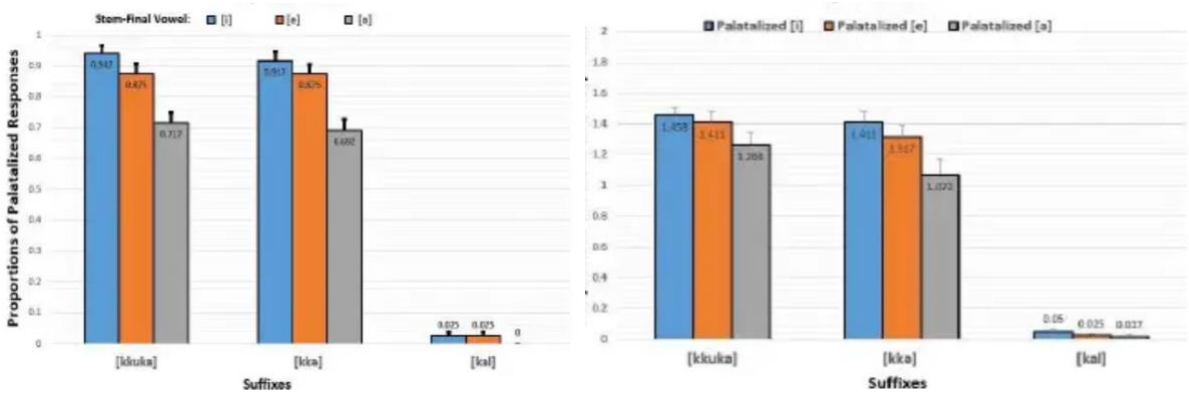


Fig. 1. (a) Forced-Choice Task (b) Modified WUG Test

Even at first glance the charts in fig.1 clearly show the kind of implicational effects Wilson (2006) discusses. In both conditions the licensed suffixes exhibit progressively greater proportions of palatalized responses as the frontness of the vowel increases from /a/ to /i/. Notice, also, that irrespective of the vowel-frontness the unlicensed suffix -kal] barely triggers any palatalization at all. In SFP the input to (UR) and output of (SR) of phonological operations are of the same *type* – substance-free, discrete feature-bundles. In other words, SFP adheres to the notion of Fodorian modularity where the representations of modules are informationally encapsulated, and use a proprietary vocabulary (Fodor, 1983). For phonology, the vocabulary consists of valued distinctive features. The alternations from the UR to the SR, consequently, only consist of insertion, deletion or changes to features or their specifications. As Volenec and Reiss,

(2020) point out, there is no notion of gradient phenomena within phonology, because the primitives of phonology and the operations defined over them are both purely algebraic. The crucial argument, however, concerns the fact that distributions of the type reflected in the charts above are not indicative of phonological representations or operations. Rather, SFP proposes a transductive theory of the relationship between phonological forms and their substantive instantiations, called *cognitive phonetics* (CP; Volenec & Reiss, 2018).

As features are substance-free neural symbols, they are not directly legible to the sensory and motor systems. Externalization of the SR as speech forms require speech motor plans encoded with temporal co-ordination of articulatory muscles and spectral configurations of the acoustic targets to be reached (Hickok & Poeppel, 2016). CP posits that the phonetics-phonology interface (PPI) consists of two distinct transduction algorithms. A *paradigmatic transduction algorithm* (PTA) scans a feature-bundle (segment) and assigns motor programs for relevant muscle control to each feature, in order to reach a desired acoustic target. Likewise, a *syntagmatic transduction algorithm* (STA) determines the temporal co-ordination of the motor activities specified by the PTA. In other words, at the PPI the PTA+STA receive the substance-free phonological SR as their input, and yields a substance-imbued *true phonetic representation* (TPR) that is capable of serving as constituents of speech motor plans. To clarify this process, consider how this applies to the Malayalam forms seen in (7). The rules in (6) take an UR as input, say /kut:ik:ə/, and yield an output SR, [kut:ikʲ:ə]. The forms enclosed within / / and [ ] are both substance-free, with the sole difference being that the underlying velar within [ ] are specified {- Continuant, +High, **-Back**} (see rule in (6) above). The PPI receives [kut:ikʲ:ə] as input, and assigns a motor-plan to each feature of each segment in its output. I will represent this form, the TPR, with angled brackets -- <kut:ikʲ:ə>. This process of transformation, from [kut:ikʲ:ə] to <kut:ikʲ:ə>, gives rise to two distinct forms of co-articulatory effects whose effects show contextual variations. For instance, in the forms [kut:ikʲ:ə] and [wekʲ:al], the triggering vowels also contain the feature {- Back}. However, while /i/ is specified {+High}, /e/ is {-High}. Thus, while the shared {-Back} means that for both segments PTA configures the extrinsic tongue muscles to pull the tongue back, for the {-High} segment (/e/) articulation would also engage the hyoglossus in order to lower the tongue, relative to /i/. Likewise, the risorius muscles required for lip spreading in unrounded vowels will be activated by both. However, given that muscles are often inter-connected and the configuration for one has an indirect effect on adjoining muscles, the final effect of a specific valued feature within a segment is a cumulative effect of that specific feature, say {-Back}, and the other features that it is bundled with. This is termed *intrasegmental co-articulation* (Volenec & Reiss, 2018). Likewise, in both [kut:ikʲ:ə] and [kanakʲam] the palatalized velar segment [kʲ:] bears the same {-Back} specification. However, in [kut:ikʲ:ə] the vowel preceding it is specified {-Back} while in [kanakʲam] it is specified {+Back}. While determining the temporal co-ordination of features in segments, the STA can temporarily extend a specification to a neighbouring segment, either preceding (anticipatory) or following (inertial). This is expected given that the transduction process is responsible for producing real motor plans that are subject to basic laws of motion, mass and inertia. Thus, while in <kut:ikʲ:ə> the palatalized velar is further influenced by the inertial extension of a preceding {-Back}, for <kanakʲam> the effect is exactly the opposite. To conclude, then, intrasegmental co-articulation implies that the inertial extension of {-Back} from the preceding vowel has distinct effects in <kut:ikʲ:ə> and <wekʲ:al>. While in the case of <kanakʲam> there is an inertial effect of a preceding {+Back}. The combined workings of the PTA+STA are clearly visible in the gradient vowel effects, with the low-back vowel /a/ showing the strongest negative impacts. Note, however, that none of these effects are directly attributable to phonological grammar in any sense. Grammar, in SFP, necessary entails *knowledge*. The transduction carried out at the PPI does not entail knowledge in any meaningful sense (Chomsky, 1980).

The transduction from the SR to TPR entails a set of automatic neuromuscular processes. Following Lenneberg (1967, section 4), CP assumes that ontogenetically the development of the PTA and STA likely followed those of other performance systems. Being automatic synergisms (Lenneberg 1967; 92), so to speak, the transduction carried out here is also completely *deterministic*. Which means that it assigns the same neuromuscular schema to any given feature irrespective of where it is encountered, including all cases of feature combination leading to inter- and intrasegmental co-articulation. To be specific, cognitive phonetics is *cognitive* precisely in this aspect. It describes the process of assigning neuromuscular schemas to susbtance-free phonological forms. These schemas, themselves, are not articulatory movements nor acoustic outputs. They are still entirely brain-internal schemas. The co-articulatory effects are derived from

the actual physiological implementation of said schemas. The deterministic aspect of CP derives straightforwardly from the fact that automatic synergisms are "whole trains of events that are preprogrammed and run off automatically" forming the "basis of all motor phenomena in vertebrates" (Lenneberg, 1967; 92). Together, SFP + CP yield a very specific modular view (fig. 2) of the PPI, and one that is consistent with emerging evidence in cognitive neuroscience (Dronkers et al., 2000; Gallistel & King, 2009; Hickok & Poeppel, 2016). Neuroimaging and electrophysiological evidence supports the idea that activity in the superior temporal gyrus (STG) and sulcus (STS) encode auditory correlates, while those in inferior frontal gyrus (IFG) encode articulatory correlates (Hickok et al., 2021; Hickok & Poeppel, 2016). Area Spt, situated in the Sylvian fissure at the boundary of the parietal and temporal lobes is responsible for integrating the two correlates into a complete feature. This is then transmitted to the anterior insula for paradigmatic transduction, and to the cerebellum and basal ganglia for syntagmatic transduction. Following this, the PTA and STA are integrated in the anterior supplementary motor area (pre-SMA) to yield TPR. The TPR is subsequently forwarded to the motor effectors to produce speech. The flowchart in fig. 3. demonstrates the outward transduction process, from mental structures to articulated speech. It follows, obviously, that transduction is a bidirectional process that also parses and repackages incoming stimuli, until specific spectro-temporal receptive fields (STRF) excite specific populations of neurons (Weerathunge et al., 2022), conjunctively capturing increasingly complex information until a feature is activated (Kemmerer, 2014).

Two crucial aspects of this framework is worth noting here. Since CP is non-language-specific and deterministic (similar to the auditory or visual system), it is not itself a source of variability in speech. However, due to CP being fed language-specific SRs via phonology, CP's outputs, TPRs, will reflect language-specific patterns via intra- and inter-segmental coarticulation. For example, Zulu might display labialized clicks before rounded vowels due to STA extending the lip rounding from [u] and [ɔ] in the anticipatory direction — a form of inter-segmental coarticulation. This effect will not be present in, for instance, Spanish for the mere fact that Spanish SRs do not have any clicks. Likewise, no two utterances are ever physically identical, even when an individual thinks they are repeatedly uttering 'the same' utterance. For example, every repetition of what is cognitively the same [+HIGH] vowel will have its F1 in a certain low region, but precise F1 measurements will show that [i_1], [i_2] and [i_3] never have equal F1 values. Thus, in addition to predicable variability (e.g., formant values predictably vary with vocal tract size), physiology also introduces truly random and intractable variability. The automatic, deterministic nature of CP's transduction is what ensures that the phonetic correlates of features fall within relatively predictable acoustic spaces, while the tremendous complexity of all other performance components gives rise to lack of invariance. The appeal, in my opinion, of this modular framework is threefold. First, it makes an imminently testable prediction that given a full and correct inventory of features, one could in principle calculate the range of transductive co-articulatory effects. Second, it affords a coherent explanation of the implicational distributions of the sort Wilson (2006) discusses without burdening the phonological grammar with substantive information. Finally, diachronic changes leading to typological trends can be accounted for through intractable/physiological variations being subject to the properties of the perceptual channel.

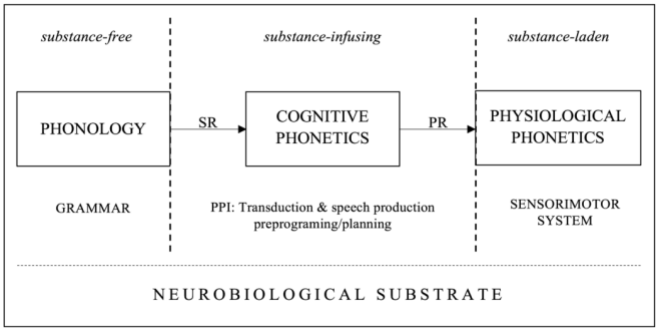


Fig. 2. Modular Phonetics-Phonology Interface

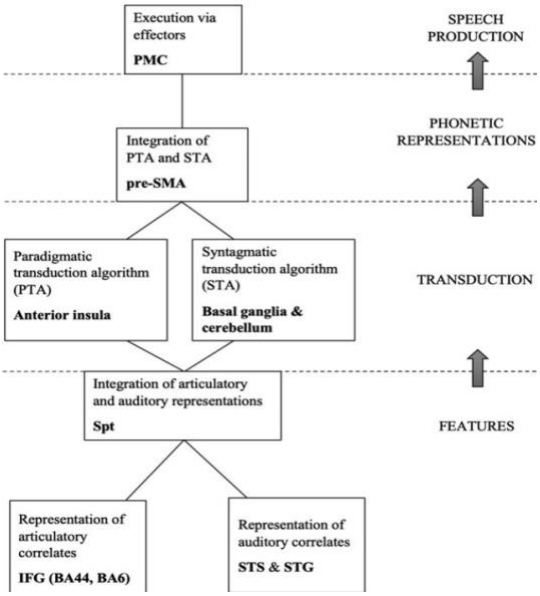


Fig. 3. Neural Substrate for Distinctive Features & CP

References

Archangeli, D. B., & Pulleyblank, D. G. (1994). *Grounded phonology* (Vol. 25). MIT Press.

Archangeli, D., & Pulleyblank, D. (2015). Phonology without universal grammar. *Frontiers in Psychology*, *6*, 1229.

Asher, R. E., & Kumari, T. (1997). *Malayalam*. Psychology Press.

Bale, A., Papillon, M., & Reiss, C. (2014). Targeting underspecified segments: A formal analysis of feature-changing and feature-filling rules. *Lingua*, *148*, 240–253. https://doi.org/10.1016/j.lingua.2014.05.015

Bale, A., & Reiss, C. (2018). *Phonology: A formal introduction*. MIT Press.

Berent, I., Platt, M., Theodore, R., Balaban, E., Fried, P. J., & Pascual-Leone, A. (2020). Speech perception triggers articulatory action: Evidence from mechanical stimulation. *Frontiers in Communication*, *5*, 34.

Best, C. T. (1995). Learning to perceive the sound pattern of English. *Advances in Infancy Research*, *9*, 217–217.

Blevins, J. (2004). *Evolutionary phonology: The emergence of sound patterns*. Cambridge University Press.

Chabot, A. (2022). On substance and Substance-Free Phonology: Where we are at and where we are going. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique*, *67*(4), 429–443.

Chabot, A. M. (2021). *Langues possibles et impossibles: Naturalité, facteurs thiers, et phonologie sans substance à la lumière des règles folles* [PhD Thesis]. Université Côte d'Azur.

Chomsky, N. (1980). Rules and representations. *Behavioral and Brain Sciences*, *3*(01), 1–15.

Chomsky, N. (2005). Three Factors in Language Design. *Linguistic Inquiry*, *36*(1), 1–22. https://doi.org/10.1162/0024389052993655

Chomsky, N. (2007). Of minds and language. *Biolinguistics*, *1*, 009–027.

Chomsky, N., & Halle, M. (1968). *The sound pattern of english*. MIT press.

Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*. Mit Press.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language*, *31*(2), 218–236.

Dronkers, N. F., Redfern, B. B., & Knight, R. T. (2000). The neural architecture of language disorders. *The New Cognitive Neurosciences*, *2*, 949–960.

Dresher, B. E. (2014). The arch not the stones: Universal feature theory without universal features. *Nordlyd*, *41*(2), 165–181.

Dresher, B. E. (2018). Contrastive hierarchy theory and the nature of features. *35th West Coast Conference on Formal Linguistics*, 18–29.

Fitch, W. T. (2018). What animals can teach us about human language: The phonological continuity hypothesis. *Current Opinion in Behavioral Sciences*, *21*, 68–75.

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. MIT press.

Fodor, J. A. (1998). *Concepts: Where cognitive science went wrong*. Oxford University Press.

Fodor, J. A. (2001). *The mind doesn't work that way: The scope and limits of computational psychology*. MIT press.

Gallistel, C. R., & King, A. P. (2009). *Memory and the Computational Brain*. Wiley-Blackwell. https://doi.org/10.1002/9781444310498

Guion, S. G. (1996). *Velar palatalization: Coarticulation, perception, and sound change*. The University of Texas at Austin.
Guion, S. G. (1998). The role of perception in the sound change of velar palatalization. *Phonetica*, *55*(1–2), 18–52.
Hale, M. (2000). *Marshallese phonology, the phonetics-phonology interface and historical linguistics*.
Hale, M., & Reiss, C. (2000). "Substance Abuse" and "Dysfunctionalism": Current Trends in Phonology. *Linguistic Inquiry*, *31*(1), 157–169. https://doi.org/10.1162/002438900554334
Hale, M., & Reiss, C. (2008). *The phonological enterprise*. Oxford University Press.
Halle, M. (2005). Palatalization/velar softening: What it is and what it tells us about the nature of language. *Linguistic Inquiry*, *36*(1), 23–41.
Hickok, G. (2022). The dual stream model of speech and language processing. *Handbook of Clinical Neurology*, *185*, 57–69.
Hickok, G., & Poeppel, D. (2016). Neural basis of speech perception. *Neurobiology of Language*, 299–310.
Hickok, G., Venezia, J., & Teghipco, A. (2021). *Beyond broca: Neural architecture and evolution of a dual motor speech coordination system*.
Jacobs, H., & Berns, J. (2013). Perception, production and markedness in sound change: French velar palatalization. In *Research on Old French: The state of the art* (pp. 107–122). Springer.
Keating, P., & Lahiri, A. (1993). Fronted velars, palatalized velars, and palatals. *Phonetica*, *50*(2), 73–101.
Kenstowicz, M. J. (1994). *Phonology in generative grammar*. Blackwell.
Kemmerer, D. (2014). *Cognitive neuroscience of language*. Psychology Press.
Kochetov, A., & Alderete, J. (2011). Patterns and scales of expressive palatalization: Experimental evidence from Japanese. *The Canadian Journal of Linguistics/La Revue Canadienne de Linguistique*, *56*(3), 345–376.
Kuhl, P. K., & Miller, J. D. (1975a). Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science*, *190*(4209), 69–72.
Kuhl, P. K., & Miller, J. D. (1975b). Speech perception in early infancy: Discrimination of speech-sound categories. *The Journal of the Acoustical Society of America*, *58*(S1), S56–S56.
Lenneberg, E. H. (1967). The biological foundations of language. *Hospital Practice*, *2*(12), 59–67.
Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*(6), 431.
Maddieson, I., & Precoda, K. (1992). Syllable structure and phonetic models. *Phonology*, 45–60.
Mandal, S, Best, C.T., Shaw, J. (forthcoming). Velar palatalization in Malayalam.
Mohanan, K. P. (1982). *Lexical phonology*.
Mohanan, K. P., & Mohanan, T. (1984). Lexical phonology of the consonant system in Malayalam. *Linguistic Inquiry*, 575–602.
Mohanan, T. (1989). Syllable structure in Malayalam. *Linguistic Inquiry*, 589–625.
Odden, D. (2022). Radical substance-free phonology and feature learning. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique*, *67*(4), 500–551.
Piattelli-Palmarini, M. (1980). *Language and learning: The debate between Jean Piaget and Noam Chomsky*.
Reiss, C. (2007). Modularity in the "Sound" Domain: Implications for the Purview of Universal Grammar. *The Oxford Handbook of Linguistic Interfaces*. https://doi.org/10.1093/oxfordhb/9780199247455.013.0003
Reiss, C. (2017). Substance free phonology. In *The Routledge handbook of phonological theory* (pp. 425–452). Routledge.
Reiss, C., & Volenec, V. (2022). Conquer primal fear: Phonological features are innate and substance-free. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique*, *67*(4), 581–610.
Samuels, B., Andersson, S., Sayeed, O., & Vaus, B. (2020). Getting ready for Primetime: Paths to acquiring substance-free phonology. *Canadian Journal of Linguistics*, *67*(3).
Samuels, B. D. (2011). *Phonological architecture: A biolinguistic perspective*. Oxford University Press.
Scheer, T. (2022). 3 x Phonology. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique*, *67*(4), 444–499.
Scheer, T., & Romance, G. (2009). Who is afraid of crazy rules. *Going Romance, Nice*, 3–5.
Tecumseh Fitch, W., & Reby, D. (2001). The descended larynx is not uniquely human. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *268*(1477), 1669–1675.
Volenec, V., & Reiss, C. (2018). *Cognitive Phonetics: The Transduction of Distinctive Features at the Phonology-Phonetics Interface* (Vol. 11).
Volenec, V., & Reiss, C. (2020). Formal generative phonology. *Radical: A Journal of Phonology*, *2*, 1–148.
Warrier, T. (1976). The Phonetics and Phonology of Malayalam and its Pedagogical Implications—A Generative Phonological Study. *Diss. CIEFL: Hyderabad*.
Weerathunge, H. R., Alzamendi, G. A., Cler, G. J., Guenther, F. H., Stepp, C. E., & Zañartu, M. (2022). LaDIVA: A neurocomputational model providing laryngeal motor control for speech acquisition and production. *PLOS Computational Biology*, *18*(6), e1010159.Weerathunge, H. R., Alzamendi, G. A., Cler, G. J., Guenther, F. H., Stepp, C. E., & Zañartu, M. (2022). LaDIVA: A neurocomputational model providing laryngeal motor control for speech acquisition and production. *PLOS Computational Biology*, *18*(6), e1010159.
Wilson, C. (2006). Learning Phonology With Substantive Bias: An Experimental and Computational Study of Velar Palatalization. *Cognitive Science*, *30*(5), 945–982. https://doi.org/10.1207/s15516709cog0000_89

The Syntax of Vocatives of Entreaty in Arabic: Functional or Lexical

Saleem Abdelhady
American University of the Middle East

This study examines the derivational structure of *nɪdaʔ ʔalʔɪstɪyaθa* 'Vocative of Entreaty' in Arabic. Vocatives of Entreaty refer to vocative phrases that a speaker employs to call an addressee to help a person who is in distress. Vocatives of Entreaty challenge previous accounts on the phrase structure of vocatives because in addition to selecting Determiner Phrases (DPs) and Noun Phrases (NPs), Vocatives of Entreaty select Prepositional Phrases (PPs) as their complement. This study examines the conditions that license a vocative head to select PPs. The findings have significant consequences for understanding the syntactic-pragmatic interface. The principal theoretical implication of this study is that the structure of vocatives is determined by its pragmatic force. If this research is to be moved forward, a better understanding of the phrase structure of vocatives needs to be established cross-linguistically.

Keywords: vocatives, vocatives of entreaty, syntax, the Minimalist Program

1 Introduction

In broad terms, vocatives can be defined as phrases that are "by default, organized around a noun (or a pronoun), either a name or a common noun, which may or may not be modified by adjectives, other nouns, […], prepositional phrases or relative clauses" (Hill, 2013, p. 42). In recent years, the structure of vocatives has seen significant advances. Work in this area is developed by investigating how vocative phrases occupy certain syntactic positions across languages. However, the structure of vocatives continues to be an open problem. The most common culprit for this issue appears in Standard Arabic. Before we proceed, it is interesting to briefly illustrate the problem by considering the following examples (1).

(1)

a. *ya* *la-rami* *lɪ-Sam.*
VOC.PART[1] to-rami.GEN to-Sam.GEN
'Hey Rami Sam (seeks your help.)'

b. *ya* *lɪ-sami* *lɪ-l-muslɪmiin.*
VOC.PART to-Sami.GEN to-DEF-Muslims.GEN
'Hey Sami Muslims (seek your help.)'

c. *ya* *lɪ-xaalɪd.*
VOC.PART to-Khalid.GEN
'Hey Khaled (I seek your help.) !

* The researcher would like to thank the committee of the Western Conference on Linguistics at California State University for organizing this conference and all the audience for sharing their ideas and commenting on this research. The researcher equally thanks the editorial team of Journal of Language Teaching and Research for providing insightful comments on an earlier version of this research.

[1] VOC: vocative, PART: particle, N: noun, CP: complementizer phrase, ADVP: adverbial phrase, A: adjective, GEN: genitive, M: Masculine, F: Feminine, DEF: definite, S: singular, PL: plural; DU: dual, VP: verb phrase, DP: determiner phrase, NP: noun phrase

d. *ya* *lɪ-hani* *lɪ-salim* *mɪn* *xaalid*.
VOC.PART to-Hani.GEN to-Salim.GEN from Khalid.GEN
'Hey Hani Salim (seeks your help) from Xhalid.'

(Arabic)
(Al-Samirai 2000, 335-336)

It should be noted that the examples in (1) are vocative phrases that are not "organized around a noun (or a pronoun), either a name or a common noun" (Hill, 2013, p. 42); the examples select PPs as their complement. Therefore, the classical definition of vocatives fails to explain those examples because the definition is in part guided by the data available for researchers. Initial observations suggest that there may be a link between the vocative head and the PP. Is this link related to a selection license? Or does the vocative structure show a case of ellipsis? To stop at those questions, the study uses the Minimalist Program (Chomsky, 1995) for data analysis.

The research has been organized in the following way. Section Two begins by laying out the theoretical dimensions of the research and explores the phrase structure of Arabic vocatives, focusing on the phrase structure of Vocative of Hailing. It is necessary, in this section, to clarify exactly what is meant by Vocative of Hailing and show how it juxtaposes with Vocatives of Entreaty. Section Three introduces the Minimalist Program (Chomsky, 1995). At its syntactic-pragmatic interface, this program is particularly useful in exploring Vocatives of Entreaty. Section Four tests the phrase structure of Vocatives of Entreaty and offers flexibility to handle the intricacies of deriving Vocatives of Entreaty. Section Five concludes the study and throws up many inquiries in need of additional investigation.

2 Theoretical Background

A search of the literature revealed few studies (cf., Abuladze & Ludden, 2013; Akkus & Hill, 2018; González López, 2020; Hill, 2022) which examine the phrase structure of vocatives. The main challenge faced by many researchers is to establish that vocatives are functional phrases. This part begins with a general overview of vocatives (e.g., Hill, 2007) and narrows it down to describe vocatives in Arabic as characterized in the literature (cf., Shormani & Qarabesh, 2018; Soltan, 2016). Taken together, these studies have substantial consequences for understanding Vocatives of Entreaty.

3 Vocatives

3.1 *Features of Vocatives* Vocatives can be described as constituents that are "by default, organized around a noun (or a pronoun), either a name or a common noun, which may or may not be modified by adjectives, other nouns, […], prepositional phrases or relative clauses" (Hill, 2013, p. 42).

(2)

a.	[Mary], when did they leave?	DP
b.	[My dear Mary], what came on you?	AP-N
c.	[Doctor Smith], may I talk to you?	DP-N
d.	[The students from France], where is your translator?	N-PP
e.	[The students who just arrived], you must stay in this room.	N-CP
f.	[You there], what do you think you're doing?	DP-AdvP
g.	[Smarty], can you keep quite now?	A/DP

(English)
(Hill 2013, p. 43)

In view of all the examples in (2), Hill (2013) proposes a generalization of the phrase structure of vocatives in English. She demonstrates that a vocative phrase is a functional projection of a vocative head that may select DPs or NPs as its complement; those selected complements can be modified by APs, PPs, CPs and the like. One of the limitations with this explanation is that it does not explain how a vocative head can select PPs as its complement (1).

3.2 *Vocatives in Arabic* A great deal of previous research on Arabic has focused on vocatives (e.g., Moutaouakil 2014; Haddad 2020; Soltan 2016; Jaradat et al., 2022; Shormani & Qarabesh 2018). The research to date has tended to focus on Vocatives of Hailing rather than Vocatives of Entreaty. Moutaouakil (2014) provides a comprehensive view of vocatives from a functional perspective. Before we precede, it is thought-provoking to present three patterns of Arabic vocatives from his perspective. This distinction is further exemplified in the following instances.

(3)

a. *Zayd-u* *nawıln-i* *al-mılħ-a.*
Zayd-NOM give-1-ACC DEF-salt-ACC
'Zayd, give me the salt.'

b. *ya* *talıɣ-a* *al-ʃdʒarah-ti,* *ʔınzıl.*
VOC-PART climbing-ACC DEF-tree-GEN come.down
'You who are climbing the tree, come down.'

c. *ħana* *waqt-u* *n-nawm-i,* *ʔayyuha* *l-tıfl-u.*
Arrived time-NOM DEF-sleep-GEN O DEF-child-NOM
'Now is the time to sleep, children.'

(Vocative of Hailing)

(4)

a. *ya* *lı-Zayd-ın* *lı-Halid-ım.*
VOC.PART to-Zayd-GEN to-Halid-GEN
'O, if only Zayd were by Halid's side.'[2]

b. *ya* *la-Amr-in* *lı-ma* *ʔsaban-a.*
VOCPART to-Amr-GEN to-what happened.PERF-1PL
'O, if only Amer were here, with all that has happened to us.'

(Vocative of Entreaty)

(5) *wa* *Zayd-ah*
VOC.PART Zayd-ah
'O Zayd, as exclaimed by a widow at the graveside.'

(Vocative of Bemoaned)
(Moutaouakil, 2014, p.139)

The three examples show three functions of vocatives in Arabic: Vocative of Hailing (3), Vocative of Entreaty (4), and Vocative of Bemoaned (5). Each pragmatic function utilizes a unique structure. This study set out with the aim of assessing the phrase structure of Vocative of Entreaty (4) because the observational data suggests that the phrase structure of Vocative of Entreaty differs from the common generalization of vocatives.

4 The Syntax-Pragmatics Interface

There is a growing body of literature that acknowledges the significance of pragmatics in analyzing syntactic phrases (e.g., Abdel-Hady, & Branigan, 2020; Abdelhady, 2020a,b). Part of the aim of this section is to shed light on the syntactic framework of this study. The section, then, aims to unravel the syntactic-pragmatic interface.

Ross (1970) claims that performative verbs govern all syntactic clauses. To say this in other words, syntactic computations operate on the idea of performing actions by utterances (Speech Acts). By way of illustration, Ross (1970) shows how the idea works in principle in (6).

(6)

a. Open the door.
b. [I order you to] open the door.

The claims of viewing syntactic structures as performative actions make an important contribution to adding a new layer for syntactic operations. This layer shows a need to be explicit about exactly what is meant by Speech Act Projections. According to Speas and Tenny (2003), Speech Act Projections are projections that depict the influence of speakers and addresses on constructions and govern clauses (i.e., Speech Act Projections head Complementizer Phrases). The applicability of Speech Act Projections has been broadened to analyze vocatives across languages.

5 Methodology

To understand the phrase structure of Vocatives of Entreaty, we collected our data from studies that reported examples of Vocatives of Entreaty (e.g., Moutaouakil, 2014). The main advantage of collecting our data in this way is that it provided access to fifty unique phrases where vocatives of Entreaty are used. The second advantage is that the examples collected in those books represent how spoken language and judged by native speakers of Arabic.

After collecting the examples, the researchers exposed the set of collected examples to an experiment. The purpose of the experiment is to test the grammaticality of the examples under syntactic tests such as coordination, substitution, and others. The next step after designing the experiment, the researchers asked five native speakers to judge the acceptability of the examples under such syntactic conditions. Finally, the researchers deployed the tools of the Minimalist Program to understand the phrase composition of Vocatives of Entreaty under current theories of language.

6 Grammaticality Assessment

Different methods have been proposed to assess vocatives. In many studies, it has become commonplace to distinguish Vocatives of Hailing from Vocatives of Entreaty on the bases of several key respects. First, while it is permissible to omit the vocative particle in vocative constructions, doing so for Vocatives of Entreaty leads to ungrammatical constructions. Second, vocatives select one argument as an addressee. However, the phrase structure of Vocatives of Entreaty requires two arguments, one for the addressee and the other for the distressed. Third, Vocatives of Hailing and Vocatives of Entreaty not only differ in the number of selected arguments, but also in the way case markings are assigned to those arguments. That is to say, the argument of a Vocative of Haling appears in a nominative or accusative case; on the other hand, contrary to expectations, the argument of a Vocative of Entreaty must show up with a genitive case. This feature is rather surprising because the argument of the vocative is part of a prepositional phrase; that is, the preposition c-commands the argument. These results provide further support for the hypothesis that a vocative head may select a prepositional phrase instead of a determiner phrase or a noun phrase. The final remarkable feature of Vocatives of Entreaty is that the arguments in this type of vocatives show fixed positions in the structure; that is, they cannot undergo any movement; if one argument precedes the other a new semantic role is assigned (addressee <> distressed).

(7)

a. *ya* *lɪ-Zayd-ɪn* *lɪ-Halid-ɪn.*
VOC.PART to-Zayd-GEN to-Halid-GEN
'O, if only Zayd were by Halid's side.'

b. *ya* *la-Amr-in* *lɪ-ma* *ʔsaban-a.*
VOCPART to-Amr-GEN to-what happened.PERF-1PL
'O, if only Amer were here, with all that has happened to us.'

(repeated)
(Moutaouakil, 2014, p.139)

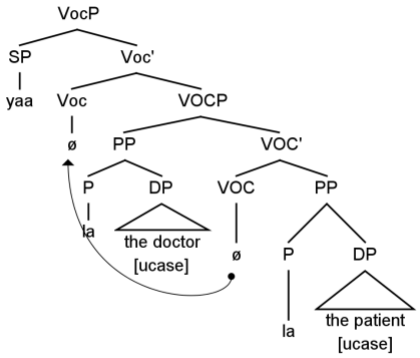
The example in (7) above illustrate the basic structure of Vocatives of Entreaty. they cannot undergo any movement; if one argument heads the other a different semantic role will be assigned for it (addressee <> distressed).

7 The Syntax of Vocatives of Entreaty

Researchers (e.g., Hill, 2007; Hill, 2022; Hill, 2013; Speas & Tenny, 2003; Abdelhady, 2020, 2021), show that a vocative phrase occupies a syntactic position in the Speech Act Projection. This phrase is made up of a functional head that an argument. The generalisability of much published research on this issue is problematic because earlier works define the configuration to revolve around a determiner phrase or a noun phrase and claim that a vocative phrase should select one argument (addressee). This study shows counterexamples for the previous studies in that it argues that VOC, in Vocatives of Entreaty, has two valences (*l-mustaɣaaθ bɪh* 'the addressee' and *l-mustaɣaaθ lah* 'the distressed'). This claim casts doubt on Hill's (2013) proposal that "vocatives are organized around a noun" since Vocatives of Entreaty select PPs as arguments.

Taken together, our results suggest that there is an association between a vocative head and its PP. I claim that the DP has an unvalued case feature [ucase]. P has a valued [genitive] feature. P probes for the closest DP to check its feature. This results in assigning the genitive case to the DPs in the PP.

(8)



One possible implication of this is that a vocative head selects two PPs as its argument. The lowest occupies its complement position and the other one occupies its specifier position. This is demonstrated in the tree diagram in (8). To account for the vocative particle, *yaa*, I argue that the vocative head moves out of a shell (similar to a vP shell (Larson, 2014)), where that vocative particle occupies the specifier position of this shell.

8 Conclusion

The purpose of the current study was to determine the phrase structure of Vocatives of Entreaty in Standard Arabic. The most obvious finding to emerge from this study is that a Vocative of Entreaty selects prepositional phrases as its arguments. The second major finding was that the addressee in Vocatives of Entreaty is marked with a genitive case as a result of being c-commanded by a preposition. This study has raised important questions about the nature of vocatives because this study shows that a vocative phrase may appear with two arguments. This questions the idea proposed in earlier studies that a vocative head is a functional head. This study shows that a vocative can be made up of a lexical head/shell. This new understanding should help to improve predictions of the phrase structure of vocatives; vocative phrases can be both lexical and functional phrases.

References

Abdelhady, Saleem. (2021). A Minimalist analysis of jā for coordination in Jordanian Arabic. *SKASE Journal of Theoretical Linguistics*. 59-79.
Abdelhady, Saleem. (2020a). *The syntax of ostensible categories in Arabic*. PhD. Memorial University of Newfoundland. Canada.
Abdelhady, Saleem. (2020a). Grounding Valuation in Minimalism: The Case of Ostensible Complementizers in Arabic (p. 1-20). In Abdelhady, Saleem, Fatema, Tenzella, & Cucinelli, Alex. *Memorial University of Newfoundland Occasional Papers in Linguistics*. Memorial University of Newfoundland.

Abdel-Hady, Saleem, & Branigan, Phil. (2020). The impact of speech act projections on categorization: Evidence from ostensible lexical categories in Arabic. *The Syntax-Pragmatics Interface in Generative Grammar*, 1-21.
Abuladze, Lia, and Andreas Ludden. (2013). The vocative in Georgian. *Vocative*s (1).25–42.
Akku, Faruk. & Virginia, Hill. (2018). Speaker Visibility in Syntax. *Proceedings of the 35th West Coast Conference on Formal Linguistics*, ed. Wm. G. Bennett et al., 49-58. Somerville, MA: Cascadilla Proceedings Project.16.
Al-Samirai, Faderl. (2000). *Kitab manyani ʔal naħu*. Jordan: Dar Al-Fikir.
Chomsky, N. (1995). *The Minimalist Program*. MIT Press.
González López, Laura. (2020). Vocatives with determiners: the case of vocatives preceded by possessives. Isogloss. *Open Journal of Romance Linguistics* (6), 1-30. doi:10.5565/rev/isogloss.59.
Haddad, Yousef. (2020). Vocatives as parenthetical adjuncts: Evidence from Arabic. *Glossa: a journal of general linguistics* 5. Ubiquity Press.132. doi:10.5334/gjgl.1302.
Hill, Virginia. (2007). Vocatives and the pragmatics–syntax interface. *Lingua* (117).2077–2105. doi:10.1016/j.lingua.2007.01.002.
Hill, Virginia. (2013). *Vocatives: How Syntax meets with Pragmatics*. Netherlands: Brill.
Hill, Virginia. (2022). The syntactization of kinship in vocative phrases. *Glossa: a journal of general linguistics* 7. Open Library of Humanities. doi:10.16995/glossa.6557. https://www.glossa-journal.org/article/id/6557/.
Jaradat, Abdulazeez Ahmad, Bassil Mashaqba, and Anas Huneety. (2022). On Pragmatics-syntax Interface: The case of vocative nominals in Jordanian Arabic. *Jordan Journal of Modern Languages and Literatures* 14.351–370. doi:10.47012/jjmll.14.2.7.
Larson, Richard. (2014). *On Shell Structure* (1 edition). New York and London: Routledge.
Moutaouakil, Ahmed. (2014). *Pragmatic functions in a functional grammar of Arabic*. Walter de Gruyter: Germany.
Ross, John Robert. "Where to do things with words." In *Speech Acts*, pp. 233-256. Brill, 1975.
Shormani, Qrabshe., & Mohammed Qrabshe. (2018). Vocatives: correlating the syntax and discourse at the interface. *Cogent Arts & Humanities* 5. doi:10.1080/23311983.2018.1469388.
Soltan, Usama. (2016). *The syntax of vocatives in Egyptian Arabic*. Middlebury College.
Speas, Peggy, and Carol Tenny. (2003). Configurational properties of point of view roles. *Linguistik Aktuell/Linguistics Today*, ed. by Anna Maria Di Sciullo, 57:315–344. Amsterdam: John Benjamins Publishing Company. doi:10.1075/la.57.15spe. https://benjamins.com/catalog/la.57.15spe.

Context Matters: The Interaction of Phonetics and Phonology in Producing Output Forms

Ankita Prasad
Indian Institute of Technology, Delhi

The velar fricative [x] in Sylheti alternates with the stop [k] in the presence of high vowels [i] and [u] (e.g., [dɛx] 'see.IMP' [dek-i] 'see.1.PRS') (Eden 2020; Prasad 2021). This paper reports the findings from a read connected speech task to test the distribution of the two allophones. The findings showed that the voiceless velar obstruent /K/ had six possible output forms: |x|, |k, |kʰ|, |g|, |ɣ|, and |h|. The surface distribution of these sounds is produced by the interaction of a phonological rule with natural phonetic processes that affect connected speech. The distributional restrictions on output forms in the high vowel and non-high vowel conditions were also used to argue that the rule is triggered by the presence of the high vowels [i] and [u], which ban an adjacent velar fricative [x], rather than the non-high set.

Keywords: Sylheti language, connected speech, fortition process, positional effects

1 Introduction

Sylheti [sileṭi] is an Eastern Indo-Aryan language spoken mainly in the Sylhet division of Bangladesh and the bordering states of Assam and Tripura in north-eastern India. Grierson (1903) and Chatterjee (1926) categorize it as a dialect of Eastern Bangla. They noted that its phoneme inventory differed significantly from Standard Bangla. Some sound changes they noted included the deaspiration and spirantization of stops. The loss of the aspiration contrast resulted in the merger of previously distinctive sounds. Spirantization further increased the perceptual difference between the voiceless and voiced obstruents as voiceless stops were the usual targets of this process.

This study looks at the fricative-stop alternation in one place of articulation: the velar. The velar stop [k] surfaces only adjacent to the high vowel [i] or [u]. It surfaces as the fricative [x] adjacent to [-high] vowels. The other voiceless fricatives ([ɸ] is used as an example in 1b. and 2b. are not affected by the vowel quality of the following vowel as illustrated below:

1. a. [ki̲ta] 'what' b. [ɸi̲az] 'onion'
2. a. [xe̲] 'who' b. [ɸa̲] 'foot'

The effect of the high vowel appears to be phonological as it can be formalized as a rewrite rule of the form A→B／_C (Chomsky and Halle 1968). In testing the application of this rule, it was found that the voiceless velar obstruent has six variants that show up consistently. These phonetic realizations are |x|, |k|, |kʰ|, |g|, |ɣ|, and |h|. In this paper, I argue that the surface distribution of these forms is determined by the interaction of the phonological rewrite rule and several phonetic processes, including voicing assimilation, spirantization, and debuccalization.

A data elicitation task with read connected speech was used. The study looked at the effect of two types of phonological environments: (i) quality of the neighboring vowel and (ii) prosodic position: word-initial

* I would like to thank the residents of Lakhipur who helped me organise and participated in the study and were invaluable collaborators. I also thank Sayantan Mandal for his helpful suggestion of an underspecification account for the data at the conference, which has been included in the paper.

onset, word-medial coda (pre-consonantal), word-medial onset (post-consonantal), intervocalic, and word-final. This study aimed to identify the phonetic outputs corresponding to /K/ and their contextual restrictions. It also aimed to determine the directionality of the process, whether it was fortition to a stop or spirantization to a plosive. In short, which value of [cont] was specified in the mental representation of the voiceless velar obstruent in the Sylheti speakers' grammar?

The paper is divided into the following sections. Section 2 introduces the data and three possible analyses of the alternation: as spirantization with output fricative, fortition with output stop, or an underspecification account with contextual allophony. Section 3 introduces the experiment, a connected speech reading task. It reports the findings of the PRAAT analysis of the recordings and graphically presents the findings. Section 4 discusses the experiment's findings in relation to the research question posed in Section 2. Section 5 concludes the paper.

2 The Problem: Directionality of the Alternation

The two allophones of the voiceless velar obstruent [x] and [k] have a complementary distribution in Sylheti. [k] occurs with high vowels, in geminates, as well as following some consonants. [x] occurs adjacent to low vowels. Table 1 summarizes the environments in which [k] and [x] can surface.

Table 1: Distributional restrictions on [k] and [x]

Allophone	[k]		[x]	
Environment	*HV*	*NHV*	*HV*	*NHV*
Initial onset	✓	×	×	✓
Medial onset	✓	×/?	×	✓
Medial coda	✓	×	×	✓
Intervocalic	✓	×	×	✓
Final coda	✓	×	×	✓
Geminate	✓	✓	×	×

There are two possible characterizations of this alternation:

3. Spirantization: k → x / [-high V]
4. Fortition: x → k / [+high V]

I had previously argued in Prasad (2021) that this was a process of fortition due to the requirement of maintaining a sonority slope within the syllable between the syllable margins and peak. However, examples like 5 demonstrate that the rule only looks at the linear string irrespective of the syllable structure. This restriction also provides evidence against the spirantization account as it clearly shows that the high vowel is the trigger for the rule. The presence of a high vowel in a preceding open syllable also leads to fortition:

5. [b**ik**el] 'evening'

A third analysis is an underspecification account with contextual allophony. In this analysis, the voiceless velar lacks an underlying specification for continuancy ([cont]), and the value of this feature is contextually determined. This analysis requires two rules:

6. Archiphoneme

a. K → [+cont] / [-high V]

b. K → [-cont] / [+high V]

This analysis contains both the fortition and spirantization rules as possible rules. I follow the archiphoneme analysis to as it does not require the input to be fully specified for continuancy. Instead, one of the rules becomes redundant and can be stated as an elsewhere condition instead. However, the important factor in deciding between the two is which of the two environments is more phonologically active and likely to trigger the realization of /K/ with the requisite feature value for [cont]. I looked at the surface realizations (phonetic form) of the underlying voiceless velar /K/ in Sylheti in different contexts in a read connected speech task to determine which of these rules could better capture the data. Although the prosodic position does not play a role in the phonological alternation, these positions are subject to other processes that affect the phonetic output.

3 The Experiment: Analysis of /K/ Tokens in Connected Speech

Data was elicited from Sylheti speakers using a connected speech reading task. Each participant was given a text consisting of Sylheti text transliterated in Bangla script. No information was given about the sounds under study. Speakers were asked to read out the two dialogues naturally. If they made a mistake, they could start again from the same line, and that portion was discarded from the analysis. They were then recorded, and analysis was conducted using PRAAT.

The text was prepared with help from a native Sylheti speaker (not part of the experiment), who checked the list of target words, some of which were drawn from the SOAS dictionary (https://sylhetiproject.wordpress.com/dictionary/), and some collected by the author during field visits to Silchar, Assam. The text consisted of texts interspersed with words containing the target segment /K/. The phonological environment of the segment was manipulated. The prosodic contexts included initial, medial (intervocalic, pre-consonantal and post-consonantal) and final position. The linear context included adjacent high and low vowels. Some combinations like low vowel+word final were missing as appropriate tokens that fit the discourse context could not be found.

Since Bangla has two graphemes, ক [kɔ] and খ [kʰɔ], that correspond to the Sylheti /K/; tokens with both graphemes were used in the task. The spellings followed the Bangla convention where applicable. Some verbal alternations were also included to test whether the allophony operated across morphological boundaries (*rakʰ-i* [rak-i] 'keep.2P.PRS.CONT' *rakʰ-so* [rax-so] 'keep.2P.PST'). The table alongside summarizes the number of tokens with orthographic ক (*kɔ*) and খ (*kʰɔ*) and the environment of the segment: high vowel (HV), non-high vowel (NHV) and geminate (G).

Table 2: Vowel context by orthography

Context	**ক ko**	**খ kho**	**total**
HV	15	17	32
NHV	28	14	42
G	1	-	1
Total	44	31	75 tokens

This experiment aimed to answer the following questions regarding the mental representation of the velars. The first question is: How productive is the velar fricative-stop alternation? Does it apply partially, or is it an inviolable rule? The second aim was to study the effect of the neighboring vowel and its features on the realization of the voiceless velar obstruent as a phonological rule that operates at the word level. The third goal was to study the realization of the velar obstruent in connected speech to study the effect of utterance-level processes like the voicing assimilation shown in example 7., which were observed in the data.

7. /mɔn-o kor-i-a/ [mɔno xoria] |mɔn-o-ɣorija|[2]
   mind-LOC do-PST.PFV
   'keep it in mind'

3.1 *Data Collection* The fieldwork was conducted in Fulertol village in the Lakhipur sub-division of Cachar district, Assam, India. Sylheti is widely spoken in the region, though the region's official language is Bangla (https://cachar.gov.in/information-services/district-at-glance). Lakhipur lies to the east of the Sylhet region in Bangladesh, about 100 kilometers from the city of Sylhet and 40 kilometers from Silchar city in Assam. This area also bordersthe neighboring states of Mizoram and Manipur. The variety spoken here is in contact with Manipuri and Naga languages, apart from the state language, Assamese.

3.2 *Recordings* The recordings were made at the site in Fulertol. The laboratory was a quiet room in a guest house. A Maono AU-PM461TR USB Condenser Microphone was used to record the speakers. The text to be read and the microphone were placed on a table Once the participant was familiar with the text, they were asked to test the microphone to make them comfortable with the set-up. They were recorded using PRAAT at a sampling frequency of 44100 Hz per second in .wav format with a bit rate of 705 kbps (16 bits). Every effort was made to reduce background noise; however, some ambient noise was present in the recordings. Techniques for analyzing the recordings included visual inspection of the sound wave, and the researcher's intuition. These were used to segment the sounds and assign each one a phonetic symbol. The data was then analyzed for the distribution of target sounds.

[1] Note that slashes are used to represent the phonological input /x/, brackets [x] for phonological output, and bars |x| for phonetic output throughout this text. These are used to signal different levels of representation.

3.3 *Participants* Four informants participated in the study (3f, 1m; Average age: 43 yrs.). All of them reported using Sylheti at home regularly and rated themselves as proficient Sylheti speakers. They were all also speakers of Bangla, as that was the medium of education. Two also spoke English, Assamese, and Hindi with varying proficiency. All resided in the Lakhipur sub-division of the Silchar district, though some were from different villages from the area. All speakers were fluent in the Bangla script, with reading fluency up to a minimum of 9th grade in Bengali. The speaker data is summarised in Table 3.

Table 3: Participant Details

	Sex	Age	Edu Qualification	Languages Known	Village
PD	M	35	MA	Sylheti, Bangla, Assamese, English	Fulertol
CD	F	47	9th grade	Bangla, Sylheti	Jirighat
CHD	F	35	10th grade	Bangla, Sylheti	Harinagar
BD	F	58	MA	Sylheti, Bangla, Assamese, English, Hindi	Fulertol

3.4 *Materials* The task involved reading out connected speech containing the target words. The text was in the form of two short scenes with dialogues between two interlocutors. Using a dialogue format, I aimed to elicit less careful, connected speech from the speakers. The speakers were presented Sylheti texts written in Bangla script. Words in common with standard Bangla were spelt according to Bangla spelling rules using the two graphs ক (*kɔ*) and খ (*kʰɔ*). The text was widely spaced and printed in a large font to accommodate easy reading.

Warner (2012) categorizes speech using three dimensions: carefulness, speech rate, and acoustic reduction. The interactions of these three and their usage by speakers to different degrees can be used to characterize most speech styles. Connected speech (both read and non-read) shows more acoustic reduction as well as a higher speech rate than unconnected speech, including read and prompted isolated words, because its carefulness value is lower. The experiments used connected speech to reduce interference from standard Bangla. In the pilot study, using word lists resulted in the production of forms that were closer to Bangla. It was hoped that lower carefulness would result in more naturalistic pronunciations, while ensuring parity between participants.

There were two dialogues. Each had a title that described the situational context within which two people were conversing. The first was between a husband and wife discussing an upcoming trip. The second was a conversation between two women at the market. The setting and content of the text was chosen to be familiar to the speakers. An extract is given below with orthographic transcription and glossing. The target words have been highlighted in (8). Expected [x] are highlighted in blue, and [k] in orange.

8. An example of the material used in the text from Dialogue 1:

রুমায়	জিগাইলা:	তুমি	কবে	যাইতায়?
ruma-e	jigai-la	tumi	kob-e	ja-i-taj?[3]
ruma-FOC	ask.PST	you	when-LOC	go-2.FUT

‘Ruma asked: when are you going?’

পরিমলএ	কৈইলা:	'কাইল	বিকালে	যাইতাম
porimol-e	koi-la	kail	bikal-e	ja-i-tam
Porimol-FOC	ask-PST	tomorrow	evening-LOC	go-1.FUT

‘Porimol said: (I am) going tomorrow evening.’

রুমা:	তুমি	হক্কলতা	রাখসো	নি?
ruma:	tumi	hok:ol-ta	rakʰ-so	ni?
Ruma:	you	all-CLF	keep-2.PST	Q

‘Ruma: Have you kept everything?’

The sentences were spoken as part of a discourse. The characters had certain social roles attached to them, and the entire text took about two minutes to read, leading to speakers moving into less careful speech over the duration of the task. They showed the effect of some reduction processes, such as assimilation and deletion that are characteristic of more casual speech styles. They also deleted some words or changed them,

[2] The transcription uses ‘j’ for the palatal affricate d͡ʒ.

leading to three missing tokens. I take this as evidence that speakers spoke more naturally or less carefully in this scenario.

3.5 *Hypothesis* The primary aim of the experiment was to study the effect of the vowel features [+high] and [-high] on the realization of the voiceless velar obstruent. It also aimed to establish whether this alternation is a spirantization or fortition rule. The target of the rule is the segment /K/. If the trigger is the [-high] vowel (i.e., spirantization rule), then we would not find any output [k]'s in this environment as they would all have been changed. However, there could be some output [x]'s in the [+high] environment as there was no rule targeting [x]'s in this environment. The reverse would be true of the [+high] vowel trigger.

9. Hypothesis 1: The [+high] vowel is the trigger for the rule. The presence of [+high vowel] predicts [-cont] output for K in this environment with no exceptions)
   Prediction: No |x| in [+high]
10. Hypothesis 2: The [-high] vowel is the trigger for the rule. The presence of [-high vowel] predicts [-cont] output for K in this environment with no exceptions)
    Prediction: No |k| in [-high]
11. Hypothesis 3: The different orthographic forms ক [kɔ] and খ [kʰɔ] map to 1 phoneme. The orthography and presence of the aspirate grapheme has no correlation with the realization of /K/. The historical aspirate does not correlate with the fricative alternant or vice versa.

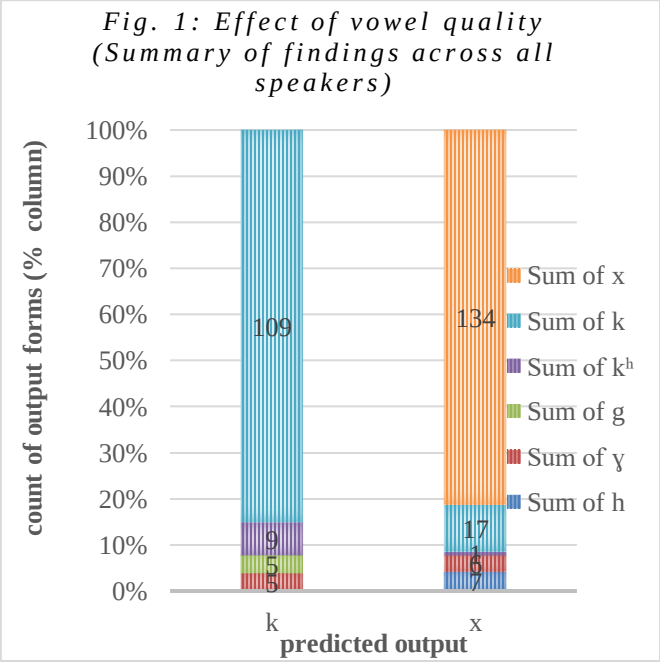
There was no hypothesis regarding the effect of preceding and following consonants (and their features) though effects like intervocalic spirantization, deletion, and voicing assimilation were expected.

3.6 *Results* We first look at the phonetic realizations of /K/ in the [+high] and [-high] vowel contexts. The list of target words included words with both [i] and [u] following and preceding the target /K/. There were 165 tokens in the [-high] context and 128 tokens in the [+high]. The main results are summarized in Fig. 1. In the [+high] context (total 128 tokens) /K/ surfaced as:

- 85.1% (109) as a stop |k|
- 7% (9) as aspirated stop |kʰ|; word-initially (e.g, |kʰuratːo| 'cousin')
- 3.9% (5 each) as voiced stop |g| following a voiced obstruent (|ʃig-bo| 'learn.FUT'), and voiced fricative |ɣ| between vowels including across word boundaries (|t̪iɣ#ase| 'okay be.PRS')

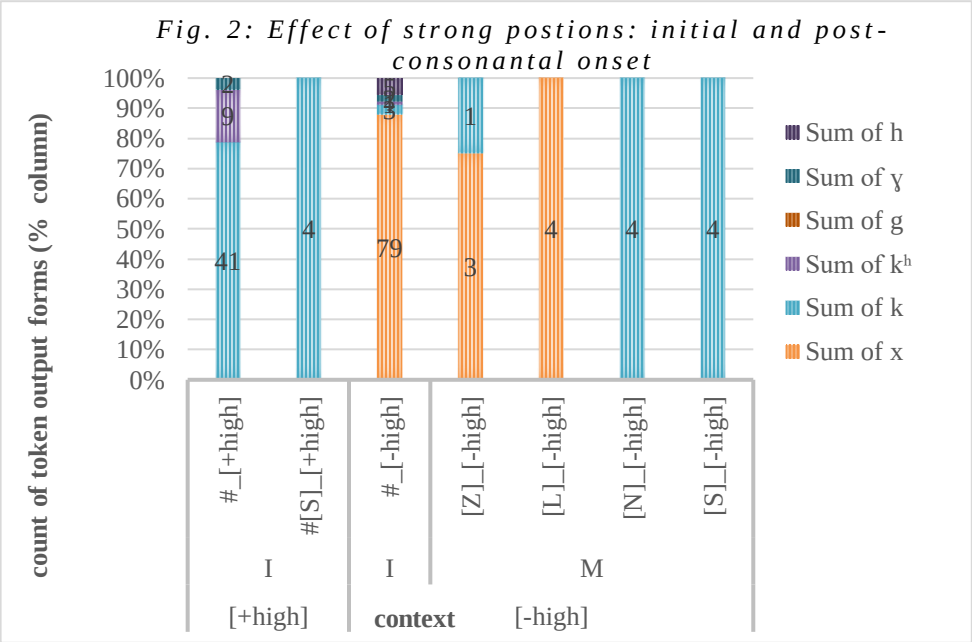
Outputs in [-high] (total 165 tokens):

- 81.2% (134) as a fricative |x|
- 10.3% (17) as a stop |k|. 8 of these instances were post-consonantal (e.g., |ɔŋko| 'number')
- 3.6% (6) as a fricative |ɣ|; following a voiced stop as well as intervocalically across word boundaries (|ɛɣ-din| 'one day')
- 4.2% (7) as |h| mostly initial with 2 word-medial tokens as well
- |kʰ| (1); stressed and word-initial



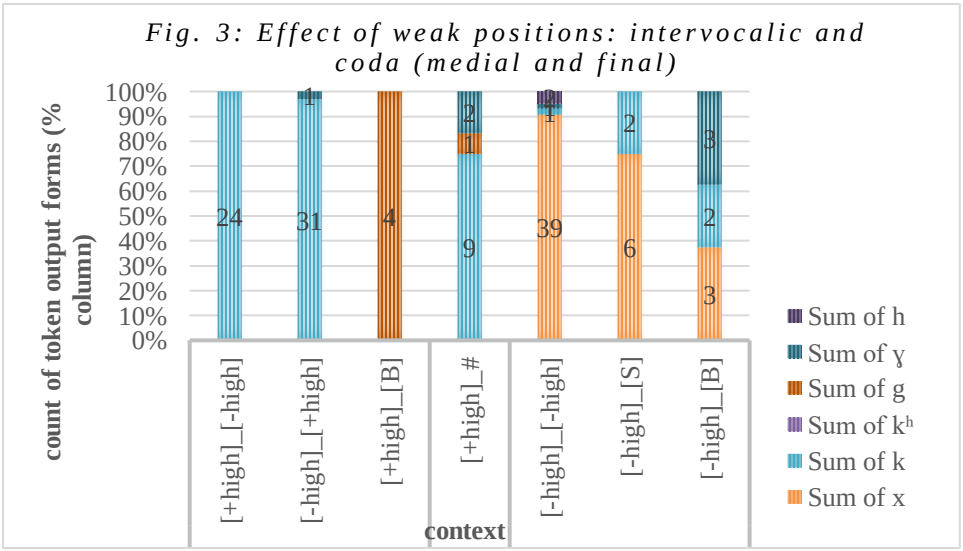
The fricatives |x| and |h| are not possible outputs for predicted [k]'s. The voiced fricative |ɣ| is the only fricative that could surface as an output for k. However, this only occurs when spirantized intervocalically across word boundaries, like the example in (3). However, output |k| never surfaces in the non-high vowel condition. This lines up with the prediction in Hypothesis 1 that the high vowel triggers a fortition process or isthe conditioning environment that requires the velar to surface as a stop. The fact that |x| and |h| do not surface in the high vowel environment shows that the position requires a [-cont] segment. However, the reverse is not true, as the stop |k| does surface in the [-high] vowel environment. Next, I look at the effect of

prosodic position on the phonetic realization of /K/. First, I look at the strong positions: word-initial onset and post-consonantal onset.



In the strong positions in Fig.2 above, the word-initial position was the most permissive, allowing output: |x|, |k|, |ɣ|, |h|, |kʰ|. Preceding a high vowel, only the two stops |k| and |kʰ| and the voiced fricative |ɣ| could surface. Post-consonantal position showed a preference for fricative |x| following a sonorant and voiced obstruent, while the position following a nasal [ŋ] and voiceless fricative [ʃ] permitted only the stop |k|..

In the weaker positions in Fig.3 below, there was only one token with word-final target /K/. This was preceded by a high vowel [i] and showed both voicing and spirantization to |ɣ| in two tokens out of 12. In the intervocalic position, the effect of the high vowel was strong with mostly stop |k| outputs. This condition shows that a preceding [+high] vowel also triggers fortition even if the stop is followed by a tautosyllabic [-high] vowel. That is, the presence of the [+high] takes precedence of a [-high] vowel in determining the output. In the pre-consonantal coda [_B], there was obligatory voicing assimilation in the high vowel context with output |g|. In the low vowel context preceding obstruents [_S], [_B], one speaker consistently used the stop |k|. There was also voicing assimilation preceding the voiced obstruent [b] with one three tokens of output |ɣ| in this context.



4 Discussion

I argue that the phonological rule, i.e., fortition, categorizes the phonological outputs of the velar into the two contextual allophones: (i) [k] in the high vowel condition and (ii) [x] elsewhere. The value of [cont] is specified at the time the word enters the phonetic module, which results in the asymmetry of forms observed on the surface. This also explains the output |k| in [-high] context as these may have been specified [-cont] at the phonological level and hence surface as stops. The phonetic module then applies to produce the variable outputs that are subject to articulatory constraints. These processes are categorized as phonetic processes in contrast to the phonological rule as they are sporadic and do not show the same degree of rule application. They also have a phonetically gounded explanation.

The initial position attracts aspiration as the |kʰ| is mostly found in the initial position. Additionally, |kʰ| occurred with both orthographic ক (*kɔ*) and খ (*kʰɔ*), though it did show a preference for orthographic *kʰɔ* in some lexical items. However, it is preferred in the high vowel environment as a variant of the [k] allophone, rather than [x].

All four speakers had surface |g| following a high vowel and usually preceding a voiced stop or word-finally. |g| was only observed in coda positions. Additionally, it only occurred with the high vowel [i]. This suggests that only allophone [k] can surface as phonetic output |g|. This variant occurred because of voicing assimilation. What is notable is that unlike the other non-canonical output forms, this output was observed in the same word (*ʃikʰ-bo* ⌊ʃigbo⌋ 'learn.FUT') across all participants, showing that this is a regular process in the language.

|ɣ̞| surfaced in two environments: intervocalic (both across and within word boundaries) and when preceded by a non-high vowel and followed by a voiced plosive. Unlike the other fricatives, it occurred with both [+high] and [-high] vowels. It occurred as a result of intervocalic spirantization as well as voicing assimilation. Its application was inconsistent as different speakers used the |ɣ| form in different words. The environment of application was always between two voiced sounds. |h| surfaces only in the non-high condition, thus, it is exclusive to the [x] allophone. It is usually initial but also surfaces intervocalically in two instances.

Output |k| in the low vowel condition rarely occurred, with some in word initial and few in intervocalic position. There were some contexts with predictable |k| in the low vowel condition, where all speakers produced a stop. The first was in the geminate condition. It also surfaced as a stop in medial consonant sequences following [ŋ] and [ʃ]. Thus, in opposition to sonorant [l] and voiced [b], these consonants cause fortition along with the high vowel. However, one speaker consistently used |k| also following these consonants. The study thus also revises the hypothesis that only the high vowel context triggers fortition, as there seems to be a clear distinction between consonants that trigger fortition and consonants that don't. The data in this study, however, was inadequate as it did not have all possible combinations of the velar in medial consonant sequences. This is a good subject for further research as the limited data exhibited interesting variation patterns between speakers.

5 Conclusion

This paper showed the restrictions on the distribution of phonetic realizations of the velar in connected speech. The aspirated stop and the glottal fricative, both [+spread glottis] sounds are restricted to the word-initial position. The most important factor determining the phonetic output is whether a high vowel is adjacent to it. An adjacent high vowel implies that the segment will be realized as a stop. This condition however, is affected by the phonetic spirantization rule between two voiced sounds. In the intervocalic context, even allophone [k] which is specified as [-cont] can be realized as a voiced fricative |ɣ| losing both its continuancy and voicing specification. However, preceding a voiced sound, it undergoes only voicing assimilation and it surfaces as |g|, preserving its [-cont] specification. Orthography does not appear to play a major role in determining the output form. Its effect is limited to certain lexical items, and it is overall not a predictor of the output form that will surface.

References

Boersma, Paul & Weenink, David (2023). Praat: doing phonetics by computer [Computer program]. Version 6.3.05, retrieved 27 January 2023 from http://www.praat.org/http://www.praat.org/.

Chatterjee, Suniti Kumar. (1926). *The origin and development of the Bengali language*. Clacutta: Calcutta University Press.

Chomsky, Noam, and Morris Halle. (1968). *The Sound Pattern of English*. New York: Harper & Row. http://web.mit.edu/morrishalle/pubworks/papers/1968_Chomsky_Halle_The_Sound_Pattern_of_English.pdf (10

Grierson, George Abraham. (1903). *Linguistic Survey of India*, vol. 5(1). Office of the Superintendent, Government Printing, India.

Eden, S.Elizabeth, (2020). Does Camden Sylheti have consonant clusters? *Language documentation and description*, *18*, pp.23-41.

Prasad, Ankita. (2021). Fortition in marked-on-marked contexts: velar stopping in Sylheti. *Formal Approaches to South Asian Languages-9,* (FASAL). Deepak Alok & Sreekar Raghotham (eds.). Portland: Reed College. Portal; University of Konstanz. https://ojs.ub.uni-konstanz.de/jsal/index.php/fasal/article/view/202 (11 September 2022)

Tunga, Sudhaṃśu Sekhara. (1995). *Bengali and other related dialects of South Assam*. Mittal Publications.

Warner, Natasha. (2012). Methods for studying spontaneous speech. In *The Oxford Handbook of Laboratory Phonology* (pp. 621-634). Oxford: Oxford University Press.

An Interplay Between Temporal Adverbials and Desire Predicatives in Bangla: Exploring the Heterogeneity

Debadatta Roychowdhury and Samir Karmakar
School of Languages and Linguistics, Jadavpur University

There is a fair share of 'heterogeneity' that prevails among the desire predicates in Bangla (a.k.a Bengali: Indo-Aryan). On the ground of semantics, this work offers a window showing the diverseness of the desire predicate *itʃtʃʰe* 'wish' depending on the selection of the light verbs. In Bangla, the predicate *itʃtʃʰe* occurs independently and also with the light verbs *kɔr-* 'do' and *hɔ-* 'happen'; forming *itʃtʃʰe kɔra* and *itʃtʃʰe hɔwa*. In this work we present a novel observation explicating the heterogeneity that exists among the members of the same set. These lexical items being under the same umbrella term 'desire predicate' explicate diverse semantics based on the temporal modifications.

Keywords: Frequency temporal adverbs, Gappy states, Interval semantics, Conjunct verbs.

1 Introduction and Research Objectives

Bangla like many other South Asian languages incorporates a strong use of light verbs. The desire predicate we are concerned about is the use of *itʃtʃʰe* 'wish' with the light verbs *kɔr-* 'do' and *hɔ-* 'happen'; forming *itʃtʃʰe kɔra* 'wish do' and *itʃtʃʰe hɔwa* 'wish happen' respectively. These three predicates share a common word *itʃtʃʰe* with a thread commonality i.e. on a basic semantic level they all signify the sense of 'desire'. This general understanding will not justify a dense semantic analysis that concerns the heterogeneity that exists among them at different levels. The first level revolves around the syntactic domain concerning the clausal dependencies. The predicates *itʃtʃʰe* and *itʃtʃʰe kɔra* / *hɔwa* employ different clausal dependencies. When *itʃtʃʰe* sits alone in a sentence, it shows alliance only with the 'subjunctive' form of the verb. The predicate *itʃtʃʰe kɔra* selects only the infinitival form of the verb and *itʃtʃʰe hɔwa* occurs both with an infinitival clausal complement and a genitive-gerundive complement. Observe the examples below.

(1) *amar* ***itʃtʃʰe*** *ami* *gan* *ʃikʰ-i*
I.GEN wish I.NOM music learn-SUBJN.1
'It is my desire that I learn music.'

(2) *amar* *gan* *ʃikʰ-ɟe* ***itʃtʃʰe*** ***kɔr-e / (hɔ-e)***
I.GEN music learn-INF wish do.PRS-3
'I wish to learn music.'

(3) *amar* *gan* *ʃekʰ-a-r* ***itʃtʃʰe*** ***hɔ-e***
I.GEN music learn-GEN-GER wish happen.PRS-3
'I wish to learn music.'

(4) * *amar* *gan* *ʃekʰ-a-r* ***itʃtʃʰe*** ***kɔr-e***
I.GEN music learn-GEN-GER wish do.PRS-3
'I wish to learn music.'

As said above, in (1) the predicate *itʃtʃʰe* sits with the subjunctive form of the verb. In (2) *itʃtʃʰe kɔra* (as

well as *it̪ʃ/t̪ʃʰe hɔwa*) occur with the infinitival clausal complement. The predicate *it̪ʃ/t̪ʃʰe hɔwa* can also occur with a genitive gerundive clausal complement, seen in (3). On the contrary *it̪ʃ/t̪ʃʰe kɔra* cannot occur with a genitive gerundive complement as seen in (4), therefore the predicate shows its dependency only with the infinitival clausal complement.

The embeddability properties of the predicate *it̪ʃ/t̪ʃʰe* are motivated by the selection of the light verbs as evident from examples (1-4). We have tried to represent this in table (1) below.

Table 1- The embeddability properties of *it̪ʃ/t̪ʃʰe* (with the light verbs)

Predicate	**Light verbs**	**Clausal complements**
it̪ʃ/t̪ʃʰe	X	Subjunctive
it̪ʃ/t̪ʃʰe	*hɔwa*	Infinitival Genitive-Gerundive
it̪ʃ/t̪ʃʰe	*kɔra*	Infinitival

In this work, we are not going to discuss the syntactic selection restrictions on the clausal complement structures offered by the predicates *it̪ʃ/t̪ʃʰe* and *it̪ʃ/t̪ʃʰe kɔra* / *hɔwa*. As the title suggests, this work will deal with the semantic diversity brought by these predicates in the context of temporal adverbials. Therefore, we need to address the peculiarity incorporated by the predicates when they sit with the 'frequency temporal adverbs'. Observe the examples below.

(5) *amar* ***madʒʰe madʒʰe*** *gan* *ʃikʰ-ɻe* ***it̪ʃ/t̪ʃʰe*** ***kɔr-e / (hɔ-e)***
I.GEN sometimes music learn-INF wish do.PRS-3
'Sometimes, I wish I could learn music.'

(6) ******* *amar* ***madʒʰe madʒʰe*** ***it̪ʃ/t̪ʃʰe*** ami *gan* *ʃikʰ-i*
I.GEN sometimes wish I.NOM music learn-SUBJN.1
'Sometimes, it is my desire that I learn music.' (A very inadequate translation)

In the examples above we have taken the frequency temporal adverb *madʒʰe madʒʰe* 'sometimes' to bring the divergence between the two predicates. As evident the *it̪ʃ/t̪ʃʰe kɔra* / *hɔwa* in (5) can sit well with the temporal adverb, but *it̪ʃ/t̪ʃʰe* in (6) when occurs independently in a sentence cannot occur with the frequency adverb.

At this juncture, we would like to address the major avenue of interest that revolves around these predicates. The research objectives concerning this work are stated below.

i) Understanding the thread of divergence created by the interaction of the temporal adverbs and desire predicates with (without) the light verbs.

ii) How do these adverbs with the association of the light verbs build a piece of strong evidence regarding the states reflecting gaps, resulting in **'gappy states'**?

2 Revising the literature

If we go by the traditional categorization then 'adverbials of time' signifies three senses of time; 'when', 'how often', and 'how long'. The adverbials that signify the meaning of 'when' are yesterday, today, tomorrow, later, now, etc. Temporal adverbs that convey 'how often' an event occurs are mainly sometimes, often, rarely, usually, regularly, etc. Adverbs that denote the sense of 'how long' are all day, for a year, since 1990, since forever, etc.

The present work will only consider 'frequency adverbials' expressing 'how often' a state holds or an event occurs. As the 'gaps' that we are interested in are implied by the frequency adverbs of time. Parsons (1990) the frequency adverbs of time provide quantification over the temporal intervals. He termed them as 'proportional adverbials'. They act as restricted quantifiers and their logical form must take restrictive clauses. A sentence like 'Mary usually walks with John' means (Most t) [Mary walks at t, Mary walks with John at t]. In this particular case the quantifier is equivalent to most and here is restricted to the times at which Mary walks. The logical form implies that the proportion of time Mary walks with John to times when she walks is high. Binnick (1991) tells that the frequency adverbs assert how many per unit of time

something occurred. Swart (1993) introduced a term called 'Quantificational Adverbs' (QADVs) and included 'iterative', ' frequentative', and 'generic' adverbials in the set. Iterative QADVs are adverbials like once, twice, and several times count events that occur in a given time frame like last month. The second type which is frequentative QADVs is adverbials like often, always, sometimes, never, and seldom. The third type is generic QADVs like generally, normally, usually, etc. Bennett and Partee (2004) when we tend to be indefinite or vague about the number of times that a generic event occurred, we refer to the plural quantifiers and use adverbials like regularly, on few occasions, sometimes, etc. They have further demarcated the adverbials concerning the units of time. Adverbs like always, regularly, continually, repeatedly, and at regular intervals are ways of expressing that a generic event occurred once for every unit of time where the unit is not specified. If we want to remain indefinite or vague about the number of repetitions for some specified unit of time by using expressions such as; a few times a week, several times a month, or many times a year. We can also be indefinite or vague about both the number of repetitions and the unit of time, and then we use expressions like seldom, occasionally, often, and frequently. Let us take the following examples (7-9) from their work. Observe below.

(7) John frequently smokes.

They regarded (7) to be asserting something like (8) below.

(8) John smokes many times each ɑ.

The variable ɑ ranges over units of time. Sentence (7) has two 'free variables' – one for what constitutes many and the other concerning what constitutes the unit of time. The sentence in (8) can be evaluated at a time interval for an assignment for the 'free variables'. Hence (9) below will be appropriate to evaluate.

(9) John smokes at least ten times each hour.

We can quite well see that the 'free variables' are assigned with expressions and now if (9) is true at an interval of time I, then (9) is true at every subinterval of I.

Bennett and Partee (2004) categorized verbs into three types; stative verb phrases, subinterval verb phrases, and nonstative, nonsubinterval verb phrases.

Stative verb phrases are verb phrases that do not take the progressive form. Verb phrases like 'be happy', 'love Marry', and 'believe that Mary walks' – are purely stative.

Subinterval verb phrases have the property that if they are the main verb phrase of a sentence which is true at some interval of time I, then the sentence is true at every subinterval of I including every moment of time in I. Examples of subinterval verb phrases are: walk, breathe, walk in the park, and push a cart.

Nonstative, nonsubinterval verb phrases are neither stative nor subinterval. Examples are 'die', 'build a house', 'catch a fish', 'walk to Rome', etc. Since they take the progressive form, they are nonstative. They are also nonsubinterval. In a situation like 'walk to Rome', if it took an hour to walk to Rome then one did not walk to Rome within the first thirty minutes of the hour.

According to Bennett and Partee (2004) verb phrases modified by the frequency adverbs are nonstative. We are going to consider this view and show that Bangla conveys a different picture regarding the stative verbs reacting differently when they interact with the frequency adverbs.

3 Data and Analysis

In Bangla 'frequency' adverbs of time are compatible with the state verbs. We have shown an example in (5) above. Here we are going to observe a few more examples with other frequency adverbs occurring with *itʃtʃʰe kɔra* / *hɔwa*. Observe below.

(10)	*amar*	***praei***	*gan*	*ʃikʰ-t̪e*	***itʃtʃʰe***	***kɔr-e / (hɔ-e)***
	I.GEN	often	music	learn-INF	wish	do.PRS-3

'Often, I wish I could learn music.'

(11) *amar* ***ʃadʰaronɔ̰to*** *gan* *ʃikʰ-ḛ* ***itʃtʃʰe*** ***kɔr-e / (hɔ-e)***
I.GEN generally music learn-INF wish do.PRS-3
'Generally, I wish I could learn music.'

The frequency adverbs *praei* 'often' and *ʃadʰaronɔ̰to* 'generally' go well with the predicates *itʃtʃʰe kɔra / hɔwa*. All frequency adverbs are compatible with these predicates.

Now we will delve into the point of heterogeneity i.e. the predicate *itʃtʃʰe* (without the light verbs) cannot sit with frequency adverbs as we see in (6). Similarly, it will not occur with any adverbs indicating the frequency. Observe below.

(12) ******* *amar* ***praei*** ***itʃtʃʰe*** ami *gan* *ʃikʰ-i*
I.GEN often wish I.NOM music learn-SUBJN.1
'Often, it is my desire that I learn music.' (A very inadequate translation)

(13) ******* *amar* ***ʃadʰaronɔ̰to*** ***itʃtʃʰe*** ami *gan* *ʃikʰ-i*
I.GEN generally wish I.NOM music learn-SUBJN.1
'Generally, it is my desire that I learn music.' (A very inadequate translation)

As seen in (12), (13), and even in (6) above the adverbs of time denoting 'frequency' will not occur with *itʃtʃʰe* when it sits alone in a sentence without the light verbs.

The point of diversity that lies here conveys the fact that the predicate *itʃtʃʰe* can take modifications that has no 'discontinuity' or 'gaps'. Adverbs like 'since forever', 'for a long time', and 'since childhood' which carry the sense of a state which is 'extended now', can cope well with *itʃtʃʰe*. Adverbs like these have a starting point that is continued till the time of utterance. Observe the examples below.

(14) *amar* ***ɔnekdin-er*** ***itʃtʃʰe*** *ami* *gan* *ʃikʰ-i*
I.GEN a long time -GEN wish I.NOM music learn-SUBJN.1
'For a long time, it has been my desire to learn music.'

(15) *amar* ***tʃʰoto*** ***tʰeke*** ***itʃtʃʰe*** *ami* *gan* *ʃikʰ-i*
I.GEN childhood since wish I.NOM music learn-SUBJN.1
'Since childhood, it has been my desire to learn music.'

Now the examples above actually imply that ever since the subject felt a love for music she carried her desire to learn the art. So the starting point of her desire is extended to the time of utterance and there is no gap or discontinuity.

At this point, the interest revolves around a state being discontinuous and implying 'gaps', as we saw in the context of *itʃtʃʰe kɔra / hɔwa*. According to Vendler (1957) states last for a period of time i.e. they show no change of phase (**q**) over a period of time (**t**). They show no behaviour of 'discontinuity'. Somewhat like the figure below.

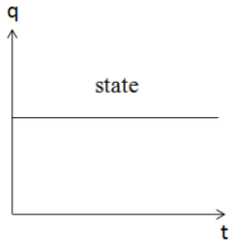


Figure 1: Vendlarian state

According to Gabbay and Moracsik (1980), there is some exceptional set of state verbs that shows discontinuities or 'gaps', they call it 'Gappy statives'. They argued verbs describing positions for instance, 'stand', 'sit', 'lie' etc and another class of verbs i.e. verbs of attention like, 'watch', 'look' and 'hope'; exhibit the feature of gapping. They considered the state of being sick and said it is not true that any two

states of sickness add up to a state of sickness. 'If someone is sick over a period of time, then he is in that state also over most if not all of the parts of the state. According to them, states can be 'broken'.

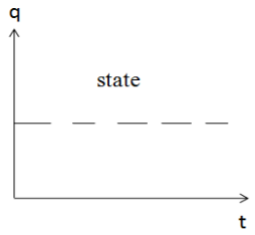


Figure 2: Gappy states

We consider the example of being sick given by Gabbay and Moracsik (1980) and we say that imagine a context where the patient is in a coma and going by the understanding of being in a coma, the person will remain in a prolonged state of deep unconsciousness. In a situation like this, the state will hold for all of its parts and subparts. Our aim is not to go by interpretations that are entirely context-dependent. That is the very reason for which we concentrate and bring the Bangla desire predicates *itʃtʃʰe kɔra* and *itʃtʃʰe hɔwa* and argue that the host noun *itʃtʃʰe* with the light verbs *kɔra* and *hɔwa* can denote 'gappy states' with the frequency adverbs.

According to the theory of Vendler (1957) which was again followed by Bennett and Partee (2004), Dowty (1977), Parsons (1990), and many others, a state is or if true for an interval then it is true for all of its subintervals. Now Bangla shows that there can be stative predicates (N+V constructions) indicating discontinuities or 'gaps' i.e. if it is true for an interval then it need not be true for all its subintervals. Consider the example below for a better understanding.

(16) Context: The subject is having quite a busy and strenuous afternoon, stuffed with work. She feels like having tea breaks in-between, but cannot due to the deadline.

amar	***madʒʰe madʒʰe***	*itʃtʃʰe*	*kor- tʃʰ-e*	*tʃa*	*kʰeye*	a*ʃ-i*
I.GEN	sometimes	wish	do-CONT-PRS.3	tea	eat.NF	go-SUBJN.1

'Sometimes I am feeling like having tea (in between work).'

If we consider the above example we can try to get a sense of the gappy state denoted by *itʃtʃʰe kɔra* (here *itʃtʃʰe hɔwa* can also be used). As evident by the context the state of the subject denotes a 'gap'. Observe the figure below.

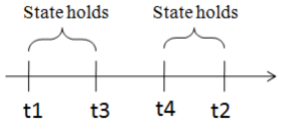


Figure 3: Gappy state through interval

So the above figure shows that if an interval I contain moments of time then the state holds for [t1,t3], [t4,t2] and [t3,t4] is the gap where the state doesn't occur. The sense of gappy state conveys that if a state holds for an interval of time then it is not the case that it holds for every moment of time in that interval. In the above example the frequency adverbial *madʒʰe madʒʰe* 'sometimes' with the desire predicate *itʃtʃʰe kɔra* delivers the sense of a gappy state i.e. if the afternoon is one interval then the desire for having tea holds for not every moment of time in that interval. As these frequency adverbs provide quantification over intervals, the sense of *itʃtʃʰe kɔra/hɔwa* with the adverbs conveys existential quantification over some subintervals (I*) that contain not all moments of time present in the interval I. So the sense of gappy state is somewhat like shown below.

(17) $[[madʒ^he\ madʒ^he]] = \lambda P_{<i,t>} . \lambda I_{<i>} . \exists I^*_{<i>} [I^* \subset I \wedge |I^*| < |I| \wedge P(I^*)]$

The frequency adverb *madʒʰe madʒʰe* (that carries the sense of gaps) takes a proposition and an interval I, so for some subintervals (I*) which contains not every moment of times present in I and I* is the subset of I and the cardinality of |I*| is less than |I| and P is true in I*

Now we should move our attention to the predicate *itʃtʃʰe* which gives a diverse reading by not combining with the frequency adverbs of time. As said above *itʃtʃʰe* will never convey discontinuities or 'gaps'. By occurring with the extended-now adverbs as seen in (14) and (15) the predicate *itʃtʃʰe* conveys that if the state holds for an interval of time then it holds for the entire interval (for every moment of time in that). As *itʃtʃʰe* signifies no 'gaps' and cannot be 'broken', it is therefore a prolonged state. The figure down below can explain further. This no-gap state of *itʃtʃʰe* triggers a universal quantification and the non-gappy sense can be observed below in (18) through the adverbial *ɔnekdiner* 'for a long time'.

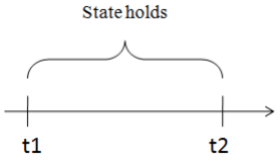


Figure 4: Continuous non-gappy state through interval

(18) $[[\textit{ɔnekdiner}]] = \lambda P_{<i,t>} . \lambda I_{<i>} . \forall I^*_{<i>} [I^* \subseteq I \wedge |I^*| = |I| \wedge P(I^*)]$

Similarly, the adverb *ɔnekdiner* (that carries the sense of non-gap) takes a proposition and an interval I, so for all subintervals (I*) which contains every moment of time present in I and I* is the improper subset of I and the cardinality of |I*| is same as |I| and P is true in I*.

4 Conclusion

This work addressed a major thread of heterogeneity that exists among the members of the same set. A minute introspection helped us to bring out the diversity among the desire predicates *itʃtʃʰe*, *itʃtʃʰe kɔra*, and *itʃtʃʰe hɔwa*. The heterogeneity that this work concentrates on revolves around *itʃtʃʰe* versus *itʃtʃʰe kɔra / hɔwa*. Apart from the different clausal dependencies, we present a novel observation showing how these predicates react differently while interacting with the temporal adverbials. The predicate *itʃtʃʰe kɔra* and *itʃtʃʰe hɔwa* develop an alliance with the frequency adverbials resulting in the phenomena of 'gappy statives'. The semantics of *itʃtʃʰe kɔra* and *itʃtʃʰe hɔwa* with the frequency adverbials convey a state delivering 'discontinuities' or 'gaps'. The sense of gappy state conveys that if a state holds for an interval of time then it is not the case that it holds for every moment of time in that interval, thus giving an existential reading. The predicate *itʃtʃʰe* on the other hand includes no modifications by the frequency adverbials. The state delivered by *itʃtʃʰe* cannot be 'broken' or include 'gaps'. It can be modified by the 'extended now' adverbials where the starting point of the state is extended to the time of utterance and there is no gap or discontinuity. If the state holds for an interval of time then it holds for the entire interval (for every moment of time in that), implying a universal reading.

References

Binnick, R. I. (1991). *Time and the Verb : A Guide to Tense and Aspect* . New York: Oxford University Press.

Bach, E. (1981). Time, Tense , and Aspect: An Essay in English Metaphysics . In P. Cole, *Radical Pragmatics* (pp. 63-81). New York: Academic Press.

Dowty, D. R. (1977). Toward a Semantic Analysis of Verb Aspect and the English 'Imperfective' Progressive. *Linguistics and Philosophy ,1* (1), 45-77.

Heny, F. (1982). Tense, aspect and time adverbials . *Linguistics and Philosophy* ,109-154.

Gronn, A. v. (2013). Tense in Adjuncts Part 2: Temporal Adverbial Clauses. *Language and Linguistics Compass ,7* (5), 311-327.

Klein, W. (1994). *Time in Language* . London and Newyork : Routledge.

Moravcsik, D. G. (1980). Verbs, events, and flow of time. In C. Rohrer, *Time, Tense and Quantifiers.* De Gruyter.

Parsons, T. (1990). *Events in the semantics of English : A study in Subatmoic Semantics.* London: The MIT Press.

Partee, M. B. (2004). Toward the Logic of Tense and Aspect in English . In B. H. Partee, *Compositionality in Formal Semantics : Selected Papers* (pp. 59-109). Blackwell Publishing Ltd.

Rathert, M. (2012). Adverbials . In R. I. Binnick, *The Oxford Handbook of Tense and Aspect.* (pp. 237-268). Oxford University Press.
Swart, H. d. (1993). *Adverbs of Quantification: A Generalized Quantifier Approach* . New York: Garland Publications.
Vendler, Z. (1957). Verbs and Times. *The Philosophical Review , 66* (2), 143-160.

A Gradient Typology of Contractions in Chinese Languages

Jarry Chia-Wei Chuang
National Chengchi University

The study examines the diverse patterns of contraction in Chinese languages. We propose a gradient account to elaborate on the typology of Chinese contractions regarding the variations of the syllable structures. It is assumed that contraction is a three-level sound change, from two syllables to one. Complete contraction is monosyllabic, while incomplete contraction should be analyzed as disyllabic, with different degrees of intervocalic reduction. Min and Rugao contractions are basically complete, while Cantonese and Taiwan Mandarin contractions are still developing at the transitional stage. They have the tendency towards complete contraction, especially in fast speech. In the gradient account, Chinese contractions can get better predicted and interpreted under broader phonological frameworks.

Keywords: contraction, Chinese languages, syllable structure

1 Introduction

Contraction is particularly prevalent in Chinese languages. Many Chinese languages have been documented to undergo contraction, particularly in spontaneous speech, including Taiwan Mandarin (Cheng & Xu, 2015; Cheng, Xu, & Gubian, 2010; Tseng, 2005a, 2005b, 2005c; Chuang, 2022), Southern Min (Chung, 1996; Hsu, 2003; Li & Myers, 2005), Hakka (Chung, 1997), Cantonese (Bauer & Benedict, 2011; Luke & Nancarrow, 1999), Tianjin (Wee, 2005, 2014), Rugao (Xu, 2020), etc. Although many studies have attempted to document the pattern of Chinese contraction, most of them failed to interpret the variations of contraction across dialects (e.g., the coexistence of full and partial contraction within/across dialects) but only focused on a specific dialect. Predictions on the grounding of evidence from a single dialect are unavailable for all the contractions in Chinese dialects. Syllable patterns may even vary within a dialect. Such diversity in a particular dialect has been widely discussed but has not been cross-dialectally compared. As a lack of adequate consideration may lead to biased predictions of contracted patterns, an exhaustive examination of Chinese contraction is necessary to help elucidate the puzzles in order to capture a more universal picture and subtle differences in dialectal variations. The present study aims to compare and discuss variations of contractions in Chinese languages, both intralingually and cross-dialectally.

2 Chinese Syllable Structure

The standard syllable in Chinese generally complies with the CGVX sequence (C= onset consonant; G= prenuclear glide; V= nucleus vowel; X= coda glide, nasal, or stop) (Duanmu, 2007); syllables of most Chinese dialects are bimoraic (e.g., Mandarin, Min, Cantonese), which are considered to be heavy syllables

* I would like to express my gratitude to Prof. Yuchau E. Hsiao and Prof. Huei-Shen Lin for constructive suggestions. Also, I would like to thank Prof. Birgit Alber, Prof. Zach Metzler, and the WECOL-2022 audiences for discussions and useful comments. Additional gratitude should go to NCCU Phonology Group and several Chinese native speakers in various dialects for language documentation and examination. The research has been supported by a grant from National Chengchi University and Ministry of Science and Technology, Taiwan. The remaining errors are my own.

Proceedings of *WECOL 2022*

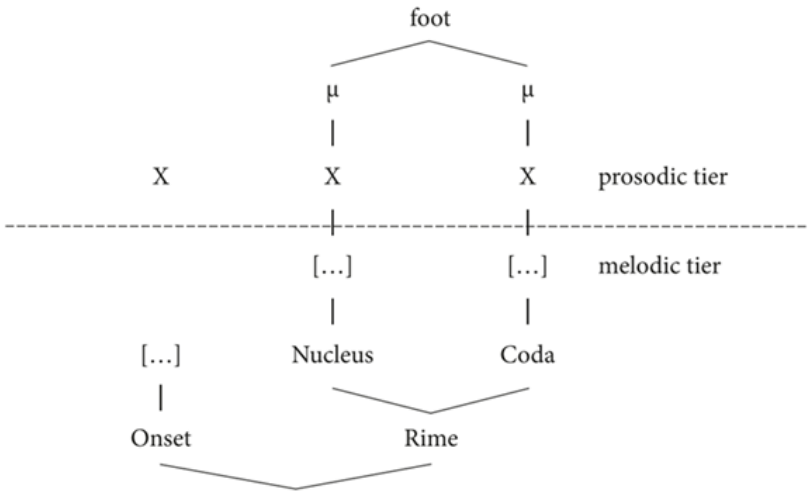
(Duanmu, 1999, 2000).

As CGVX is the general template for Chinese syllables, different points of view have been held for the basic syllable structures in Chinese. The major concern that gave rise to the debate is the affiliation of the prenuclear glide. We conclude the to-date classifications into two groups of models: One is the traditional OR and IF models, which attempt to divide a syllable into two constituents; the other is the μ-model, which assigns syllable weight to the weighted segments. It is convinced that the latter has better interpretability for Chinese languages.

Under the μ-model, the segment-melodic complex, proposed by Srinivas (2016), can explain cases of Chinese contraction well. As previous discussions adopted the idea of the skeleton for analysis, the complex can also associate it with the μ-model. As shown in Figure 1, the prosodic tiers include the skeletal tier as well as the phonological hierarchy (e.g., counting of syllable weight, moraic foot, syllable foot, etc.); the melodic tiers consist of segments within the OR model.

Figure 1

Projection of the Segment-melody Complex (Liang & Wee, 2022; Srinivas, 2016)



Two tiers are linked with the segments as the connecting points. The projection, with the mora-based model and the OR model, has good suitability in Chinese syllable structures (Liang & Wee, 2022). We then work on Chinese contractions under the complex.

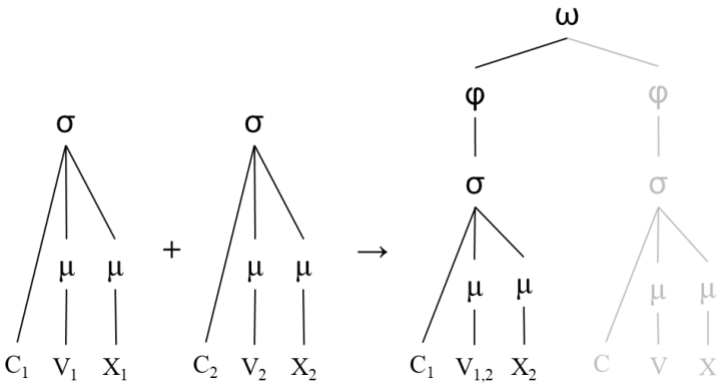
3 Interaction between Syllable Weight and Contraction

Identifying different degrees of contraction, the correspondence between syllable weight and the output segments should be treated differently. At least, full (complete) and partial (incomplete) contractions may have their own structure. To avoid the ambivalent status of the medial glide, we temporally adopt the case of CVX sequence for a standard syllable.

3.1 *Complete Contraction* First, complete contraction refers to the total integration of two syllables into one. Chung's (1996) skeletal account used the concept of the skeletal tier for syllable-timing, under *Autosegmental Phonology*. Adopting the μ-model, we alter Chung's (1996) skeletal account into the mora-based structure with phonological hierarchy. As in Figure 2, two bimoraic syllables can be converted into a bimoraic syllable. Note that the output C and X would come from the first and the final source syllable, thus being C_1 and X_2, which is constrained by Edge-in Association (Yip, 1988). The output nucleus may be co-decided by the V_1 and V_2 and the more sonorous source vowel would be selected as the output nucleus. A complete contraction case will find another applicable neighbor syllable, if any, to form a new syllabic

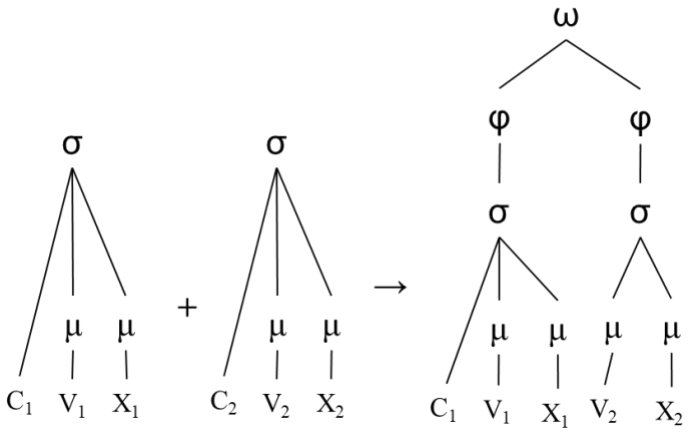
trochee and a phonological word[2], which can be as productive as a new foot and can be viewed as a type of word formation accordingly (as shown in gray color).

Figure 2
Complete Contraction in CVX (Chung, 1996; Hsu, 2003)



3.2 *Incomplete Contraction* Second, incomplete contraction addresses the partial deletion of the intervocalic segments (X_1+C_2), which is much more complicated than the complete one. In incomplete contractions, the possible structures can be $C_1V_1X_1$+ V_2X_2 and C_1V_1+V_2X_2. For $C_1V_1X_1$+ V_2X_2, it is doubtless that V_1X_1 and V_2X_2 belong to different syllables, σ_1 and σ_2, as illustrated in Figure 3. CVX.VX also complies with the general possible syllable sets, where the onset consonant C in a heavy CVX syllable is optional; thus, the deletion or reduction of C is fairly sensible for the universal analysis.

Figure 3
Incomplete Contraction in CVX+VX

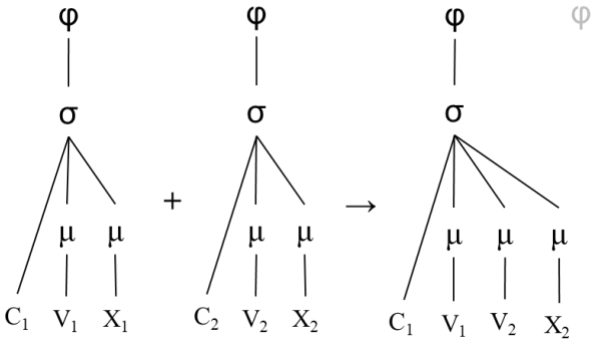


As for C_1V_1+V_2X_2, the resyllabification can be questionable. There are two possible considerations for the output syllabification. One is that C_1V_1 and V_2X_2 are belonging to one syllable, which can be superheavy with 3μ. The other possibility we argue is that C_1V_1 and V_2X_2 belong to two different syllables.

Chen (2020) measured the duration of C_1V_1+V_2X_2 and found the rime duration is about 1.5 times longer than a standard CVX syllable, thus suggesting the moraic structure of the $C_1V_1V_2X_2$ set should be CV.VX, as shown in Figure 4.

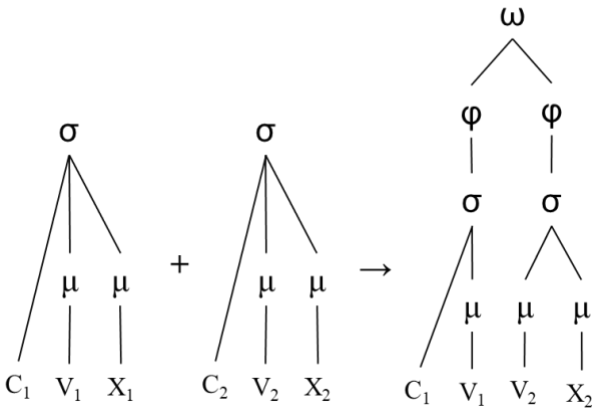
[2] According to Duanmu (1999), two mora forms a foot (M-foot), and two syllables form a syllabic foot (S-Foot). As the minimal word should contain a disyllabic trochee, a S-foot equals to a phonological word.

Figure 4
Trimoraic Analysis of Incomplete Contraction in CVVX (Chen, 2020)



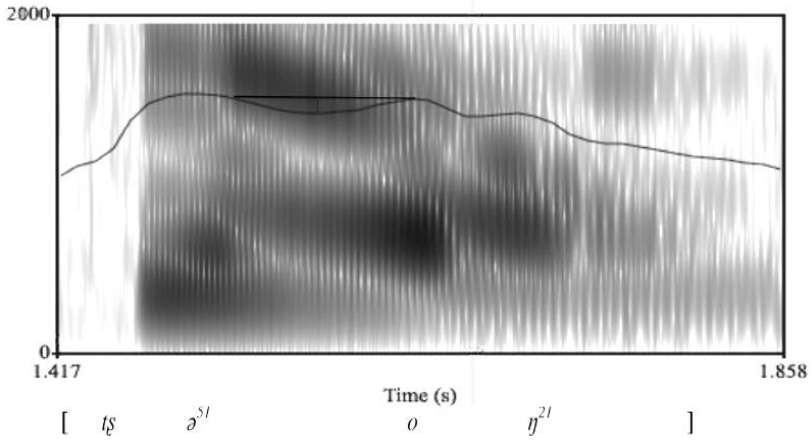
However, we cast a doubt on the possibility of the overloaded syllable weight, for the following reasons: First, the violability of the well-formedness of a syllable, where the maximal size of syllable weight for a syllable should be bimoraic, may grant the costly permission for many other cases. In the trimoraic analysis, $*\sigma_{\mu\mu}$ is violated, while it is not preferred under the universal grammar. It will contribute to the loss of a syllable and a corresponding foot, in which Max-σ-SO and MAX-φ-SO are violated meanwhile. As we allow CVVX to appear, many constraints (e.g., $*\sigma_{\mu\mu}$, Max-σ-SO, and MAX-φ-SO) are violated accordingly. This makes the rankings even more byzantine. Therefore, to solve the challenges mentioned above, we propose an alternative analysis for the resyllabification. It is argued that C_1V_1 and V_2X_2 are affiliated with two syllables, as shown in Figure 5. The analysis prevents MAX-φ-SO and Max-σ-SO from getting violated.

Figure 5
Disyllabic Analysis of Incomplete Contraction in CVVX



Evidence for the present analysis can also be solicited from acoustic measurement. As observed by Kuo (2010), an amplitude drop is sometimes possible between V_1 and V_2. This explains that though two sets are connected closely, there remains to be a distance of intimacy in between. Take [tʂə: + tʂoŋ] 'this kind' for example. As in Figure 6, [tʂə: + tʂoŋ] will become [tʂə.oŋ], where the lengthened [ə] is canceled and the onset [tʂ] is deleted. An intensity fall can be found between [ə] and [o]. This can be strong support for the syllabification, as they are belonging to two syllables.

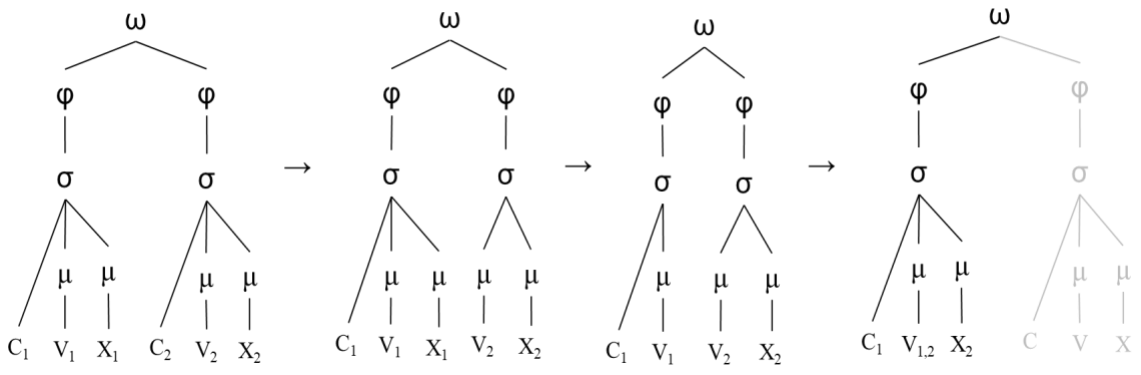
Figure 6
Intensity Contour in CVVX (Kuo, 2010)



4 A Gradient Account

4.1 *Modeling Phonological Structures of Chinese Contractions* Intricate as syllable structures are in Chinese contractions, we summarize the discussions above, proposing a gradient account for these variations. The gist of the gradient account is that contraction is a gradual sound change. We consider contraction to be two syllables gradually merging into one, along with a decrease in prosodic counts like the skeleton and the more number. There can be three levels of contractions. The gradient change generally follows the paradigm as in Figure 7.

Figure 7
The Gradual Change of Chinese Contraction



4.2 *Intralingual and Cross-dialectal Variations* As contractions are various in Chinese languages, we select a few notable cases from certain Chinese dialects to demonstrate the variations with the gradient account. First, in Table 1, Taiwan Mandarin contractions are mainly CVX.VX or CV.VX, such as [tʰɔŋ.(h)aŋ] or [tʰɔ̃.aŋ] for [thɔŋ + ʈʂʰaŋ] 'usually.' As CVX.VX is the initial stage of contraction, CV.VX is also the common contraction form in Taiwan Mandarin. In contrast, complete contractions within the CVX sequence are infrequent. Only for some lexicalized items like [njaŋ] 'that way' and [tɕjɛn] 'today,' complete contraction is possible.

Table 1
Taiwan Mandarin Contraction

Uncontracted	**Incomplete contraction**		**Complete contraction**	**Gloss**
CVX+CVX	**CVX.VX**	**CV.VX**	**CVX**	
tʰɔŋ + ʈʂʰaŋ	tʰɔŋ.(h)aŋ	tʰɔ̃.aŋ	—	‘usually’
kʰə: + ʂɨ:	kʰə:.(h)ɨ:	kʰə.ɨ:	—	‘but’
tan + ʂɨ:	tan.(h)ɨ:	tɑ̃.ɨ:	—	‘but’
tʰjɛn + tʰjɛn	tʰjɛn.jɛn	tʰjɛ̃.jɛn	—	‘every day’
na: + jaŋ	na:.jaŋ	na.jaŋ	njaŋ	‘that way’
tɕin + tʰjɛn	tɕin.jɛn	tɕĩ.jɛn	tɕjɛn	‘today’

In Taiwan Southern Min Contraction, complete contraction CVX is the majority, as in Table 2. Some of them are even highly lexicalized, such as [sjaŋ] for [sjo: + kaŋ] ‘the same.’ For non-violability of phonotactics, there may be some cases that are in an atypically disyllabic template. For example, [ki.ai] for [ke: + lai] is one of the cases.

Table 2
Taiwan Southern Min Contraction

Uncontracted	**Incomplete contraction**		**Complete contraction**	**Gloss**
CVX+CVX	**CVX.VX**	**CV.VX**	**CVX**	
ke: + lai	—	ki.ai	kai	‘come over’
sjo: + kaŋ	—	—	sjaŋ	‘the same’
bo: + e:	—	—	bwe:	‘unable’
tsai + kʰi:	—	—	tsai	‘morning’
bo: + ai	—	—	bwai	‘not want’

In Rugao contraction, most of the cases we solicit are complete contractions in the CVX template, as in Table 3. For example, [ɕjæn + ɕjən] is completely merged as [ɕjæn] ‘believed,’ in which two C(G)VX syllables are completely fused into one.

Table 3
Rugao Contraction

Uncontracted	**Incomplete contraction**		**Complete contraction**	**Gloss**
CVX+CVX	**CVX.VX**	**CV.VX**	**CVX**	
ɕjæn + ɕjən	—	—	ɕjæn	‘believe’
tsa + kow	—	—	tsaw	‘this’
sɨ +xej	—	—	sej	‘time’
ɹin+ xej	—	—	ɹej	‘then’
ɕjən +jɔn	—	—	ɕjɔn	‘credit’

In Cantonese contraction, CV.VX is the most common template. [dzi: + gej] ‘know’ is partially merged as [dzi.ej], as in Table 4. In fast speech, CVX is applicable (Hsu, 2005). [ji: + ga:] ‘oneself’ is accordingly entirely fused as [ja:]. CV.VX and CVX are common in Cantonese.

Table 4
Cantonese Contraction

Uncontracted	**Incomplete contraction**		**Complete contraction**	**Gloss**
CVX+CVX	**CVX.VX**	**CV.VX**	**CVX**	
dzi: + gej	—	dzi.ej	dzej	'know'
ji: + ga:	—	ji:.a:	ja:	'oneself'
bɐt + gɔ:	—	bɐ.ɔ:	bɔ:	'however'
ji: + geŋ	—	ji.eŋ	jeŋ	'already'

In the gradient model, we consider the intralingual and cross-dialectal variations to be a gradual process of sound change. It is syllable timing that leads to the change, from the dissyllabic syllables to the monosyllabic ones. As the standard syllables in three languages are bi-moraic, the uncontracted forms are CVX.CVX (4μ; 3X+3X). In TM, contractions present a gradient change, in which intervocalic segments may not be totally reduced. In the semi-contracted stage, the unweighted onset of the 2nd source syllable is possibly contracted or deleted first. This contraction does not involve the change of the moraic structure but the duration of the dissyllabic word (4μ; 3X+2X). In the extreme reduction, the weighted coda glide/nasal of the 1st source syllable is then removed. As for Min and Cantonese contractions, two syllables can only be, to the fullest, contracted into one CVX syllable (2μ; 3X), with X_1 and C_2 deleted. It should be noted that the output nucleus in Min contractions is merged from two source syllables, thus co-decided by V_1 and V_2, which may reflect the sonority preference (Hsu, 2003).

5 Conclusion

The study examines the diverse patterns of contraction in Chinese languages. We propose a gradient account to elaborate on the typology of Chinese contractions regarding the variations of their syllable structures. We assume contraction is a gradual sound change, from two bimoraic syllables to one bimoraic syllable. Min and Rugao contractions are mostly in a clear dichotomized distribution, with uncontracted CVX+CVX or contracted CVX available. Their contracted forms comply with the syllable structures and phonotactic constraints, better viewed as syllable contraction. Taiwan Mandarin and Cantonese contractions are developing at the transitional stage, where intervocalic segments may not be totally reduced and two syllables haven't been fully contracted into one sometimes. Cantonese contraction has a strong tendency to complete contraction, for several cases supporting the idea. Taiwan Mandarin contractions are mostly incomplete contractions, except for some lexicalized examples.

In the gradient account, Chinese contractions can get better predicted and interpreted under broader phonological frameworks. Future studies are suggested to examine more data from other dialects and conduct an empirical study for our assumption of contraction caused by the sharp prosodic drop.

6 References

Bauer, R. S., & Benedict, P. K. (2011). Modern cantonese phonology. In *Modern Cantonese Phonology*: De Gruyter Mouton.

Chen, Y.-J. (2020). *The Moraic Structure of Taiwan Mandarin Contractions.* Paper presented at the The 32nd North American Conference on Chinese Linguistics, University of Connecticut.

Cheng, C., & Xu, Y. (2015). Mechanism of disyllabic tonal reduction in Taiwan Mandarin. *Language and speech, 58*(3), 281-314.

Cheng, C., Xu, Y., & Gubian, M. (2010). Exploring the mechanism of tonal contraction in Taiwan Mandarin.

Chuang, J. C. (2022). *Place Features in Phonetic Contraction in Taiwan Mandarin: Evidence from Spontaneous Speech*. Paper presented at the the 23rd National Conference on Linguistics, Taipei, Taiwan.

Chung, R.-f. (1996). *The segmental phonology of Southern Min in Taiwan*: Crane Publishing.

Chung, R.-F. (1997). Syllable contraction in Chinese. *Chinese languages and linguistics III: Morphology and lexicon*, 199-235.

Duanmu, S. (1999). Metrical structure and tone: evidence from Mandarin and Shanghai. *Journal of East Asian Linguistics*, *8*(1), 1-38.

Duanmu, S. (2000). Stress in Chinese. In *Chinese phonology in generative grammar* (pp. 117-138): Brill.

Duanmu, S. (2007). *The phonology of standard Chinese*: OUP Oxford.

Hsu, H.-c. (2003). A sonority model of syllable contraction in Taiwanese Southern Min. *Journal of East Asian Linguistics*, *12*(4), 349-377.

Hsu, H.-c. (2005). An optimality-theoretic analysis of syllable contraction in Cantonese. *Journal of Chinese Linguistics*, *33*(1), 114-139.

Kuo, G. C. H. (2010). Production and perception of Taiwan Mandarin syllable contraction. *The Journal of the Acoustical Society of America*, *127*(3), 1954-1954.

Li, Y.-S., & Myers, J. (2005). Modeling variation in Taiwan Southern Min syllable contraction. *Taiwan Journal of Linguistics*, *3*(2), 79-118.

Liang, Y., & Wee, L.-H. (2022). Melodic-prosodic duality of the syllable: An application to Chinese. *Language and Linguistics*, *23*(1), 82-115.

Luke, K., & Nancarrow, O. (1999). Short forms in Hong Kong Cantonese. *第七屆國際粵方言研討會*.

Srinivas, S. K. (2016). *The mora-constituent interface model.* Hong Kong Baptist University,

Tseng, S.-C. (2005a). Contracted syllables in Mandarin: Evidence from spontaneous conversations. *LANGUAGE AND LINGUISTICS-TAIPEI-*, *6*(1), 153.

Tseng, S.-C. (2005b). Monosyllabic word merger in Mandarin. *Language Variation and Change*, *17*(3), 231-256.

Tseng, S.-C. (2005c). Syllable contractions in a Mandarin conversational dialogue corpus. *International Journal of Corpus Linguistics*, *10*(1), 63-83.

Wee, L. H. (2005). Elision and tone sandhi in the Tianjin dialect.

Wee, L. H. (2014). Casual speech elision and tone sandhi in Tianjin trisyllabic sequences. *International Journal of Chinese Linguistics*, *1*(1), 71-95.

Xu, C. (2020). *The Phonology and Phonetics of Rugao Syllable Contraction: Vowel Selection and Deletion*: Michigan State University.

Yip, M. (1988). Template morphology and the direction of association. *Natural Language & Linguistic Theory*, *6*(4), 551-577.

Processes in the Ogbia Numeral System

Merosobo Lawson-Ikuru
Port Harcourt Polytechnic

The paper examines the formation of numerals in Ogbia with particular interest in the morphological and phonological processes involved in the formation of these numerals. Firstly, the paper looks at the basic numerals of the language, stating that basic numerals are those that do not undergo any process to attain their present form. The basic numerals in the language are also seen as cardinal numbers. The ordinal numbers can be used to describe positions of objects or numerals occurring in a sequential order and they can also serve as modifiers. They use the morpheme 'olelemən' when describing positions except for the numeral 'one'. Distributive numerals can be derived through a partial or complete reduplication. Apart from the basic numerals, other numerals in Ogbia are derived. Some of the derivational processes observed are compounding, contraction, insertion, elision, addition, subtraction and multiplication. These processes are grouped as morphological and phonological. The phonological processes exhibit both underlying and surface forms. The vowel harmony phenomenon is also seen as playing a role in the numeral system of the language. The language exhibits an almost complete vowel harmony system with rare exceptions. The paper concludes that, from available linguistic literature, this paper seems to be the first description of the numeral system of Ogbia, therefore there will be the need for more observations and analysis.

Keywords: Ogbia, Numerals, Processes, Morphological, Phonological

1 Introduction

The numeral systems of languages have been described by linguistic scholars as an important aspect of linguistics that has not been given close attention or totally ignored until lately. A numeral system is the set of counting words and the principles or rules utilized in a particular language for constructing these words (Trask 1997). The numeral system according to Comrie (2005) is as important as other aspects of the grammar of a language. He says that languages that are considered not to be endangered as it were, could have its numeral system endangered if it is not properly documented and that is most dangerous, hence the need for an urgent attention to this aspect of language. The significance of the numeral system of a language cannot be overemphasised as some authors have reiterated. For example, Omachonu (2013) expresses that counting and numeracy are an integral aspect of language because there is rarely any linguistic discourse in a language that does not make reference to quantity, size, time, distance, and weight in definite numbers or numerals. For Blažek (1999) in Mbah and Uzoigwe (2013:64) , numerals exist in all known languages (both living and dead). According to him, it is possible that numerals are same age as the idea of counting. Therefore, they could have to search no later than the beginning of the late Paleolithic. In the linguistic literature of number systems, it is observed that certain linguistic processes aid the realisation of certain numbers in languages. This paper sets out to examine those processes in the number system of Ogbia.

The Ogbia language is linguistically grouped as a Central Delta Language, spoken in Southern Nigeria with an estimation of about 200,000 speakers. The language has a linguistic dearth, which is generally the linguistic status of the Central Delta Languages (kari 2009:8). The documented linguistic information of the language can be described as scanty and available linguistic information show that the number system of the language has not been analysed from a linguistic perspective. Comrie (2005) stresses that, numerical

data from most languages of the world have not been documented, therefore it is not surprising that the numeral system of the Ogbia language is lacking in the linguistic literature. This study observes that there both morphological and phonological processes involved in the realisation of numerals in Ogbia.

2 Basic Numerals

Basic numerals are generally seen as those numerals that do not undergo any morphological or phonological process to attain their present form. Ajiboye (2013:2) refers to them as numerals whose forms cannot be broken down into identifiable morphemes. In Ogbia, these can be grouped into two. The first group is made up of numerals from one to ten.

(1) 1. ònín ‘one’
2. wàlì ‘two’
3. sàrì ‘three’
4. ìŋə̀ ‘four’
5. òwù ‘five’
6. òdìn ‘six’
7. òɗùə̀l ‘seven’
8. íŋàn ‘eight’
9. ísíó ‘nine’
10. ɗìòβ ‘ten’

Another basic numeral other than 1-10 is twenty. This number is also not derived from any process.

(2) 20 sùβ or ípó ‘twenty’

It is important to state that, the numerals; íŋàn (8), ísíó (9) and sùβ (20) are used for counting but when they have to be used to derive bigger numerals, they appear as wàĺìbádìòβ (8), òníbádìòβ (9) and ípó (20). In this case (8) an (9) are also derived through a subtraction process where wàĺìbádìòβ simply means two out of ten and òníbádìòβ means one out of ten. However ípó simply represents the function of twenty in the realisation of other numbers.

2.1 *Cardinal Numerals*

Cardinal numbers are used attributively and some of the basic numerals in Ogbia can be considered as cardinal numbers. They are used for counting and describing quantity. Except for number one, every other basic numeral is used for counting and describing quantity. There are basically no distinct forms for counting other than using basic numerals. In the case of “number one”, which is called ‘onin’ in its basic form, becomes ‘èkpò’ in describing quantity. Hence we cannot find *ònìn òtù rather we can have ‘èkpó òtù’ meaning one house while for counting; onin is used to refer to number (1). These simple cardinal numerals can be used to derive complex numerals.

2.2 *Ordinal Numerals*

Crystal (2008:344) defines ordinal as a term used in some models of grammatical description referring to the class of numerals first, second etc. Obikudo (2013:35) also says the ordinal numeral as describing positions of objects in a sequential order. According to Stolz and Veselinova (2005), cardinals have two major functions. They are distinguishing hierarchies and ordering sequential events. The ordinal numerals in Ogbia can perform attributively as adjectives in sentences as well as act as modifiers. They can be used to indicate positions but not without adding the morpheme ‘òlélémə́n’. This morpheme indicates number when describing position and can be used with any numeral with exception of number (1). In describing position, (1) becomes ɔ̀pɔ̀rɔ̀ ‘first’, every other number goes with ‘òlélémə́n’. It is therefore semantically wrong to say ‘onin ɔ̀ŋì as first child. This would rather mean, ‘the same child’. For example;

(3) a. ɔ̀pɔ̀rɔ̀ ɔ̀ŋì ‘first child’
b. òlélémə́n wàlì àŋwì’ ‘second child’
c. òlélémə́n’sàrì àŋwì ‘third child’
d. òlélémə́n ɗìòβ àŋwì ‘tenth child’

2.2.1 *Ordinal Numerals as Nominal Modifiers*

Ordinal numerals in Ogbia can modify nouns together with adjectives alone and also with the definite article. In order of occurrence, the ordinal numeral comes first before other modifiers.

(4) a. ɔpɔ̀rɔ̀ óɓéβ ɔ̀tù-ə̀
First beautiful house the → 'the first beautiful house'
b. ɔ̀lélémən wàlì ɔ̀ɓéβ ɔ̀tù-ə̀
number two beautiful house the →'the second beautiful house'
c. ɔ̀lélémən ɗìɔβ ɔ̀ɓèβ ɔ̀tù-ə̀
number ten beautiful house the → 'the tenth beautiful house'

2.3 *Distributive Numerals*

In Ogbia, a distributive numeral could be derived by complete or partial reduplication of a cardinal numeral as it is with most sub-Saharan African languages. Reduplication tends to be the most morphological method applied in showing distributive expressions. Gil (2005) says the term distributive numeral is reserved for numerals which allow NPs to be chosen as distributive key. The cardinal numerals in the data below show alternations in some of the distributive forms.

(5)

Cardinal Numeral		Distributive Numeral	
èkpò	'one'	èkpèkpò	'one by one'
wàlì	'two'	wàlìwàlì	'two by two'
sàrì	'three'	sàrìsàrì	'three by three'
ìŋə̀	'four'	ìŋìŋə̀	'four by four'
òwù	'five'	òwòwù	'five by five'

The data in (5) show that the numerals 'one','four' and 'five' are partially reduplicated in their distributed forms while the others have complete reduplication. The numerals 'one' 'four' and 'five' are partially reduplicated as a result of the elision process of contiguous vowels in the distributive forms.

3 Derivation of the Ogbia Numerals

Derivation is a very common linguistic concept which tends to occur in most levels of linguistic analysis and in many languages as well. Crystal (2008: 138) defines derivation as a term used in morphology to refer to one of the two main processes or categories of word formation and O' Grady & Archibald (2008:116) also see derivation as an affixational process that forms a word with a meaning or category distinct from that of its base. As mentioned earlier in this work, that apart from the basic numerals 1-7, 10 and 20, all others are derived. They can be separated into identifiable morphemes; whether bound or free. The derived numerals in the language undergo certain processes which are discussed below.

3.1 *Derivational Processes of Numerals in Ogbia*

The derivational processes of numerals in Ogbia can classified as both morphological and phonological. The processes evident in the language are compounding, reduplication, contraction, insertion and elision as well as mathematical operations like addition, subtraction and multiplication.

3.1.1 *Compounding*

Compounding is a morphological process that basically involves the merging of two independent units to have a single meaning. In this context, compounding of numerals is a process of joining two already existing numerals to form a new one. In the language of study, it is observed that the numeral on the right usually has a strong numerical strength than the second. A typical example in Ogbia is the numbers 11-17, 21-27 and the likes. For example;

(6) a. ɗìɔβ ɗĩnì ɔ̀nín → ɗìɔβ nɔ̀nín
ten and one 'eleven'
10+1 11

b. ɗìòβ ɗìnì wàlì → ɗìòβ nìwàlì
ten and two ‘twelve’
10+2 12

c. ɗìòβ ɗìnì sàrì → ɗìòβ nìsàrì
ten and three ‘thirteen’
10+3 13

d. ɗìòβ ɗìnì ìŋə̀ → ɗìòβ nìŋə̀
ten and four ‘fourteen’
10+4 14

e. ɗìòβ ɗìnì òwù → ɗìòβ nòwù
ten and five ‘fifteen’
10+5 15

f. ɗìòβ ɗìnì òdìn → ɗìòβ nòdìn
ten and six ‘sixteen’
10+6 16

g. ɗìòβ ɗìnì òɗùə̀l → ɗìòβ nòɗùə̀l
ten and seven ‘seventeen’
10+7 17

The output of the numbers in the example in 6, shows that the first syllable 'ɗìni' which represents 'and' deletes and the vowel in the second syllable assimilates to the initial vowel of the next word. This usually happens during fast speech and it has become the standard way of saying these numbers. Numerical compounding is also observed in bigger numerals. The numerals forty, fifty, sixty, seventy, eighty, ninety and hundred are derived through compounding. Here, the compounding involves two numerals with equal strength or a numeral with a smaller strength on the left and the higher one on the right. For example;

(7) a. wàlì ípó → wàlpó
two and twenty ‘forty’
2 x 20 40

b. wàlì ípó ɗĩnì ɗìòβ → wàlpó nì ɗìòβ
two twenty and ten ‘fifty’
2 x 20 + 10 50

c. sàrì ípó → sàrpó
three twenty ‘sixty’
3 x 20 60

d. sàrì ípó ɗĩnì ɗìòβ → sàrìpó nì ɗìòβ
three twenty and ten ‘seventy’
3 x 20 +10 70

e. ìŋə̀ ípó → ìŋìpó
four twenty ‘eighty’
4 x 20 80

f. ìŋə̀ ípó ɗìnì ɗìòβ → ìŋípó nì ɗìòβ
four twenty and ten ‘ninety’
4 x 20 + 10 90

g. òwù ípó → òwùpó
five twenty 'hundred'
5 x 20 100

Numerals above hundred take longer expressions to show how they are derived, hence they are not included in the analysis above. From the data above, the numerals without the addition of 'ten' are compounded underlyingly, but they occur as single units in their surface forms. This process will be discussed later.

3.1.2 *Addition*

Addition is a major derivation process of numerals in Ogbia. Many numerals in the language are derived through the addition of two independent units. The examples in 4 are all derived through the process of addition. The addition process in the language involves the use of the conjunction ɗini'. Apart from the examples given in 4, other numbers derived from the additional process are;

(8) a. súβ díní ònín → suβ nonin
twenty and one 'twenty one'
20 +1 21

b. súβ díní wàlì → súβ ní wàlì
twenty and two 'twenty two'
20 +2 22

c. súβ díní sàrì → súβ ní sàrì
twenty and three 'twenty three'
20 + 3 23

d. súβ ɗíní ìŋə → súβ ní ìŋə
twenty and four 'twenty four'
20 + 4 24

e. súβ díní òwù → súβ ní òwù
twenty and five 'twenty five'
20 + 5 25

f. súβ díní òdìn → súβ ní òdìn
twenty and six 'twenty six'
20 +6 26

g. súβ díní òɗùəl → súβ ní òɗùəl
twenty and seven twenty seven
20 + 7 27

Another group of numerals that are derived from addition are some numbers in tens. These numerals are fifty,seventy , ninety, and others above hundred. (see example 7). From the data shown so far, one begins to wonder what happens to the derivation of 18, 19, 28 and 29. These are rather derived through the subtraction process as presented below.

3.1.3 *Subtraction*

As stated earlier, subtraction is also a major derivational process of numerals in Ogbia. The numerals; eight, nine, eighteen, nineteen, twenty eight, twenty nine, thirty eight, thirty nine etc. are derived through this process. It involves the use of 'ba' which literarily means 'not' and indicates minus too.

(10) a. wàlì bá ɗìòβ
two out of ten → ten minus two → 'eight'
10 – 2 8

b. ónìn bá ɗìòB
one out of ten → ten minus one → 'nine'
10 – 1 9

c. wàlì bá súβ
two out of twenty → twenty minus two → eighteen'
20 -2 18

d. ònín bá súβ
one out of twenty → twenty minus one → 'nineteen'
20 – 1 19

e. wàlì bá ɔ̀dáβàr
two out of thirty → thirty minus two → 'twenty eight'
30 – 2 28

f. ònín bá ɔ̀dáβár
one out of thirty → thirty minus one → 'twenty nine'
30 – 1 29

g. wàlì bá wàlìpó
two out of forty → forty minus two → 'thirty eight'
40 – 2 38

h. ònìn bá wàlìpó
one out of forty → forty minus one → 'thirty nine'
40 – 1 39

3.1.4 *Multiplication*

In Ogbia, some numerals in tens are also derived through multiplication. The output involves a compounding of two numerals. The multiplicand in this case is usually twenty. The numerals forty, sixty, eighty and hundred are derived from a multiplication process involving twenty as the multiplicand.

(11) a. wàlì ípó → wàlpó
'two twenty' 'forty'
2 x 20 40

b. sàrì ípó → sàrìpó
'three twenty' 'sixty'
3 x 20 60

c. ìŋɔ́ ípó → ìŋìpó
'four twenty' 'eighty'
4 x 20 80

d. òwù ípó → òwùpó
'five twenty' 'hundred'
5 x 20 100

The surface realisation of 100 is òwùpó and it is used as the multiplicand in the derivation of numbers in hundreds. This is shown in the example below.

(12).a. wàlì èsì òwùpó
'two place of hundered' → 'two hundred'
2 x 100 200

b. sàrì èsì òwùpó
'three places of hundred' → 'three hundred'
3 x 100 300

This is the pattern for which numerals in hundreds are derived up to nine hundred.

4 Phonological Processes in the Derivation of Numerals

Yul-Ifode (1999:141) defines phonological processes as those changes which segments undergo, that result in the various phonetic realisations of underlying phonological segments. A lot of the derived numerals in Ogbia, undergo some phonological processes to have their surface forms. Some of these processes include; contraction, insertion and elision.

4.1 *Contraction*

Contraction as a phonological process involves the reduction of a sequence of two identical segments or fusing a sequence of forms so that they appear as a single form. Most derived numerals in the language are fused to form a single unit. For example, the contraction of the conjunction 'dìni' which means 'and' to 'ni' is fused progressively to the adjacent root numeral to form a single form as well as the fusing of ípó with other numerals to form a single unit.

(13)a. súβ dínì wàlì → súβ nìwàlì
'twenty and two' 'twenty two'
20 + 2 22

b. súβ dínì sàrì → súβ nìsàrì
'twenty and three' 'twenty three'
20 +3 23

c. wàlì ípó → wàlpó
'two twenty' 'forty'
2 x 20 40

d. sàrì ípó → sàrpó
'three twenty' 'sixty'
3 x 20 60

The fusion of ípó and the other numerals shown in example (13), result to consonant clusters. Consonant clusters are not common occurrences in Ogbia.

4.2 *Elision*

In Ògbíà, sounds may get elided during fast or connected speech. Elision can occur at word boundary or at intervocalic positions. This process can affect both consonants and vowels (Lawson-Ikuru 2014: 45). Most of the derivation processes observed in this study involve elision.

4.3 *Insertion*

Consonant clusters rarely occur in Ogbia. Hence, speakers tend to insert vowels to break clusters. In the pronunciations of compounded numerals, speakers of the Anyama variety of Ogbia and some uneducated people insert vowels to break up the clusters in the derived forms. For example;

(14) a. ɗìòβ ɗìnì ònìn → ɗìòβ nònìn → ɗìòβù nònìnì
'ten and one' 'eleven'
10 + 1 11

b. dìòβ dìnì walı → dìòβ nìwàlì → dìòβù nìwàlì
‘ten and two’ ‘twelve’
10 + 2 12

c. dìòβ dìnì sàrı → dìòβ nìsàrì → dìòβù nìsàrì
‘ten and three’ ‘thirteen’
10 + 3 13

d. súβ dìnì wàlì → súβ nìwàlì→ súβú nìwàlì
‘twenty and two’ ‘twenty two’
20 + 2 22

e. súβ dìnì sàrì → súβ nìsàrì → súβú nisarì
‘twenty and three’ ‘twenty three’
20 + 3 23

5 Vowel Harmony

It is important to state the role of vowel harmony in any phonological discussion in Ogbia because of its prominence in the data. Williamson (1984:22) is the first to observe that the vowels in Ogbia show an almost complete vowel harmony system. The vowels in the language are grouped into two sets, such that vowels in one set cannot co-occur with vowels in the other set in a given word and across word boundary in some cases. This phenomenon is observed in the numeral system discussed here. This is seen in both basic numerals and in the derived forms of the numerals. The feature used to distinguish the vowels in Ogbia, is the Advanced Tongue Root (ATR) feature.

6 Conclusion

This paper examines the formation of numerals and the processes involved in deriving them in the Ogbia language. These processes are classified as phonological and morphological. The numeral system of a language can also be viewed beyond its phonological and morphological making. That is, it can be examined from a syntactic and semantic perspective, which is obviously not captured in this work, thereby giving room for more description on the numeral system of the language. The sketchy description in this paper has shown that apart from the basic or root numerals, all others numerals are derived. A major derivational process in the numeral system is compounding. This paper classifies the mathematical concepts; addition, subtraction and multiplication as derivational processes. It is important to note that rules apply underlyinly to generate the surface forms of the Ogbia numerals, even though this is not linguistically expressed here. From available literature, this seems to be the first linguistic work on the number system of the language, therefore it is hopeful that there will be more observations by linguists and those interested in the study of the Ogbia language in the future. This study can serve as a basis for any further view into the numeral system of the language.

References

Ajiboye,Oladipo. 2013. The Yoruba Numeral System. In Ozo-Mekuri Ndimele and Eugene, S.L Chan. (eds.). *The Numeral Systems of Nigerian Languages.* Vol. 2:2-25. M&J Grand Orbit Communications Ltd.

Comrie, Bernard. 2005. Endangred Numeral Systems In Jan Wohlgemuth and Tyko Dirksmeyer (eds.).*Bedrohte Vielfalt: Aspekte des Sprach(en)tods [Endangered diversity: Aspects of Language Death*], 203-230. Berlin.

Crystal, David. 2008. *A Dictionary of Linguistics and Phonetics*. Malden:Blackwell Publishers.

Gil, David. 2005. Numeral Classifier In Haspelmath, Martin., Dryer, Mathew.S., Gil, David. and Comrie Bernard (eds). The World Atlas of Language Structure, 222-225. Oxford.

Kari, Ethelbert. 2009. *A Grammatical Description of the Odual Language*. Research Institute for World Languages: Osaka University.

Lawson-Ikuru, Merosobo. 2014. A Phonology of Ogbia.’’ M.A Thesis. Department of Linguistics and Communication Studies, University of Port Harcourt.

Mbah, Evelyn. And Uzoigwe, Benita. 2013. Elements in Traditional and Modern Numerals in Nsukka Igbo. In Ozo-Mekuri Ndimele and Eugene S.L Chan. (eds.). *The Numeral Systems of Nigerian Languages.*Vol. 2:2-25. M&J Grand Orbit Communications Ltd.

Obikudo, Ebitare. 2013. The Numeral System of Nkoroo. In Ozo-Mekuri Ndimele and Eugene S.L Chan. (eds.).*The Numeral Systems of Nigerian Languages.* Vol. 2:27-39. M&J Grand Orbit Communications Ltd.

O'Grady,William. and Archibald, John. 2008. *Contemporary Linguistic Analysis*. Sixth edition. Pearson Longman. Toronto.

Omachonu, Gideon. 2013. Derivational Processes in the Igala Numeral System. In Ozo-Mekuri Ndimele and Eugene S.L Chan. (eds.). *The Numeral Systems of Nigerian Languages.*.Vol.2:2-25. M&J Grand Orbit Communications Ltd.

Stolz, Thomas. and Veselino, Ljuba.N. 2005. Ordinal Numerals In Haspelmath, Martin., Dryer, Mathew.S.,Gil, David. and Comrie Bernard (eds). The World Atlas of Language Structure, 218-224. Oxford.

Trask, R.L. 1996. A Dictionary of Grammatical Terms in Linguistics. Routlrdge.

Williamson, Kay. 1984. *Practical Orthography in Nigeria*. Ibadan: Heineman Education Books.

Yul-Ifode, Shirley. 1999. *A Course in Phonology*. Port Harcourt: University of Port Harcourt.

Perceptual Discrimination of Proper Names and Homophonous Common Nouns

Chelsea Sanker
Stanford University

Various characteristics of homophone mates can result in acoustic differences in how they are produced, e.g., frequency, part of speech, and morphological breakdown. Some of these acoustic cues influence perceptual decisions, though resulting accuracy is usually quite low. In this study, I examine homophone mate pairs in which one of the words is a proper name, analyzing both the acoustic characteristics that differ between proper names and homophonous common nouns as well as how these characteristics influence listeners' decisions in identifying these items. Proper names are produced with longer duration, higher F0, and higher intensity than common nouns. All of these characteristics have a corresponding influence on how listeners identify a stimulus; longer duration, higher F0, and higher intensity increase the likelihood that a stimulus will be identified as being a proper name. I explain these acoustic differences as resulting from proper names receiving sentential stress more often than other words and listeners having corresponding expectations that proper names are more likely to be stressed.

Keywords: homophones, perception, proper names, sentential stress

1 Introduction

Homophone mates often exhibit acoustic differences when produced in natural speech. There are multiple ways that these patterns in production might be explained. On the one hand, they might reflect phonetic details that are an inherent part of the representation of each word, consistent with an exemplar model (cf. e.g., Goldinger 1998, Pierrehumbert 2002). On the other hand, they might be explained by effects of the environment and the process of lexical access, e.g., higher predictability in context facilitating lexical retrieval and subsequently characteristics like duration (Gahl et al. 2012, Kahn & Arnold 2012), and part of speech aligning with prosodic differences based on the position that a word typically occurs in (Sorensen et al. 1978, Conwell 2017).

In contrast to some of the relatively robust differences found in production, perceptual identification of homophone mates generally has low accuracy (e.g., Bond 1973), even under facilitating conditions, e.g., after recent exposure to the same words, with all items produced by the same speaker (Sanker 2022). While low accuracy in perceptual identifications might support the analysis that production patterns are merely effects of the environment, the fact that accuracy is ever above chance might be interpreted as favoring the analysis that phonetic details are part of the representation of individual words.

In this paper, I examine perceptual identifications of proper names and homophonous common nouns. Based on listeners' overall accuracy and the acoustic cues that predict their decisions, I argue that there are predictable prosodic tendencies of proper names. They are more likely to receive sentential stress than other words are, and listeners make use of these acoustic characteristics because they expect proper names to exhibit correlates of stress.

1.1 *Acoustic details in production* Homophone mates can differ in the acoustic characteristics that they are produced with. For example, lower frequency words have longer durations and larger vowel spaces

* I would like to thank the audience at WECOL 2022 for their thought-provoking questions and comments.

than higher frequency words (Gahl 2008, Guion 1995). Nouns have longer durations and larger vowel spaces than verbs (Lohmann 2018, Conwell 2017). Some studies find that segment duration is influenced by the morphological breakdown of a word, but results vary across studies and across different morphemes (Plag et al. 2017, Seyfarth et al. 2018).

There are two main possibilities for how these phonetic differences are represented. The first possibility is that they are an inherent part of the representation of specific words. The second possibility is that they are driven by syntactic and pragmatic context and cognitive processes in lexical access. In addition to the question of how these differences are represented, it is relevant to consider how these differences arise. Even if the acoustic patterns become part of the inherent part of specific words, effects of contextual factors are necessary in order to explain how the differences originate and why they have the observed patterns, e.g., why lower frequency words have longer durations rather than the opposite.

Many of the acoustic differences between homophone mates disappear when environmental factors are controlled for, suggesting that they are effects of the environment rather than being inherent characteristics in the representation of individual words. For example, Jurafsky et al. (2002) demonstrate that the frequency-based differences between homophone mates within a speech corpus are largely eliminated when factors like speech rate, predictability based on neighboring words, and surrounding segments, are included as factors. Guion (1995) finds that frequency-based differences are only present for words produced in meaningful sentences, and that they are eliminated when words are elicited in a frame sentence. Differences based on part of speech are also reduced when position in the sentence is controlled (Sorensen et al. 1978, Conwell 2017), though there are still some differences predicted by part of speech when position in the sentence is controlled (Conwell 2017, Lohmann & Conwell 2020).

Some work suggests that phonetic patterns caused by the context that a word frequently occurs in can become part of the representation of that word; some effects of word-specific informativity are significant even when the environmental factors are accounted for (e.g., Tang & Shaw 2021, Seyfarth 2014). Sóskuthy and Hay (2017) find that words which are often lengthened due to occurring utterance-finally are also longer than other words when they occur elsewhere. However, these results might reflect indirect effects, rather than the phonetic details becoming part of the representation of specific words. Informativity might influence ease of lexical retrieval, which in turn results in acoustic differences (Gahl et al. 2012, Kahn & Arnold 2012). The relationship between how frequently a word occurs utterance-finally and the typical duration of that word after accounting for position might also be an effect of word-specific informativity: A word might have longer duration due to low overall informativity, and words with lower informativity might also be more likely to occur in prominent positions, such as utterance-finally. Effects of informativity do not necessarily require word-specific phonetic targets, even when informativity are associated with the particular word rather than its context.

Contextual effects associated with part of speech are similar for real words and for nonce words (Conwell & Barta 2018). Given that nonce words do not have pre-existing representations, these patterns in nonce words must be attributed to the environment rather than being inherent to the (nonce) word's representation. Other environmental effects are also similar in real words and nonce words. For example, the first mention of a word has a longer duration than the second mention (Fowler & Housum 1987, Clopper & Turnbull 2018), which is also observed in nonce words (Keung 2013). These effects must be due to context rather than being inherent to particular words, since the variation is within the realization of an individual word. Goldinger (1998) demonstrates that the number of repetitions of nonce words can create patterns similar to the lexical frequency of real words, so the results for effects of repetition on acoustic characteristics may suggest that effects of lexical frequency could be explained in the same way. If these patterns can be predicted without word-specific phonetic details, using word-specific phonetic details to account for the same patterns in real words is unnecessary.

In convergence experiments, shifts are generalized to words that were not heard during exposure and also to sounds that were not heard during exposure but which have features shared with the exposure items, e.g., exposure to lengthened VOT in /p/ results in lengthened VOT in /p/ in novel words and also lengthened VOT in /k/ (Nielsen 2011). While generalization of a shift across words does not exclude the possibility that listeners also have word-specific phonetic targets, it raises the question of the weighting that each association would have, e.g., how strongly weighted word-specific phonetic details would need to be in order to outweigh category-level phonetic details, given that speakers encounter far more instances of each phonological category than of a specific word containing that phoneme.

1.2 *Acoustic details in perception* Patterns in perception provide a separate line of evidence that can help shed light on the status of phonetic details as they relate to individual words. There is evidence that listeners do have acoustically detailed memories, e.g., more accurately remembering that they have heard a word if it is presented again in the same voice (e.g., Hintzman et al. 1972) and more accurately identifying familiar tokens (Chiu 2000). However, these acoustically detailed memories are not limited to speech characteristics; listeners identify a word more accurately when it is presented with the same background noise (e.g., a phone ringing) the second time (Pufahl & Samuel 2014). This sensitivity to non-linguistic acoustic details might suggest that these studies are capturing something about the broad range of details retained in short-term memory, rather than indicating that phonetic details of particular speech events are integrated into the representation of each word.

Some of the tendencies that are present in production seem to influence listeners' expectations. For example, speakers make more accurate identifications for stimuli that exhibit reduction patterns that align with the reduction that they are typically produced with, e.g., deletion of /ə/ (Connine et al. 2008) and realization of underlying /t/ (Pitt et al. 2011). The existence of these perceptual effects might provide evidence in favor of phonetic details being part of word-specific representations. However, these expectations do not need to be based on word-specific phonetic representations; it is possible that listeners' decisions are based on expectations about broader patterns. For example, listeners might expect high frequency words to be reduced, rather than having separate expectations about the reduction of each particular word. A potential parallel comes from reduction with repetition; listeners have above-chance accuracy in deciding whether a stimulus was the speaker's first or second time saying that word (Fowler & Housum 1987), which cannot be due to phonetic details inherent in the representation of each word.

Listeners are also sensitive to the acoustic correlates of part of speech. Conwell (2015) demonstrates that noun vs. verb uses of polysemous real words like *hug* elicit different neural responses, but finds no significant effect for nonce words produced in the same sentences. While this could be interpreted as suggesting that acoustic differences are encoded in the representation of words, the results for real words and nonce words in this study might differ because nonce words do not activate a stage of processing in which they would be linked with part of speech. Infants habituated to noun forms of phonologically ambiguous items (e.g., *dance*) preferentially look towards stimuli of verbs, suggesting that they are sensitive to the acoustic patterns that are shared within each category (Conwell & Morgan 2012). Notably, this learning is at the level of the part of speech category, rather than being associated with individual words, because the infants were being habituated to the broader part-of-speech categories rather than one part of speech just for a particular word.

Listeners are also sensitive to correlates of emotional valence, e.g., duration, F0 mean, and F0 range. They can identify the emotion being conveyed by a speaker (Nygaard & Lunders 2002) and also will use these cues to emotion to identify the meaning of nonce words (Nygaard et al. 2009) and homophones (Nygaard & Lunders 2002). The use of these cues in nonce words indicates that listeners have associations between emotional valence and acoustic characteristics that are independent of the representation of specific words.

Although accuracy for distinguishing between most homophone mates is low, some homophone mates seem to have higher discriminability based on having strong tendencies in their prosody. Martinuzzi and Schertz (2022) demonstrate that listeners can distinguish between the attention-seeking vs. apology functions of "sorry" with high accuracy (64.7%). Several of the prosodic cues that distinguish these two functions in production are predictive of listeners' identifications: Duration, mean F0, intensity, and F0 contour. Accuracy in these decisions doesn't necessarily require word-specific memories (cf. intonational patterns for questions vs. statements); the prosodic differences between each function of "sorry" may be a by-product of syntax and pragmatics, rather than being inherent parts of the representation of the word. In part, listeners' accuracy for these items may be influenced the fact that both functions of "sorry" often occur in isolation, which could help listeners map the prosodic patterns from production onto the stimuli being heard in isolation.

1.3 *Proper names* Homophone mate pairs with proper names (e.g., *Phoenix, phoenix*) may be a useful group to examine, because proper names differ from other words in a range of ways. Little previous work has examined whether proper names and common nouns exhibit systematic phonetic differences, though there is some evidence for such differences, e.g., in duration (Whalen & Wenk 1994). There are several

reasons why differences might be expected.

Proper names can differ from other nouns syntactically; for example, determiners cannot combine with proper names in English (Longobardi 1994). Such patterns might suggest that proper names have different syntactic structure than other nouns, e.g., forming DPs on their own. Different phrasal structure could result in different prosodic patterns for proper names than for other nouns. If those prosodic differences set expectations that listeners use in identifying words, homophone mate pairs including proper names might produce perception results similar to what was found for the attention-seeking vs. apology functions of "sorry" (Martinuzzi & Schertz 2022). Use of prosodic cues might interact with the type of context expected for different types of proper name; personal names might be easier to identify than other words because they appear in isolation as vocatives, while most nouns usually do not appear in isolation.

Proper names might be processed differently than other words are. They differ from other words semantically because they have no clear meaning as such; they just function as reference to particular entities (e.g., Yasuda et al. 2000). This relative lack of meaning may underlie why people tend to remember names less accurately than other words, and why names with no associated meaning are remembered less accurately than words which do have an associated meaning, e.g., the name *Baker* is associated with the noun *baker* (Cohen 1990). The differences in processing of proper names are also reflected in different neural activity (Yasuda et al. 2000, Desai et al. 2023) and can produce differences in aphasia, with some patients exhibiting greater impairment for proper names than common nouns and others exhibiting greater impairment for common nouns than proper names (Semenza 2006). Differences in lexical access might result in phonetic differences, as discussed above (Gahl et al. 2012, Kahn & Arnold 2012), and the lack of semantic connections might impact predictability and subsequently the acoustic correlates of predictability.

Personal names are likely to be less predictable in context than other words; when the referent is predictable from the discourse context, a name is likely to be replaced by a pronoun. Other words may have high predictability in context based on their semantic connections, but names lack this semantic network (Yasuda et al. 2000). Less predictable words are more likely to be stressed (Pan & Hirschberg 2000), so proper names might be more likely to receive sentential stress than other words are. If proper names are more likely to be stressed, they should exhibit the characteristics of sentential stress, including longer duration, higher F0, and greater intensity (Breen et al. 2010).

Lexical frequency might also contribute to phonetic differences between proper names and other words. Proper names usually have lower frequency than other words: Among words in SUBTLEX that are listed as just having a noun usage and a proper name usage, the proper names have a median log frequency of 1.1, while the nouns have a median log frequency of 2.5. As discussed above, lexical frequency is correlated with duration and other reduction patterns (Gahl 2008, Guion 1995, Clopper & Turnbull 2018).

1.4 *This study* This study examines the perceptual identification of homophone mate pairs in which one is a proper name, using several categories of proper nouns with a range of lexical frequencies. What acoustic cues do listeners use to try to distinguish between these homophone mates, and are these the same acoustic differences that are present in production? The results can shed some light on the status of phonetic details in the representation of proper names and other words.

2 Methods

2.1 *Participants* The participants were 22 native speakers of American English, who were members of the Brown University community. Data was also collected from 2 additional participants, but they were excluded based on providing the same response for almost all of the trials for homophone mate pairs; that uniformity created convergence issues in models testing predictors of how a stimulus was identified, in addition to suggesting that those participants were not completing the task as intended.

2.2 *Stimuli* Stimuli were produced by two native speakers of American English, one male and one female, in meaningful sentences. The sentences were constructed so that homophone mates occurred in environments that were syntactically and phonologically as similar as possible (e.g., "John likes the Phoenix painting", "John likes the phoenix painting"). These sentences were elicited orthographically in randomized order using the software PsychoPy (Peirce 2007) in a sound-attenuated room with a stand-mounted Blue Yeti microphone using the Audacity software program, and digitized at a 44.1 kHz sampling

rate. The target items were extracted from these sentences and presented in pairs.

There were 5 categories of proper names: Brands (e.g., *Bobcat*), cities (e.g., *Buffalo*), human names (e.g., *Holly*), possessives (e.g., *Poppy's*), and teams/bands (e.g., *Dolphins*). There was also a category of definite phrases (e.g., *The Creature*).

2.3 *Procedure* The study was run in-person in a quiet room using PsychoPy (Peirce 2007); the acoustic stimuli were presented to participants over headphones. Participants were instructed that they would hear pairs of words, and that in each pair one item would be a word that occurs with a capitalized first letter, and one would be a word that occurs in lowercase. For half of the participants, the framing of the instructions was that they were deciding whether the "capitalized word" was the first word or the second word (e.g., *Phoenix, phoenix* or *phoenix, Phoenix*). For the other half of the participants, the framing of the instructions was that they were deciding whether the "plain/lowercase word" was the first word or the second word. Responses were given with the left and right arrow keys on the keyboard; instructions remained on the screen indicating which arrow corresponded to the "capitalized" or "lowercase" word being first and second.

There were 33 pairs of target items and 29 pairs of filler items. Filler trials contained unambiguous pairs (e.g., *Seattle, unicorn*); filler trials are not included in the analysis. Each pair appeared in both orders, for a total of 124 stimuli in each block. Each participant heard one block of items from each of the two speakers, with the same pairs in both blocks; the order of the speakers was balanced across participants.

Results come from mixed-effects regression models calculated with the lme4 package in R (Bates et al. 2015); p-values were calculated with the lmerTest package (Kuznetsova et al. 2015).

3 Results

3.1 *Accuracy* Table 1 presents the summary of an intercept-only mixed effects logistic regression model for accuracy in identifications of homophone mates. There were random intercepts for participant and for word pair.

Table 1 Regression model for accuracy.

	Estimate	SE	z-value	p-value
(Intercept)	0.24	0.060	4.0	< 0.0001

Overall accuracy was significantly higher than chance (56%); listeners could distinguish between proper names and homophonous common nouns, though accuracy was low as compared to decisions about the phonologically distinct filler items (86%).

Including the category of proper name did not significantly improve the model ($\chi^2 = 1.7$, df = 4, p = 0.78), so it was not included as a factor. However, Figure 1 presents the accuracy of identifications for homophone mate decisions divided by category and by word. The accuracy is for both words in each homophone mate pair (e.g., both *apple* and *Apple*), even though the words are organized based on the category of the proper name. Some words potentially could fall into multiple categories (e.g., *Harmony* was categorized as a city, but can also be a human name).

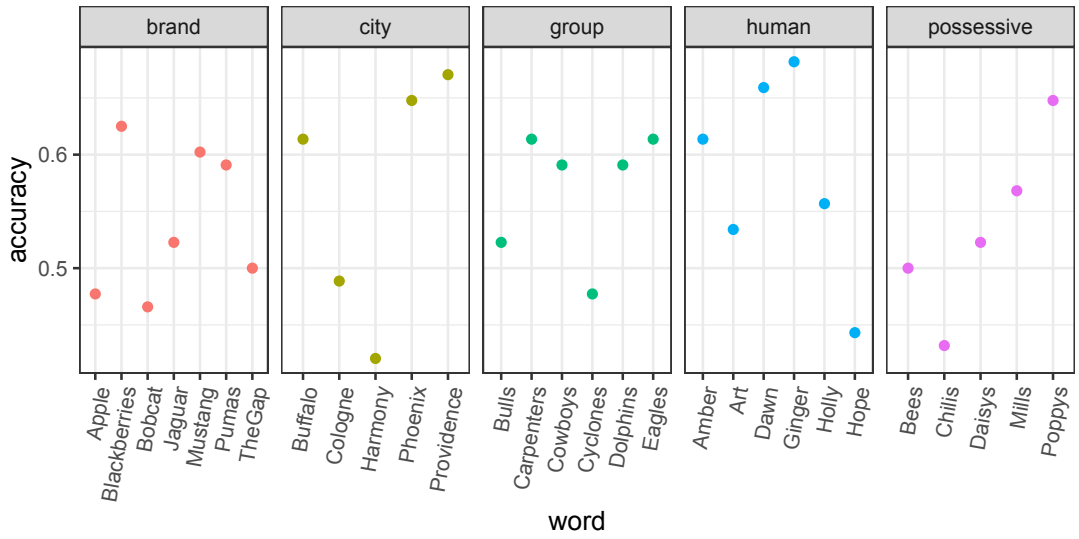
Adding trial number and block number marginally improved the model, suggesting a slight improvement with subsequent trials within a block ($\beta = 0.0029$, SE = 0.0017, z = 1.8, p = 0.077) but a decrease in accuracy in the second block, perhaps due to the switch to a different speaker ($\beta = -0.21$, SE = 0.12, z = -1.8, p = 0.074).

Lexical frequency was considered as a possible predictor, using the frequency of each word within the SUBTLEX corpus for US English (Brysbaert & New 2009). Analyses used log(1+Frequency), to handle words with a frequency of 0. Accuracy was not predicted by lexical frequency; adding log lexical frequency to the model did not significantly improve the fit ($\chi^2 = 0.97$, df = 1, p = 0.33). For a model restricted to accuracy just for identifications of the proper names, log lexical frequency also did not provide a better fit than a model without it ($\chi^2 = 0.26$, df = 1, p = 0.61). Given the low frequency of many of the proper names, it is likely that some of the proper names were not familiar to all of the listeners, particularly the names of bands and sports teams. The lack of effect of lexical frequency suggests that listeners'

identifications are not based on memories of word-specific acoustic details.

Some of the possessives could not be familiar to participants because they were not based on real businesses. Chili's is a well-known chain and Bee's is a restaurant in Providence (where the study was run), while the other three possessives were invented as plausible business names; some participants might have been familiar with actual businesses by these names, but most of them probably were not. Notably, these three were the possessives with the highest accuracy.

Figure 1 Accuracy of identifications by category and by word.



3.2 *Acoustic correlates of proper names* What are the acoustic correlates of proper names vs. common nouns that might influence decisions? Several acoustic differences might be predicted, based on differences in typical lexical frequency and also differences in how often proper names and common nouns receive sentential stress. Table 2 presents the mean values for five acoustic characteristics of the homophone mates used as stimuli, divided by whether they were proper names or common nouns: Proper names had longer duration, higher mean F0, larger F0 range, higher intensity, and lower spectral tilt.

Table 2 Acoustic characteristics of homophone mates based on whether they were proper names or common nouns.

	Word Duration	F0 mean	F0 range	Intensity	Spectral Tilt
Proper name	457 ms	153 Hz	73.6 Hz	56.4 dB	-4.1
Common noun	441 ms	142 Hz	64.6 Hz	55.9 dB	-3.0

Table 3 presents the summary of a mixed effects logistic regression model for "capitalized" identifications (vs. "lowercase") as predicted by acoustic characteristics of the stimulus relative to the paired item. The fixed effects were word duration ratio, F0 mean ratio, intensity ratio, and spectral tilt ratio; all were centered. There were random intercepts for participant and for word pair.

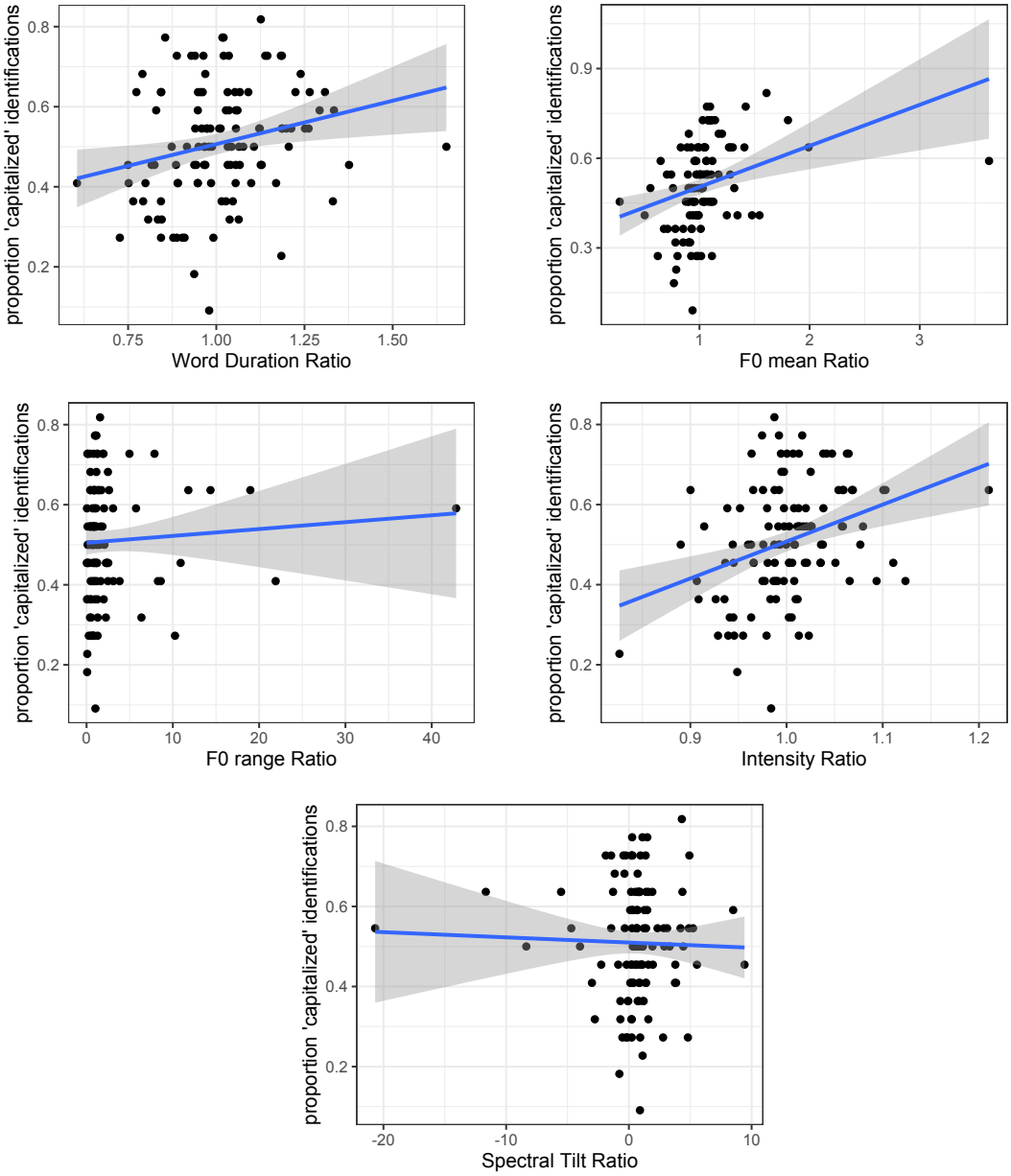
Table 3 Regression model for "capitalized" identifications.

	Estimate	SE	z-value	p-value
(Intercept)	0.071	0.096	0.73	0.46
Word Duration Ratio	1.1	0.27	4.2	< 0.0001
F0 Mean Ratio	0.48	0.14	3.5	0.00039
Intensity Ratio	4.1	0.85	4.8	< 0.0001
Spectral Tilt Ratio	0.013	0.013	1.0	0.31

Several acoustic characteristics of the stimuli were significant predictors of how a stimulus was identified. Listeners were more likely to identify a stimulus as being a proper name if it had longer duration, higher mean F0, or higher intensity. Figures 2a-e illustrate the relationship between acoustic characteristics and how listeners identified each stimulus.

A model including F0 range ratio was tested, but the strong correlation between F0 mean ratio and F0 range ratio within the stimuli ($r(56) = 0.78$, $p < 0.0001$) makes a model including both factors unreliable. In that model, a larger F0 range predicted significantly fewer "capitalized" identifications, which is the opposite of the relationship observed in production. In a model including F0 range and excluding F0 mean, there is no evidence for an effect of F0 range.

Figure 2 The proportion of "capitalized" identifications as predicted by the acoustic characteristic of each stimulus relative to its paired homophone mate: (a) Word duration ratio, (b) F0 mean ratio, (c) F0 range ratio, (d) Intensity ratio, (e) Spectral tilt ratio.



There is no evidence for response time influencing use of acoustic cues. No interactions between the acoustic predictors and log response time (as measured from the beginning of the first item of the pair) produced a model with a significantly better fit than a model without interactions.

The proper names were from five different categories: Brands (e.g., *Bobcat*), cities (e.g., *Buffalo*), human names (e.g., *Holly*), possessives (e.g., *Poppy's*), and teams/bands (e.g., *Dolphins*). Cue usage could potentially differ based on the type of proper name. Adding an interaction between construction and F0 mean ratio provides a significantly better fit than a model that includes both factors but no interaction (χ^2 = 11.4, df = 4, p = 0.022). The effect of F0 mean is strongest for city names and human names, and weakest for team names and possessives; Figure 3 illustrates. Adding an interaction between construction and word duration ratio also provides a significantly better fit than a model without the interaction (χ^2 = 9.9, df = 4, p = 0.042). The effect of word duration is strongest for city names and weakest for human names; Figure 4 illustrates. Note that the figures are based on the raw data, not the model output.

Figure 3 The proportion of "capitalized" identifications as predicted by F0 mean ratio, by type of pair.

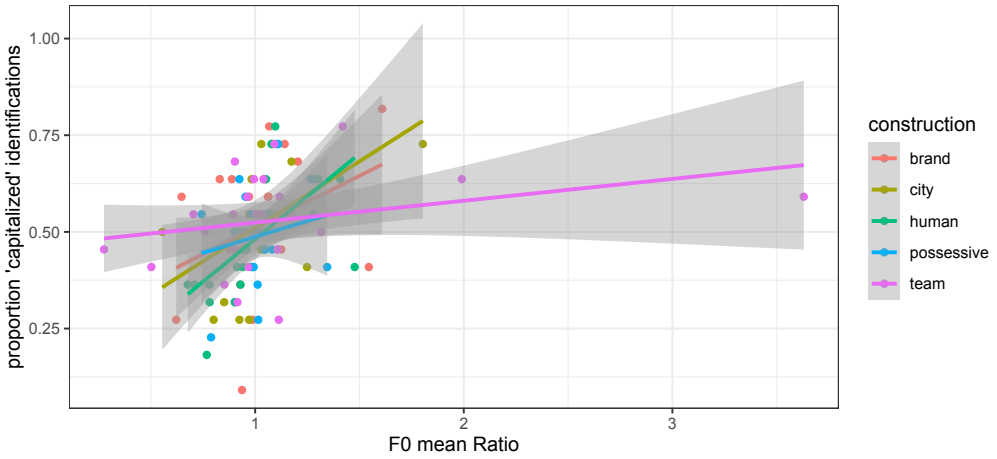
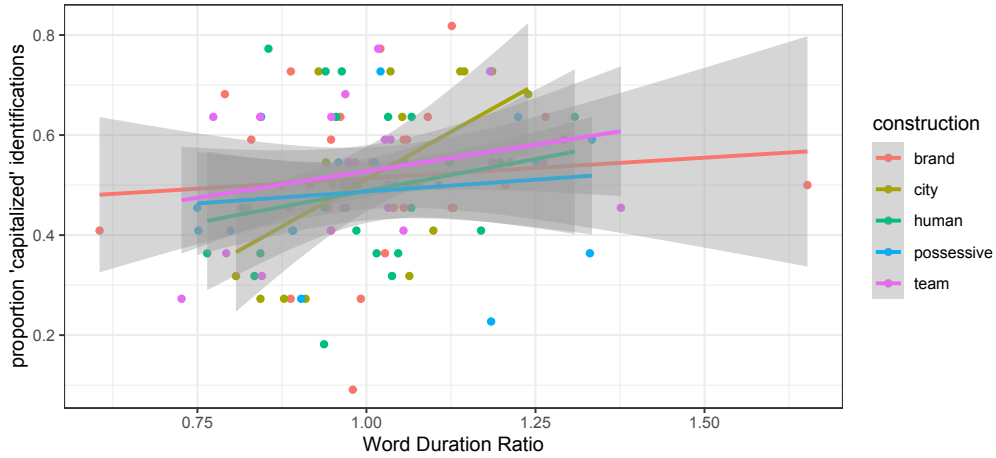


Figure 4 The proportion of "capitalized" identifications as predicted by word duration ratio, by type of pair.



3.3 *Lexical frequency* One of the potential factors driving the acoustic characteristics of proper names is lexical frequency. Previous work has demonstrated that lexical frequency is a predictor of word duration (Guion 1995, Gahl 2008), though it isn't as clear that it predicts the other acoustic characteristics that differ between proper names and common nouns.

Within the stimuli used in this experiment, log lexical frequency is a predictor of some acoustic characteristics: Word duration is negatively correlated with frequency (r(114) = -0.4, p < 0.0001), and F0

range is also negatively correlated with frequency ($r(114) = -0.18$, $p = 0.058$). The correlations between lexical frequency and F0 mean, intensity, and spectral tilt did not approach significance.

There is a confound between lexical frequency and being a proper name in this dataset; proper names have a much lower mean frequency than other words do (the mean log counts from SUBTLEX are 2.6 vs. 5.6). A similar difference is found in the lexicon more generally, as described above. Being a proper name seems to be a clearer predictor of the acoustic characteristics than lexical frequency.

Lexical frequency only predicts listeners' identifications to the extent that it is a predictor of acoustic characteristics. While lexical frequency does significantly predict identification decisions in a model with no acoustic predictors ($\beta = -0.047$ SE = 0.017, $z = -2.8$, $p = 0.0057$), adding lexical frequency to the model in Table 3 does not significantly improve the model ($\chi^2 = 1.4$, $df = 1$, $p = 0.24$).

3.4 *Sentential stress* One of the potential factors driving acoustic differences between proper names and homophonous common nouns is sentential stress. To examine this, the stimuli included four items that were capitalized vs. non-capitalized definite phrases (e.g., *the creature* vs. *The Creature*).

Table 4 presents the mean values for five acoustic characteristics of capitalized vs. non-capitalized definite phrases. While several of the effects are the same as what is observed in proper names vs. common nouns, F0 mean and F0 range are lower for the capitalized phrases than non-capitalized phrases, while they were both higher for proper names than for common nouns, suggesting a different form of emphasis.

Table 4 Acoustic characteristics of capitalized and not capitalized definite phrases.

	NP Duration	F0 mean	F0 range	Intensity	Spectral Tilt
Capitalized	509 ms	153 Hz	84.6 Hz	54.1 dB	-1.4
Not capitalized	500 ms	161 Hz	102.6 Hz	53.3 dB	1.7

Accuracy of identification of these items was higher than chance (60%). Table 5 presents the summary of a mixed effects logistic regression model for accuracy. The only fixed effect was type of pair (Definite Phrases, Bare Nouns). There were random intercepts for participant and for word pair.

Table 5 Regression model for accuracy among all homophone mates, comparing pair types. *Reference levels: Type = Definite Phrases.*

	Estimate	SE	z-value	p-value
(Intercept)	0.43	0.18	2.4	0.017
Type Bare Nouns	-0.19	0.19	-1.0	0.31

Accuracy for identifications of capitalized vs. non-capitalized definite phrases was significantly above chance. Accuracy for bare nouns (e.g., *Phoenix, phoenix*, as discussed in the previous sections) was slightly but not significantly lower than accuracy of identification of the definite phrases.

Table 6 presents the summary of a mixed effects logistic regression model for "capitalized" identifications (vs. "lowercase") for capitalized vs. non-capitalized definite phrases, as predicted by acoustic characteristics of the stimulus relative to the paired item. The fixed effects were word duration ratio, F0 mean ratio, intensity ratio, and spectral tilt ratio; all were centered. There were random intercepts for participant and for word pair.

Table 6 Regression model for "capitalized" identifications in capitalized vs. non-capitalized phrases.

	Estimate	SE	z-value	p-value
(Intercept)	0.068	0.13	0.54	0.59
Word Duration Ratio	3.1	1.0	3.0	0.0026
F0 Mean Ratio	-1.2	0.61	-2.0	0.048
Intensity Ratio	9.0	2.7	3.3	0.00092
Spectral Tilt Ratio	0.063	0.049	1.3	0.2

Listeners were more likely to identify a stimulus as being the capitalized phrase if it had longer duration, lower F0, or higher intensity.

4 Discussion

Listeners had above chance accuracy at distinguishing between proper names and homophonous nouns and at distinguishing between capitalized and non-capitalized definite phrases. Responses were strongly predicted by several acoustic cues: Word duration, mean F0, and intensity, which are also all correlates of these different categories in production. These results can be explained by listeners using prosodic cues based on expectations set by syntactic, semantic, and pragmatic characteristics.

The main factor that seems to drive the prosodic differences between proper names and common nouns is sentential stress. Proper names are probably more likely to be stressed than other words are, because they are more likely to be unpredictable in context and are often key elements of the utterances containing them. Their low lexical frequency may also contribute to how likely they are to receive stress, though the results suggest that lexical frequency is not directly driving the effects in this study. The primary differences observed between proper names and common nouns are characteristics of sentential stress: Longer duration, higher F0, greater intensity (Breen et al 2010).

The capitalized vs. non-capitalized definite phrases used in this study (e.g., *the creature* vs. *The Creature*) seem to exhibit a different type of intonational difference than proper names vs. common nouns, based on their acoustic characteristics. Adding capitalization to definite phrases does not seem to make them into proper names or create sentential stress, at least for the phrases used in this study. Using word-initial capitalization in phrases that are not proper names has been analyzed as indicating that the phrase refers to a well-established or prominent meaning (Linden 2020), which may be the function that has been captured in this study.

Listeners make use of the acoustic correlates of proper names and capitalized phrases, which is apparent both in overall above-chance accuracy and also the relationship between acoustic characteristics of a stimulus and how it was identified. Speakers and listeners use prosodic cues in a range of ways, such as indicating phrase boundaries and other syntactic contrasts (Shattuck-Hufnagel & Turk 1996, Cho et al. 2007). Thus, prosodic structure can set some expectations about syntactic, semantic, and pragmatic information; in this study, those expectations may be directly about proper names, or may be about sentential stress, which in turn is linked to the categories of words that are more likely to be stressed. Direct association of phonetic details with particular words is not necessary for listeners to make use of prosodic information. This study was set up to encourage decisions based on the broader group rather than considering each pair separately, as every trial asked listeners to identify whether the "capitalized" word (or "lowercase" word) was first or second in the pair, rather than asking listeners to identify the ordering of the particular words.

The results suggest that use of the prosodic cues associated with proper names is based around proper names as a broad category. Accuracy is above chance not just for names that are likely to be familiar but also for invented business names and for low-frequency names that are likely to be unfamiliar to many of the participants; there is no evidence that the use of acoustic cues is stronger for real names than for invented names or stronger for higher frequency names than for lower frequency names. This generalized cue usage supports the analysis that listeners' expectations about acoustic characteristics that distinguish between proper names and homophonous common nouns are set by syntactic and pragmatic factors rather than by phonetic details that are associated with the representations of particular words.

There was some evidence for differences in cue usage for different types of proper names (e.g., city names vs. team names), though the limited number of items in each category makes it somewhat unclear what might drive these differences, e.g., variation in word length, lexical frequency, or different typical prosodic environments (e.g., the fact that human names can appear in vocatives, while the other categories of proper names generally will not). In this study, the only pairs that differed morphologically were the possessive vs. plural pairs (e.g., *poppies* vs. *Poppy's*). If listeners are sensitive to differences in duration between different morphemes (cf. Plag et al. 2017), the effect of duration might be a stronger predictor for these items. However, this category did not exhibit a stronger effect of word duration than other categories; perhaps duration of the word is not sufficient to capture an effect of the duration of particular segments.

Use of prosodic cues associated with capitalized phrases is necessarily capturing a contextually-driven pattern based on factors outside of the individual word, because the capitalized and non-capitalized phrases contain exactly the same words. Expectations about prosody seem to be set by a particular pragmatic usage of word-initial capitalization. One way that this form of capitalization is used informally in written English

is to draw attention to the intended meaning being the most prominent or well-established meaning rather than something that is more limited to the particular discourse context (Linden 2020).

5 Conclusions

Proper names and homophonous common nouns exhibit several systematic phonetic differences, which seem to align with correlates of phrasal stress: Word duration, mean F0, and intensity. Listeners make use of these acoustic characteristics when identifying these items, demonstrating expectations for the acoustic characteristics of proper names vs. common nouns. The results can be explained based on connections between prosodic structure and semantic categories like proper names; similar results for real and invented names suggests that the results are not driven by phonetic details associated with individual words.

References

Bates, Douglas, Martin Mächler, Ben Bolker, & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1-48.

Bond, Zinny S. 1973. The perception of sub-phonemic phonetic differences. *Language and Speech* 16. 351-355.

Breen, Mara, Evelina Fedorenko, Michael Wagner & Edward Gibson. 2010. Acoustic correlates of information structure. *Language and Cognitive Processes* 25. 1044-1098.

Brysbaert, Mad & Boris New. 2009. Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods* 41(4). 977-990.

Chiu, Ching-Yiu Peter. 2000. Specificity of auditory input and explicit memory: Is perceptual priming for environmental sounds exemplar specific? *Memory & Cognition* 28(7). 1126-1139.

Cho, Taehong, James M. McQueen & Ethan A. Cox. 2007. Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics* 35. 210-243.

Clopper, Cynthia G. and Rory Turnbull. 2018. Exploring variation in phonetic reduction: Linguistic, social, and cognitive factors. In Francesco Cangemi, Meghan Clayards, Oliver Niebuhr, Barbara Schuppler & Margaret Zellers (eds), *Rethinking reduction: Interdisciplinary perspectives on conditions, mechanisms, and domains for phonetic variation*, 25-72. Berlin: Mouton de Gruyter.

Cohen, Gillian. 1990. Why is it difficult to put names to faces? *British Journal of Psychology* 81. 287-297.

Connine, Cynthia M., Larissa J. Ranbom & David J. Patterson. 2008. Processing variant forms in spoken word recognition: The role of variant frequency. *Perception & Psychophysics* 70(3). 403-411.

Conwell, Erin. 2015. Neural responses to category ambiguous words. *Neuropsychologia* 69. 85-92.

Conwell, Erin. 2017. Prosodic disambiguation of noun/verb homophones in child-directed speech. *Journal of Child Language* 44(3). 734-751.

Conwell, Erin & James L. Morgan. 2012. Is it a noun or is it a verb? Resolving the ambicategoricality problem. *Language Learning and Development* 8. 87-112.

Conwell, Erin & Kellam Barta. 2018. Phrase position, but not lexical status, affects the prosody of noun/verb homophones. *Frontiers in Psychology* 9. Article 1785.

Desai, Rutvik H., Usha Tadimeti & Nocholas Riccardi. 2023. Proper and common names in the semantic system. *Brain Structure and Function* 228. 239-254.

Fowler, Carol & Jonathan Housum. 1987. Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language* 26. 489-504.

Gahl, Susanne. 2008. Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language* 83(4). 474-498.

Gahl, Susanne, Yao Yao & Keith Johnson. 2012. Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *Journal of Memory and Language* 66. 789-806.

Goldinger, Stephen D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105(2). 251-279.

Guion, Susan G. 1995. Word frequency effects among homonyms. In *Texas linguistic forum,* vol. 35, 103-116.

Hintzman, Douglas, Richard A. Block, & Norman R. Inskeep.1972. Memory for mode of input. *Journal of Verbal Learning and Verbal Behavior* 11(6). 741-749.

Jurafsky, Daniel, Alan Bell & Cynthia Girand 2002. The role of the lemma in form variation. In Carlos Gussenhoven &

Natasha Warner (eds.), *Laboratory phonology VII*, 3-34. Berlin: Mouton de Gruyter.
Kahn, Jason M. & Jennifer E. Arnold. 2012. A processing-centered look at the contribution of givenness to durational reduction. *Journal of Memory and Language* 67. 311-325.
Keung, Lap-Ching. 2013. *Effects of discourse status and planning difficulty on acoustic variation*. Chapel Hill, NC: University of North Carolina at Chapel Hill Bachelor of Science thesis.
Kuznetsova, Alexandra, Per Bruun Brockhoff & Rune Haubo Bojesen Christensen. 2015. *lmerTest: Tests in linear mixed effects models.* https://CRAN.R-project.ord/package=lmerTest. R package version 2.0-29.
Linden, Josh. 2020. Contrastive Focus Capitalization: Nonstandard usages of capital letters in web-based English and their capital-I Implications. *Studies in the Linguistic Sciences: Illinois Working Papers 2020*. 116-138.
Lohmann, Arne. 2018. *Cut* (N) and *cut* (V) are not homophones: Lemma frequency affects the duration of noun-verb conversion pairs. *Journal of Linguistics* 54(4). 753-777.
Lohmann, Arne & Erin Conwell. 2020. Phonetic effects of grammatical category: How category-specific prosodic phrasing and lexical frequency impact the duration of nouns and verbs. *Journal of Phonetics* 78. Article 100939.
Longobardi, G. 1994. Reference and proper names: A theory of N-movement in syntax and logical form. *Linguistic Inquiry*, 25(4), 609-665.
Martinuzzi, C. & Schertz, J. 2022. Sorry, not sorry: The independent role of multiple phonetic cues in signaling the difference between two word meanings. *Language and Speech*, 65(1), 143-172.
Nielsen, Kuniko. 2011. Specificity and abstractness of VOT imitation. *Journal of Phonetics* 39. 132-142.
Nygaard, Lynne C., Debora S. Herold & Laura L. Namy. 2009. The semantics of prosody: Acoustic and perceptual evidence of the prosodic correlates to word meaning. *Cognitive Science* 33. 127-146.
Nygaard, Lynne C. & Erin R. Lunders. 2002. Resolution of lexical ambiguity by emotional tone of voice. *Memory & Cognition* 30(4). 583-593.
Pan, Shimei & Julia Hirschberg. 2000. Modeling local context for pitch accent prediction. In *Proceedings of the 38th annual meeting of the Association for Computational Linguistics*, 233-240.
Peirce, Jonathan W. 2007. PsychoPy–Psychophysics software in Python. *Journal of Neuroscience Methods* 162(1-2). 8-13.
Pierrehumbert, Janet. 2002. Word-specific phonetics. In Carlos Gussenhoven & Natasha Warner (Eds.), *Laboratory Phonology VII*, 101-139. Berlin: Mouton de Gruyter.
Pitt, Mark A., Laura Dilley & Michael Tat. 2022. Exploring the role of exposure frequency in recognizing pronunciation variations. *Journal of Phonetics* 39(3). 304-311.
Plag, Ingo, Julia Homann & Gero Kunter. 2017. Homophony and morphology: The acoustics of word-final S in English. *Journal of Linguistics* 53(1). 181-216.
Pufahl, April & Arthur G. Samuel. 2014. How lexical is the lexicon? Evidence for integrated auditory memory representations. *Cognitive Psychology* 70. 1-30.
Sanker, Chelsea. 2022. Homophone discrimination based on prior exposure. *Journal of Phonetics* 95. Article 101182.
Semenza, Carlo. 2006. Retrieval pathways for common and proper names. *Cortex* 42. 884-891.
Seyfarth, Scott. 2014. Word informativity influences acoustic duration: Effects of contextual predictability on lexical representation. *Cognition* 133(1). 140-155.
Seyfarth, Scott, Marc Garellek, Gwendolyn Gillingham, Farrell Ackerman & Robert Malouf. 2018. Acoustic differences in morphologically-distinct homophones. *Language, Cognition and Neuroscience* 33(1). 32-49.
Shattuck-Hufnagel, Stefanie & Alice E. Turk. 1996. A prosodic tutorial for investigators of auditory sentence processing. *Journal of Psycholoinguistics Research* 25(2). 193-247.
Sorensen, John M., William E. Cooper & Jeanne M. Paccia. 1978. Speech timing of grammatical categories. *Cognition* 6(2). 135-153.
Sóskuthy, Márton & Jennifer Hay. 2017. Changing word usage predicts changing word durations in New Zealand English. *Cognition* 166. 298-313.
Tang, Kevin & Jason A. Shaw. 2021. Prosody leaks into the memories of words. *Cognition* 210. Article 104601.
Whalen, Douglas H. & Heidi E. Wenk. 1994. Durational characteristics of proper names common words. *Journal of the Acoustical Society of America* 95. 2924.
Yasuda, Kiyoshi, Tetsuo Nakamura & Bobbie Beckman. 2000. Brain processing of proper names. *Aphasiology*,14(11). 1067-1089.

All-or-nothing Long-Distance Metathesis in Italiot Greek

Eirini Apostolopoulou
Aristotle University of Thessaloniki

This paper investigates the conditions under which diachronic long-distance metathesis (LDM) of rhotics takes place in the two Italiot Greek (IG) dialects, i.e. Apulian Greek and Calabrian Greek. In IG, the process of LDM is manifested as the displacement of a rhotic from a non-initial complex onset to the leftmost syllable of the root. The phonological analysis, couched within Optimality Theory (Prince & Smolensky 1993/2004), accounts for the phenomenon in the dialects at hand, and, importantly, captures cases of metathesis blocking, which have drawn little, if any, attention in previous literature. Special emphasis is placed on the cross-dialectal variation observed between Apulian and Calabrian Greek with respect to the evolution of specific non-initial configurations.

Keywords: long-distance metathesis, Italiot Greek, Optimality Theory

1 Introduction

Long-distance metathesis (LDM) of liquids is robustly attested in the diachrony of several Romance languages, such as Sardinian (Wagner 1941; Bolognesi 1998; Molinu 1999; Lai 2013, 2015), Gascon (Grammont 1905; Blevins & Garrett 1998), Alguerese Catalan (Torres-Tamarit et al. 2012), Judeo-Spanish (Lipski 1990; Bradley 2007), and Italo-Romance (Rohlfs 1966; Kilpatrick 2010). Moreover, LDM is found in Italiot Greek (IG; Rohlfs 1930, 1950), namely the two Modern Greek dialects still spoken in Apulia and Calabria (see Manolessou 2005 for a historical and sociolinguistic overview). As a general rule, this type of metathesis involves the intersyllabic displacement of a post-consonantal liquid and the creation of a consonant–liquid sequence closer to the left edge of the word, with two distinct patterns emerging cross-linguistically (for a typological analysis see Apostolopoulou 2022):

(a) LDM to the leftmost syllable
(e.g. IG, Sardinian, Gascon, Southern Italo-Romance);

(b) LDM to the adjacent syllable to the left
(e.g. Judeo-Spanish, Alguerese Catalan, Northern Italo-Romance).

Although LDM in IG has attracted the attention of recent literature (Blevins & Garrett 2004; Canfield 2015; Chandlee 2014; Coffman 2013), certain idiosyncrasies remain underexplored. The remainder of the paper is organized as follows. In section 2, I present the LDM pattern that IG displays as well as the conditions under which LDM fails to apply. In section 3, I propose a phonological analysis within Optimality Theory (OT; McCarthy & Prince 1995; Prince & Smolensky 1993/2004). In section (4), I briefly discuss the biases that underlie the process of LDM and I argue that language change is the product of constraint re-ranking. Finally (section 5), I round up the paper by summarizing its key points.

* I am indebted to my informants in Apulia and Calabria, my PhD thesis advisors, Birgit Alber and Martin Krämer, and the audience of *WECOL 2022*. All errors are my own.

Proceedings of *WECOL 2022*

2 LDM and metathesis blocking in IG

The historical evolution of IG from Medieval Greek involved, among other changes, LDM of rhotics,[1] from a non-initial complex onset /Cr/ to the leftmost syllable of the root (Blevins & Garrett 2004; Coffman 2013; Rohlfs 1950; Apostolopoulou 2022). As a result, innovative forms emerged that contained a root-initial complex onset /Cr/, as shown in (1a–f) (first column: Medieval Greek; second column: IG; unless specified otherwise, the contemporary data are taken from own fieldwork; see also Rohlfs 1930; Karanastassis 1984–1992; Greco & Lambrogiorgou 2001; Squillaci & Squillaci 2016).

(1)	a.	pán.d**r**em.ma	>	p**r**án.dem.ma	‘marriage’
	b.	kó.p**r**i.a	>	k**r**ó.pi.a	‘manure’
	c.	vó.θ**r**a.ko	>	v**r**ú.θa.ko	‘frog’
	d.	gam.b**r**ó	>	g**r**am.bó	‘son-in-law’
	e.	xon.d**r**ó	>	x**r**on.dó	‘fat’
	f.	ka.pís.t**r**i	>	k**r**a.pís.ti	‘halter’

Unlike rhotics in non-initial complex onsets, rhotics in coda position are unaffected by metathesis (2a–b).

(2)	a.	xó**r.**to	*x**r**ó.to	‘grass’
	b.	ka**r.**pó	*k**r**a.pó	‘fruit’

Systematic exceptions to the general rule that non-initial post-consonantal rhotics are attracted by the first syllable have been pointed out in the existing literature, the main focus being on unfitting landing sites (Blevins & Garrett 2004: 130–131). First, no movement is triggered when the root begins with a vowel – in other words, with an onsetless syllable (3a–b) – despite the fact that the result, i.e. a root beginning with /rV/, is grammatical (3a′–b′).

(3)	a.	a.lé.v**r**i	***r**a.lé.vi	‘flour’
	a′.		**r**ad.dí	‘stick’
	b.	é.xen.d**r**a	***r**é.xen.da	‘grass snake’
	b′.		**r**é.ma	‘sea’

Even if the C_V environment is available, /r/ does not move next to another sonorant (4a–b), a sibilant (4c), or an affricate (4d), as the resulting sequences (e.g. [mr], [sr]) clash with the phonotactics of IG (Blevins & Garrett 2004).

(4)	a.	má.v**r**o	*m**r**á.vo	‘black’
	b.	ne.f**r**ó	*n**r**e.fó	‘kidney’
	c.	sá.g**r**i.pa	*s**r**á.gi.pa	‘type of goat’
	d.	t͡ʃén.d**r**o	*t͡ʃ**r**én.do	‘center’

Moreover, LDM fails to apply when the first onset is occupied by a coronal consonant (5a–b), even though roots of the same shape are well-documented (5a′–b′) (Blevins & Garrett 2004).

(5)	a.	ð**r**á.k**r**i	*ð**r**á.ki	‘tear’
	a′.		ð**r**á.ko	‘dragon’
	b.	ta.v**r**í	*t**r**a.ví	‘bull’
	b′.		t**r**í.pi	‘hole’

[1] Blevins & Garrett (2004) and Coffman (2013) refer to metathesis of *liquids* in IG, based on rare examples of Latin loans, e.g. *ko.nú.kla* > *klo.nú.ka* ‘distaff’ (found in the local Romance dialects). However, the only native root where LDM could have applied does not, in fact, display movement, i.e. *ka.vlí* **kla.ví* ‘stem’ (Karanastassis 1984–1992; Squillaci & Squillaci 2016). Given that the vast majority of words containing /Cl/ fall into categories in which LDM is blocked, e.g., *a.vlí* **la.vi* ‘yard’, *mú.xla* **mlú.xa* ‘mold’, *di.pló* **dli.pó* ‘double’ (see 3–4 below), the discussion is confined to LDM of rhotics, which is abundantly attested.

Aside from the limitations with respect to the landing site, an additional exception constitutes the immunity /r/ belonging to non-initial /tr/ clusters exhibit to LDM, even in the presence of a suitable host in the first syllable, as shown in (6a–c) (Apostolopoulou 2022).

(6) a. pé.t**r**a *p**r**é.ta ‘stone’
b. fi.t**r**ía *f**r**i.tí.a ‘potato bud’
c. ko.t**r**o.ní.t͡ʃi *k**r**o.to.ní.t͡ʃi ‘rock partridge’

However, /tr/ preceded by a sibilant /s/ appears to have evolved variably. In the Apulian variety (7a–b) (data from Greco & Lambrogiorgou 2001; Karanastassis 1984–1992), the Medieval /str/ usually remains intact, whereas in Calabria it tends to split (7c–d) (Squillaci & Squillaci 2016). Free alternation is reported in both dialects (7f) (Karanastassis 1984–1992; Squillaci & Squillaci 2016).

(7) a. pás.t**r**i.ko *p**r**ás.ti.ko ‘clean’
b. ka.pís.t**r**i *k**r**apís.ti ‘halter’
c. pás.t**r**i.ko > p**r**ás.ti.ko ‘clean’
d. ka.pís.t**r**i > k**r**a.pís.ti ‘halter’
f. gás.t**r**a > gás.t**r**a ~ g**r**ás.ta ‘type of vase’

Notably, LDM is an *all-or-nothing* process (Coffman 2013): unless it results in an initial [Cr], LDM does not take place at all. For example, in the absence of a suitable landing site, we do not encounter metathesis to a post-vocalic position, e.g. *a.lé.v**r**i, *a.lé**r.**vi* (3a), *má.v**r**o*, **má**r.**vo* (4a), *ðá.k**r**i*, **ðá**r.**ki* (5a), even though /VrC/ is allowed, e.g. *xó**r.**to* (2a). Finally, metathesis does not create non-initial [Cr], e.g. *é.xen.d**r**a*, *é.x**r**en.da* (3b).

It should be pointed out that LDM is not associated with metrical considerations. The movement of the rhotic occurs independently from the position of the main stress and does not serve to optimize the foot structure in any way. Consider, for instance, the first two examples: in *pán.d**r**em.ma* > *p**r**án.dem.ma* (1a), the rhotic migrates from an unstressed syllable to a stressed one, whereas in *gam.b**r**ó* > *g**r**am.bó* (1d) the opposite is observed. Moreover, in *ka.pís.t**r**i* > *k**r**a.pís.ti* (1f) none of the syllables involved bears stress. Additionally, the rhotic is not attracted exclusively by closed, i.e. potentially heavy, syllables: in (1a), it departs from and lands in closed syllables, while in *kó.p**r**i.a* > *k**r**ó.pi.a* (1b) both the source and the target syllable are open. Given the irrelevance of stress in the process, in what follows I omit the stress symbol in the input/output forms.

In the next section, I outline a phonological account that captures (a) the LDM pattern found in IG, (b) the underapplication of LDM due to the lack of a suitable phonological environment in the first syllable, and (c) the immunity of /tr/ as well as the variable fate of /str/. The proposed analysis is couched within the framework of OT (Prince & Smolensky 1993/2004; McCarthy & Prince 1995).

3 Phonological analysis

3.1 *LDM in IG* Concurring with the view that prominent positions such as the left edge of a word may accommodate greater complexity (Beckman 1998; Zoll 1996, 1998; de Lacy 2001; Smith 2005; Walker 2011; see also Alber 2001; Coffman 2013), I maintain that LDM is driven by the need to avoid complex onsets in non-initial, i.e. non-prominent, positions. Along these lines, the IG LDM pattern, i.e. the displacement of post-consonantal rhotics all the way to the first syllable and the formation of a novel Cr there, is captured as the interaction of a positional markedness constraint that disfavors non-initial complex onsets,[2] on the one hand, and a faithfulness constraint penalizing the disruption of the linear order of the input segments in the output, on the other hand. Specifically, I employ COINCIDE(Complex Onset, Left Edge) (COINCIDE; 8a) (Zoll 1998; see also Alber 2001; Coffman 2013), which is violated when a complex onset is not found at the left edge, and LINEARITY (8b) (McCarthy & Prince 1995), which returns a single violation mark when the precedence relations in the input are not preserved in the output. Additionally, I use *RC, which militates against pre-consonantal codas (8c) (Torres-Tamarit et al. 2012).

[2] Note that only obstruent-liquid clusters qualify to form a complex onset in IG (Apostolopoulou 2022).

(8) a. COINCIDE(Complex Onset, Left Edge)
Assign a violation for every complex onset that is not found in the first syllable

b. LINEARITY
Assign a violation if the precedence relations in the input are not preserved in the output (no metathesis)

c. *RC
Assign a violation for every pre-consonantal coda.

The constraint hierarchy that yields the IG pattern is given in (9). LINEARITY outranks *RC, therefore pre-consonantal rhotics stay put. Crucially, COINCIDE dominates LINEARITY, thus allowing for metathesis from /Cr/ environments to arise.

(9) COINCIDE >> LINEARITY >> *RC

As illustrated in Tableau (10), the faithful realization of an input containing a /Cr/ in the third syllable (candidate a) is ruled out by the high-ranking COINCIDE. Instead, the migrating rhotic lands in the first onset (candidate d). Docking on a non-initial onset (candidate b) or a coda position (candidate c) both do not constitute viable options, as fatal violations of COINCIDE and *RC, respectively, are incurred.

(10) Tableau /kapistri/

	/kapistri/	COINCIDE	LINEARITY	*RC
	a. ka.pis.tri	*!		
	b. ka.pris.ti	*!	*	
	c. kar.pis.ti		*	*!
☞	d. kra.pis.ti		*	

On the other hand, etymological codas remain intact, as any movement incurs a fatal violation of LINEARITY (Tableau 11).

(11) Tableau /karpo/

	/karpo/	COINCIDE	LINEARITY	*RC
☞	a. kar.po			*
	b. kra.po		*!	

3.2 *Unsuitable landing sites* LDM does not apply blindly to all roots containing a non-initial complex onset /Cr/. As illustrated above, /r/ does not move to onsetless syllables (see 3) as well as next to certain consonantal segments, either because of the ill-formedness of the resulting cluster, e.g. $C_{[son]}$r (see 4), or despite the well-formedness of the outcome, i.e. a $C_{[cor]}$r sequence (see 5). Interestingly, metathesis to alternative positions does not take place. In order to accommodate these cases of LDM blocking in the analysis, additional constraints are employed. First, for the sake of brevity, all possible ill-formed clusters based on the phonotactic rules of IG are taken to violate the constraint *ILLICITCLUSTERS (12a). Second, OCP[cor] disallows the emergence of consecutive coronal consonants (12b) (Anttila 2008; Pater & Coetzee 2005). Third, *PW]r penalizes prosodic words that begin with a liquid (12c) (Bolognesi 1998). Furthermore, the local conjunction LINEARITY&*RC (see Smolensky 1995; Lubowicz 2003) is violated when metathesis results in a novel coda (12d)

(12) a. *ILLICITCLUSTERS
Assign a violation for every phonotactically inadmissible cluster (incl. nasal-liquid, sibilant-liquid, affricate-liquid)

b. OCP[cor]
Assign a violation for every coronal-coronal sequence

c. *PW[r
Assign a violation for every word-initial liquid

d. LINEARITY&*RC
Assign a violation for every candidate in which the disruption of the linear order from the input to the output results in a liquid syllabified in a pre-consonantal coda

Tableau (13) explicates the absence of metathesis in roots with an initial vowel. The violation of COINCIDE by the winning candidate (a) is not enough to motivate metathesis due to the workings of higher-ranked constraints. Candidate (b), which contains a novel simplex onset [r] in the first syllable, fatally violates *PW[r. Alternative movement to non-initial onsets is also excluded: candidate (c), although the second onset constitutes a fitting host, is ruled out by LINEARITY. In essence, it includes an unnecessary movement that does not result in the avoidance of a non-initial complex onset. Finally, candidate (d), which involves metathesis to coda position, is eliminated by the conjunction LINEARITY&*RC.

(13) Tableau /exendra/

	/exendra/	LINEARITY&*RC	*PW[r	COINCIDE	LINEARITY	*RC
☞	a. e.xen.dra			*		
	b. re.xen.da		*!		*	
	c. e.xren.dra			*	*!	
	d. er.xen.dra	*!			*	*

However, roots beginning with /rV/ do not undergo metathesis (or another repair process), since the violation of *PW[r is not fatal (Tableau 14). Alternatives such as movement to a coda (candidate b) are eliminated due to the ranking LINEARITY&*RC >> *PW[r.

(14) Tableau /rema/

	/rema/	LINEARITY&*RC	*PW[r	COINCIDE	LINEARITY	*RC
☞	a. re.ma		*			
	b. er.ma	*!			*	

Similarly, due to the ranking LINEARITY&*RC >> OCP[cor] >> COINCIDE, LDM as well as alternative metathesis to e.g. a coda fail to apply when the root involved has an initial coronal (Tableau 15). In this case, given that /r/ is a coronal segment, LDM creates a coronal-coronal cluster that incurs a fatal violation of OCP[cor] (candidate b). By contrast, etymological coronal-coronal clusters in the first syllable are allowed, as potential repair strategies translate into fatal violations of constraints dominating OCP[cor] (e.g. Tableau 16, candidate b).

(15) Tableau /ðakri/

	/ðakri/	LINEARITY&*RC	OCP[cor]	COINCIDE	LINEARITY	*RC
☞	a. ða.kri			*		
	b. ðra.ki		*!		*	
	c. ðar.ki	*!			*	*

(16) Tableau /ðrako/

	/ðrako/	LINEARITY&*RC	OCP[cor]	COINCIDE	LINEARITY	*RC
☞	a. ðra.ko		*			
	b. ðar.ki	*!			*	*

Furthermore, LDM is not triggered when the process would create an illicit cluster, e.g. a sonorant-liquid sequence, in the first syllable. Such candidates violate the undominated constraint *ILLICITCLUSTER (Tableau 17, candidate b). Once again, other options, e.g. metathesis to coda position (candidate c), are excluded by LINEARITY&*RC.

(17) Tableau /mavro/

	/mavro/	LINEARITY&*RC	*ILLICITCLUSTERS	COINCIDE	LINEARITY	*RC
☞	a. ma.vro			*		
	b. mra.vo		*!		*	
	c. mar.vo	*!			*	*

3.3 *The curious case of /tr/* The last and most underexplored instance of LDM blocking concerned the non-initial /tr/. Independently of the structure of the first syllable, medial /tr/ resists splitting in both IG dialects (6). On the other hand, rhotics found in the configuration /str/ were subject to LDM in Calabrian Greek.

Interestingly enough, certain clusters, among which /tr/, were affected by retroflexion and, subsequently, affrication in all dialects of the Extreme South, including IG. Specifically, /tr/ seems to sound more closely to [tˢ] and to exhibit the behavior of an affricate, i.e. a single complex segment (Romano 1999; Loporcaro 2001; Celata 2006, 2010; Romano & Gambino 2010). Relying on these findings and considering the great influence Romance exerted on IG at multiple levels, it is viable to hypothesize that, being an inseparable unit, the IG /tr/ (henceforth represented as *t͡r*) is not subject to the restrictions imposed on a complex onset [tr] (see van de Weijer 1996). In other words, an output containing a non-initial [t͡r] does not violate COINCIDE (compare candidates a and b in Tableau 18). Thus, there is no motivation for LDM to be triggered; in fact, candidates involving metathesis are excluded by LINEARITY (candidate c).

(18) Tableau /pe*t͡r*a/

	/pe*t͡r*a/	COINCIDE	LINEARITY
☞	a. pe.*t͡r*a		
	b. pe.tra	*!	
	c. pre.ta		*!

According to the above-cited studies, retroflexion and affrication also affected /str/. A solution to the puzzle of the cross-dialectal variation between Calabrian and Apulian Greek with respect to the evolution of /str/ is to postulate that the processes that turned the cluster /tr/ into a complex segment /t͡r/ delayed in the context of /s/ in Calabria. Consequently, since the /tr/ at hand continued violating COINCIDE, LDM applied (consider Tableau 19).

(19) Tableau /pastriko/ (Calabrian Greek)

	/pastriko/	COINCIDE	LINEARITY
	b. pas.tri.ko	*!	
☞	c. pras.ti.ko		*

In Apulia, though, /str/ sequences were targeted by affrication early on. The complex segment was thus not forced to split (consider Tableau 20).

(20) Tableau /pas*t͡r*iko/

	/pas*t͡r*iko/	COINCIDE	LINEARITY
☞	a. pa.s*t͡r*i.ko		
	b. pas.tri.ko	*!	
	c. pras.ti.ko		*!

4 Diachronic LDM: perceptual bias and the role of grammar

The vast participation of liquids in long-distance metathetic processes (Ultan 1971) is by no means accidental. Liquids, and especially rhotics, are particularly susceptible to displacement (Kelly & Local 1986; Ladefoged & Maddieson 1996; Russell-Webb 2002; Blevins & Garrett 2004; Russell-Webb & Bradley 2009), due to their elongated phonetic cues, that can span even five syllables away from the original position (Coffman 2013 and discussion therein). According to non-teleological accounts, this long-distance effect paves the way for perceptual metathesis, i.e. "a 'mistake' from the point of view of the

previous linguistic system" (Blevins & Garrett 2004: 128) which stands high chances to result in the "false" reinterpretation of the segment at hand in a non-etymological position within this elongated span. This reanalysis of the liquid, that ends up surfacing in a new position, gives the impression of movement.

In this vein, the leftward directionality of LDM and the profound preference for the leftmost onset is also no accident. A misperceived liquid is more likely to be reanalyzed as originating in a position of enhanced saliency, such as the first syllable (Beckman 1998; Zoll 1996, 1998; de Lacy 2001; Smith 2005; Walker 2011). Moreover, the prominence of the first syllable may be further boosted by the accrual of segmental material, which additionally has advantages for word recognition (Alber 2001 and references therein). Thus, LDM can be viewed as a mechanism that optimizes perception.

Even though there is little doubt that factors associated with the acoustic profile and the enhancement of perception play a pivotal role in LDM, the ultimate selection among possible variants (recall that not all languages display LDM to the absolute initial onset) seems to be assigned to grammar (Coffman 2013; see also Lai 2013, 2015; Tifrit 2020; Apostolopoulou 2022; Alber & Apostolopoulou 2023). In light of this, in a constraint-based analysis such as the one pursued here, this type of sound change results from constraint reranking (see Sloos & van Oostendorp 2010; Alber & Meneguzzo 2016). The reconstruction of grammar leads to the avoidance of certain marked structures that are rendered inadmissible in a given historical stage.

For instance, the grammar of Medieval Greek permitted non-initial [Cr] onsets and prohibited all sorts of metathesis. Based on the proposed constraints, it is mandatory that LINEARITY dominates both markedness constraints COINCIDE and *RC (20) so that the Medieval Greek pattern is generated. The ranking between COINCIDE and *RC is not crucial.

(21) LINEARITY >> COINCIDE, *RC

Via this ranking, any movement, regardless of the phonological environment, is prevented due to the penalty the relevant candidates incur with respect to LINEARITY. Inescapably, non-initial [Cr] emerged intact. Compare Tableau (20) with Tableau (10) (repeated here for convenience).

(22) Tableau /kapistri/ (Medieval Greek)

	/kapistri/	LINEARITY	COINCIDE	*RC
☞	a. ka.pis.tri		*	
	b. ka.pris.ti	*!	*	
	c. kar.pis.ti	*!		*
	d. kra.pis.ti	*!		

(10) Tableau /kapistri/ (Calabrian Greek)

	/kapistri/	COINCIDE	LINEARITY	*RC
	a. ka.pis.tri	*!		
	b. ka.pris.ti	*!	*	
	c. kar.pis.ti		*	*!
☞	d. kra.pis.ti		*	

Once the ranking between LINEARITY and COINCIDE was reversed, the grammar of IG ceased to tolerate non-initial [Cr] (compare 23a–b). The innovative ranking (23b) allows for LDM, which targeted all post-consonantal rhotics in medial positions, unless the prospective landing site did not meet the requirements for hosting a rhotic.

(23) a. Medieval Greek LINEARITY >> COINCIDE, *RC
b. IG COINCIDE >> LINEARITY >> *RC

As noted above, LDM is considerably common in Romance languages, among which Southern Italo-Romance (Rohlfs 1966). By contrast, this type of metathesis, at least in a systematic fashion, is absent in all other Greek dialects. It is thus highly likely that LDM is contact-induced in IG due to the centuries-long language coexistence of IG with the local Romance dialects. In particular, IG converged with the Italo-Romance phonological grammar with respect to the ranking between COINCIDE and LINEARITY (COINCIDE

>> LINEARITY), rather than retaining the ranking found in the ancestral variety, i.e. Medieval Greek (LINEARITY >> COINCIDE), as well as in the Greek dialects outside of Italy (Apostolopoulou 2022).

5 Conclusions

The LDM pattern observed in IG is, as alluded to above, part of a broader typology which was not discussed in detail in the present paper (see Apostolopoulou 2022). Here the focus was on the application and blocking of LDM from a language-specific point of view, i.e. in IG. The paper offered a formal account of LDM in IG framed within OT. The proposed constraint hierarchy that gives rise to LDM is COINCIDE >> LINEARITY >> *RC. However, phonotactic restrictions formalized as constraints dominating COINCIDE may inhibit the application of LDM. In particular, the creation of ill-formed clusters (e.g. sonorant-liquid) is blocked by the top-ranked constraint *ILLICITCLUSTERS (17). The rankings LINEARITY&*RC >> PW[r >> COINCIDE (13–14) and LINEARITY&*RC >> OCP[cor] >> COINCIDE (15–16) ensure that metathesis does not form novel [#rV] (13) and [$C_{[cor]}C_{[cor]}$] (15) clusters, respectively, and, furthermore, they disallow alternative movement to a coda position. At the same time, these rankings permit the faithful surfacing of /#rV/ (14) and /$C_{[cor]}C_{[cor]}$/ (16).

In neither of the two IG dialects did LDM target intervocalic /tr/ sequences (18). It was shown that in fact the old cluster /tr/ gave its place to a complex segment /t͡r/ through retroflexion and affrication processes. Cross-dialectal variation regarding LDM affecting /str/ was attributed to different representations. Apulian Greek has acquired a retroflexed configuration /st͡r/ that seems to behave as an inseparable unit (24a), whereas Calabrian Greek retains a tri-segmental cluster /str/ (24b).

(24)		*Dialect*	*Optima*	*Representation*
	a.	Apulian	/#C…str/ → [#C…str]	/st͡r/
	b.	Calabrian	/#C…str/ → [#Cr…st]	/str/

Language change was accounted for in terms of constraint re-rankings. The ranking LINEARITY >> COINCIDE, *RC (23a) held in Medieval Greek. When COINCIDE outranked LINEARITY (COINCIDE >> LINEARITY >> *RC, 23b), this grammar change led to the rise of LDM. The latter ranking is also found in local Romance varieties. Given the intense contact between Romance and IG, it is rather reasonable to assume that the above re-ranking is an instance of contact-induced change.

References

Alber, Birgit. 2001. Maximizing first positions. In Caroline Féry, Antony Dubach Green & Ruben van de Vijver (eds.), *Proceedings of HILP5*, 1–19. University of Potsdam.

Alber, Birgit & Eirini Apostolopoulou. 2023. Language change as minimal change of property values. Paper presented at the *20th Old World Conference on Phonology 2023* (*OCP20*), Université de Tours, the Laboratoire Ligérien de Linguistique (UMR 7270 CNRS) and Structures Formelles du Langage (UMR 7023). 25–27 January 2023, Tours.

Alber, Birgit & Marta Meneguzzo. 2016. Germanic and Romance onset clusters - how to account for microvariation. In Ermenegildo Bidese, Federica Cognola & Manuela Caterina Moroni (eds.), *Theoretical Approaches to Linguistic Variation*. 25–51. Amsterdam: John Benjamins.

Anttila, Arto. 2008. Gradient phonotactics and the Complexity Hypothesis. *Natural Language and Linguistic Theory* 26(4): 695–729.

Apostolopoulou, Eirini. 2022. Typological variation in language contact. A phonological analysis of Italiot Greek. Doctoral dissertation, University of Verona / University of Tromsø.

Beckman, Jill. 1998. Positional faithfulness. Doctoral dissertation, University of Massachusetts Amherst.

Blevins, Juliette & Andrew Garret. 1998. The origins of consonant-vowel metathesis. *Language* 74: 508–556.

Blevins, Juliette & Andrew Garret. 2004. The evolution of metathesis. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.), *Phonetically Based Phonology*. 117–156. Cambridge: Cambridge University Press.

Bolognesi, Roberto. 1998. The phonology of Campidanian Sardinian. Doctoral dissertation, University of Amsterdam.

Bradley, Travis G. 2007. Constraints on the metathesis of sonorant consonants in Judeo-Spanish. *Probus* 19: 171–207.

Canfield, Tracy. 2015. Metathesis is Real, and it is a Regular Relation. Doctoral dissertation, Georgetown University.

Celata, Chiara. 2006. Analisi dei processi di retroflessione delle liquide in area romanza, con dati sperimentali dal còrso e dal siciliano. Doctoral dissertation, Scuola Normale Superiore di Pisa.

Celata, Chiara. 2010. Rhotic retroflexion in Romance. Acoustic data for an articulation-driven sound change. In Daniel Recasens, Fernando Sánchez Miret & Kenneth J. Wireback (eds.), *Experimental Phonetics and Sound Change* [LINCOM studies in phonetics 5]. 36–60. Munich: LINCOM Europa.

Chandlee, Jean. 2014. Strictly Local Phonological Processes. Doctoral dissertation, University of Delaware.
Coffman, Ian. 2013. Explaining Long-Distance Liquid metathesis: misperception vs. Optimization. *North East Linguistics Society* 41: 113–126.
de Lacy, Paul. 2001. Markedness in prominent positions. In Ora Matushansky, Albert Costa, Javier Martin-Gonzalez, Lance Nathan & Adam Szczegielniak (eds.), *HUMIT 2000, MIT Working Papers in Linguistics 40*. 53–66. Cambridge, MA: MITWPL.
Grammont, Maurice. 1905–1906. La métathèse dans le parler de Bagnères-de-Luchon. *Mémoires de la Societé de Linguistique de Paris* XIII: 73–90.
Greco, Carmine & Georgia Lambrogiorgou. 2001. *Lessico di Sternatia*. Lecce: Edizioni di Grifo.
Karanastassis, Anastassios. 1984–1992. *Ιστορικόν Λεξικόν των Ελληνικών Ιδιωμάτων της Κάτω Ιταλίας [Historical Dictionary of the Greek Dialects of Southern Italy]*. Athens: Academy of Athens.
Kelly, John & John Local. 1986. Long-domain resonance patterns in English. In *International Conference on Speech Input/Output: Techniques and Applications. Conference*, publication no. 258. London: Institute of Electrical Engineers.
Kilpatrick, Jeff. 2010. The development of Latin post-tonic /cr/ clusters in select Northern Italian dialects. Doctoral dissertation, University of Georgia.
Ladefoged, Peter & Ian Maddieson. 1996. *The Sounds of the World's Languages.* Oxford: Blackwell.
Lai, Rosangela. 2013. *Positional Effects in Sardinian Muta cum Liquida*. *Lenition, Metathesis, and Liquid Deletion*. Alessandria: Edizioni dell'Orso.
Lai, Rosangela. 2015. Lateral relations in Sardinian metathesis: A unified account. In Domenico Russo (ed.), *The Notion of Syllable across History, Theories and Analysis.* 264–293. Newcastle upon Tyne: Cambridge Scholars Publishing.
Lipski, John. 1990. Metathesis as template matching: A case study from Spanish. *Folia Linguistica Historica* 11: 89–104.
Loporcaro, Michele. 2001. Le consonanti retroflesse nei dialetti italiani meridionali: Articolazione e trascrizione. *Bollettino del Centro di Studi Linguitici e Filogici Siciliani* 19: 207–233.
Lubowicz, Ania. 2003. Local conjunction and comparative markedness. *Theoretical Linguistics* 29: 101–112.
Manolessou, Io. 2005. The Greek dialects of Southern Italy: An overview. *KAMPOS: Cambridge Papers in Modern Greek* 13: 103–125.
McCarthy, John & Alan Prince. 1995. Faithfulness and reduplicative identity. In Jill Beckman, Suzanne Urbanczyk & Laura Walsh Dickey (eds.), *Occasional Papers in Linguistics 18: Papers in Optimality Theory*. 249–384. University of Massachusetts.
Molinu, Lucia. 1999. Métathèse et variation en sarde. *Cahiers de Grammaire* 24: 153–181.
Pater, Joe & Andries Coetzee. 2005. Lexically specific constraints: Gradience, learnability, and perception. In *Proceedings of the 3rd Seoul International Conference on Phonology*. 85–119. Seoul: The Phonology-Morphology Circle of Korea.
Prince, Alan & Paul Smolensky. 1993/2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden, MA & Oxford: Blackwell.
Rohlfs, Gerhard. 1930. *Etymologisches Wörterbuch der Unteritalienischen Gräzität*. Halle (Saale): Niemeyer.
Rohlfs, Gerhard. 1950. *Historisches Grammatik der Unteritalienischen Gräzität*. Munich: H. Beck.
Rohlfs, Gerhard. 1966. *Grammatica Storica della Lingua Italiana e dei suoi Dialetti. Vol. 1: Fonetica*. Turin: Einaudi.
Romano, Antonio. 1999. A phonetic study of a Sallentinian variety (southern Italy). *Atti del XIV Congresso Internazionale di Scienze Fonetiche*: 1051–1054.
Romano Antonio & Francesco Gambino. 2010. Cacuminali calabresi: modi e luoghi d'articolazione alla luce di misurazioni acustiche e indagini per risonanza magnetica (IRM). In Franscesco Cutugno, Pietro Maturi, Roberta Savy, Giovanni Abete & Iolanda Alfano (eds.), *Parlare con le Macchine, Parlare con le Persone*. 505–513. Torriana (RN): EDK.
Russell Webb, Eric. 2002. The relational /R/: Three case studies of rhotic integrity and variation. Doctoral dissertation, University of Texas.
Russell-Webb, Eric & Travis G. Bradley. 2009. Rhotic metathesis asymmetries in Romance: Formalizing the effects of articulation and perception on sound change. In Pascual J. Masullo, Erin O'Rourke & Chia-Hui Huang (eds.), *Romance Linguistics 2007: Selected Papers from the 37th Linguistic Symposium on Romance Languages (LSRL)*. 321–337. Amsterdam: John Benjamins.
Sloos, Marjoleine & Marc van Oostendorp. 2010. The relationship between phonological and geographical distance. Umlaut on the diminutive in Dutch dialects. *Taal en Tongval* 62(2): 204–234.
Smith, Jennifer. 2005. *Phonological Augmentation in Prominent Positions*. New York: Garland.
Smolensky, Paul. 1995. On the structure of the constraint component Con of UG. Handout, University of California, Los Angeles.
Squillaci, Maria Olimpia & Tito Squillaci. 2016. Glossa Greka. [Mobile application software]
Tifrit, Ali. 2020. Obstruent liquid clusters: Locality, projections and percolation. In Kuniya Nasukawa (ed.), *Morpheme-Internal Recursion in Phonology, Studies in Generative Grammar*. 317–368. Boston & Berlin: Mouton de Gruyter.

Torres-Tamarit, Francesc, Clàudia Pons-Moll & Maria Cabrera-Callís. 2012. Rhotic metathesis in Algherese Catalan: A Harmonic Serialism account. In Kimberly Geeslin & Manuel Díaz-Campos (eds.), *Selected Proceedings of the 14th Hispanic Linguistics Symposium*. 354–364. Somerville, MA: Cascadilla Proceedings Project.

Ultan, Russell. 1971. A typological view of metathesis. In Joseph H. Greenberg, Charles A. Ferguson & Edith A. Moravcsik (eds.), *Universals of Human Language. Vol. 2: Phonology*. 367–402. Stanford: Stanford University Press.

van de Weijer, Jeroen. 1996. *Segmental Structure and Complex Segments*. Tübingen: Niemeyer.

Wagner, Max Leopold. 1941. *Historische Lautlehre des Sardischen*. Halle (Saale): Max Niemeyer Verlag.

Walker, Rachel. 2011. *Vowel Patterns in Language*. Cambridge: Cambridge University Press.

Zoll, Cheryl. 1996. Parsing below the segment in a constraint-based framework. Doctoral dissertation, University of California, Los Angeles.

Zoll, Cheryl. 1998. Positional asymmetries and licensing. Ms., MIT.

Negation in Wagdi: Negators or Polarity Makers?

Vyom Sharma
Indian Institute of Technology Delhi

This paper gives an analytical description of negation and polarity in Wagdi, a variety of Bhil Languages (Western Indo-Aryan) spoken in Southern Rajasthan and adjoining areas. Wagdi has 5 types of negation markers: *ne, nathi, na, nakke,* and *nak*. Each negation marker is used to express negation in a different context and is ungrammatical in other contexts. *ne* expresses negation of a proposition with an event. *nathi* is used in negated copular sentences or Copular Clauses (CC).*na* is used specifically to mark a negative answer to a yes-no type of question (YNQ) in Wagdi, optionally followed by the negated proposition. *nakke* is used for imperative sentences, which express a command or order that must be followed immediately. Whereas nakk is used with the polite imperative, to express a warning or a request for future action.
Given this data, it can be observed that the contexts of these lexical items are all different, and all of them but one (propositional negator *ne*) have a discourse flavor to them. They do not have a truth value to be 'negated', so can these lexical items even be termed 'negators'? This paper will look at the contexts of each of the given lexical items and come to a conclusion that Wagdi has one negator, the propositional negator *ne*, and the other items are polarity markers.

Keywords: Negation, Polarity, Indo-Aryan

1 Introduction to Negation and Polarity Markers in Wagdi

Wagdi is a dialect of Bhil Languages (BL), a group of Western Indo-Aryan (WIA) languages spoken by the people of Bhil tribes. It is spoken in Dungarpur and Banswara districts in Southern Rajasthan. Wagdi has many sub-dialects and micro-variations based on geography, social, cultural, and economic classes, and communities. The data used in this study is primarily from the urban speakers of the district headquarters of Dungarpur and Banswara.

There are five different types of negators in Wagdi: *ne, nathi, na, nakke,* and *nak*. They are each used in separate and independent contexts as it will be clear with the description given below:

1.1 *Propositional Negator ne* *ne* is used to express what has been referred to as sentential negation in literature. The negator *ne* is used in Wagdi to negate declarative sentences with a verb. I call *ne* a propositional negator, because its basic function is to perform the semantic operation of negation (¬) on the semantic *proposition* in the given sentence. A *proposition* is the meaning of an utterance under consideration. Every proposition is either true or false in the world in which it exists. The value of the relation of a proposition to the truth is called its *truth value.*

ne occurs only with propositions which express a semantic *event*. An event is an action situated in space and time and it is represented by the precence of a lexical verb in the syntax. Thus, all sentences with verbal predicates are negated by the propositional negator *ne*. Let us see a set of examples:

(1) me keri khaa-di (Wagdi)
1SG.ERG mango eat-PERF
'I ate a mango.'

(2) Me keri **ne** khaa-di
1SG.ERG mango NEG eat-PERF
'I did not eat a mango.'

In the above example, the proposition p in (1) is 'I ate a mango'. (2) represents ¬p, that is 'I did *not* eat a mango'. The proposition in example (1) has a semantic event. *ne* is the default negator in Wagdi, unless there is a special negator for a given context, like the negators discussed ahead.

1.2 *Copula Negator nathi* *nathi* is used in a negated copular sentences or Copular Clauses (CC). Following is an example of a CC in Wagdi, and its negation as shown below:

(3) mu doctor hũ (Wagdi)
1SG doctor be.PRS
'I am a doctor.'

(4) mu doctor **nathi/*ne**
1SG doctor CN
'I am not a doctor.'

Since *nathi* occurs in the context of CC in Wagdi, it is a Copula Negator (CN). *nathi* shows some interesting properties. In the following example, let us look at a negated CC utterance set in a contrastive context:

Context: Lala had gone to the hospital as he was not feeling well. He met Ram, who treated him well. Lala, thus, started feeling better in a day and was praising Ram to Sita the following day:

(5) * ram DAAKTAR **nathi**, bhagwaan he (Wagdi)
Ram doctor-F CN god be.PRS
'Ram is not only a doctor. He is God!'

(6) ram DAAKTAR **ne** he, bhagwaan he
Ram doctor NEG be.PRS god be.PRS
'Ram is not only a doctor. He is God!'

As we can see in example (5) and (6), the CN *nathi* is absent in contrastive focus-marked sentences. In section 3, I discuss why the CPM *nathi* shows these properties.

1.3 *Anaphoric Negator na* *na* is used specifically to mark a negative answer to a yes-no type of question (YNQ) in Wagdi, optionally followed by the negated proposition. Let us look at the following example:

(7) A: Te khavaanu khaadu ke?
2SG.ERG food eat.PERF Q
'Have you had lunch/dinner?'

B: **na**, (me khavaanu ne khaadu he)
AN 1SG.ERG food NEG eat.PERF be.PRS
'No, I haven't had lunch/dinner'

YNQs are also called polar questions as they must be answered either with a yes or a no, optionally followed by an antecedent sentence which is the actual answer to the question. I will use the term polar question in this study. In the above example, Speaker A asks a YNQ *have you had your lunch*? to B. B replies negatively, that is, they have not had their lunch. However, in the reply by B, the negated proposition *I haven't had lunch* is in parentheses, preceded by a negator *na.*

It is clear from the pattern of the example (7), that *na* occurs in a context where there is a conversation or a *discourse*. *na* occurs only when a person asks the speaker a YNQ, and the speaker replies negatively. The reader will also notice in example 4B, that the propositional negator *ne* is present separately from the anaphoric negator in the following proposition. These properties are important as they will form the base of the main proposal of this dissertation, that negation in natural language as a concept exists at two levels: the propositional level and discourse level. I elaborate more on this distinction in section 3.

1.4 *Imperative and Polite Imperative Negators: nak and nakke* *nakke* is used for imperative sentences, which express a command or order that must be followed immediately. Following is an example of an imperative sentence and its negation:

(8) aaNaaa kamraa-me jaa/jo (Wagdi)
This room-LOC go.IMP
'Go inside this room!'

(9) aaNaaa kamraa-me **nakke** jaa/jo
This room-LOC IN go.IMP
'Don't go inside this room!'

As we can see above, (8) is an *imperative proper* sentence. It expresses the speaker's direct command or order to the listener. (9) is the negation of (8), and it uses the imperative negator *nakke* instead of the propositional negator *ne*.

Wagdi has two types of imperatives. The second type, the polite imperative uses a different negator *nakk.* A polite imperative sentence expresses a warning or a request for future action, instead of a command. Following is an example of a polite imperative and its negation:

(10) aaNaaa kamraa-me jaa-je (Wagdi)
This room-LOC go-SUB
'Please! go inside this room.'

(11) aaNaaa kamraa-me **nakk** jaa-je
This room-LOC PIN go-SUBJ
'Please do not go inside this room.'

(10) is a polite imperative sentence. It expresses the speaker's request or warning to the listener instead of a command, as in (8). (11) is the negation of (10) and it uses the polite imperative negator *nakk* instead of standard negator *ne*. It can be called a subjunctive imperative as well because it occurs with a verb which has a subjunctive suffix.

The imperative negator is infelicitous in a polite imperative sentence and vice versa. It is shown in the following example:

(12) # aaNaaa kamraa-me **nakke** jaa-je (Wagdi)
This room-LOC IN go-SUBJ
'Please do not go inside this room.'

(13) # aaNaaa kamraa-me **nakk** jaa/jo
This room-LOC PIN go.IMP
'Don't go inside this room!'

The imperative and polite imperative are semantically related. It is interesting that Wagdi has separate negators for each of them. It can also be noticed that there are lexical similarities between the two imperative negators *nakk-e* and *nakk*.

2 Negation and Polarity

In order to understand the concept of polarity, it is necessary to understand the difference between propositional negation and discourse-level negation. A declarative sentence carries a semantic proposition p. The application of the logical negation operator (¬) on p yields ¬p. Propositional negation is ¬p in natural language. As mentioned in an earlier section it is represented by the feature [NEG] in syntax. Propositional negation operates on the propositional domain (over FinP).

However, in the discourse domain, the speaker and listener of p/¬p will either believe it to be true or false. In other words, the discourse participant either *agrees* or *denies* with p/¬p. Another possibility is that the discourse participant may be neutral towards the proposition. This relation of the discourse with the proposition is called its *polarity*. When the speaker does not agree with the proposition, it is said to have negative polarity. Negative polarity stands for discourse-level negation in natural language. Apart from the propositional negator, the four lexical items introduced in the first section are polarity markers in Wagdi . In my proposed system, the value of the polarity of a sentence can range on a spectrum as shown below:

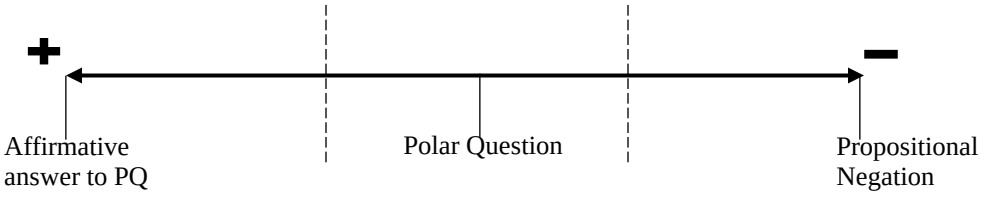


Fig 1

The lower end of the polarity spectrum is the positive polarity, which is the polarity of affirmative answers to a PQ, when the speaker agrees with the antecedent proposition. Thus, in my proposed system and the data presented only the sentence with the *yes* or *howe/ho* affirmation markers have positive polarity.
On the upper end of the polarity spectrum is negative polarity. In the current proposal, an example of a sentence that has negative polarity is a negative response to a PQ, with the *no* or *na* polarity markers. Apart from *na*, non-propositional sentences in Wagdi like a negated copular clause (with *nathi*) or negative imperatives (with *nakke* and *nakk*) also have negative polarity.

In the middle of the polarity spectrum is the neutral polarity. It stands for all the cases in which the sentence has no polarity or unvalued polarity. A sentence with non-negative polarity does not mean it has positive polarity. In the sense of discourse, neutral polarity means that the speaker is ambivalent about his belief of the utterance under the given contextual evidence. Declarative sentences, interrogatives including PQs, imperatives, and copular clauses have neutral polarity.

3 Polarity Markers: Evidence from The Context

In this section I show that *nathi*, *na*, *nakke* and *nakk* are polarity markers in Wagdi, as the contexts associated with them cannot be semantically 'negated', but they can have negative polarity.

3.1 *Nathi as a Polarity Marker* As mentioned before, *nathi* is present in negated Copular Clauses (CCs). CCs do not have a proposition with a semantic event. In this study, CCs are assumed to be *stative constructions*, which are different from events in that they are not bound by the time variable. A stative construction expresses a state or a condition instead of an action or an event. Example (14) is a CC, which expresses the state of Lala, i.e., 'beautiful'. It does not talk about the *when* or the time-period of him being beautiful. The CC in (14) is negated with *nathi* as shown in (15) and the past tense of (14) is (16). But as it can be seen in (17), *nathi* is incompatible with past tense, and the negation of 8c is done with propositional negator *ne*.

(14) Lala rupaaro he (Wagdi)
Lala beautiful be.PRS
'Lala is beautiful'

(15) Lala rupaaro **nathi**
Lala beautiful CN
Intended: 'Lala was not beautiful'

(16) Lala rupaaro hato
Lala beautiful be.PAST
'Lala was beautiful'

(17) Lala rupaaro **ne (*nathi)** hato
Lala beautiful NEG be.PAST
'Lala was not beautiful'

Let us say that the speaker utters the CC in (14) at a time instance t, then:

P: Lala is beautiful
P is true if,
P is true for T>t,
P is true for T<t,
P is true for T=t.

It is assumed that P is true irrespective of the value of T. Hence, the meaning of the stative sentence P is unbound in time. However, in natural language, it is possible to isolate the meaning of the following:

P is true for T>t

Although *nathi* can convey the meaning of 'Lala is not beautiful', it cannot convey the meaning of 'Lala was not beautiful', as shown in in example (17). This shows that CCs are not events with a propositional truth value as a standalone stative construction. The CC in past tense is negated (example (17)) with a propositional negator. The meaning conveyed by (17) is 'it is not the case that 'Lala is beautiful' is true for the time t<T'.

The explanation above shows that CCs are not propositions with an event to negate. However, it does not automatically imply that *nathi* has a discourse-related item. Let us look at another property of negated CCs in Wagdi. *nathi* does not occur in contrastive contexts. Let us set up a context, which creates alternatives for the predicate in a CC to induce contrast:

Context: Lala had gone to the hospital as he was not feeling well. He met Ram, who inspected Lala and gave him proper treatment. Lala talks to Sita the other day, praising what a great doctor Ram is. However, Sita knows Ram from before and knows that ram is not a doctor but a nurse at the hospital. The following conversation ensues (words in caps lock are focus-marked):

(18) Lala: kevu padega, ram khub asal daaktar he!
Say.INF MOD.FUT Ram very good doctor be.PRS
'Have to say, Ram is a very good doctor!'

(19) Sita: #ram daaktar **NATHI**, nars he

(20) Sita: ram DAAKTAR **ne** he, nars he
Ram doctor NEG be.PRS nurse be.PRS
'Ram is not only a doctor, he is a nurse!'

In the above example, we see that the context of usage makes one negator (*nathi*) infelicitous while another negator (*ne*) is felicitous. The context given above introduced *alternatives* to the predicate NP *doctor*. Thus, *doctor* is focus-marked in the (19) and (20), and *nathi* is incompatible with these sentences. This is because when *nathi* is present with a CC, it is focus-marked. In the presence of another focus-marked entity, *nathi* cannot occupy the same position in the structure, hence it does not get articulated.

The context discussed above demonstrates the discourse-related properties of *nathi*. As shown through examples (14)-(17), the CC is a stative with no event, and hence no proposition. When there is no truth value to change, what is being 'negated'? As shown in examples (18)-(20), *nathi* is related to the belief system of the speaker engaged in a discourse. Thus, *nathi* represents the negative polarity of the CC. This makes *nathi* a polarity marker instead of a negator.

3.2 *Nathi as a Polarity Marker* There have been two approaches towards the analysis of yes-no particles, depending on whether *yes-no* are part of the antecedent structure or not. In the first approach, Krifka (2013) treats yes-no particles as discourse-level anaphors that pick up the salient proposition referent from the PQ and assert it in the discourse. Under Krifka's analysis example (21) would be analyzed as:

(21) A: Is Pari coming?
simplified LF: [$_{TP}$ Pari is coming] → d

a. *Yes* = ASSERT(d) [meaning: Pari is coming]
b. *No* = ASSERT(¬d) [meaning: Pari is not coming]

According to Krifka (2013), the particle *yes* can pick up discourse referent d to assert in the discourse, but it cannot assert ¬d. The particle *no* negates d, asserting ¬d in the discourse. *no* cannot assert d. On the other hand, a negative PQ like example (22) under Krifka's analysis would be analyzed as below:

(22) a. Is Pari not coming?
simplified LF: [$_{NegP}$ not [$_{TP}$ Pari is coming] → d] → d'

b. Yes = ASSERT(d) [meaning: Pari is coming]
c. No = ASSERT(¬d) [meaning: Pari is not coming]
d. Yes = ASSERT(d') [meaning: Pari is not coming]
e. No = ASSERT(¬d') [meaning: Pari is coming]

As it can be seen above in example (22)a, according to Krifka (2013), there are two antecedents available in a negative PQ for the *yes-no* particles to pick up, d and d', which is ¬d. When *yes* particle picks d, we get (22)b, and when it picks d' we get (22)d. When the no particle picks d, we get ¬d, which is (22)c. When it picks d', we get ¬(¬d), which is d in (22)e. Thus, Krifka's system explains the interchangeability of *yes-no* particles in negative PQs. In conclusion, Krifka (2013) treats *yes-no* as discourse-level anaphors which are structurally unrelated to the antecedent.

In the other approach to the analysis of yes-no particles, they are treated as adverbs hosted in the specifier of a projection in the left periphery of the clausal spine, scoping over TP. Kramer and Rawlins (2009), Holmberg (2013,2016), and Roelofson and Farkas (2015) all agree and have called it the **Polarity Phrase (PolP)**. The yes-no particles represent the polarity feature.

What is common among the separate accounts in both approaches is that *yes-no* particles encode polarity. The polarity of a sentence is related to both the propositional information of the utterance as well as the speaker's belief about the contextual evidence. As it has also been noted in the above examples, *no* is present in the utterances beside the propositional negator. This shows that *no* does not express propositional negation. Thus, given all the arguments, *yes-no* are polarity markers. Extending the argument to Wagdi, the equivalent lexical items, i.e., *na* (negative) and *howe-ho* (affirmative) are polarity markers as well.

3.3 *Nakke and nakk as Polarity Markers* Imperatives can span over a range of different meanings from a direct command to a deferred instruction, request, or even non-performation[1]. There have been different approaches to the semantics of imperatives, which can be divided into two types based on how imperatives are related to the notion of a proposition. One approach, proposed by Searle (1969), Han (2000), Barker (2005) and Takahashi (2012) says that imperatives an independent prototype from a proposition in semantics. It encodes a 'To-Do' list (Portner, 2004, 2007) or a PLAN SET (Han, 2000),

[1] For more on non-performative imperatives see Kaufmann & Poschmann 2013; Medeiros 2015

which is a set of partially ordered propositions which express the speaker's wish or desire that they be fulfilled by the other participant(s) in the discourse. For example, the command *Shut the door!* is a set of propositions as follows:

Step 1: Hearer H moves from their location to the location of the door.

Step 2: H holds the door.

Step 3: H pushes the door until it is shut.

The semantic interpretation of an imperative like *Shut the door*! in terms of a T0-Do List looks like below (Portner, 2004):

(23)

$$\|Shut\ the\ door\|^{w*,c} = \begin{bmatrix} \lambda w.\lambda x: x = \text{addressee(c)}.x \text{ moves to location of door in w} \\ \lambda w.\lambda x: x = \text{addressee(c)}.x \text{ holds the door in w} \\ \lambda w.\lambda x: x = \text{addressee(c)}.x \text{ pushes the door until it is shut in w} \end{bmatrix}$$

In the above notation, w refers to a world among the set of possible worlds w*, where the entire set of propositions is true. C refers to the context of utterance. Given that w is just a subset of the possible worlds w*, there may also be worlds where the set of propositions might not be all true or be partially true. The fulfilment of the set of propositions will depend on the commitment of the hearer H. So, although there is propositional content in the imperatives, they do not have a certain truth value.

The To-Do List kind of semantic interpretation of an imperative is similar to the semantic interpretation of an interrogative, which has also been analyzed as a Question Set by Roberts (1996), a set of propositions. Based on earlier approaches to the semantic analysis of imperatives, I have shown above that they are not propositions and thus do not have a truth value. However, they can have negative polarity. The lexical items specific for negating imperatives in languages like Wagdi, must be hosted at the polarity projection, and are therefore polarity markers.

4 Conclusion

Wagdi has five kinds of negation markers: *ne*, *nathi*, *na*, *nakke* and *nakk*. Out of these, *ne* is a propositional negator. The other four markers stand for discourse-level negation which is called polarity. I showed in this paper that the contexts associated with nathi, na, nakke and nakk do not have a proposition with a semantic event. They represent polarity and thus, they are polarity markers.

References

Barker, Chris (2012). Imperatives denote actions. In *Proceedings of sinn und bedeutung* (Vol. 16, No. 1, pp. 57-70).

Han, Chung-Hye (2000). *The structure and interpretation of imperatives: mood and force in Universal Grammar*. Psychology Press.

Holmberg, Anders (2016). *The Syntax of Yes and No*. Oxford University Press.

Holmberg, Anders (2013). The syntax of answers to polar questions in English and Swedish. *Lingua*, *128*, 31-50.

Kaufmann, Magdalena, & Poschmann, Claudia (2013). Embedded imperatives: Empirical evidence from colloquial German. *Language*, 619-637.

Kramer, Ruth, & Rawlins, Kyle (2009). Polarity particles: an ellipsis account. In *Proceedings of NELS* (Vol. 39, pp. 479-92).

Krifka, Manfred (2013, August). Response particles as propositional anaphors. In *Semantics and linguistic theory* (Vol. 23, pp. 1-18).

Medeiros, David J. (2015). Embedded Ancient Greek imperatives: A feature transfer analysis. Syntax, 18(2), 124-156.

Portner, Paul (2004, September). The semantics of imperatives within a theory of clause types. In *Semantics and linguistic theory* (Vol. 14, pp. 235-252).

Portner, Paul (2007). Imperatives and modals. *Natural language semantics*, *15*(4), 351-383.

Roberts, Craige (1996). Information structure in discourse: Toward a unified theory of formal pragmatics. *Ohio State University Working Papers in Linguistics*, *49*, 91-136.

Roelofsen, Floris, & Farkas, Donka F. (2015). Polarity particle responses as a window onto the interpretation of questions and assertions. *Language*, 359-414.

Searle, John R., & Searle, John R. (1969). *Speech acts: An essay in the philosophy of language* (Vol. 626). Cambridge university press.

Takahashi, Hidemitsu (2012). *A cognitive linguistic analysis of the English imperative: With special reference to Japanese imperatives* (Vol. 35). John Benjamins Publishing.

On the Categorial Status of (è)wo in Yorùbá

Ọlaide Akinwande
University of Lagos

This study investigates the syntax and semantics of *(è)wo,* which is believed to be a nominal item in the linguistic literature in Yorùbá and it is used in content interrogative. The article contends that *èwo* is not a single word morpheme but two, which its morphological make-up is {è-*}* and [wo], where {è-*}* is a prefix and [wo] is the root/stem/host word. On the syntax of *(è)wo,* the study establishes that *èwo* can occur before the subject or object DP, where {è-*}* in *èwo* is obligatorily pronounced and if the nominal word in question is to occur after the subject or object DP, {è-*}* in *èwo* is optionally pronounced. In this case, the study contends that the prefix {è-*}* in *èwo* is just an emphatic marker as the main question element there is *wo.* The study also claims that basic clause is the appropriate answer when *èwo* features in rightward but focus construction is the acceptable answer when (*è)wo* occurs in discourse position. The study conclusively construes (*è)wo* as specificity, emphatic, numerical and negative.

Keywords: Content-question, Nominal-word, Semantics, Subject-DP or Object-DP, Syntax

1 Introduction

Yorùbá is one the Kwa languages. The language belongs to the West Benue-Congo of the Niger-Congo phylum of African languages (Williamson/Blench 2000: 31). Majority of the speakers of the language reside in the Southwestern part of Nigeria with a population of about sixty million. Yorùbá is regarded as one of the three major languages in Nigeria. The language has more than twenty sub-languages which differ from one another. Aside from Nigeria, the language is also spoken in countries like Republic of Bénin, Togo, Ghana, Cote D'ivoire, Sudan, Sierra-Leone, etc. However, just like other languages of the world, interrogative construction is also expressed in Yorùbá, most especially, content interrogative. Thus, content question has been expressed as questions that attract some more specific answers (Ajiboye, 2013:1). *(È)wo,* which has the feature of [± human] is one of the question morphemes used in content questions in Yorùbá (Bamgbose 1966, 1990; Awobuluyi 1978). Among other question elements used in content questions in Yorùbá are *kí 'what', ta 'who'. ibo 'where'*, *èló 'how-much', mélòó 'how-many', báwo 'how'*, etc. The content question items, alias wh-operators are shown in ex-situ in (1).

(1ai) *Èwo* ni Adé rà?
Which FOC Ade buy
'Which one did Ade buy?'

(ii) Ìwé *(è)wo* ni o rà?
Book which FOC 2sg buy
'Which book did you buy?'

(b) *Kí* ni Adé jẹ?
What FOC Ade eat
'What did Ade eat?'

(c) *Ta* ni Adé pè?
Who FOC Ade call
'Who did Ade call?'

* I really appreciate the insightful comments of Daniel Arẹmu of Goethe University on this paper. All the remaining errors therein are mine

(d) *Ibo* ni Adé lọ?
Where FOC Ade go
'Where did Ade go?'

(e) *Èló* ni o gbà?
How-much FOC 2sg take
'How-much did you collect?'

As revealed in (1a-e), all the content question morphemes show up in clause-initial and immediately followed by the focus marker *ni* except in (1aii), where the wh-word features after the object-DP to express its content questions. However, the main contributions of this paper is to show that (i) {è-} in *èwo* is an emphatic marker and to show such projection, (ii) propose structures for the syntactic distributions of *(è)wo,* (iii) give the likely semantic readings of *(è)wo*. This paper is organized into seven (7) sections. Section one is introduction; section two explains the theoretical framework to the study; section three discusses morphological formation of *èwo*; section four deliberates on the syntactic distributions of *(è)wo*; section five explains the earliest position of (*è)wo;* section six talks about the semantic interpretations of *(è)wo* and section seven, which is the last section gives the concluding remarks.

2 Theoretical Framework

The theoretical framework adopted for this work is the Minimalist Program (MP). Minimalist Program (MP henceforth) was propounded by Chomsky (1993, 1995, 1998, and 2002). MP is assumed that every clause formation starts with the selection of lexical items contained within the lexicon (***Operation Select***). Every word within the lexicon is believed to possess three universal sets of features, which are <Sem, Syn, Phon> i.e. Semantic, Syntactic and Phonetic features. Syntactic derivation is formed by a simple mathematical operation called ***Merge***. ***Merge*** is an operation which combines words selected from the lexicon with an already formed elements or syntactic objects. Operation merge can be internal or external. It is known as ***External Merge*** when the operation combines words selected directly from the lexicon with an already formed elements or syntactic objects; while it is known as ***Internal Merge*** when it recombines elements within an already formed syntactic object. These and other operations such as ***Attract*** and ***merge*** within the MP are based on binary principle. Based on this, every operation of the type ***x, y*** is the set of elements that includes ***x*** and ***y***. It means then that the binary nature of merge rules out ternary branching. MP ensures that the output of a syntactic derivation is not an arbitrary set but is derived from the input labels. For instance, the product of merge (α, β) can only be either an α-phrase or a β-phrase but not γ-phrase since γ was not part of the input in the first place (cf.Ilọri and Oyebade 2012).

In MP, it is also assumed that every product of operation merge moves towards the interface level for ***Spell-Out***. Spell-Out therefore, is the output of derivation at the level of pronunciation; i.e. ***Phonetic Form*** (PF). MP also ensures that every syntactic derivation converges at ***Spell-Out*** based on the compatibility of features of words employed in the derivation which must be checked against one another in the course of the process of derivation; otherwise, the derivation crashes.

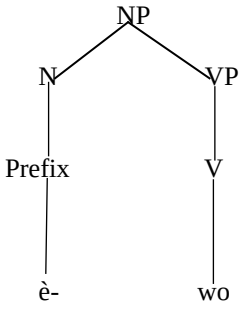
Other MP assumptions are ***Split-CP Hypothesis*** of Rizzi (1997) which assumes the splitting of CP projection into a number of different projections, specifically ForceP (Force Projection), FocP (Focus Projection), InterP (Interrogative Projection), TopP (Topic Projection), etc. in order to accommodate other elements that could move with the head in the process of ***Internal Merge*** **and** ***External Merge***. The Split-CP analysis therefore provides a principled account of both overt and covert/null spell-out of all forms of syntactic heads in the CP layer. Hence, our study relies solely on these for the analysis of content interrogative projections of *(è)wo* in Yorùbá.

3 Morphological Make-up of *èwo*

Generally, in Yorùbá linguistic literature, it is believed that vowels begin every nominal word (Awobuluyi, 2013:17) and where the case is otherwise, a phonological process, e.g. deletion has taken place or an intentional process has been carried out, e.g. optional usage, i.e. empty prefix. Ilọri (2010:408) quoted Akinlabi and Liberman (2000:17) as saying "in the case of a consonant-initial word, we are tempted

to think of it as an empty prefix". The scholars further gave such items as *ìjẹ* 'bait', *àlọ* 'act of going', *ilé* 'house', etc. where vowel begins all the items mentioned. Another fascinating report in the linguistic literature is that either low-tone or mid-tone begins every nominal item in Yorùbá and never high-tone. If the case is what has been claimed above, *èwo* is not exempted from all these criteria mentioned. A careful look at the *èwo*, one can observe that it is vowel that begins the nominal word and the vowel has low-tone. Hence, I argue that *èwo* is a nominal word and it consists of two morphemes, i.e. *{è-}* and [*wo*], where *{è-}* is a prefix attached to the root word, [*wo*] to realize the nominal word, *èwo*. Knowing fully well that *èwo* is an NP, I sketch out this in (2).

(2)



As shown in the diagram above, *wo* is a VP which is externally merged with prefix {*è*-} to derive the nominal word.

4 Syntactic Distributions of (è)wo

In this section, I propose two syntactic positions for the occurrence of *(è)wo*. The first position is *èwo* occurring before the subject or object DP, depending on the clause construction and the latter is immediately after the subject or object DP. This is discussed in the following sub-sections.

4.1 *Occurrence of (è)wo Before the Subject or Object DP* In this type of position where *èwo* shows up before the subject or object DP, I claim that *{è-}* in *èwo* is obligatorily pronounced, as only [wo] cannot realize the desired interrogative alone without the presence of *{è-}* otherwise, it results to ill-formed constructions. This is illustrated in (3).

(3a) *Èwo* ni Adé jẹ?
Which FOC Ade eat
'Which one did Ade eat?'

(b) *Èwo* ni o fẹ́?
Which FOC 2sg want
'Which one did you want?'

(c) **Wo* ni Adé jẹ?
Which FOC Ade eat

(d) **Wo* ni o fẹ́?
Which FOC 2sg want

As seemingly evident in (3a-b), we have well-formed constructions because {è-} in *èwo* occurred before the subject DP and it is obligatorily pronounced. The case is different in (3c-d) where {è-} in *èwo* is elided, thereby resulting to ill-formed constructions, though it shows up before the subject DP.

4.2 *Occurrence of (è)wo After the Subject or Object DP* In the following contexts, where *(è)wo* occurs after the subject or object DP, I argue that {è-} in *èwo* is optionally pronounced or may not be pronounced at all, as only *wo* there can realize meaningful and acceptable interrogative without the presence of {è}, as shown in (4).

(4a) Iṣẹ́ *(*è*)wo* ni Adé ń ṣe?
Work which FOC Adé Prog do
'Which work is Ade doing?'

(b) Ọmọ *(*è*)wo* ni o fẹ́?
Child which FOC 2sg want
'Which child did you want?'

(c) Àwọn akẹ́kọ̀ọ́ (è*)wo* ni ó ń sọ̀rọ̀?
3pl students which FOC 3sg Prog talk
'Who are the students talking?'

A careful look at the examples in (4a-c), *(*è*)wo* occurs after the object DP, it is apparent that they are well-formed structures even without {è-} being pronounced. If that is the case, it means that *wo* is actually the question word there. Though, some speakers may pronounce {è-} there but it is not common for a native and typical Yorùbá speaker to pronounce {è-} in *èwo* in expressions like (4). A logical question one needs to raise here is that, what is the syntactic function of {è-} in *èwo* in this type of contexts? I postulate that {è-} in *èwo* in expressions like (4) is an emphatic marker, which emphasizes on the preceding DP (subject or object in the contexts like (4)). More details would be discussed on this under the semantic interpretations of (*è*)*wo* in section (6).

5 The Earliest Position of *(è)wo*

I discuss original syntactic position of (è*)wo* in this section. Firstly, it is logical to account for how (è*)wo* is used to project interrogative phrase (InterP), i.e. content question. Thus, I argue that (è*)wo*-expression is projected by an Inter head. This Inter head obligatorily selects FocP as complement to project its content question. However, (è*)wo* actually occupies clause-final position (object of the verb) but moves to the Spec-InterP, its landing site, leaving its trace behind in order to project question. In contexts where *èwo* occurs before the subject or object DP, I have the following examples, (5).

(5a) *Èwo*$_i$ ni Adé jẹ [t_i]?
Which FOC Ade eat
'Which one did Ade eat?'

(b) *Èwo*$_i$ ni ó fẹ́ [t_i]?
Which FOC 3sg want
'Which one did s/he wants?'

(c) *Èwo*$_i$ ni ó rà [t_i]?
Which FOC 3sg buy
'Which one did s/he buy?'

Evidently in (5a-c), *èwo* moves from clause-final position to the clause-initial position, leaving its trace (t_i) behind and at the same time becoming the head of **InterP**. Though, *èwo* featuring in the clause final does not mean it is not a wh-question. The semantic interpretation of *èwo* showing up in rightward indicates that is a content question, as shown in (6).

(6) Adé ra èwo?
Ade bought what
'Which one did Ade buy?'

But something quickly fished out here. The matter is that focus construction cannot answer the content question raised in (6). It is a well known fact in linguistic literature that focus expression is the answer to wh-question, as expressed in (7) and (8).

(7a) Èwo ni Adé rà? [Content Question]
Which Foc Ade buy
'Which one did Ade buy?'

(b) Dúdú ni Adé rà. [Answer/Focus Construction]
Black FOC Ade buy
'It was back that Ade bought'

(8a) Ìwé (è)wo ni Adé rà? [Content Question]
Book which FOC Ade buy
'Which book did Ade buy?'

(b) Ìwé ńlá ni Adé rà. [Answer/Focus Construction]
Book big FOC Ade buy
'It was big book that Ade bought'

It appears in (7) & (8) that (7b) & (8b) which are focus contructions provide direct and acceptable answers to (7a) & (8a). But the case of (6), focus construction is not fit to answer the content question raised in (6).

Cf. (9) Adé ra èwo? [Content Question]
Ade bought what
'Which one did Ade buy?'

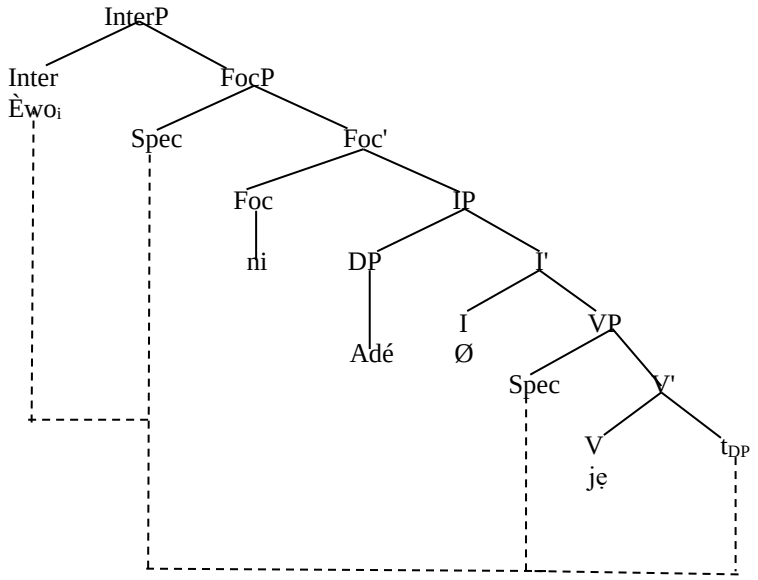
(10a) Dúdú ni Adé rà. *[Answer/Focus Construction]
Black FOC Ade buy
'It was back that Ade bought'.

(b) Adé ra dúdú [Basic clause/Answer]
Ade buy back
'Ade bought black'.

It is obvious in the above examples that (10a) is inappropriate to answer the question in (9) because it does not provide direct answer to the question. This signals that basic clause is the best and appropriate answer to such echo question, where *èwo* features at the right periphery. Thus, for focus construction to answer *èwo*-expression, *èwo* has to be imported to the Spec-InterP. Doing this C-commands the focus element **ni,** as requirement for wh-question to converge well in Yorùbá. However, it could be construed that the echo question in (6) & (9) occurs when there have been previous conversations or dialogues between the speaker and the hearer, with which the speaker desires to affirm what the hearer (respondent) actually did. The echo question may be labeled affirmative wh-question. Thus, I propose enquiry wh-question and affirmative wh-question. Enquiry wh-question is the case that the speaker (interrogator) is innocent of the current situation and desires to know by raising question. No former background or knowledge is involved in this type of question. This question type is exemplified in (6) and (9). Affirmative wh-question is said to establish a known event by asking the interrogatee. Interrogator may like to demonstrate negative side or oppose an action or event carried out by the interrogatee, as shown in (7a) and (8a). In some contexts (not in all), when *èwo* is raised from clause-final position to clause-initial, a resumptive pronoun **ó** (which some scholars like Awóbùlúyì described as High-Tone Syllable (HTS)) is inserted to play the role of subject as evident in (5b-c). We show the structural representation of (5) in (6).

[2]Àkànbí (2016:419-420) claims that *wo* can not independently convert declarative to interrogative but together with NP or PP

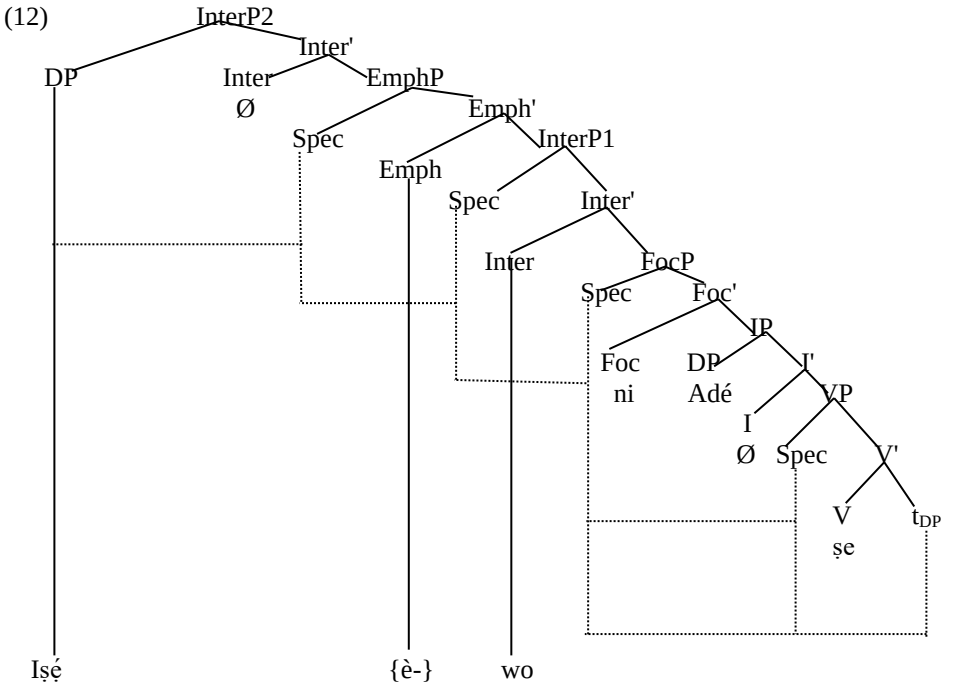
(11)



Apparently in (6), the object DP *èwo,* which is used as question word moves from its extraction site to clause-initial, i.e. Spec-InterP (ex-situ) to satisfy the specifier feature requirement. The constituent leaves a trace at its extraction site that is co-indexed with the question item at the clause-initial (landing site). This is in line with what Rizzi (2001) and Aboh (2004) proposed. This also implies that two functional heads – Inter and Foc, feature to encode a single discourse information, i.e. interrogative force (Aboh & Pfau 2011:96). It should also be noted that IP attracts the DP Adé (name of a person) to satisfy the EPP.

However, on the account of the occurrences of content question words in interrogative clauses in languages, Siemund (2001) notes that "language can be described in terms of the position of occurrence of its interrogative words as fronting, in-situ or optional fronting respectively, as the case may be." However, in contexts where (è)*wo* shows up after the subject or object DP, I assume that {è-} in è*wo* is an emphatic marker while [*wo*] is the authentic question word there. The structure of such proposition is shown in (12).

(12)



Apparently in the structure above, {è-} in *èwo* there indicates emphatic while [*wo*] is the question word. In this type of projection, I propose multiple layers of InterP, i.e. InterP1 and InterP2, where InterP1 projects EmphP and EmphP projects InterP2. It should be noted that the InterP2 has null phonetic manifestation (head) there. However, the DP iṣẹ́ 'work', which is the complement of the V scoped across and began to search for its landing site. It first landed at the Spec-VP, which makes the verb automatically becomes a probe. It later moved from there to the Spec-FocP, touching ground at the Spec-InterP1, Spec-EmphP and finally landed at the Spec-InterP2, to check and satisfy the specifier requirement. It should also be noted that IP and InterP2 attract the DP Adé (name of a person) and **iṣẹ́** 'work' respectively to satisfy the EPP.

In this type of derivation, {è-} and [*wo*] were first inserted before the internal merge, so that the structure could converge accordingly. More so, since the Spec-InterP1 of such structure has question feature, there is need to check this feature against any question element. This position parallels (Chomsky 2000) assertion that "movement must be triggered by a feature on a functional head." This Chomsky's assertion is in line with the proposition of Rizzi (1997:282) that "syntactic movement…must be triggered by the satisfaction of certain quasi-morphological requirement of heads. The table below shows the distributions of (è)wo.

	Clause-Initial		**Clause-Final**	
	Before the Subject or Object DP	**After the Subject or Object DP**	**Before the Verb**	**After the Verb**
Emphatic	√	X	X	√
Non-emphatic	X	√	√	X

Table 1: The Distribution of (è)wo

The keys: √ simply means the question element is present in the syntactic position, while x connotes that the question morpheme is absent in the syntactic vicinity.

The table simply interpreted as if *(è)wo* shows up before the DP (subject or object premises), *{*è-*}* in *èwo* is obligatorily pronounced, i.e. [+emphatic] but after the DP (subject or object), *{*è-*}* in *èwo* is optionally pronounced, i.e. [+non-emphatic]. In other words, if *{*è-*}* in *èwo* features after the V/VP (clause-final), *{*è-*}* in *èwo* is obligatorily pronounced, i.e. [+emphatic] but if the same element surfaces linearly before the verb, it is optionally pronounced [+non-emphatic].

6 How can (è)wo be construed?

In this section, we take a look at the likely readings that could be deduced from *(è)wo*.

6.1 *Specificity* The use of *èwo* can be construed as a way for a speaker to indicate to the hearer that in a set of possible items, only one is appropriate or the hearer is permitted to take only one, i.e. the speaker uses *èwo* as an open cheque for the hearer to pick or select his/her choice and thereby connotes strong obligation to select one and not two, as seen in (13).

(13a)	*Èwo*	ni	o	fẹ́	nínú		aṣọ,	bàtà	àti	fìlà?
	Which	FOC	2sg	want	LOC-inside		cloth	sandal	CONJ	cap
	'Which one did you want among the cloth(es), sandal and cap?'									
(b)	*Èwo*	ni	Adé	fẹ́ràn	nínú	ìrẹsì,		ẹ̀wà	àti	iṣu?
	Which	FOC	Ade	like	LOC-inside	rice		beans	CONJ	yam
	'Which one did Ade like among the rice, beans and yam?'									
(c)	*È*wo	ni	ó	wu	ẹ	nínú	ilé,	ilẹ̀	àti	ọkọ̀?
	Which	FOC	HTS	like	2sg	LOC-inside	house	land	CONJ	motor
	'Which one did you want among the house, land and motor?'									

[3]Note that if *èwo* shows up before the subject or object DP, we refer to it as non-emphatic structure but if *(è)wo* features after the subject or object DP, it is referred to as emphatic structure

In (13a), the speaker uses a set of aṣọ 'cloth(es)', bàtà 'sandal' and fìlà 'cap', which is obligatory for the hearer to specify the one he/she likes. The same scenerio happens in (13b) where the hearer is permitted to pick one item among ìrẹsì 'rice', ẹ̀wà 'beans' and iṣu 'yam'. In (13c), the addressee is also allowed to select one among ilé 'house', ilẹ̀ 'land', and ọkọ̀ 'motor'. All these express specificity.

6.2 *Emphatic* In this kind of interpretation, the addresser has only one person or item which has varieties in his/her custody but wants the addressee to be definite about the one he/she prefers. Examples (14) below are declarative sentences, which are interrogated in (15).

(14a) Wọ́n ń pe Adé.
3pl Prog call Adé
'They are calling Ade'

(b) Wọ́n ti mú ẹja.
3pl Perf take fish
'They have taken fish'

(15a) Adé *(è)wo* ni wọ́n ń pè?
Adé which FOC 3pl Prog call
'Which of the Ades' are they calling?'

(b) Ẹja *(è)wo* ni wọ́n mú?
Fish which FOC 3pl take
'Which of the fishes do they take?'

Adé as used in (15a) has varieties; it may be short, tall, black, light or Adé with different surnames but the addresser wants the person asking for Adé to be definite on the type of Adé. In (15b), ẹja 'fish'could be any type (as we have them in the market) but the addresser wants the addressee to be definite about the type of fish.

6.3 *Numerical* Numerically, if *(è)wo* is used with long pronoun **àwọn** 'third person plural' in Yorùbá, the reading is always expressing number, as illustrated in (16).

(16a) [Àwọn ọmọ *(è)wo*] ni wọ́n ń pariwo?
3pl child which FOC 3pl Prog make-noise
'Who are the children that are making noise?'

(b) [Àwọn akẹ́kọ̀ọ́ *(è)*wo] ni wọ́n wọlé?
3pl students which FOC 3pl enter-house
'Who are the students that entered the house?'

---

[4]Ajíbóyè (2010:148) reported that ***àwọn*** '3pl pronoun'is used to mark plural in Yorùbá. For instance,
(i). [Àwọn obìnrin] wa tún pín sí ọ̀wọ́ méjì.
PL 1pl-Gen again divide to group two
'The women again divided into two groups.'
(Ajíbóyè, 2010:148)

[5]Bámgbóṣé (1990:214) identified **o** with mid-tone as emphatic marker but due to the tone interference, **o** in example (17) bears low tone. For instance,
(ii). Mò ń bọ̀ o.
1sg Prog to-come Emph
'I am coming.'
(iii). A ti ṣetán o.
1pl Perf do-finish Emph
'We are done.'
(Bámgbóṣé, 1990:214)

All the examples in (16a&b) appear to indicate the number of the children or students in question, i.e. plural form. Àwọn ọmọ 'the children'as used in (16a), may be construed as two, three, four or more children. In the same vein, in (16b), àwọn akẹ́kọ̀ọ́ 'the students' as used there maybe interpreted as brilliant, obedient, hardworking, diligent students, etc. Both examples (16a-b) show the number of persons under consideration.

6.4 *Negative* In expressions where *(è)*wo takes emphatic marker **ò**, the readings are constantly negative. This type of expression is common among the Yorùbá to show that somebody or something has no value. This is shown in (17).

(17a) Ọmọ *(è)wo* ò?
Child which Emph
'Which child/children?'

(b) Ilé *(è)wo* ò?
House which Emph
'Which house(s)?'

(c) Òbí *(è)wo* ò?
Parent which Emph
'Which parent(s)?'

Ọmọ 'child', ilé 'house(s)'and òbí 'parent(s)' as used in (17a-c) connote negative. The child maybe a bastard, disobedient, slotful, prodigal, or lazy child. The house could mean it is not well built or no one lives there and achieve better things while the parent(s) could be interpreted as irresponsible ones. All construed negative meanings because of the presence of emphatic marker. When the speakers of Yorùbá are expressing this type of notion, they frown their face and as well turn their mouth badly (to one side) to demonstrate that such thing or fellow is unvalued.

7 Conclusion

The article has examined *(è)wo*, which is believed to be one of the content question elements in the linguistic literature in Yorùbá. It contended that the morphological make-up of *èwo* is prefix {è-} and VP *wo* (root or stem). It proposed two syntactic positions for *(è)wo*: before and after subject or object DP with the emphasis on {è-} in *èwo* as emphatic marker when it shows up immediately after the subject or object DP. The paper also claims that basic clause is the acceptable and appropriate answer when *èwo* features in rightward (10b) (= echo question) but focus expression is the rightful answer when *èwo* occurs in discourse position (7b&8b). The paper did not fail to show the original syntactic position of *(è)wo* before being scoped to the Spec-InterP. Finally, it discussed the semantic interpretations of *(è)wo* in Yorùbá.

References

Aboh, E. O. (2004). The Morphosyntax of Complement-Head Sequences: Clause Structure and Word Order Patterns in Kwa. Oxford University Press, New York

Aboh, E. O. & Pfau, R. (2011). What's a Wh-word Got to Do with It? P. Benincà (ed.) Mapping the Left Periphery: the Cartography of Syntactic Structures, Volume 5, pp. 91-124

Akanbi, T. A. (2016). On the WH Question Particle [Wo] in Yoruba. European Scientific Journal, Vol. 12, No 17

Akinlabi, A. & Liberman, M. (2000). The Tonal Phonology of Yorùbá Clitics; B. Gerlach & J. Grijzenhout (Eds.), Clitics in Phonology, Morphology and Syntax. John Benjamins, Amsterdam

Ajiboye, O. (2010). Plural Strategies in Yoruba[1]. Studies in African Linguistics, Vol. 39, No 2

Ajiboye, O. (2013). A Comparative Study of Yes-No Questions in Àwórì, Ìgbómìnà, Ìkálẹ̀, Mọ̀bà and Owé Dialects of Yorùbá in Fábùnmi, F.A. and Sàláwù (eds.). Readings in African Dialectology and Applied Linguistics. Pp. 3-24

Awóbùlúyì, O. (1978). The Essentials of Yorùbá Grammar. Oxford University Press, Nigeria

Awóbùlúyì, O. (2001a). Arọ́pò orúkọ Kúkúrú Ẹnì-Kẹta Ẹyọ Aṣolùwà, Yorùbá: Journal of the Yorùbá Studies Association of Nigeria, Vol. 2, No 1, pp. 1-8

Awóbùlúyì, O. (2006). 'Ó Kì í Ṣe Arọ́pò orúkọ ní Èdè Yorùbá', Seminar Paper, Department of Linguistics and Languages, Adekunle Ajasin University, Akungba-Akoko, Ondo State, Nigeria

Awóbùlúyì, O. (2008a). Ẹ̀kọ́ Ìṣẹ̀dá-Ọ̀rọ̀ Yorùbá. Montem Paperbacks, Akurẹ
Bámgbóṣé, A. (1966). A Grammar of Yorùbá. Cambridge University Press, London
Bámgbóṣé, A. (1990). Fonọ́lọ́jì àti Gírámà Yorùbá. University Press Plc
Chomsky, N. (1993). A Minimalist Program for Linguistic Theory, Hale and Keyser (Eds.) View from Building 20: Essay in Linguistics in Honor of Sylvain Bromberger pp 1-52
Chomsky, N. (1995b). Minimalist Program. MIT Press, Cambridge, Mass
Chomsky, N. (1998). The Minimalist Enquiries: The Framework, MIT Occasional Papers in Linguistics, No 15
Chomsky, N. (2002). On Nature and Language. Cambridge University Press, Cambridge
Collins, C. (2013). Introduction to Minimalist Syntax. African Linguistics School Handout, Ibadan, Nigeria
Ilọri, J.F. (2010). Nominal Constructions in Igálà and Yorùbá. Ph.D. Dissertation, Adékúnlé Ajásin University, Ondo
Ilọri, J.F. (2017). Interrogative Projections in Yoruboid Languages. Journal of West African Languages, Vol. 44, No. 1
Ilọri, J. F. & Oyebade, F. O. (2012). Negation in Igala. Entrepalavras, Fortaleza-ano, 2(1)
Ọláògún, S. (2018). Yes/No and Wh-Question in Ǹjò-kóó: A Unified Analysis. Corela Cognition, Représentation Language, Vol. 16, No. 1
Radford, A. (2009). An Introduction to English Sentence Structure. Cambridge: Cambridge University Press
Williamson, Kay/Blench, Roger (2000): Niger-Congo. In Heine, Bernd/Nurse, Derek(eds.): African Languages: An Introduction. Cambridge
Rizzi, L. (1997). The Fine Structure of Left Periphery. In Liliane Haegeman (ed.) Element of Grammar. Dordrecht: Kluwer
Rizzi, L. (2001). On the Position of Inter(rogative) in the Left Periphery of the Clauses, G. Cinque and G. Salvi (eds.), Current Studies in Italian Syntax: Essays Offered to Lorenzo Renzi, Elsevier-North Holland, Amsterdam, pp. 287-296