UNIVERSITY OF CALGARY

ISSN 2371-2643

CALGARY WORKING PAPERS IN LINGUISTICS

Volume 31, Fall 2020

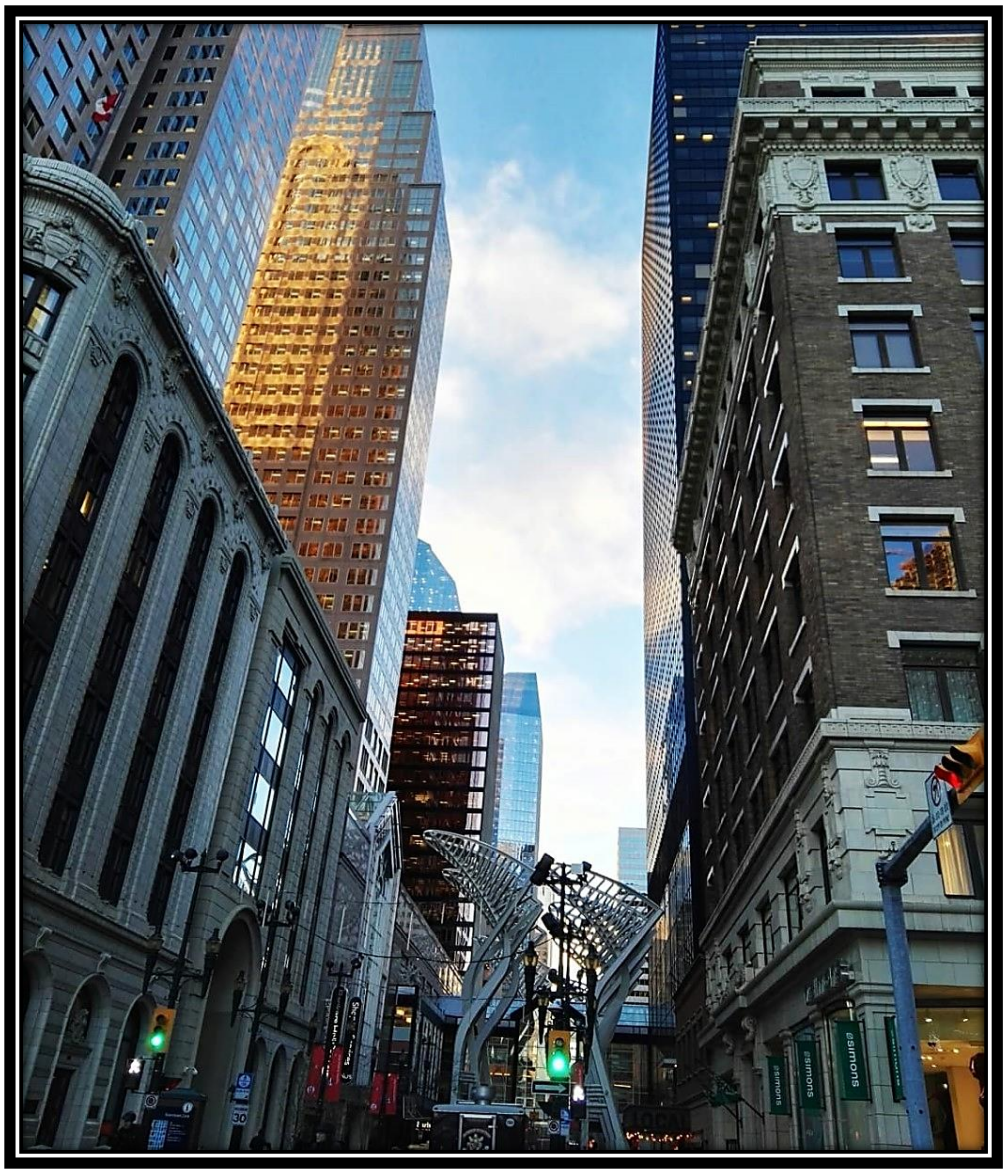


TABLE OF CONTENTS

FOREWORD

The editors of this volume are pleased to present the thirty-first volume of the Calgary Working Papers in Linguistics (CWPL). CWPL is a publication affiliated with the School of Languages, Linguistics, Literatures, and Cultures (SLLLC) at the University of Calgary, focusing on the most recent contributions in linguistics and related disciplines by researchers affiliated with the University of Calgary. All the previous volumes of CWPL from Vol. 1 (published in 1975) are stored in PRISM: The University of Calgary Digital Repository and can be accessed at: http://dspace.ucalgary.ca/handle/1880/51236

Each paper submitted to Volume 31 has been reviewed and edited by two editors, all graduate students of linguistics at the University of Calgary. It should be noted that the papers published in CWPL represent works in progress and should not be considered as final or definitive papers. Therefore, publication in CWPL does not preclude further submission of further revisions of the same papers to another journal.

Volume 31 consists of 9 papers and it is one of the most substantial volumes edited, rich in its diversity of linguistic topics it explores. The topics in this volume range from Syntax, Phonetics, and Phonology to First and Second Language Acquisition. In addition to the regular submissions, this volume contains three papers that were presented at the *Northwest Phon{etics;ology} Conference* held on September 20th, 2019, at the University of Calgary. We would like to express our gratitude to all the attendees and organizers of this conference, and all the authors of papers in this volume.

Finally, we would like to take this opportunity to sincerely thank the members of the editorial board who are instrumental in providing a rapid and efficient editorial process and maintaining high standards within our publication.

Editors of CWPL - Vol. 31,

Dušan Nikolić
Adam D. Daniel
Brett C. Nelson
Metehan Oguz
Kang Xu

October, 2020

Front page photo courtesy of Milica Nikolić, 2020.

Acquisition of Tense and Lax Vowels by Persian Learners of English

Elias Abdollahnejad

University of Calgary

Abstract

This study investigates the qualitative and quantitative features of English tense (/i: & u:/) and lax (/ɪ & ʊ/) vowels in the L2 production of advanced L1 Persian speakers. The first two formants (F1 & F2) are used to determine the spectral quality of vowels, while the vowel length (in milliseconds) represents the vowel quality. Unlike English, which has two high front (/i:/ & /ɪ/) and back (/u:/ & /ʊ/) vowels, Persian has only one high front vowel /i/ and one high back vowel /u/. Considering this difference, the main questions of this study are whether L1 Persian speakers have separate representations of English tense and lax vowels in their L2 phonological inventory and, if yes, whether they differentiate these vowels based on their qualitative and/or quantitative features. According to the Equivalence Classification Hypothesis (Flege, 1987), L1 Persian speakers are expected to assimilate/merge English front vowels /i: & ɪ/ to their existing L1 single front vowel /i/ and English /u: & ʊ/ to their existing single Persian back vowel /u/. A group of ten female advanced L1 Persian speakers of English are compared with a control group of ten female native English speakers in their production of these vowels in open and closed syllables. Results show that while both quality and quantity are used by English speakers, the Persian speakers have difficulties in acquiring this distinction in their L2 English on both measures. However, the difficulty seems to be mostly in the acquisition of L2 phonetic features (i.e. gradual) than phonological ones (i.e. categorical).

Key words: tense vowels, lax vowels, vowel quality, vowel quantity, Persian

1 Introduction*

First language (L1) transfer has always been considered as one of the most probable sources of interference in the process of second language (L2) acquisition. As Odlin (2003) points out, "all linguistic subsystems including pragmatics and rhetoric, semantics, syntax, morphology, phonetics, and orthography" can be influenced by language transfer (p. 437). For instance, it has been observed that L1 Arabic learners of English tend to transfer their L1 vowel features to produce target vowels (Flege & Port, 1981; Munro, 1993; among others). This study aims to investigate the extent to which phonological characteristics of L1 interfere with L2. To do so, it compares and contrasts the qualitative and quantitative features of English tense (/i:/ & /u:/) and lax (/ɪ/ & /ʊ/) vowels in the interlanguage of advanced L1 Persian speakers of English and native English speakers. Unlike English, which has a pair of high front (/i:/ & /ɪ/) and back (/u:/ & /ʊ/) vowels, each differentiated according to [±Tense] feature (or advanced versus retracted tongue root feature, i.e. [±ATR], according to some studies)[1], Persian has only one high front [i] and only one high back [u] vowel. Considering this difference between the two languages and according to equivalence classification hypothesis (Flege, 1987), L1 Persian speakers are expected to face difficulties in distinguishing the tense/lax contrast within each of these pairs in their L2 English.

In most English L2 classes, the students are taught that the only difference between tense and lax vowels is in their length, i.e. tense vowels are longer than lax vowels (Wang & Munro, 2004). For instance, in order to differentiate between the tense /i/ and the lax /ɪ/, Korean learners of English are asked to focus on length (quantity) rather than qualitative phonetic differences (Flege et al., 1997). Yet, the latter has been reported to be a more reliable factor for English speakers to differentiate between the tense and lax counterparts in their native language (Flege et al., 1997; Iverson & Evans, 2009; Kondaurova & Francis, 2010; McAllister et al., 2002; among others).[2]

Whether to determine the contrast in a vowel system through quantity or quality has been largely discussed in the literature (van Oostendorp, 1995; Odden, 2011; among others). By quantity, we mean the length (duration) of each vowel. As Gordon (2004)

*This study is approved by the University of Calgary Conjoint Faculties Research Ethics Board (Approval number: REB17-1315). The paper was presented at the *Northwest Phon{etics;ology} Conference* held on September 20th, 2019, at the University of Calgary.

[1] "[ATR] in African languages is often claimed to be the 'same thing' as [tense] in European languages" (Carr & Honeybone, 2007, p. 139). However, Durand (2005) points out that "this feature does not cross-classify sounds in exactly the same way as tense/lax" (p. 4).

[2] It should be pointed out that there is an ongoing debate in the literature on whether English vowel system is purely qualitative-based (e.g. Jakobson et.al., 1952; Jakobson & Waugh, 1979), purely quantitative-based (e.g. Lass, 1976; Durand, 2005; Kwon, 2011), or depends on both quality and quantity (e.g. Chomsky & Halle, 1968; Halle, 1977; Lee, 2003). However, this issue is above the scope of this research and does not majorly influence the main aims of this paper. The main aims of the present research are the possible effects of L1 phonological system on L2 acquisition in general and whether adult L2 learners can adapt their existing phonological system (representations) according to different/new features in the phonological system of their L2 in specific. More specifically, whether L1 Persian speakers can acquire the high tense/lax counterparts in English not existing in their L1.

points out “Contrasts in segmental length are represented by assuming that long segments are associated with two weight units, while short segments are associated with one unit of weight” (p.3). Therefore, any vowel containing more than one mora is considered long (Perlmutter, 1995; Tranel, 1995). In other words, the long vowels are bimoraic and the short ones are monomoraic. The vowel system of a language is considered quantitative if it is the number of morae in each vowel which makes the contrast between different vowels. Quality, on the other hand, refers to the phonological features (i.e. ±High, ±Back, etc.) causing contrasts among different vowels of a language. The first two formants (F1 & F2) are commonly used to determine the spectral quality of vowels, while duration gives a good measure of the quantity aspect of the vowel.

This study investigates the qualitative and quantitative features of L2 English tense (/i:/ & /u:/) and lax (/ɪ/ & /ʊ/) vowels. The main questions are whether L1 Persian speakers have separate representations of L2 English tense and lax vowels in their phonological inventory and, if they do, whether they differentiate these vowels based on quantity and/or quality. Considering the simpler Persian vowel inventory (only 6 vowels) compared to English (15 vowels), L1 Persian speakers of English are expected to assimilate/merge English front vowels /i:/ and /ɪ/ with their existing single L1 front vowel [i] and English /u:/ and /ʊ/ with their existing Persian back vowel [u]. A group of advanced L1 Persian speakers of English are compared with a control group of native English speakers in their production of these vowels in open and closed syllables. Results show that while both quality (F1 & F2) and quantity (length) are used by English speakers, the Persian speakers have difficulties in acquiring this distinction in their L2 English on both measures. However, the difficulty seems to be mostly in the acquisition of L2 phonetic features (i.e. gradient) than phonological ones (i.e. categorical). In other words, although they are able to distinguish and categorize L2 tense/lax counterparts, their performance is not native-like. They were different from the native group in their F1s of both front and back tense/lax counterparts in closed syllables, and the lengths of /ɪ/ and both back vowels in both syllable types. The syllable type (#CVC# vs. #CVσ) was also observed to cause quality and quantity distinctions for both L1 groups, although no clear L1 effect was observed for the L1 Persian speakers regarding the environment (closed vs. open syllable). In other words, they do not treat L2 vowels quite the same way they treat their similar L1 vowels.

This paper is organized as follows. Section 2 provides a brief overview of Persian vowel system, explaining the controversy over its sensitivity to quantitative and/or qualitative features. Section 3 is focused on the acquisition of L2 phonological system, discussing equivalence classification and markedness differential hypotheses and their predictions regarding the acquisition of English tense/lax counterparts by L1 Persian speakers. The material, design and analyses used in the study are explained in Section 4. A summary of the main results is provided in Section 5. Section 6 discusses the results and their interpretations considering the role of L1 on the acquisition of L2 phonological system and the prediction of the two main hypotheses introduced in Section 3. Finally, the concluding remarks are provided in Section 6.

2 Persian vowels

Modern Persian has six vowels: [i, e, a, u, o, and ɑ] (Windfuhr & Perry, 2010; Mahootian & Gebhardt, 1997; among others).[3] The distinguishing (i.e. active) phonological features of vowels in Modern Persian has been a controversial topic in the literature. Some researchers (e.g. Hayes, 1979) have considered it a quantity-only differentiation system, while others (e.g. Samareh, 1977) have argued for a quality-only system.[4] Additionally, Toosarvandani (2004) has argued for the system to be sensitive to both quality and quantity features in its vowel distinction.[5]

According to the quantity-only view of Modern Persian vowels (e.g. Hayes, 1979; Windfuhr, 1979), quantity is the active feature in the system, i.e. the vowels [a], [e], and [o] are phonologically short, while [ɑ], [i], and [u] are long. According to this view, the system contains three vowels, each with two different phonetic representations caused by the length contrast, as in Table (1) (Rohany-Rahbar, 2009).

Table 1: Quantity-based distinction of Persian vowels

Main Vowel	Phonetic representation	
	Long	Short
i	[i]	[e]
u	[u]	[o]
a	[ɑ]	[a]

However, as pointed out by Mahootian and Gebhardt (1997), length is a redundant feature in the case of Persian vowels and they cannot be distinguished only based on this feature. In other words, although some vowels are long and others are short, this feature is not contrastive in Persian. Thus, there are some "quality-only" based alternative analyses of Persian vowels, categorizing them based on the features such as [±High] and [±Back], as in Table (2).[6]

Table 2: Persian Vowel Inventory Mahootian and Gebhardt (1997)

	Front		Back	
High	i	[long, tense, closed]	u	[long, tense, closed, round]
Mid	e(ɛ)	[short, lax, slightly open]	o	[short, lax, slightly open, round]
Low	a*	[short, lax, open]	ɑ**	[long, tense, open]

*They use the phonetic symbol [æ]. **They use the phonetic symbol [a].

[3] It should be noted that "all Persian vowels are voiced" (Mahootian & Gebhardt, 1997, p. 294).

[4] Some researchers (e.g. Pisowicz, 1985 as cited in Rohany-Rahbar, 2009) argue for a change from a quantity-only vowel system in Middle Persian to a quality-only system in Modern Persian.

[5] According to this synthetic analysis (i.e. both quantity and quality features), "the vowel system of Modern Persian is considered to be in a 'transition state' between the purely quantitative system of classical Persian and the system of future Persian which will eliminate any phonological evidence for quantity altogether and keep quality as the only distinguishing feature" (Rohany-Rahbar, 2009, p. 2).

[6] Mahootian and Gebhardt (1997) also proposes the presence of /ow/ as the only phonemic diphthong in Persian.

Regarding the environments in which different types of vowels can occur, the long vowels are more frequent word-finally, whereas the short counterparts are not so commonly used in this position (Mahootian & Gebhardt, 1997).[7] Additionally, in modern Persian, the length distinction between vowels is neutralized in most environments (Windfuhr, 1979). In other words, the length differentiation between vowels is only possible in limited environments (Toosarvandani, 2004). More specifically, the long vowels behave consistently regarding their length in different environments, but the short vowels have an inconsistent behavior. Therefore, the long vowels are referred to as "stable" and the short vowels are called "unstable". Specifically, long vowels are always long, but "short vowels in Persian are long in closed or word-final syllables and short elsewhere" (Toosarvandani, 2004, p. 250).

Toosarvandani (2004) points to the necessity of a lowering rule to get [e] and [o] from [i] and [u], respectively, and also a fronting rule to get [a] from [ɑ]. He argues that "though the 'quantity only' analysis is able to generalize unstable vowel lengthening as a group process, its qualitative opacity requires us to reject it" (p. 244). He also rejects the efficiency of a purely-qualitative analysis and argues that such view cannot account for the unstable behavior of some vowels (i.e. short vowel lengthening) and needs "re-categorisation of short vowels in closed syllables as long" (p. 251).[8] Hence, Toosarvandani (2004) suggests that a quality-only analysis does not give a concise view of the surface vowel length variations, and a quantity-only view provides an excessively abstract picture of the system. He proposes a synthetic view of Modern Persian vowel system and points to the necessity of the "oppositions of both quantity and quality in the underlying vowel system in order to describe the observed distributional facts and alternations adequately" (p. 250). Table (3), illustrates his feature specifications of Persian vowels.

Table 3: Feature specifications of Persian vowels (Toosarvandani, 2004, p. 245)

	e	o	A	i:	u:	ɑ:
High				+	+	
Low			+			+
Back		+			+	+

Considering the clear differences between Persian and English in their vowel inventories and their active features, L1 Persian speakers of English are expected to face some difficulties acquiring their L2 vowel system. The following section will discuss some of these expected difficulties through *equivalence classification* and *markedness differential* hypotheses and will end with some research questions.

[7] Although all these vowels can occur word initially, [e] and [æ] are the most common vowels, [a] is also common and [u] is not commonly used in this position (Mahootian & Gebhardt, 1997).

[8] Specifically, Toosarvandani (2004) posits that "the unstable vowels do not form a natural class distinguishable from the stable vowels. For the unstable vowels: 1) *e* and *a* are [−back] while *o* is [+back]; 2) *e* and *o* are [−low] while *a* is [+low]; and 3) *e*, *o*, and *a* share the [−high] feature with the stable vowel *a*. Conversely, for the stable vowels: 1) *u* and *a* are both [+back] while *i* is [−back]; 2) *u* and *i* are [−low] while *a* is [+low]; and 3) *i* and *u* share the [+high] feature while *a* does not." (p. 245).

3 The acquisition of L2 phonetic system

It is generally believed that L2 learners use their established L1 phonetic categories to identify L2 phones. Considering the differences between languages in their categorization of contrastive sound units, Flege (1987) suggests that the L2 phones are acoustically classified as "identical", "similar", or "new" in comparison to the L1 phones. Trubetzkoy (1969, as cited in Flege, 1987) proposes a "phonological filtering" process in L2 acquisition which will filter out the perceptually acoustic differences not established in L1. For instance, Japanese learners of English have difficulties in differentiating between [l] and [r]. As Flege (1987) points out, the same filtering is caused in the case of "similar" L1 and L2 sounds.

Flege (1987) introduces *equivalence classification* as a basic cognitive mechanism allowing "humans to perceive constant categories in the face of the inherent sensory variability found in the many physical exemplars which may instantiate a category" (p.49). It is this phenomenon which allows children to realize if the phones produced by different speakers and in different phonetic contexts belong to the same category (Flege, 1987). However, the same phenomenon may prevent the older children and adult L2 learners from "making effective use of auditorily accessible acoustic differences between phones in L1 and L2" (p. 50). The idea is that since the L2 learners already have a well-established inventory of phonetic categories, the equivalence classification prevents them "from establishing a phonetic category for similar but not new L2 phones" (p. 50). In other words, Flege (1987) argues for equivalence classification as a phenomenon constraining L2 learners from reaching L2 phonetic norms in the case of L2 phones "similar" to their existing L1 phonetic categories. He argues that equivalence classification does not occur in the case of "new" L2 phones not existing in L1 system.[9]

Since in the case of contrast between the existing vowels and the new ones, L2 learners need to employ new phonological patterns typically different from the existing ones (Strange & Shafer, 2008), it is likely for the L2 learners to categorize the new vowels according to their closest counterparts in the existing L1 vowel system (Flege et al., 1997). For instance, since Arabic has /i/ and /u/ but not /ɪ/ and /ʊ/, L1 Arabic speakers of English are expected to face difficulties acquiring tense/lax distinction of high vowels (/i/ vs. /ɪ/ and /u/ vs. /ʊ/) in their L2 (Flege, 1995; Alzahrani, 2014). In their study of L1 Arabic learners of English, Flege and Port (1981) noticed such use of L1 values to produce L2 sounds. Additionally, Flege et al. (1997) observed the absence of /ɪ/ in the vowel system of Spanish, Korean, and Mandarin learners whose L1 contain only /i/ not /ɪ/. Thus, these learners are more likely to classify both /i/ and /ɪ/ under their L1 /i/ sound.

Considering the *Equivalence Classification hypothesis* and that Persian has single high front and high back vowels, L1 Persian speakers of English are expected to overlook tense/lax contrast of English high vowels and merge them into only one front and one back

[9] He also mentions a restructuring of the phonetic space influencing L1 categories "similar" to L2 ones as a result of L2 exposure. He calls it "merger process" because "similar" L1 and L2 categories Merge in a new category approximately different from their relevant categories in both languages.

vowel category. In other words, they are expected to fail to notice spectral differences between /i:/ and / ɪ/, and /u:/ and /ʊ/ in English.

Eckman (1977), in his *Markedness Differential Hypothesis*, considers the acquisition of marked phenomena more difficult than the less marked ones. Since cross-linguistically the lax vowels are believed to be more marked than the tense vowels, they are expected to be more difficult to acquire (Maddieson, 1984). In the literature (Munro, 1993; Wang & Munro, 2004; Munro & Derwing, 2008; among others), it has been observed that the L2 front vowels are acquired earlier and with less difficulty than the back vowels. Moreover, according to Markedness Differential Hypothesis, increase in the degree of markedness of L2 sounds causes them to become more difficult to acquire. Therefore, the back vowels are generally considered more marked than the front vowels. Considering the Markedness Differential Hypothesis, the L1 Persian speakers are expected to have more difficulty in the native-like production of English lax vowels compared to tense vowels in general, be even more challenged with the English back lax vowel /ʊ/ than the front lax vowel /ɪ/.

The present study investigates the qualitative and quantitative differences between L1 Persian speakers and native speakers in the productions of English tense (/i:/ & /u:/) and lax (/ɪ/ & /ʊ/) vowels. The main goal is to see if L1 Persian speakers have separate representations of English tense and lax vowels in their phonological inventory and differentiate them based on quantity and/or quality. Therefore, the main research questions of this study are:

1. Do advanced L1 Persian speakers have separate representations of English tense and lax vowels in their L2 English phonological inventory?
2. Do advanced L1 Persian speakers and native English speakers use different strategies (quality and/or quantity) to differentiate English tense and lax vowels?
3. Do advanced L1 Persian speakers merge English lax vowel /ɪ/ (marked) with tense /i:/ (unmarked) and merge English lax vowel /ʊ/ (marked) with tense /u:/ (unmarked)?
4. Do advanced L1 Persian speakers of English consider the lax vowel /ɪ/ and / ʊ / as short vowels and, similar to their L1 short vowels, lengthen them in closed syllables?

4 Methodology

4.1 Participants

Ten female English native speakers and ten female advanced L1 Persian speakers of English participated in this study. Being exposed to English in daily life at least from the age of three and using English as the primary/dominant language of conversation while growing up and in the present time were the criteria to be considered native speaker of English. The native English speakers were all Canadian and the Persian speakers were Iranian graduate students at the University of Calgary in Canada.

4.2 Material

Sixteen English words (four instances of each target vowel) were selected to examine the four target English vowels /iː/, /ɪ/, /uː/, and /ʊ/. In order to test each vowel in two different phonological environments, i.e. closed (CVC) and open (CV) syllables, the hosting words were either monosyllabic (CVC) or multisyllabic with the target vowel in the first syllable (CV). There were two #CVC# and two #CVσ instances of each target vowel, i.e. sixteen target syllables containing the target vowels. To make the experiment more natural, the words were used in different carrier sentences, instead of a stable carrier sentence. Moreover, to minimize the prosodic effects on the pronunciation of the target vowels, the hosting words were all selected to be nouns or adjectives following the verb of the sentence and were also followed by at least one word to prevent being in the final position.[10] There were also 21 filler sentences. All the 37 sentences were randomly ordered and put in a list. Table (4) illustrates the hosting words containing the target vowels.

Table 4: The hosting words used for the target vowels in this study

Syllable Type	Tense		Lax	
	/iː/	/uː/	/ɪ/	/ʊ/
#CVC#	seat /ˈsiːt/	rude /ˈruːd/	big /ˈbɪg/	cook /ˈkʊk/
	cheap /ˈʧiːp/	juice /ˈjuːs/	rich /ˈrɪch/	full /ˈfʊl/
#CVσ	fever /ˈfiː·vər/	lunar /ˈluː·nər/	city /ˈsɪ·tiː/	butcher /ˈbʊ·ʧər/
	teacher /ˈtiː·ʧər/	noodle /ˈnuː·dəl/	wizard /ˈwɪ·zərd/	sugar /ˈʃʊ·gər/

4.3 Procedure

Each participant did the experiment individually over one session (about 15 minutes) in one of the language labs at the University of Calgary. After filling in a consent form and a voluntary demographic questionnaire, the participants were given the list of sentences and were asked to read the sentences out loud as natural as they speak English in their daily life while their voice being recorded. For the sake of consistency, all the recordings were done by the same digital recorder.

4.4 Analysis

All the recorded audio files were analyzed by the phonetic analysis software PRAAT version 6.0.14. The target vowels in each case were extracted by the software and analyzed for their length, first (F1) and second formant (F2). The formants were used to determine the tongue position (i.e. F1 for [±High] and F2 for [±Back] features) while pronouncing each target vowel. For the sake of consistency, all the phonetic analysis and vowel extractions was done by the researcher himself using PRAAT. To neutralize the effects of anatomical differences on the formants in each speaker's production, all the statistical tests were run on the Z-scores of the formant values. Generalized Estimating Equation (i.e., GEE under Genlin procedures in SPSS v.20) was used to analyze the data to investigate whether

[10] Refer to Appendix A for a list of the carrier sentences containing the target vowels.

there was any significant effect of vowel type (tense vs. lax), first language (Persian vs. English), syllable type (CVC vs. CV), and also their possible interactions. The α level of 0.05 was applied for all the statistical tests.

As in figure (1), since the *y* axis in the formant plot indicates first formant (F1) and the *x* axis indicates the second formant (F2), the Pythagorean Theorem ($c^2 = \sqrt{a^2 + b^2}$) was used to determine the distance between tense/lax counterpart vowels produced by each language or vowel group.

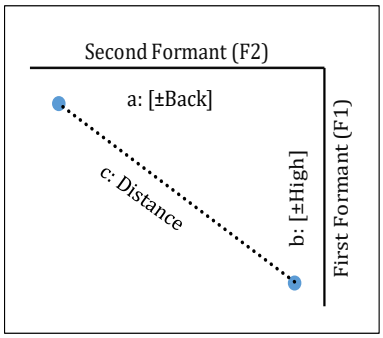


Figure 1: Distance between vowels on the formant plot through the Pythagorean Theorem

Thus, the equations below (1) were used to measure the vertical, horizontal, and overall distance between each target pair of vowels. Z-score formant values were also used in these formulae.

1) a) Vertical distance = $F1_{tense} - F1_{lax}$
   b) Horizontal distance = $F2_{tense} - F2_{lax}$
   c) Euclidean Distance = $\sqrt{(F1_{tense} - F1_{lax})^2 + (F2_{tense} - F2_{lax})^2}$

5 Results

The main focus of this study is on the analysis of high front tense and lax vowels (/iː/ and /ɪ/) and high back tense and lax vowels (/uː/ and /ʊ/) produced by L1 English and L1 Persian speakers to see if there are qualitative and/or quantitative variations in their accounts of these vowels in the two target environments (CVC vs. CV syllables).

5.1 High Front vowels (/iː/ versus /ɪ/)

5.1.1 First formant (F1)

Figure (2) below depicts the raw mean z-scores of different conditions of the study regarding the first formant of tense/lax front vowels /iː/ and /ɪ/.

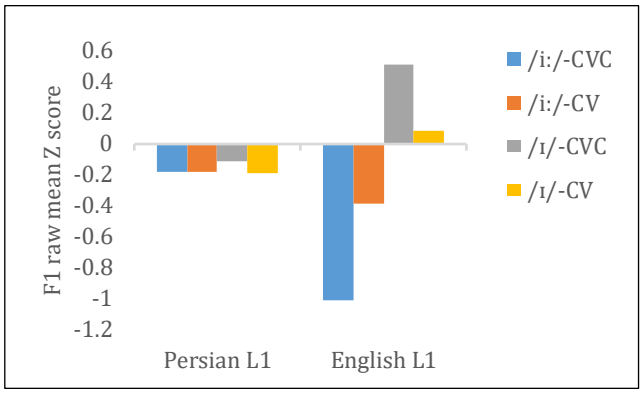


Figure 2: Raw mean z-scores of different conditions regarding F1 of /i:/ and /ɪ/

The results of the GEE analysis show that there is statistically significant two-way Language*Vowel ($\chi^2(1)=19.581, p<0.001$) and Vowel*Syllable ($\chi^2(1)=5.85, p=.016$) interaction effect, which means that the Language (Vowel) effect varies with different levels of Vowel (Language) and the Vowel (Syllable) effect varies with different levels of Syllable (Vowel), respectively. Thus, there is no need to interpret the main effects and we will directly perform the testing of simple effects, i.e. different levels of each variable broken down by the other variable. Regarding Vowel*Language interaction, L1 English group has a significantly greater mean for /ɪ/ than the L1 Persian group ($p<0.001$), while the L1 Persian group has a significantly greater mean for /i:/ than L1 English group ($p<0.001$). For Language*Vowel interaction, L1 English group has a greater mean of /ɪ/ than /i:/ ($p<0.001$). Additionally, for Syllable*Vowel interaction, /ɪ/ is significantly larger than /i:/ in CVC syllables ($p<0.001$). Table (5) illustrate the ZF1 means of these two-way interactions.

*Table 5: ZF1 means of Language*Vowel and Vowel*Syllable two-way interactions for /i:/ and /ɪ/*

	Persian L1	English L1	CV	CVC
/i:/	-.1803	-.7585	-.3398	-.6081
/ɪ/	-.1400	.3378	-.0190	.2168

Moreover, the results also show that there is a statistically significant three-way Language*Vowel*Syllable interaction effect($\chi^2(1)=4.118, p=.042$), which means that the Syllable (Vowel, Language) effect varies with different levels of Vowel (Language, Syllable) and Language (Vowel, Syllable). Testing of simple effects is performed to look at each variable effect broken down by the two other variables. Table (6) illustrates the estimated marginal means calculated through the GEE model.

Table 6: Estimated Marginal Means calculated through the GEE model for F1 of /i:/ and /ɪ/

	Persian L1		English L1	
	CVC	CV	CVC	CV
/i:/	-.1741	-.1865	-1.0421	-.4750
/ɪ/	-.0924	-.1875	.5261	.1495

The Bonferroni pairwise comparisons revealed significance in the case of the tense vowel /i:/ (p=.000) in CVC syllable types between the two language groups (red arrow in figure 3). Moreover, the pairwise comparisons revealed significance between the two vowels in CVC syllable types (p=.001) within the L1 English group (orange arrow in figure 3). Additionally, the Bonferroni pairwise comparisons showed significance of the two syllable types in the case of /i:/ (p=.000) within the L1 English group (green arrow in figure 3).[11]

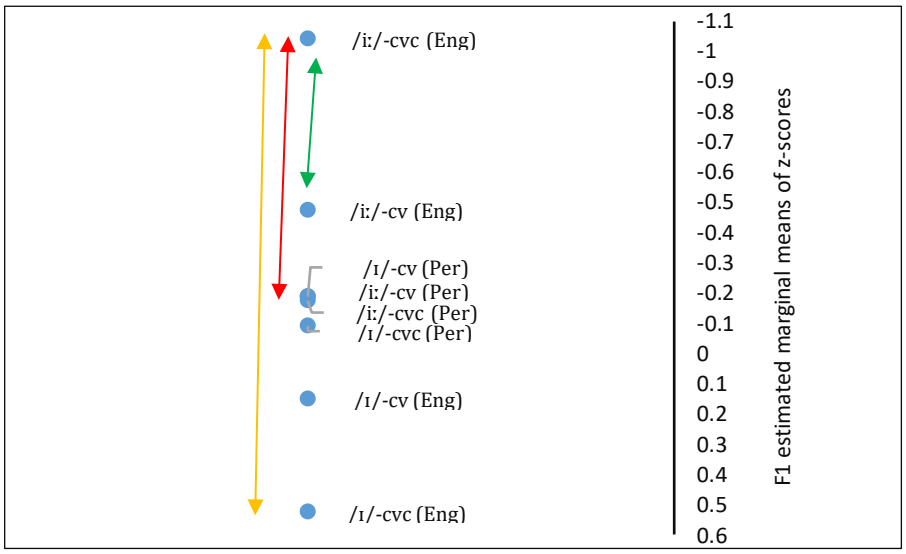


Figure 3: Differences between conditions regarding the first formants of /i:/ and /ɪ/

5.1.2 Second formant (F2)

Figure (4) below depicts the raw mean z-scores of different conditions of the study regarding the second formant of tense/lax front vowels /i:/ and /ɪ/.

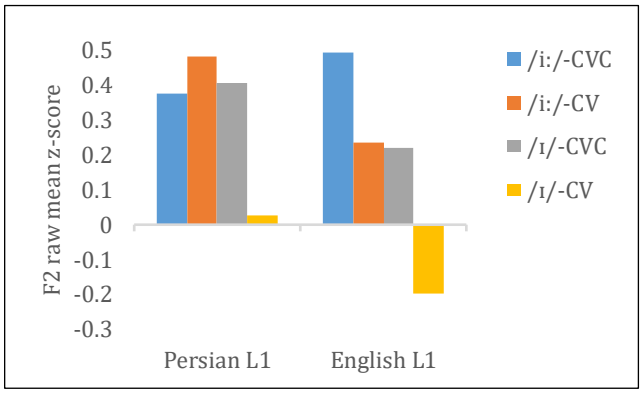


Figure 4: Raw mean z-scores of different conditions regarding F2 of /i:/ and /ɪ/

The results of the GEE analysis did not show any significance in different (two-way or three-way) interactions among different levels of the variables. Moreover, there was no

[11] Refer to Appendix B for the relevant SPSS output tables

significant effect of any of the main effects, i.e. L1 group, syllable type, and vowel type.[12] Figure (5) depicts the relative position of each case regarding its ZF2.

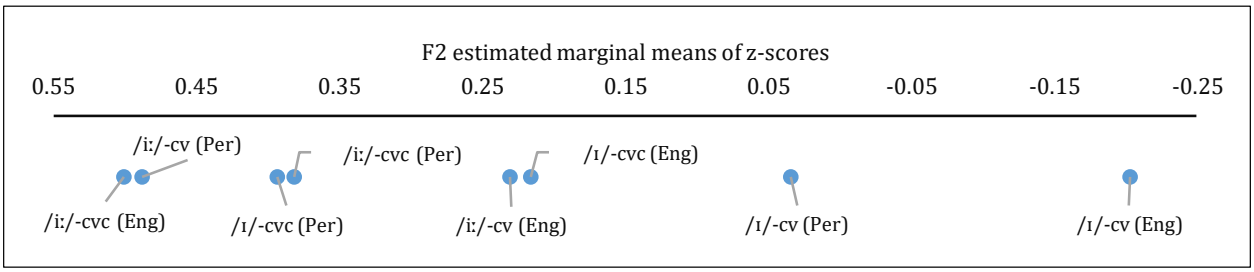


Figure 5: Differences between conditions regarding the second formants of /i:/ and /ɪ/

5.1.3 /i:/ and /ɪ/ distances on the formant plot in different condition

As indicated in Table (7) below, results of the Pythagorean Theorem considering means of F1 and F2 in each case determined that the distance between /i:/ and /ɪ/ was significantly shorter when produced by L1 Persian speakers compared to L1 English speakers, especially in the case of CVC syllables (.08 vs. 1.59).

Table 7: Z-score Means of each formant of /i:/ and /ɪ/ and their distances for each L1 group

Syllable Type	L1 Group	Mean $F1_{/i:/}$	Mean $F1_{/ɪ/}$	Mean $F2_{/i:/}$	Mean $F2_{/ɪ/}$	Vertical distance	Horizontal distance	Distance
CVC	Persian	-.1741	-.0924	.3814	.3932	-.0817	-.0118	.0825
	English	-1.0421	.5261	.5005	.2156	-1.5682	.2849	1.5939
CV	Persian	-.1865	-.1875	.4876	.0338	.001	.4538	.4538
	English	-.4750	.1495	.2309	-.2032	-.6245	.4341	.7606

5.1.4 Vowel length /i:/ vs. /ɪ/

Figure (6) below depicts the raw mean vowel length (in milliseconds) of different conditions of the study regarding the tense/lax front vowels /i:/ and /ɪ/.

[12] Refer to Appendix C for the relevant SPSS output tables

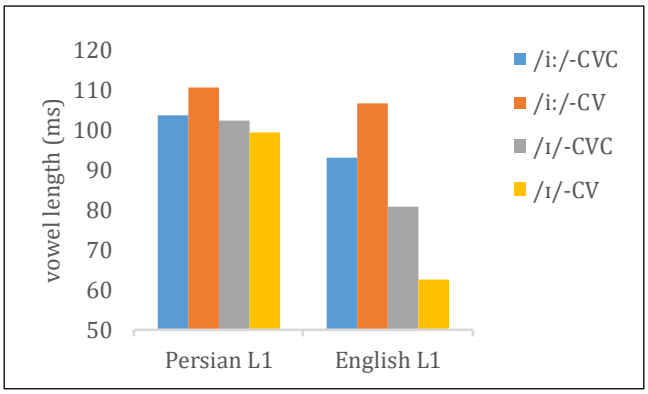


Figure 6: Raw mean vowel lengths of different conditions of vowels /i:/ and /ɪ/ (milliseconds)

The results of the GEE analysis show that there is statistically significant two-way Language*Vowel ($\chi^2(1)=12.69, p<0.001$) and Vowel*Syllable ($\chi^2(1)=31.078, p<0.001$) interaction effect, which means that the Language (Vowel) effect varies with different levels of Vowel (Language) and the Vowel (Syllable) effect varies with different levels of Syllable (Vowel), respectively. Thus, there is no need to interpret the main effects and we will directly perform the testing of simple effects, i.e. different levels of each variable broken down by the other variable. Regarding Vowel*Syllable interaction, /ɪ/ is significantly greater in CVC than CV (p=.001), while /i:/ is significantly greater in CV than CVC (p<0.001). For Syllable*Vowel interaction, /i:/ is significantly larger than /ɪ/ in CV syllables (p<0.001). Additionally, for Language*Vowel interaction, L1 English group has a greater mean of /i:/ than /ɪ/ (p<0.001), while for Vowel*Language interaction, in the case of /ɪ/, L1 Persian group has a larger mean than L1 English group. Table (8) illustrate the vowel length means of these two-way interactions. However, the results did not show any statistically significant three-way Language*Vowel*Syllable interaction effect($\chi^2(1)=3.681, p=.055$).[13]

*Table 8: Vowel length means of Language*Vowel and Vowel*Syllable interactions for /i:/ and /ɪ/ (milliseconds)*

	Persian L1	English L1	CV	CVC
/i:/	108.7	101.9	112.5	98.2
/ɪ/	101.4	70.4	79.0	92.9

One important thing here is that for the L1 English group there is a clearer length difference within each tense and lax counterpart pair than the L1 Persian group. This reveals the role of quantity properties in distinguishing between these vowels for English native speakers, especially in the case of CV syllables. Moreover, to clarify the role of quantity properties in each situation for the two groups, the duration ratio of tense and lax vowels (i.e. tense/lax) was calculated for both language groups regarding their production of these two vowel types. As Table (9) below shows, the Persian speakers had generally a lower ratio than the English group in both syllable types.

[13] Refer to Appendix D for the relevant SPSS output tables

Table 9: Duration Ratios of /i:/ and /ɪ/ vowels produced by English and Persian speakers

First Language	i:/ɪ Ratio	
	CVC	CV
Persian	0.9819	1.1694
English	1.1545	1.8414

5.2 High back vowels (/u:/ versus /ʊ /)

5.2.1 First formant (F1)

Figure (7) below depicts the raw mean z-scores of different conditions of the study regarding the first formant of tense/lax back vowels /u:/ and /ʊ/.

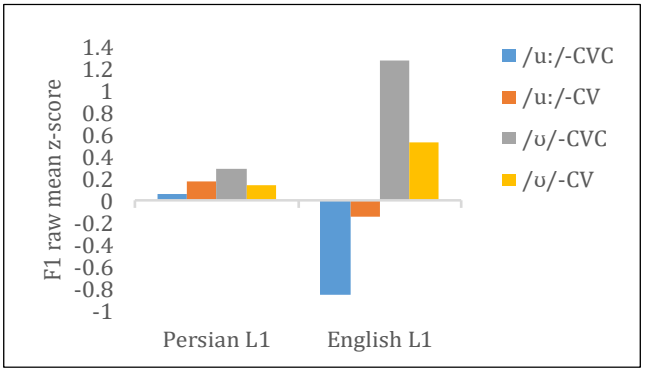


Figure 7: Raw mean z-scores of different conditions regarding F1 of /u:/ and /ʊ/

The results of the GEE analysis show that there is statistically significant two-way Language*Vowel ($\chi^2(1)=16.986, p<0.001$) and Vowel*Syllable ($\chi^2(1)=10.015, p=.002$) interaction effects, which means that the Language (Vowel) effect varies with different levels of Vowel (Language) and the Vowel (Syllable) effect varies with different levels of Syllable (Vowel), respectively. Thus, there is no need to interpret the main effects and we will directly perform the testing of simple effects, i.e. different levels of each variable broken down by the other variable. Regarding Vowel*Syllable interaction, /ʊ/ is significantly greater in CVC than CV (p=.006), while /u:/ is significantly greater in CV than CVC (p=.028). For Syllable*Vowel interaction, /ʊ/ is significantly larger than /u:/ in CVC syllables (p<0.001). Moreover, regarding Vowel*Language interaction, L1 English group has a significantly greater mean for /ʊ/ than the L1 Persian group (p<0.001), while the L1 Persian group has a significantly greater mean for /u:/ than L1 English group (p<0.001). Additionally, for Language*Vowel interaction, L1 English group has a greater mean of /ʊ/ than /u:/ (p<0.001). Table (10) illustrate the ZF1 means of these two-way interactions.

*Table 10: ZF1 means of Language*Vowel and Vowel*Syllable two-way interactions for /u:/ and /ʊ/*

	Persian L1	English L1	CV	CVC
/u:/	.1111	-.5027	.0021	-.3937
/ʊ/	.2172	.9061	.3466	.7767

Moreover, the results also show that there is a statistically significant three-way Language*Vowel*Syllable interaction effect($\chi^2(1)=5.192, p=.023$), which means that the syllable (vowel, language) effect varies with different levels of Vowel (language, syllable) and Language (Vowel, Syllable). Testing of simple effects is performed to look at each variable effect broken down by the two other variables. Table (11) illustrates the estimated marginal means calculated through the GEE model.

Table 11: Estimated Marginal Means calculated through the GEE model for F1 of /i:/ and /ɪ/

	Persian L1		English L1	
	CVC	CV	CVC	CV
/u:/	.0619	.1604	-.8492	-.1561
/ʊ/	.2835	.1508	1.2698	.5423

The Bonferroni pairwise comparisons revealed significance between the two language groups in the case of both tense vowel /u:/ (p=.001) and lax vowel /ʊ/ (p=.000) used in CVC syllables types (red arrows in figure 8). Moreover, the pairwise comparisons revealed significance between the two vowels in both syllable types CVC (p=.000) and CV (p=.004) within the L1 English group (orange arrows in figure 8). Additionally, the Bonferroni pairwise comparisons showed significance between the two syllable types in the case of both /u:/ (p=.000) and /ʊ/ (p=.011) within the L1 English group (green arrows in figure 8).[14]

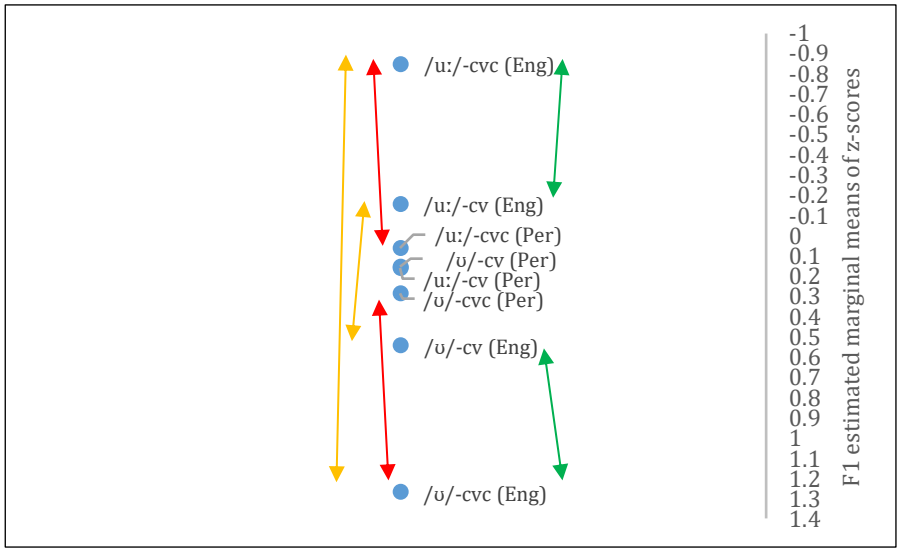


Figure 8: Differences between conditions regarding the first formants of /u:/ and /ʊ/

5.2.2 Second formant (F2)

Figure (9) below depicts the raw mean z-scores of different conditions of the study regarding the second formant of tense/lax back vowels /u:/ and /ʊ/.

[14] Refer to Appendix E for the relevant SPSS output tables.

The results of the GEE analysis show that there is statistically significant two-way Language*Vowel ($\chi^2(1)=12.889, p<0.001$) and Vowel*Syllable ($\chi^2(1)=57.785, p<0.001$) interaction effect, which means that the Language (Vowel) effect varies with different levels of Vowel (Language) and the Vowel (Syllable) effect varies with different levels of Syllable (Vowel), respectively. Thus, there is no need to interpret the main effects and we will directly perform

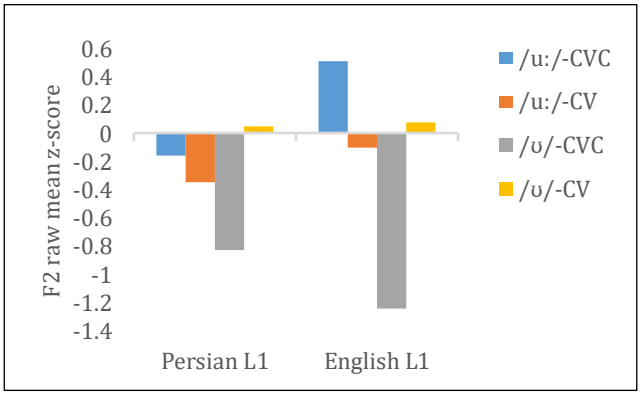


Figure 9: Raw mean z-scores of different conditions regarding F2 of /u:/ and /ʊ/

the testing of simple effects, i.e. different levels of each variable broken down by the other variable. Regarding Vowel*Syllable interaction, /ʊ/ is significantly greater in CV than CVC ($p<0.001$), while /u:/ is significantly greater in CVC than CV ($p<0.001$). Moreover, regarding Syllable*Vowel interaction, /ʊ/ is significantly larger than /u:/ in CV syllables ($p=.031$), while /u:/ is significantly larger than /ʊ/ in CVC syllables ($p<0.001$). For Language*Vowel interaction, L1 English group has a greater mean of /u:/ than /ʊ/ ($p<0.001$). Additionally, for Vowel*Language interaction, L1 English group has a significantly greater mean for /u:/ than the L1 Persian group ($p=.01$). Table (12) illustrate the ZF1 means of these two-way interactions.

*Table 12: ZF2 means of Language*Vowel and Vowel*Syllable two-way interactions for /u:/ and /ʊ/*

	Persian L1	English L1	CV	CVC
/u:/	-.2509	.2011	-.2223	.1725
/ʊ/	-.3829	-.5656	.0743	-1.0228

Moreover, the results also show that there is a statistically significant three-way Language*Vowel*Syllable interaction effect($\chi^2(1)=5.151, p=.023$), which means that the Syllable (Vowel, Language) effect varies with different levels of Vowel (Language, Syllable) and Language (Vowel, Syllable). Testing of simple effects is performed to look at each variable effect broken down by the two other variables. Table (13) illustrates the estimated marginal means calculated through the GEE model.

Table 13: Estimated Marginal Means calculated through the GEE model for F2 of /u:/ and /ʊ/

	Persian L1		English L1	
	CVC	CV	CVC	CV
/u:/	-.1620	-.3398	.5069	-.1047
/ʊ/	-.8173	.0514	-1.2284	.0972

The Bonferroni pairwise comparisons showed significance between the two L1 groups in the case of CVC syllables for /u:/ (p=.005) (red arrow in figure 10). Moreover, the pairwise comparisons showed significance between the two vowels in the case of CVC syllables within both English (p<0.001) and Persian groups (p=.001) (orange arrows in figure 10). Additionally, the Bonferroni pairwise comparisons showed significance between the two syllable types in the case of both /u:/ (p=.000) and /ʊ/ (p=.000) within the English group and for /ʊ/ (p=.002) within the Persian group (green arrows in figure 10).[15]

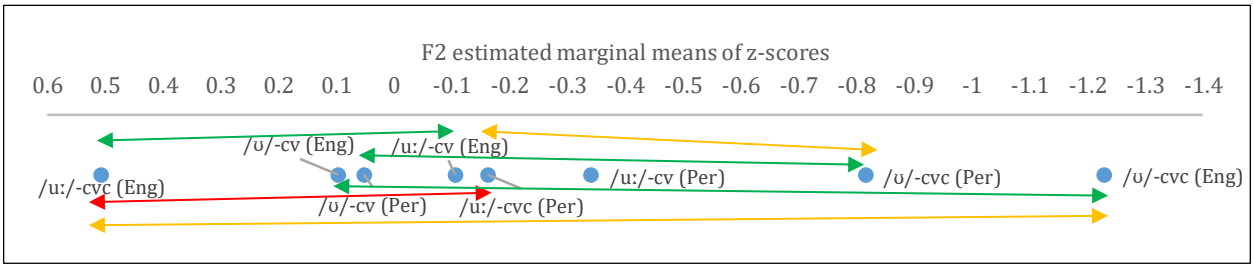


Figure 10: Differences between conditions regarding the second formants of /u:/ and /ʊ/

One thing grabbing attention here is the relative positions of tense and lax back vowels in CVC syllables for both L1 groups. Specifically, for the L1 English speakers the tense vowel is expected to be on the right side of the diagram and the lax on the left. One possible cause of such unexpected results can be the local anticipatory coarticulatory effects of the coda in the hosting syllables, e.g. vowel nasalisation before nasal consonants. The hosting words for the tense vowel are "rude" and "juice" and the words used for its lax counterpart are "full" and "cook". Thus, the coarticulation effects of the coronal coda consonants (/d/ & /s/) in the former words and dorsal coda consonants (/k/ & /l/) in the latter ones influence the positions of our target tense (more to the front, i.e. higher F2) and lax vowels (more to the back, i.e. lower F2), respectively.[16]

[15] Refer to Appendix F for the relevant SPSS output tables.

[16] It should be mentioned that this unexpected data does not affect the main goals of this study which is to investigate the strategies used by L1 English and L1 Persian speakers to differentiate English tense and lax counterparts. Also, the same pattern observed for both L1 groups provides a very interesting result which will be discussed in the discussions section. In other words, the focus of this study is on possible asymmetries between L1 Persian speakers and L1 English speakers in differentiating English tense/lax counterparts and possible effects of L1 on L2 production, so the non-standard locations of such vowels on the formant plot does not influence the findings.

5.2.3 /u:/ and /ʊ/ distances on the formant plot in different condition

As indicated in table (14), results of the Pythagorean Theorem considering means of F1 and F2 in each case determined that the distance between /u:/ and /ʊ/ was significantly shorter when produced by L1 Persian speakers compared to L1 English speakers, especially in the case CVC syllables (.69 vs. 2.73).

Table 14: Z-score Means of each formant of /u:/ and /ʊ/ and their distances for each L1 Group

Syllable Type	L1 Group	Mean $F1_{/u:/}$	Mean $F1_{/ʊ/}$	Mean $F2_{/u:/}$	Mean $F2_{/ʊ/}$	Vertical distance	Horizontal distance	Distance
CVC	Persian	.0619	.2835	-.1620	-.8173	-0.2216	-.0118	.6918
	English	-.8492	1.2698	.5069	-1.2284	-2.119	1.7353	2.7389
CV	Persian	.1604	.1508	-.3398	.0514	0.0096	-0.3912	.3913
	English	-.1561	.5423	-.1047	.0972	-0.6984	-0.2019	.7270

5.2.4 Vowel length

Figure (11) below depicts the raw mean vowel length (in milliseconds) of different conditions of the study regarding the tense/lax back vowels /u:/ and /ʊ/.

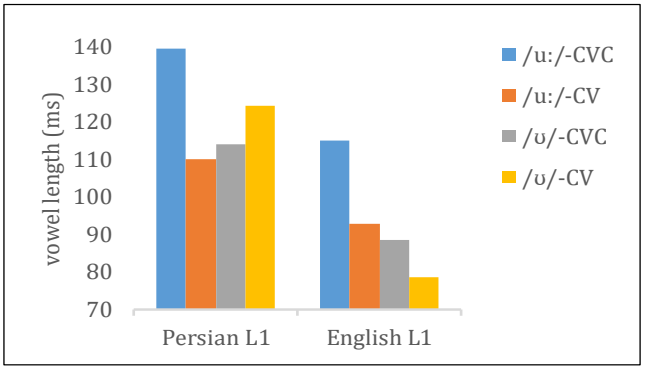


Figure 11: raw mean vowel lengths of different conditions of vowels /u:/ and /ʊ/ (milliseconds)

The results of the GEE analysis show that there is statistically significant two-way Language*Vowel ($\chi^2(1)=7.251, p=.007$) and Vowel*Syllable ($\chi^2(1)=14.934, p<0.001$) interaction effect, which means that the Language (Vowel) effect varies with different levels of Vowel (Language) and the Vowel (Syllable) effect varies with different levels of Syllable (Vowel), respectively. Thus, there is no need to interpret the main effects and we will directly perform the testing of simple effects, i.e. different levels of each variable broken down by the other variable. Regarding Vowel*Language interaction, L1 Persian group has a significantly greater mean than the L1 English group for both /ʊ/ ($p<0.001$) and /u:/ ($p=.015$). For Language*Vowel interaction, L1 English group has a greater mean of

/u:/ than /ʊ/ (p<0.001). For Syllable*Vowel interaction, /u:/ is significantly larger than /ʊ/ in CVC syllables (p<0.001). Regarding Vowel*Syllable interaction, /u:/ is significantly greater in CVC than CV (p<0.001). Table (15) illustrate the vowel length means of these two-way interactions.

*Table 15: Vowel length means of Language*Vowel and Vowel*Syllable interactions for /u:/ and /ʊ/ (milliseconds)*

	Persian L1	English L1	CV	CVC
/u:/	124.3	104.8	102.2	126.9
/ʊ/	119.1	83.8	99.8	103.1

Additionally, the results also show that there is a statistically significant three-way Language*Vowel*Syllable interaction effect($\chi^2(1)=7.815$,p=.005), which means that the Syllable (Vowel, Language) effect varies with different levels of Vowel (Language, Syllable) and Language (Vowel, Syllable). Testing of simple effects is performed to look at each variable effect broken down by the two other variables. Table (16) illustrates the estimated marginal means calculated through the GEE model.

Table 16: Estimated Marginal Means calculated through the GEE model for vowel lengths of /u:/ and /ʊ/

	Persian L1		English L1	
	CVC	CV	CVC	CV
/u:/	139.0	109.6	114.7	94.9
/ʊ/	115.4	122.8	90.8	76.9

The Bonferroni pairwise comparisons revealed significance between the two language groups in the case of lax vowel /ʊ/ for both CVC (p=.008) and CV (p=.000) syllable types. Moreover, the pairwise comparisons revealed significance between the two vowels within the L1 English group for both CVC (p<0.001) and CV (p=.003) syllable types, and also within the L1 Persian group for CVC syllable types (p=.001). Additionally, the Bonferroni pairwise comparisons showed significance between the two syllable types in the case of /u:/ within both L1 English (p=.006) and L1 Persian (.009) groups.[17]

However, the important thing here is that the L1 English group generally produce these two vowel types shorter than the Persian group, although both groups showed a clear influence of the syllable type on the vowel length. Also, to clarify the role of quantity properties in each situation for the two groups, the duration ratio of tense and lax vowels (i.e. tense/lax) was calculated for both Language groups regarding their production of these two vowel types. As Table (17) below shows, the Persian speakers had the ratio of less than 1 in the case of CV syllable types.

[17] Refer to Appendix G for the relevant SPSS output tables.

Table 17: Duration Ratios of /u:/ and /ʊ/ vowels produced by English and Persian Speakers

First Language	u:/ ʊ Ratio	
	CVC	CV
Persian	1.2045	0.8925
English	1.2632	1.2341

6 Discussion

In this study, the vowel productions of L1 Persian learners of English were compared and contrasted with native English speakers to determine their difficulties acquiring the tense/lax distinctions in English high vowels. The main aim of the study was to investigate the ability of L2 learners in creating separate categories for tense and lax vowels of the target language. The results of this study showed that the L1 English speakers use both quality (F1 & F2) and quantity (vowel length) to differentiate between the English tense/lax counterparts in their native language. These distinctions, although clearly causing difficulties, were also observed to some extent in the production data of the L1 Persian speakers of English (Figures 12 & 13).

The results of the study are somewhat in line with the Flege's (1987) speech learning model (Equivalence Classification hypothesis) which predicts difficulty in the acquisition of similar L2 sounds. L1 Persian group were not observed to make clear (i.e. statistically significant) quality (F1) distinctions between the English tense/lax counterparts in general. However, although quality does not seem to play a statistically significant role in distinguishing the tense/lax counterparts for L1 Persian group, the formant plots showed that at least in some cases they differentiate between these English vowel counterparts in terms of their qualities. As depicted in Figures (11) and (12) above, the major difference between the Persian and English speakers seems to be the relative distance between these counterpart vowels. As observed in these figures, the Persian speakers' vowels are located

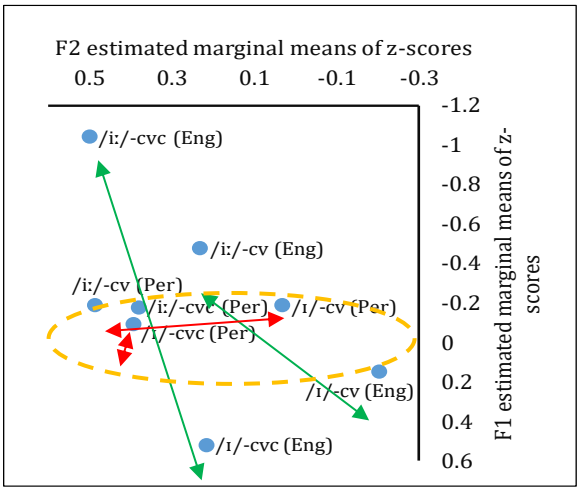


Figure 12: Differences in the distance between /i:/ and /ɪ/ for L1 English vs. L1 Persian speakers

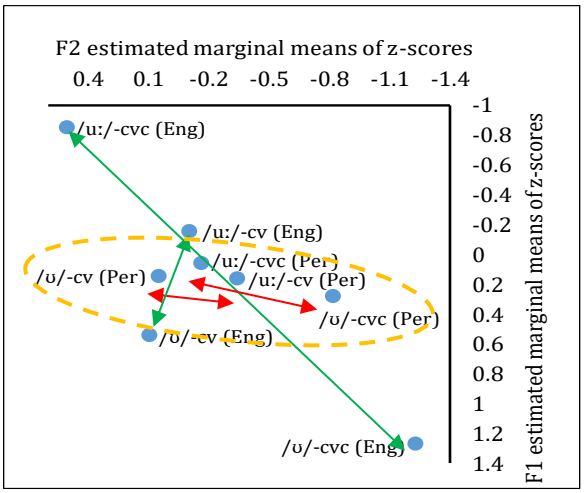


Figure 13: Differences in the distance between /u:/ and /ʊ/ for L1 English vs. L1 Persian speakers

in an area somehow inside the whole realm of the English ones. In other words, although the L1 Persian learners have difficulties in acquiring tense/lax contrasts, phonologically speaking they have acquired this abstract contrast in the target language. However, there are still some phonetic properties which differentiate them from the native speakers of English regarding the formants and duration of tense/lax vowels. The formant plot results are also, to some extent, in line with Munro's (1993) study on Arabic accent of English which indicated spectral differences between Arabic learners' accent of English and the native English speakers. However, the results did not support his findings on the exaggerated tense/lax length differences by Arabic L2 speakers of English.

Considering all the aspects of quality and quantity, no clear evidence was observed in support of the Markedness Differential Hypothesis (Eckman, 1977) predicting the lax vowels (i.e. the marked form) be more difficult to acquire than the tense vowels. In fact, for both front and back vowels the tense vowels (compared to their lax counterparts) were more deviant from the native speakers' data. In the same line, there was also no clear evidence for back vowels (as marked forms) causing more difficulties to the L2 learners compared to the front vowels. Additionally, in terms of quantity, although the tense/lax length ratio showed that the Persian learners had a more native-like performance in the case of CVC than CV syllables for both vowel pairs, there was no clear evidence in support of the Markedness Differential Hypothesis. In fact, the results of tense/lax length ratios showed that at least in the case of relative lengths of the counterpart vowels, the Persian speakers had a more native-like performance for back vowels than the front ones, especially in CVC syllables.

In terms of the role of syllable type (#CVC# vs. #CVσ) on the quality/quantity distinction between and within the tense and lax counterpart vowels, it was observed that both L1 groups, especially L1 English speakers, differentiate vowels in these environments. However, no clear L1 effect regarding such syllable type effect was observed for L1 Persian speakers. In other words, they do not treat L2 vowels quite the same way they treat their similar L1 vowels, especially in the case of English lax vowels. As pointed out earlier Persian short vowels are "unstable" regarding their length depending on their environment. They are long in closed and word final syllables and short in other environments (Toosarvandani, 2004). Therefore, if L1 Persian speakers were considering English lax vowels in the same way as Persian short vowels, they were expected to follow a similar pattern, which was not observed.

7 Conclusion

This study investigated different features (quality & quantity) used by L1 English and L1 Persian speakers to distinguish the front (/i:/ & /ɪ/) and back (/u:/ & /ʊ/) tense/lax counterparts in English. It was observed that both quality (F1 & F2) and quantity (length) are used by English speakers. In the case of Persian speakers, it was observed that they are clearly facing difficulties in acquiring these distinctions in their L2 English both in terms of quality and quantity. However, this difficulty mostly seems to be the acquisition of phonetic features of L2 than the phonological features. In other words, it seems that they are able to distinguish these vowel counterparts, but as showed in the formant plots, their

distinguishing features are not as distinct as they are for English speakers. Also, the syllable type (#CVC# vs. #CVσ) was also observed to cause quality and quantity distinctions for both L1 groups, especially for English speakers, but no clear L1 effect was observed for the L1 Persian speakers regarding this element.

References

Alzahrani, D. S. M. (2014). *The acquisition of tense/lax distinction by Arabic speakers learning English as a second language*. United States: Ann Arbor.

Carr, P., & Honeybone, P. (2007). English phonology and linguistic theory: an introduction to issues, and to 'Issues in English Phonology'. *Language Sciences*, *29*(2), 117-153.

Chomsky, N. & Halle. M. (1968). *The Sound Pattern of English*. New York: Harper & Row.

Durand, P. (2005). Tense/lax, the vowel system of English and phonological theory. In P. Carr, J. Durand and C. Ewen (eds.) *Headhood, Elements, Specification and Contrastivity*. 77-98. Amsterdam: John Benjamins.

Eckman, F. R. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning*, *27*(2), 315-330.

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, *15*(1), 47-65.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (ed.) *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, 233-277. Timonium: York Press.

Flege, J. E., & Port, R. (1981). Cross-language phonetic interference: Arabic to English. *Language and Speech*, *24*(2), 125-146.

Flege, J. E., Bohn, O. S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, *25*(4), 437-470.

Gordon, M. (2004). Syllable weight. In B. Hayes, R. Kirchner, and D. Steriade (eds.), *Phonetic Bases for Phonological Markedness*, 277-312. Cambridge: Cambridge University Press.

Halle, M. (1977). Tenseness, vowel shift and the phonology of back vowels in modern English. *Linguistic Inquiry*, 8, 611-625.

Hayes, B. (1979). The rhythmic structure of Persian verse. *Edebiyat*, 4, 193-242.

Jakobson, R., Fant, C. G. M., & Halle, M. (1952). Preliminaries to speech analysis. *Technical Report (No. 13).* Cambridge: Acoustics Laboratory, M.I.T.

Jakobson, R., & Waugh, L. R. (1979). *The sound shape of language*. Brighton: Harvester Press.

Kondaurova, M. V., & Francis, A. L. (2010). The role of selective attention in the acquisition of English tense and lax vowels by native Spanish listeners: Comparison of three training methods. *Journal of Phonetics*, *38*(4), 569-587.

Kwon, Y. K. (2011). Doing without [tense/lax] in English. *Linguistic Research*, *28*(3), 605-624.

Lass, R. (1976). *English phonology and phonological theory: synchronic and diachronic studies*. Cambridge: Cambridge University Press.

Lee, J.Y. (2003). Tense and length in English phonology. *English Language and Literature,* 49, 1093–1144.

Maddieson, I., & Disner, S. F. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.

Mahootian, S., & Gebhardt, L. (1997). *Persian: Descriptive grammars*. London: Routledge

McAllister, R., Flege, J. E., & Piske, T. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics*, *30*(2), 229-258.

Munro, M. J. (1993). Productions of English vowels by native speakers of Arabic: Acoustic measurements and accentedness ratings. *Language and Speech*, *36*(1), 39-66.

Munro, M. J., & Derwing, T. M. (2008). Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning*, *58*(3), 479-502.

Odden, D. (2011). The representation of vowel length. In M. van Oostendorp, C. Ewen, E. Hume, and K. Rice (eds.), *Companion to Phonology*, *1*, 465-490. Wiley-Blackwell.

Odlin, T. (2003). Cross-linguistic influence. In C. J. Doughty & M. H. Long (eds.), *Handbook of Second Language Acquisition*, 436-86. Oxford: Blackwell.

Oostendorp, M. V. (1995). Vowel quality and syllable projection. Dissertation. *Catholic University of Brabant, Tilburg*.

Perlmutter, D. (1995). Phonological quantity and multiple association. *The Handbook of Phonological Theory. Cambridge & Oxford: Blackwell*,*307317*.

Pisowicz, A. (1985). *Origins of the New and Middle Persian phonological systems* (No. 101). Nakł. Uniwersytetu Jagiellońskiego.

Rohany-Rahbar, E. (2009). On contrasts in the Persian vowel system. *Toronto Working Papers in Linguistics*, *31*..

Samareh, Y. (1977). *The arrangement of segmental phonemes in Persian*. Tehran: University of Tehran Press.

Strange, W., & Shafer, V. L. (2008). Speech perception in second language learners: The re-education of selective perception. *Phonology and second language acquisition*, 153-191.

Toosarvandani, M. D. (2004). Vowel length in modern Farsi. *Journal of the Royal Asiatic Society*, *14*(03), 241-251.

Tranel, B. (1995). On phonetic evidence for the phonological mora: comments on Hubbard. *In B. Connell and A. Arvanti (eds.), Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV, 188-201. Cambridge: Cambridge University Press*.

Trubetzkoy N. (1939). Grundzüge der phonology. In C. Baltaxe (Ed.), Travaux du Cercle Linguistique de Prague, 7, University of California Press, Berkeley.

Wang, X., & Munro, M. J. (2004). Computer-based training for learning English vowel contrasts. *System*, *32*(4), 539-552.

Windfuhr, G. (1979). *Persian grammar: History and state of its study* (Vol. 12). Mouton: The Hague.

Windfuhr, G., & Perry, J. R. (2010). Persian and Tajik. In G. Windfuhr (ed.), *The Iranian Languages.* 416-544. London: Routledge.

Appendix A: List of test sentences

	Sentence	Target vowel	Target syllable	IPA pronunciation
1	He was very **r<u>u</u>de** at the party.	/uː/	CVC	/ruːd/
2	He drank some **j<u>ui</u>ce** before lunch.	/uː/	CVC	/dʒuːs/
3	They use the **l<u>u</u>nar** calendar.	/uː/	CV	/ˈluː·nər/
4	He likes **n<u>oo</u>dle** soup.	/uː/	CV	/ˈnuː·dəl/
5	The tank is **f<u>u</u>ll** of gas.	/ʊ/	CVC	/fʊl/
6	He is a good **c<u>oo</u>k** in Calgary.	/ʊ/	CVC	/kʊk/
7	He put some **s<u>u</u>gar** in his coffee.	/ʊ/	CV	/ˈʃʊ.gər/
8	They hired a **b<u>u</u>tcher** to prepare the meat.	/ʊ/	CV	/ˈbʊ.tʃər/
9	You can have a **s<u>ea</u>t** over there.	/iː/	CVC	/sit/
10	I found a **ch<u>ea</u>p** house.	/iː/	CVC	/tʃip/
11	I had a **f<u>e</u>ver** last night.	/iː/	CV	/ˈfiː·vər/
12	He was a **t<u>ea</u>cher** in Canada.	/iː/	CV	/ˈtiː·tʃər/
13	They moved to a **b<u>i</u>g** city.	/ɪ/	CVC	/bɪg/
14	He is from a **r<u>i</u>ch** family.	/ɪ/	CVC	/rɪtʃ/
15	He grew up in a small **c<u>i</u>ty** near Paris.	/ɪ/	CV	/ˈsɪ.tiː/
16	They found a **w<u>i</u>zard** in the town.	/ɪ/	CV	/ˈwɪ.zərd/

Appendix B: ZF1 /i:/ and /ɪ/ statistics

Tests of Model Effects

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	32.198	1	.000
Language	.592	1	.442
Vowel	22.689	1	.000
Syllable	.019	1	.891
Language * Vowel	19.581	1	.000
Language * Syllable	.243	1	.622
Vowel * Syllable	5.850	1	.016
Language * Vowel * Syllable	4.118	1	.042

Dependent Variable: ZF1
Model: (Intercept), Language, Vowel, Syllable, Language * Vowel, Language * Syllable, Vowel * Syllable, Language * Vowel * Syllable

Two-way interactions:

		CV vs CVC	/ɪ/ vs /i:/	Persian vs. English
Vowel*Syllable	/ɪ/	.229		
	/i:/	.109		
Syllable*Vowel	CV		.068	
	CVC		.0	
Language*Vowel	Persian		.736	
	English		.0	
Vowel*Language	/ɪ/			.0
	/i:/			.0

Three-way interactions:

Syllable*Language*Vowel			
CV vs. CVC	Persian	/ɪ/	1
		/i:/	1
	English	/ɪ/	.382
		/i:/	.0

Vowel*Language*Syllable			
/ɪ/ vs. /i:/	Persian	CV	1
		CVC	1
	English	CV	.055
		CVC	.001

Language*Vowel*Syllable			
Persian vs. English	/ɪ/	CV	.866
		CVC	.056
	/i:/	CV	1
		CVC	.0

Appendix C: ZF2 /i:/ and /ɪ/ statistics

Tests of Model Effects

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	12.436	1	.000
Language	.912	1	.340
Vowel	3.304	1	.069
Syllable	1.591	1	.207
Language * Vowel	.188	1	.664
Language * Syllable	.340	1	.560
Vowel * Syllable	1.539	1	.215
Language * Vowel * Syllable	.408	1	.523

Dependent Variable: ZF2
Model: (Intercept), Language, Vowel, Syllable, Language * Vowel, Language * Syllable, Vowel * Syllable, Language * Vowel * Syllable

Appendix D: Vowel length /i:/ and /ɪ/ statistics

Tests of Model Effects

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	626.645	1	.000
Language	6.150	1	.013
Vowel	32.766	1	.000
Syllable	.005	1	.942
Language * Vowel	12.690	1	.000
Language * Syllable	.484	1	.487
Vowel * Syllable	31.078	1	.000
Language * Vowel * Syllable	3.681	1	.055

Dependent Variable: Duration
Model: (Intercept), Language, Vowel, Syllable, Language * Vowel, Language * Syllable, Vowel * Syllable, Language * Vowel * Syllable

Two-way interactions:

		CV vs CVC	/ɪ/ vs /i:/	Persian vs. English
Vowel*Syllable	/ɪ/	.001		
	/i:/	.0		
Syllable*Vowel	CV		.0	
	CVC		.217	
Language*Vowel	Persian		.063	
	English		.0	
Vowel*Language	/ɪ/			.0
	/i:/			.498

Appendix E: ZF1 /u:/ and /ʊ/ statistics

Tests of Model Effects

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	34.031	1	.000
Language	.359	1	.549
Vowel	22.969	1	.000
Syllable	.025	1	.873
Language * Vowel	16.986	1	.000
Language * Syllable	.000	1	1.000
Vowel * Syllable	10.015	1	.002
Language * Vowel * Syllable	5.192	1	.023

Dependent Variable: ZF1
Model: (Intercept), Language, Vowel, Syllable, Language * Vowel, Language * Syllable, Vowel * Syllable, Language * Vowel * Syllable

Two-way interactions

		CV vs CVC	/ʊ/ vs /u:/	Persian vs. English
Vowel*Syllable	/ʊ/	.006		
	/u:/	.028		
Syllable*Vowel	CV		.130	
	CVC		.0	
Language*Vowel	Persian		.711	
	English		.0	
Vowel*Language	/ʊ/			.0
	/u:/			.0

Three-way interactions:

Syllable*Language*Vowel			
CV vs. CVC	Persian	/ʊ/	1
		/u:/	1
	English	/ʊ/	.011
		/u:/	.0

Vowel*Language*Syllable			
/ʊ/ vs. /u:/	Persian	CV	1
		CVC	1
	English	CV	.004
		CVC	.0

Language*Vowel*Syllable			
Persian vs. English	/ʊ/	CV	.464
		CVC	.0
	/u:/	CV	1
		CVC	.001

Appendix F: ZF2 /u:/ and /ʊ/ statistics

Tests of Model Effects

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	11.866	1	.001
Language	.864	1	.353
Vowel	25.848	1	.000
Syllable	18.453	1	.000
Language * Vowel	12.889	1	.000
Language * Syllable	.005	1	.944
Vowel * Syllable	57.785	1	.000
Language * Vowel * Syllable	5.151	1	.023

Dependent Variable: ZF2
Model: (Intercept), Language, Vowel, Syllable, Language * Vowel, Language * Syllable, Vowel * Syllable, Language * Vowel * Syllable

Two-way interactions:

		CV vs CVC	/ʊ/ vs /u:/	Persian vs. English
Vowel*Syllable	/ʊ/	.0		
	/u:/	.0		
Syllable*Vowel	CV		.031	
	CVC		.0	
Language*Vowel	Persian		.331	
	English		.0	
Vowel*Language	/ʊ/			.266
	/u:/			.010

Three-way interactions:

Syllable*Language*Vowel			
CV vs. CVC	Persian	/ʊ/	.002
		/u:/	1
	English	/ʊ/	.0
		/u:/	.0

Vowel*Language*Syllable			
/ʊ/ vs. /u:/	Persian	CV	.454
		CVC	.001
	English	CV	1
		CVC	.0

Language*Vowel*Syllable			
Persian vs. English	/ʊ/	CV	1
		CVC	.124
	/u:/	CV	1
		CVC	.005

Appendix G: Vowel length /u:/ and /ʊ/ statistics

Tests of Model Effects

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	887.802	1	.000
Language	14.246	1	.000
Vowel	19.959	1	.000
Syllable	10.742	1	.001
Language * Vowel	7.251	1	.007
Language * Syllable	.478	1	.490
Vowel * Syllable	14.934	1	.000
Language * Vowel * Syllable	7.815	1	.005

Dependent Variable: Duration
Model: (Intercept), Language, Vowel, Syllable, Language * Vowel, Language * Syllable, Vowel * Syllable, Language * Vowel * Syllable

Two-way interactions:

		CV vs CVC	/ʊ/ vs /u:/	Persian vs. English
Vowel*Syllable	/ʊ/	.479		
	/u:/	.0		
Syllable*Vowel	CV		.563	
	CVC		.0	
Language*Vowel	Persian		.316	
	English		.0	
Vowel*Language	/ʊ/			.0
	/u:/			.015

Three-way interactions:

Syllable*Language*Vowel			
CV vs. CVC	Persian	/ʊ/	1
		/u:/	.009
	English	/ʊ/	.238
		/u:/	.006

Vowel*Language*Syllable			
/ʊ/ vs. /u:/	Persian	CV	.234
		CVC	.001
	English	CV	.003
		CVC	.0

Language*Vowel*Syllable			
Persian vs. English	/ʊ/	CV	0
		CVC	.008
	/u:/	CV	.419
		CVC	1

Contact Information

Elias Abdollahnejad

elias.abdollahnejad@ucalgary.ca
School of Languages, Linguistics, Literatures, and Cultures, University of Calgary
CHC 211, 2500 University Drive NW
Calgary, Alberta, T2N 1N4
Canada

N-Behavior in Quranic Reading

Ahmed Alnuqaydan

University of Utah, Department of Linguistics

Abstract

*Idgham is a Quranic-reading rule that governs how the coronal /n/ is pronounced when it is followed by a set of segments: /j, w, r, l, n and m/. According to Quranic scholars, when /n/ is followed by a glide /famə**n** **j**aʔməl/ or a nasal /mɪ**n** **m**al/, it deletes and the [+nasal] feature moves to the following segment yielding [famə **j̃**aʔməl] and [mɪ **m̃**al], respectively. On the other hand, when /n/ precedes a liquid, both the /n/ and the [+nasal] feature are phonetically unrealized: /mɪ**n** **l**adunh/ →[mɪ **l**adunh]. Idgham only applies when /n/ occurs word-finally and the triggering segments occupy the initial onset position of the following word. It does not occur word-medially: [qɪ**nw**an].The present paper provides a unified OT account for the phenomenon illustrated above. Since, in most cases, the [+nasal] feature sticks around, I argue that Idgham is a fusion process not a deletion process. This paper also explores the vulnerability of /n/ and the immunity of /m/ to Idgham: /la**m** nara/ "we did not see" → [la**m** nara] not *[la ñ$_{1,2}$ara]. I argue that in Quranic reading, the more marked /m/ is exempt from fusion while the less marked /n/ is not because IDENT constraints for the more marked segment /m/ outrank IDENT constraints for the less marked segment /n/ (De Lacy, 2002). Finally, in answering why /n/ only fuses with sonorants, the reason is attributed to faithfulness.*

Key words: Idgham, fusion, coronal nasal, Quranic reading

1 Introduction*

Given its idiosyncratic interaction with some segments in particular environments, the phonological behavior of the coronal nasal /n/ in different languages has been given an ample attention in the literature (Halle & Clements, 1983; Herrick, 1999; Lombardi, 1998; Pater, 1999). In the religion of Islam, the holy Quran (the sacred book of Muslims) should be read according to a set of rules, and mistakes in recitation are forbidden (Muhammad, R, Muhammad, A & Martinez-Enriquez, 2010). The way Quran should be read is governed by a set of rules referred to as "Tajweed"[1]. One of those rules is *Idgham* which is a rule that governs how the coronal nasal /n/ is pronounced when it is followed by a set of segments: /*j*, *w*, *r*, *l*, *n* and *m*/.

According to Qamawee (1985) and Nassr (1994), when /n/ is followed by a glide or a nasal, it deletes and the [+nasal] feature moves to the following segment. On the other hand, when /n/ precedes a liquid, both the segment and the feature are phonetically unrealized. *Idgham* only applies when /n/ occurs word-finally and the triggering segments occupy the initial onset position of the following word. The co-occurrence of /n/ with one of the triggers word-medially does not trigger *Idgham*. Finally, *Idgham* does not occur when names of Quranic Surahs[2] participate in providing an *Idgham* environment.

Idgham is prescriptive and not a natural part of speakers' native grammars. We can view it as akin to a language game (Gotowski, 2019). This makes the present analysis similar to formal analyses of other language games. The main purpose of the present paper is to provide a unified OT account for the phenomenon illustrated above with all its different facets. I argue that *Idgham* is a coalescence process derived by an interaction between well-attested markedness constraints and faithfulness constraints from Correspondence Theory (McCarthy and Prince, 1995). In particular, it is a fusion process driven by an interaction between a constraint against a specific sequence of consonants and faithfulness constraints for particular positions and particular features. This paper also tackles and develops an OT account for the vulnerability of /n/ and the immunity of /m/ to *Idgham*. The coronal nasal undergoes *Idgham* when it is followed by one of the triggers. *Idgham*, on the other hand does not affect /m/ when it is followed by the same triggers. The immunity of /m/, I argue, is due to a highly ranked faithfulness constraint that holds only for labials. The paper is concluded by throwing some light on why /n/ only coalesces with sonorants.

This paper is organized as follows: section 2 provides a unified OT account for *Idgham*. It is divided into five subsections: 2.1 deals with examples in which /n/ is followed by a glide, 2.2 tackles the tolerance of /n/ followed by one of the *Idgham* triggering segments word-medially, 2.3 explains what happens when /n/ is followed by a nasal, 2.4 is devoted for the co-occurrence of /n/ with a liquid, 2.5 demonstrates the blocking of *Idgham* in the names of Quranic Surahs where the environment is *Idgham*-motivated. Section 3 dwells on the nature of segments involved in *Idgham*. It is divided into two subsections: 3.1 explores the

* The paper was presented at the *Northwest Phon{etics;ology} Conference* held on September 20th, 2019, at the University of Calgary.

[1] Tajweed is an Arabic word that means proper pronunciation during recitation (Ibrahim, Razak, Yusoff, Idris, Tamil, Noor, & Naemah, 2008).

[2] Surah is the term used for a chapter in the Quran.

vulnerability of /n/ and the immunity of /m/ to *Idgham* while 3.2 touches on sonorants as *Idgham* triggering segments. Section 4 concludes.

2 A unified OT account for *idgham*

2.1 /n/ followed by a glide

In Quranic reading, when /n/ is followed by a glide, it deletes and the [+nasal] feature moves to the following glide[3] (Qamawee, 1985 and Nassr, 1994).

1) a. **nj**

/wa barqə**n j**ad͡ʒʔalun/	[wa barqə **j̃**ad͡ʒʔalun]	"and thunder they make"
/famə**n j**aʔməl/	[famə **j̃**aʔməl]	"and he who works"
/mə**n j**aqul/	[mə **j̃**aqul]	"who says"
/wa mə**n j**utəʔ allah/	[wa mə **j̃**utəʔ allah]	"and he who obeys Allah"
/wud͡ʒuhu**n j**awmaʔðın/	[wud͡ʒuhu **j̃**awmaʔðın]	"faces on that day"
/jawmaʔðı**n j**ataðakər/	[jawmaʔðı **j̃**ataðakər]	"on that day he/ she will remember"

b. **nw**

/mı**n w**alli/	[mı **w̃**alli]	"apart from God you have no guardian"
/jawmaʔðı**n w**ahijah/	[jawmaʔðı **w̃**ahijah]	"on that Day it will be frail"
/mı**n w**al/	[mı **w̃**al]	"apart from God you have no guardian"
/wa walıdə**n w**a ma walad/	[wa walıdə **w̃**a ma walad]	"and by a father and what he fathered"
/mı**n w**aq/	[mı **w̃**aq]	"no defender against God"

I argue that the phonological process that /n/ undergoes in the abovementioned examples is not deletion, but rather coalescence. This is supported by the fact that the [+nasal] feature appears on the glide indicating that /n/ is not deleted but fused with the glide. This argument, as will be discussed in detail in the subsequent sections, holds whenever /n/ is followed by any of the 6 *Idgham* triggering segments */j*, *w*, *r*, *l*, *n* and *m*/. The difference lies in the preservation or loss of the coronal /n/'s [+nasal] feature depending on the nature of the following segment[4].

[3] (All data translations are cited from https://www.clearquran.com/)

[4] It is possible to come up with an account that treats the behavior of /n/ when it is followed by a liquid differently from when it is followed by the rest of the triggers. However, in this paper, it is argued that the loss

Prior to exploring what motivates *Idgham*, let's touch on the nature of the triggering segments. Scrutinizing the nature of the triggering segments */j*, *w*, *r*, *l*, *n* and *m*/, it is clear that all of them are sonorants. It seems that in Quranic reading, the sequence (n + sonorant) is not allowed. The ban on [nS] (where S = sonorant) sequence is resolved by fusion. When a word underlyingly contains a sequence of /nS/, these two segments coalesce surfacing as one output segment standing in correspondence to the two input segments /nS/.

In Optimality Theoretic terminology, fusion occurs when the anti-fusion faithfulness constraint UNIFORMITY[5] is dominated by the markedness constraint *NS and the other faithfulness constraints, namely, MAX-IO (C) and DEP-IO (V). The definitions of the constraints are given below.

2) UNIFORMITY: Nothing in the output can have more than one correspondent in the input; penalizes coalescence.

3) *NS: The sequence nasal + sonorant must not be allowed; penalizes consecutive [nasalSonorant][6].

4) MAX-IO (C): Every consonant in the input must have a correspondent in the output; penalizes consonant deletion

5) DEP-IO (V): Every vowel in the input must have a correspondent in the output; penalizes vowel deletion.

Tableau 1: Motivating Idgham

famən$_1$ j$_2$aʔməl	MAX-IO (C)	DEP-IO (V)	*NS	UNIFORMITY
a.→famə j̃$_{1,2}$aʔməl				*
b. famən$_1$ j$_2$aʔməl			*!	
c. famənə jaʔməl		*!		
d. famə jaʔməl	*!			

The most faithful candidate (b) is ruled out by incurring a fatal violation of *NS. Candidates (c) and (d) are ruled out by violating the highly-ranked faithfulness constraints DEP-IO (V) and MAX-IO (C), respectively. The winner (a) satisfies the undominated constraints by fusion of the unpermitted sequence; the cost, though, is a violation of UNIFORMITY which is ranked low in the hierarchy.

---

of /n/'s [+nasal] feature when it is followed by a liquid stems from a lower ranking of the constraint responsible for the preservation of the feature as will be seen in 2.4.

[5] Unless stated otherwise, the constraints used in this paper are from Correspondence Theory (McCarthy and Prince, 1995). When the source of a constraint is not stated, it means that it is developed specifically for purpose of accounting for the present data.

[6] The examples show that the sequence (n/+ sonorant) is banned which is resolved by fusion. However, it is worth noting that this constraint bans the sequence nasal + sonorant. Broadening the constraint this way will come in handy later when we deal with the immunity of the bilabial nasal /m/.

There are still two more important suboptimal forms that need to be incorporated in our analysis: *[famənʲ$_{1,2}$ aʔməl] and *[famə j$_{1,2}$aʔməl]. In *[famənʲ$_{1,2}$ aʔməl], the coalescence goes the other way around resulting in a palatalized [nʲ] whereas in *[famə j$_{1,2}$aʔməl], the two input segments /n j/ coalesce yielding a nonnasalized [j$_{1,2}$] which indicates that the [+nasal] feature is deleted. No constraint in the analysis developed thus far can rule out these two suboptimal forms.

Dealing with *[famənʲ$_{1,2}$ aʔməl] where the resultant of the fusion process is a palatalized [nʲ], it is noticeable that both the surface form [famə j̃$_{1,2}$aʔməl] and the candidate with a palatalized [nʲ] only violate UNIFORMITY out of the constraints developed so far. This means that there is a tie between these two candidates. To break the tie, we need a constraint that is violated by *[famənʲaʔməl], but not by the winner.

In all of the examples mentioned above, the 6 triggering segments occupy a position where they are followed by a vowel which is not the case for /n/. This environment can be used to the winner's advantage. In [famə j̃$_{1,2}$aʔməl], the glide precedes a vowel and all of the features of the glide are preserved plus an addition of the feature [+nasal]. On the other hand, in *[famənʲ$_{1,2}$ aʔməl], all of the features of the glide are lost except for [dorsal] while all the features of the /n/ are preserved with the addition of [dorsal]. What we need here is an output segment that is faithful to the input segment preceding the vowel. In other words, when there is a conflict between two consonants, the features of the one preceding the vowel must be preserved.

In order to do that, we need to go through the feature mismatches between the glide and the /n/ and turn them into IDENT constraints. These IDENT constraints are going to be bundled in a cover constraint: Faith-C/ _V (cf. Becker, 1999). The same set of IDENT constraints are going to be bundled in another cover constraint: Faith-C. The latter is more general as it is not restricted to a specific environment. Faith-C/ _V is violated when one of the IDENT constraints is violated by the consonant preceding the vowel. When Faith-C/ _V is violated, Faith-C is also violated but not the reverse.

The features in which the glide and the /n/ mismatch are: [vocalic], [approximant], [continuant], [labial], [dorsal], [nasal], and [coronal]. We are going to set the nasal feature aside because we will need it later to rule out the candidate *[famə j$_{1,2}$aʔməl]. All of the remaining feature mismatches are going to be turned into IDENT constraints and grouped into Faith-C/ _V and Faith-C.

Tableau 2: [nʲ] blocking

famən$_1$ j$_2$aʔməl	**Faith-C**	**Faith-C/ _V**
a.→famə j̃$_{1,2}$aʔməl	*	
b. famənʲ$_{1,2}$ aʔməl	*	*

Candidate (b) is harmonically bound by candidate (a). The winner does not violate Faith-C/ _V because all of the features of the consonant that precedes a vowel which is in our case the glide are preserved with the exception of [nasal], whose Ident constraints are not included in our cover constraints. The reason the winner violates Faith-C is that the fusion resultant [j̃] is [-coronal] which violates IDENTCoronal for /n/, and /n/ is not followed by a vowel. That is

why only Faith-C is violated but not Faith-C/ _V. Candidate (b), on the other hand, violates both constraints. The fusion resultant [nʲ] violates all of the IDENT constraints of the glide except for IDENTDorsal, and the glide precedes a vowel. This incurs a violation of Faith-C/ _V which means that Faith-C is also violated.

Turning to the other suboptimal form that needs to be ruled out by our analysis *[famə $j_{1,2}$aʔməl], in this form the two input segments /n j/ coalesce yielding a nonnasalized [$j_{1,2}$] indicating that the [+nasal] feature is deleted. As mentioned above, the nasal feature is not included in the cover constraints in tableau (2). This gives us the liberty to state that the [+ nasal] feature cannot be deleted due to the activation of the faithfulness constraint MAX-NAS. The ranking of this constraint with respect to Faith-C/ _V and Faith-C does not need to be specified yet.

6) MAX-NAS: If the input has a [+nasal] feature, it must not be deleted in the output.

Tableau 3: [+nasal] preservation

famən$_1$ j$_2$aʔməl	**Faith-C**	**Faith-C/ _V**	**MAX-NAS**
a.→famə $\tilde{j}_{1,2}$aʔməl	*		
b. famənʲ$_{1,2}$ aʔməl	*	*!	
c. famə j$_{1,2}$aʔməl	*		*!

Candidate (c) violates Faith-C for the same reasons candidate (a) violates it. However, Candidate (c) is now harmonically bounded by the winner since it violates MAX-NAS which is not violated by the winner.

2.2 NS tolerance word-medially

Idgham only applies when /n/ occurs word-finally and the triggering segments occupy the initial onset position of the following word. The co-occurrence of /n/ with one of the triggers word-medially does not trigger *Idgham* (7). In Quran, there are only four words where /n/ is followed by one of the 6 *Idgham* triggers word-medially, and in each case the *Idgham* trigger is a glide (Qamawee, 1985 and Nassr, 1994).

7) a. **Root-internal nw**

/qɪ**nw**an/	[qɪ**nw**an]	"hanging clusters of palm trees"
/si**nw**an/	[si**nw**an]	"from the same root"

b. **Root-internal nj**

/addu**nj**a/	[addu**nj**a]	"the life"
/bu**nj**an/	[bu**nj**an]	"in ranks"

From an Optimality Theoretic perspective, the tolerance of the sequence [nS] root-internally occurs when the anti-fusion faithfulness constraint LINEARITY (8) dominates the markedness constraint *NS. It worthwhile to highlight that words are assumed to be unordered in the input, so Linearity doesn't apply to them and doesn't block coalescence across word boundaries.

8) LINEARITY: If A comes before B in the input, A should come before B in the output

Tableau 4: Root-internal NS tolerance

qɪnwan	**LINEARITY**	***NS**
a. →qɪnwan		*
b. qɪw̃an	*!	

Candidate (b) loses because it incurs a fatal violation of the higher raked constraint LINEARITY. The winner violates *NS but it is not a fatal violation because this constraint is outranked by LINEARITY which is violated by the suboptimal form.

2.3 /n/ followed by a nasal

In Quranic reading, when /n/ is followed by a nasal, it deletes and the [+nasal] feature moves to the following segment (Qamawee, 1985 and Nassr, 1994).

9) a.**nn**

/wa la**n** **n**uʃrɪk/	[wa la **ñ**uʃrɪk]	"we will never associate anyone with our Lord"
/mɪ**n** **n**aʔməh/	[mɪ **ñ**aʔməh]	"seeking no favor in return"
/la**n** **n**adxulaha/	[la **ñ**adxulaha]	"we will never enter it"
/amʃad͡ʒɪ**n** **n**abtalih/	[amʃad͡ʒɪ **ñ**abtalih]	"we created man from a liquid mixture, to test them"

b. **nm**

/mɪ**n** **m**al/	[mɪ **m̃**al]	"any money"
/ʔaðabu**n** **m**uqim/	[ʔaðabu **m̃**uqim]	"a lasting torment"
/mɪ**n** **m**aʔən dafɪq/	[mɪ **m̃**aʔən dafɪq]	"from gushing liquid"
/sɪrata**n** **m**ustaqɪm/	[sɪrata **m̃**ustaqɪm]	"a straight path"

It is argued, in this paper, that /n/ coalesces with the following nasal. Example set (9, a) shows that /n/ fuses with the following /n/ yielding an output segment [n] which corresponds to both of the input segments /nn/. In example set (9, b), /n/ fuses with the following /m/ resulting in [m] which stands in correspondence to both of the input segments /nm/.

As for the nasal feature of the input segment at the end of the first word, Quranic scholars including Qamawee, (1985) and Nassr, (1994) argue that in Quran reciting, when /n/ is followed by a nasal, it deletes and the following nasal is overnasalized[7]. The overnasalization (represented by tilde [~] in the examples) of the fusion resultant indicates two things: 1- it is a fusion process 2- the overnasalization of the following nasal is a result

[7] This might be tested acoustically which is out of the scope of this study. It is worth highlighting that overnasalization is different from lengthening. Overnasalization here simply means that the resultant consonant is produced with a lower velum and with much more air coming out through the nose than it is with usual nasal consonants. The motivation for overnasalization stems from the desire to preserve the [+nasal] feature of the fused segment.

of having two [+nasal] features. The analysis developed so far coincides with what Quranic scholars have proposed in terms of feature preservation as will be shown momentarily. Our analysis preserves both [+nasal] features of the two input segments. It already predicts the right outcome: MAX-NAS preserves both [+nasal] features. The following tableau includes all the relevant constraints discussed thus far as well as all the possible candidates. The tableau will be discussed in detail below.

Tableau 5: n + nasal candidates and preservation of nasal features

mɪn$_1$ m$_2$al	**MAX-IO (C)**	**DEP-IO (V)**	***NS**	**Faith-C**	**Faith-C/ _V**	**MAX-NAS**	**UNIFORMITY**
a.→ mɪ m̃$_{1,2}$al				*			*
b. mɪn mal			*!				
c. mɪ mal	*!						
d. mɪnə mal		*!					
e. mɪñ$_{1,2}$ al				*	*!		*
f. mɪ m$_{1,2}$al				*		*!	*

The ranking between MAX-IO (C), DEP-IO (V), *NS and Faith-C needs to be specified. The constraints MAX-IO (C), DEP-IO (V) and *NS must dominate Faith-C because the winner violates both Faith-C and UNIFORMITY whereas candidates (b), (c), and (d) only violate one constraint each: *NS, MAX-IO (C) and DEP-IO (V), respectively. Candidate (b) is ruled out by incurring a fatal violation of the markedness constraint *NS. The suboptimal forms (c) and (d) are ruled out by violating the highly-ranked faithfulness constraints MAX-IO (C) and DEP-IO (V), respectively.

Candidates (e) and (f) are harmonically bound by the winner. Comparing the winner with candidate (f), we can see that both of them violate UNIFORMITY because the output segment [m$_{1,2}$] stands in correspondence to the input segments /n$_1$ m$_2$/. However, candidate (f) violates MAX-NAS since the [+nasal] feature of the input segment /n/ disappeared. The winner does not violate MAX-NAS because the fusion resultant output segment [m̃] preserves both its own [+nasal] feature and the /n/'s [+nasal] feature shown by the nasalization diacritic [~]. This parallels with what has been proposed by Quranic scholars regarding feature preservation.

Comparing candidate (e) with the winner, both violate UNIFORMITY since in both cases one output segment corresponds to two input segments. The winner only violates Faith-C while candidate (e) violates both Faith-C and Faith-C/ _V. The feature mismatches between /m/ and /n/ are: [coronal] and [labial] where /m/ is [+labial] while /n/ is [+coronal]. These two features are turned into IDENT constraints and bundled into the two faith constraints: Faith-C and Faith-C/ _V. The winner preserves all the features of consonant preceding a vowel which is /m/ in our case with the addition of the [+nasal] feature of /n/ which is not among the IDENT constraints grouped into Faith-C and Faith-C/ _V as explained earlier. The winner only preserves the [+nasal] feature of the input segment /n/. The [+coronal] feature of the underlying segment /n/ is lost which in turn violates Faith-C but not Faith-C/ _V because /n/ is not followed by a vowel. Candidate (e), on the other hand, preserves all the features of /n/ with the addition to the [+nasal] feature of the underlying segment /m/. The

[+ labial] feature of /m/ is lost which incurs a violation of Faith-C/ _V because /m/ underlyingly precedes a vowel. It is worth repeating that incurring a violation of Faith-C/ _V will automatically violate Faith-C but not the other way around. This makes candidate (e) violate both constraints, and thus candidate (e) loses.

2.4 /n/ followed by a liquid

In Quranic reading, when /n/ is followed by a liquid, both the segment and the [+nasal] feature disappear (Qamawee, 1985 and Nassr, 1994). This makes it different from the previous triggering segments since with liquids the [+nasal] feature is phonetically unrealized.

10) a. **nl**

/jawmaʔðı**n** **l**axbir/	[jawmaʔðı **l**axbir]	"their Lord, on that Day, is fully informed of them"
/fasalmu**n** **l**ak/	[fasalamu **l**ak]	"then, Peace upon you"
/mı**n** **l**adunh/	[mı **l**adunh]	"from him"
/malla**n** **l**ubada/	[malla **l**ubada]	"so much money"
/huda**n** **l**ılmutaqin/	[huda **l**ılmutaqin]	"these are upon guidance from their lord"

b. **nr**

/mı**n** **r**abbık/	[mı **r**abbık]	"from your god"
/ʁafuru**n** **r**aħim/	[ʁafuru **r**aħim]	"God is Forgiving and Merciful"
/wa θamaratu**n** **r**isqa/	[wa θamaratu **r**isqa]	"fruit there from as sustenance"
/mı**n** **r**abbıhım/	[mı **r**abbıhım]	"from their god"
/mı**n** **r**asul/	[mı **r**asul]	"any prophet"

The ranking so far predicts fusion, and the only task is to prevent the [+nasal] feature from surviving in this case. What prevents the [+nasal] feature from appearing on the fusion resultant output liquid is a highly ranked markedness constraint that militates against nasalized liquids: *L̃ (11). Given the marked status occupied by nasalized liquids, this constraint Penalizes nasalized liquids.

11) *L̃: Nasalized liquids must not be allowed.

This constraint interacts with the faithfulness constraint that militates against deleting the [+nasal] feature: MAX-NAS. The deletion of the [+nasal] feature when /n/ fuses with a following liquid occurs when the liquid anti-nasalization constraint *L̃ dominates MAX-NAS.

Tableau 6: Nasalized liquid blocking

mɪn$_1$ l$_2$adunh	***L̃**	**MAX-NAS**
a. → mɪ l$_{1,2}$adunh		*
b. mɪ l̃$_{1,2}$adunh	*!	

Candidate (b) is ruled out by incurring a violation of *L̃ which outranks the constraint violated by the winner MAX-NAS.

No ranking between MAX-NAS and Faith-C/ _V has been established yet; this must change now. The reason the ranking between these two constraints must be specified is that the winner violates MAX-NAS while this candidate *[mɪn$_{1,2}$ adunh] does not. If MAX-NAS which is violated by the winner happens to be above Faith-C/ _V which is violated by *[mɪn$_{1,2}$ adunh] but not by the winner, *[mɪn$_{1,2}$ adunh] would win. This is because all of the other constraints violated by this form is also violated by the winner as the next tableau shows. Therefore, Faith-C/ _V must outrank MAX-NAS. The following tableau shows how the relevant possible candidates fail when they are passed through the constraints. Only the surface form survives through the hierarchy. For time and space purposes, the most faithful candidate will not be included in the following tableau as it is ruled out by the highly ranked *NS as illustrated earlier. The candidates that satisfy *NS by deleting a consonant or epenthesizing a vowel will not be included either as they are ruled out by the highly ranked faithfulness constraints MAX-IO (C) and DEP-IO (V), respectively. The tableau will be discussed in detail below.

Tableau 7: n + liquid candidates and deletion of nasal features

mɪn$_1$ l$_2$adunh	***L̃**	**Faith-C/ _V**	**MAX-NAS**	**Faith-C**	**Uniformity**
a.→ mɪ l$_{1,2}$adunh			*	*	*
b. mɪ l̃$_{1,2}$adunh	*!			*	*
c. mɪn$_{1,2}$ adunh		*!		*	*

Candidate (b) is ruled out early by incurring a fatal violation of the liquid anti-nasalization constraint *L̃. Comparing candidate (c) with the winner, the winner violates MAX-NAS because the [+nasal] of the underlying /n/ is deleted. It also violates Faith-C but not Faith-C/ _V. The reason the winner only violates Faith-C but not Faith-C/ _V is because all of the features of the segment preceding a vowel which is the liquid are preserved on the fusion resultant [l$_{1,2}$] (the features are [+approximant] and [+continuant]. The feature [+lateral] is not included because /r/ is an *Idgham*-triggering liquid and it is not [+lateral]). This does not hold for /n/ as it has lost all of its features by fusing with the following liquid which causes the winner to violate Faith-C (the features lost are [-approximant] and [–continuant]. Remember: the nasal feature is not included in the faith constraints). Candidate (c) violates both of the faith constraints because all of the features of the segment preceding a vowel which is the liquid are lost on the fusion resultant output segment [n$_{1,2}$] (note that violating

Faith-C/ _V means violating Faith-C too). The fact that Faith-C/ _V is ranked higher than MAX-NAS rules out candidate (c).

2.5 NS tolerance in the names of Quranic surahs

Quran is divided into Surahs and verses. In Quranic terminology, Surah is the term used for a chapter in the Quran. Each Surah has a name. In some cases, the name of the Surah is used inside some verses. Only found in two cases of the whole Quran that the last segment of the name of the Surah is /n/ and the following onset of the subsequent word is one of the *idgham* triggering segments. In these two cases, *Idgham* is not triggered. In other words, when a name of a Surah participates in making an *idgham*-motivated environment, fusion is not triggered (Qamawee, 1985 and Nassr, 1994). The underlying /n/ occupying the last coda position does not fuse with the following onset of the subsequent word because it is advised by Quranic scholars that names of Surahs are read carefully, clearly and slowly without any fusion of any segments (Qamawee, 1985 and Nassr, 1994).

12) /jasi**n w**a alquran/ "jasin [name of Surah]. By the wise Quran"
/nu**n w**a alqalam/ "Nun [name of Surah]. By the pen"

The blocking of fusion when a name of a surah participates in making an *idgham*-motivated environment occurs when Faith-Surah outranks the constraint that bans the sequence [nS]: *NS.

13) Faith-Surah: A cover constraint for all the faithfulness constraints, relativized for Surahs.

Tableau 8: Fusion blocking in names of Surahs

nun wa	**Faith-Surah**	***NS**
a. → nun wa alqalam		*
b. nu w̃a alqalam	*!	

Candidate (a) wins because candidate (b) applies fusion in a name of a Surah which violates the higher ranked Faith-Surah.

3 The nature of segments involved in *idgham*

3.1 The vulnerability of /n/ and the immunity of /m/ to *idgham*

This subsection tackles the following question: why is /n/ vulnerable to *Idgham* while /m/ is immune to it? Put differently, both /n/ and /m/ are nasals, why is /m/, unlike /n/, *Idgham* resistant? The coronal /n/ undergoes fusion when it is followed by the fusion triggering sonorants discussed above. However, the labial /m/ is exempt from fusion when it is followed by the same set of segments. The underlying phrase /lam nara/ "we did not see" would surface faithfully without fusing /m/ with /n/: [lam nara].

The coronal nasal fusion with a following sonorant is not trigged by a shared place of articulation. If that was the reason, the labial nasal would be better-equipped to fuse with /w/. This also would go against fusing /n/ with /m/ since they do not share the same place of articulation.

De Lacy (2002) proposed a Major Place of Articulation Scale in which dorsals are more marked than labials while labials are more marked than coronals. In the scale, glottals are the least marked. Some languages tend to preserve more marked segments but not less marked segments. Less marked segments undergo certain phonological processes while more marked segments are exempt from these processes. In Catalan, the coronal /n/ place-assimilates to the following segment whereas the dorsal /ŋ/ and the labial /m/ do not (Herrick, 1999). A similar process occurs in Yamphu. Coda coronal stops /t/ and /t:/ undergo debuccalization and become [ʔ] whereas the more marked /p/ and /k/ do not (De Lacy, 2002). I argue that the same applies to the labial /m/ and the coronal /n/ in Quranic reading. The more marked /m/ is exempt from fusion while the less marked /n/ is not.

Following De Lacy's (2002) analysis but with a simpler version, I propose that IDENT constraints for the more marked segment /m/ must outrank IDENT constraints for the less marked segment /n/. These constraints interact with *NS. IDENTLabial (14) which prevents changing the feature values of a labial must outrank *NS. On the other hand, IDENTCoronal (15) which militates against changing the feature values of a coronal is ranked low in the hierarchy: IDENTLabial >> *NS >> IDENTCoronal.

14) IDENTLabial: Penalizes changing the feature values of a labial.

15)IDENTCoronal: Penalizes changing the feature values of a coronal.

The following tableaux show how the ranking between these constraints predicts the right outcomes.

Tableau 9: Tolerance of changing coronal feature values

mɪn$_1$ m$_2$al	**IDENTLabial**	***NS**	**IDENTCoronal**
a. → mɪ m̃$_{1,2}$al			*
b. mɪn mal		*!	

The winner violates IDENTCoronal which is ranked low in the hierarchy because the coronal feature values of the underlying coronal /n/ are changed. Candidate (b) loses because it incurs a fatal violation of *NS.

Tableau 10: Intolerance of changing labial feature values

lam$_1$ n$_2$ara	**IDENTLabial**	***NS**	**IDENTCoronal**
a. → lam nara		*	
b. la ñ$_{1,2}$ara	*!		

The fusion resultant output segment of candidate (b) [ñ$_{1,2}$] is [-labial], therefore, IDENTLabial is fatally violated which rules out this candidate. The winner has a sequence of m + sonorant,

but, given the constraints ranking provided, this violation is not as serious as fusing /m/ with /n/.

There is still one more candidate to take into consideration: *[lam̃$_{1,2}$ ara]. This candidate only violates IDENTCoronal. Given the constraints we have so far, *[lam̃$_{1,2}$ ara] would win. To rule out this candidate, the constraint Faith-C/ _V must come into play. This constraint, as illustrated earlier, militates against changing the feature values of the consonant preceding a vowel. In order to produce the right outcome, it must outrank *NS. This ranking does not contradict the ranking established earlier.

Tableau 11: Intolerance of changing word-initial consonant's place feature

Lam$_1$ n$_2$ara	**Faith-C/ _V**	**IDENTLabial**	***NS**	**IDENTCoronal**
a. → lam nara			*	
b. lam̃$_{1,2}$ ara	*!			*

Candidate (b) loses by incurring a fatal violation of Faith-C/ _V. It is violated because the place feature of the underlying coronal /n/ which precedes a vowel is lost. The fusion resultant segment [m̃$_{1,2}$] is [+labial] and [-coronal].

Preserving the more marked segment in Quranic reading is also supported by another Tajweed rule: *Iqlab*. *Iqlab* is a rule that governs how the coronal /n/ is pronounced when it precedes the voiced labial stop /b/. According to Qamawee, (1985) and Nassr, (1994) and many other Quranic scholars, when /n/ precedes /b/, it assimilates in place to /b/: /mɪ**n** **b**aʔdi/ [mɪ**m** **b**aʔdi] "after" (for more examples on this phenomenon see Qamawee, 1985 and Nassr, 1994). The undelaying coronal /n/ place-assimilates to the following /b/ and becomes [m]. However, going along the lines with Catalan place assimilation highlighted earlier, an underlying /m/ would not assimilate in place with a following voiced coronal stop /d/: /θu**m** **d**ana/ →[θu**m** **d**ana] not *[θu**n** **d**ana] "then he came closer". This indicates that our analysis is on the right track and it works across all Tajweed rules.

3.2 Sonorants as *idgham*-triggering segments

This subsection attempts to briefly explain why the coronal /n/ fuses with sonorants but not with other classes of segments. Fusion with sonorants seems to be determined by faithfulness. When /n/ fuses with a sonorant, the resultant output segment would be a sonorant since both the /n/ and the trigger are sonorants. However, if /n/ fused with a nonsonorant segment, the resultant output segment would be faithful to only one of the involved segments in terms of the feature [sonorant]. In other words, the resultant output of fusing /n/ with a nonsonorant would be either a sonorant /n/ which would not be faithful to the other nonsonorant involved in the fusion process or a nonsonorant which would be unfaithful to the sonorant /n/ involved in the fusion process. Ident(sonorant) seems to be high-ranking which is, I believe, why /n/ only fuses with sonorants.

4 Conclusion

This paper thoroughly explores *idgham* and briefly touches on *Iqlab*; two Tajweed rules that govern how the coronal /n/ is pronounced in Quranic reading. The subsections that fall under section 2 focus on *idgham* (the way /n/ is pronounced when it is followed by */j,w, m, n, r* and *l/*) and provide a unified OT account for *idgham* regardless of what segment follows /n/. I argue that *Idgham* is a fusion process driven by an interaction between a constraint against a specific sequence of consonants and faithfulness constraints for particular positions and particular features. Section 3 covers the nature of the segments involved in *idgham*. The subsection 3.1 discusses the vulnerability of /n/ and the immunity of /m/ to *Idgham* and develops an OT account for that. I propose that IDENT constraints for the more marked segment /m/ must outrank IDENT constraints for the less marked segment /n/. That is why the more marked segment /m/ is exempt from fusion. This is supported by *Iqlab* where /n/ assimilates in place to the following /b/ which is not the case when /m/ is followed by /d/. The subsection 3.2 explains briefly why /n/ fuses only with sonorants but not with other segments. This fact is attributed to faithfulness. When /n/ fuses with a sonorant, the resultant is a sonorant which is not the case if /n/ fused with a nonsonorant segment because the resultant would be faithful to only one of the involved segments.

References

Beckman, J. (1999). *Positional faithfulness*. New York: Garland.

De Lacy, P. V. (2002). The formal expression of markedness*. PhD dissertation*.

Gotowski, M. (2019). La règle du jeu: Verlan as an example of antifaithfulness. *Linguistica Atlantica, 37(2).*

Halle, M., & Clements, G. N. (1983). *Problem book in phonology: a workbook for introductory courses in linguistics and in modern phonology*. MIT Press.

Herrick, D. (1999). Catalan Cluster Simplification and Nasal Place Assimilation. *Phonology at Santa Cruz (PASC)* 6:25- 37.

Ibrahim, N., Razak, Z, Yusoff, M., Idris, I., Tamil, M., Noor, M., & Naemah, N. (2008). Quranic verse recitation recognition module for support in j-QAF learning: A review. *International Journal of Computer Science and Network Security (IJCSNS)*, *8(8), 207-216*.

Lombardi, L. (1998). Evidence for MaxFeature constraints from Japanese. *ms*.

McCarthy, J, & Prince, A. (1995). Faithfulness and reduplicative identity. *Linguistics Department Faculty Publication Series, 10,* Rutgers University.

Muhammad, M., Muhammad, R., Muhammad, A., & Martinez-Enriquez, M. (2010). Voice content matching system for quran readers. *Ninth Mexican International Conference on Artificial Intelligence*, 148-153.

Nasser, A. (1994). *Ghayat Al-Mareed fee Alm AL-Tagweed*. Fourth Edition. Cairo Modern Library.

Pater, J. (1999). Austronesian nasal substitution and other NC effects. *Prosody-morphology interface*, 310-343.

Qamawee, M. (1985). *Alburhan in Tagweed Al-Quran*. First Editin. Saudi Islamic Affairs.

Quranic data translations are taken from https://www.clearquran.com/

Contact Information

Ahmed Alnuqaydan

u1215249@utah.edu
University of Utah
2873 S Eugene Lane #181
Salt Lake City, UT 84115
United States of America

Exploring L2 English Learners' Articulatory Problems Using a Read-Aloud Task

Shuo Kang

University of Calgary

Abstract

Although the irregular pause in oral production is a common problem in the speech of language learners, it is unclear what factors cause the pauses and to what extent learners' native language affects the oral fluency of their target language. This study investigated the anomalous pauses made by English learners who speak Mandarin Chinese as their mother tongue, with attempts of finding out why these pauses occur in the articulation process of their speech productions. 36 learners and 36 native English speakers participated in a read-aloud task (124 English texts with 26081 words for each participant), which was recorded for further analysis. Based on previous research into oral fluency, the anomalous pause in this study is operationalized to have a long duration (over 1s) and occur in the formulaic sequences or in a single sentence. That is, any pauses that conform to both long silent intervals in formulaic sequence and long silent intervals in a sentence where there is no formulaic sequence, count as anomalous pauses in this study. The results demonstrated that L2 learners made significantly more anomalous pauses when reading aloud English texts and they had difficulty articulating the phonemes /b/, /d/, /g/, /dʒ/, /z/ and /ð/. More importantly, no difference was found in the frequency of anomalous pauses between advanced and less advanced learners, indicating that even proficient learners have difficulties in L2 pronunciation. The reasons why these articulatory problems cause irregular pauses and related pedagogical implications are discussed.

Key words: Articulatory problems, anomalous pauses, phonological contrasts

1 Introduction

High-level oral fluency is many second language (L2) students' learning target (Tavakoli, 2011). However, L2 speaking has been considered a very difficult skill that is affected by a series of factors, such as "learning curriculum, language and socio-cultural background, interpersonal and institutional contexts in which L2 students found themselves" (Gan, 2012, p. 3). The aim of the present study is to investigate linguistic influences on articulatory problems, therefore the factors that are non-linguistic are beyond the scope of this study.

While the L2 oral fluency has been operationalized in different aspects based on specific research purposes (Ellis, 2009), this study adopts a 'narrow' definition of fluency that learners are able to speak with fewer pauses at a regular speech rate (Fillmore, 1979). That is, the frequency of speech pauses serves as a strong indicator of learners' speaking proficiency levels. The frequent pauses in learners' speech normally suggest their lack of oral fluency. However, the anomalous pauses that hinder the speaking process should be differentiated from those naturally occurring pauses, as the latter usually appear even in native speakers' extemporaneous speech. To differentiate the natural and anomalous pauses, it is rational to start from the well-known filled and unfilled pauses as shown in the following examples (1) and (2), respectively:

1) This paper investigates, um, you know, the anomalous pauses.
2) This paper investigates...(short silence)...the anomalous pauses.

Unlike writing, oral production is a continuous process that normally requires speakers to rapidly conceptualize what to say, formulate their ideas by choosing proper words, encode the words based on the phonological and grammatical rules and articulate each word of the message in a phonetically acceptable way (Levelt, 1995). In this process, both native speakers and L2 learners tend to have fillers, false starts, repetitions and self-corrections in their speech. As illustrated in example (1), both 'um' and 'you know' are filled pauses with the latter being called lexical filler. Since the filled pause normally reflects the authentic and natural features of speaking (Gilmore, 2004), it is not reasonable to consider filled pause as a lack of oral fluency particularly when they do not occur many times in the speech.

The current study centers on the unfilled pauses made by L2 learners in the articulation process (Levelt, 1995). According to Riggenbach (1991), the unfilled pause or silent interval serves as a very strong indicator of lack of fluency. However, it is noteworthy that not all silent intervals necessarily count as anomalous pauses. While some short unfilled pauses, as shown in the above example (2), are simply juncture pauses by which speakers highlight the importance of information following the pauses, being silent for a long time without junctures is normally unacceptable. As such, the duration, function, and place of occurrence for the silent pauses should be taken into account to judge whether or not some unfilled pauses are anomalous. Many minimum

cut-off points for the pause duration are used to distinguish between naturally occurring and anomalous pauses. While there is no clear consensus on the appropriate cut-off point, many researchers argue that silent intervals that are over 400 milliseconds (ms) count as irregular unfilled pauses (Derwing et al., 2004). More recently, Yuan, Xu, Lai and Liberman (2016) classified the silent pause into three categories: short pause (0-200 ms), intermediate pause (200-1000 ms) and long paude (> 1000 ms). To make sure the identification of unfilled pauses as anomalous is consistent, the author adopts 1000 ms (one second) as the cut-off point in this study and those longer than this duration count as irregular pauses.

Another characteristic of anomalous pauses can be found in the production of formulaic sequences. Because empirical studies have shown that the multiword units are processed as a whole rather than word by word (Jiang, 2018), L2 learners are less likely to pause when speaking formulaic sequences (Tavakoli, 2011). Thus, if the pauses occur in the formulaic expressions, they count as anomalous pauses. According to Grant and Nation (2006), the formulaic language is classified into *collocation*, *figurative expression*, and *core idiom*. Collocations are those frequently used expressions where the meaning of the component clearly makes up the meaning of the whole, as shown in the following example (3):

3) You *are supposed to* write your exam, so *go to school* now.

The figurative expressions, in this narrow sense, are items with fixed patterns that combine literal meanings with figurative meanings. Please refer to the example given in (4):

4) *Everything comes to them who wait*.

Most figurative expressions from this category (Grant & Nation, 2006) could be found in the idiom dictionaries. Typically, the figurative meaning is to some degree related to the literal meaning in the multiword unit. The core idioms, on the other hand, are items where the meaning cannot be inferred by the components within them, as illustrated in the following example (5):

5) I understand the lecture. It's *a piece of cake* for me.

Finally, the L2 anomalous pauses are more likely to occur in a single sentence rather than between sentences (Tavakoli, 2011). Although the example (2) illustrates a deliberate pause in a sentence for the purpose of emphasis, more natural pauses occur between two sentences and function as a syntactic boundary in a similar way as punctuation does in the writing. As such, long pauses in a sentence might count as irregular pauses.

Taken together, the anomalous pause in this study is operationalized to have a long duration (over 1s) and occur in the formulaic sequences or in a single sentence. To ensure that no irregular pauses are missed for data analyses, any pauses that conform to both long

silent intervals (over 1s) in a formulaic sequence and long silent intervals in a sentence where there is no formulaic sequence, count as anomalous pauses in this study.

2 Literature review

Although this study centers on the articulatory pauses, it is necessary to review the research on general oral fluency to gain more insights. Many studies in this line have investigated the comprehensibility of L2 speech (Suzuki & Kormos, 2019). By using either native or L2 speakers' judgement on L2 speech in terms of accent, accuracy, fluency and linguistic complexity, these studies demonstrate that the L2 oral fluency is strongly related to comprehensibility with fewer unfilled pauses indicating higher comprehensibility (Derwing et al., 2004). In addition, the comprehensibility varies across different speech activities or learning tasks, with dialogues and monologues including fewer unfilled pauses and narratives involving more silent intervals. This suggests that the comprehensibility of narratives is normally lower than other speech activities. Finally, learners with higher language proficiency produce fewer unfilled pauses. Issacs and Trofimovich (2012) found that advanced L2 learners' oral expressions were more fluent as they made fewer silent pauses than intermediate L2 learners. Consequently, the comprehensibility of proficient learners' speech is at a higher level.

While these studies demonstrate the relationships among unfilled pauses, comprehensibility and language proficiency, it is unclear whether there are differences in the general pausing pattern of L2 learners and native speakers. To fill this research gap, Tavakoli (2011) employed four narrative tasks to elicit the oral performance of both L2 learners and native speakers of English. After reading the picture stories in the tasks, two groups of participants were given three minutes for memorization, conceptualization and formulation. Then, they were instructed to narrate the stories. In line with previous comprehensibility studies, narratives are more difficult for L2 learners and it is easier to find out the different pausing patterns between learners and native speakers by using narrative tasks. The results suggested that L2 learners paused more frequently and had a longer duration of silence than native speakers did. More importantly, learners tended to pause in a sentence rather than between sentences. This research finding echoes what is discussed earlier – that natural pauses normally occur between sentences, and L2 learners are more likely to have anomalous pauses that occur in a single sentence. Finally, both L2 and native speakers have very few pauses in formulaic sequences in which every word is intimately connected to another. While Tavakoli' s study (2011) provided the first attempt to compare the pausing patterns between L1 and L2 speech and shed light on the studies on unfilled pauses in oral production, the participants in this study were all intermediate L2 English learners which is why it is unclear whether these research findings could be generalized to other L2 learners with different proficiency levels.

Given that the focus of the current study is on the articulation process of speech production, it is reasonable to review the studies on L2 learners' speech pauses in read-aloud tasks that do not involve speakers' conceptualization and formulation processes.

Kowal, O' Connell, O'Brien and Bryant (1975) conducted three experiments to examine the pausing patterns of both read-aloud and story-telling tasks. The results demonstrated that participants were less likely to pause in the reading task, which is not surprising given that reading aloud allows speakers to skip the conceptualization and formulation processes of speech production (Levelt, 1995). Consistent with the research findings of Tavakoli (2011) that L2 learners paused in the middle of a sentence of their narratives, learners in this study had significantly more unfilled pauses in a sentence in the read-aloud tasks than native speakers did, although not as many as those in story-telling tasks. Excluding the factors of conceptualization and formulation, only articulatory reasons could explain this research finding. More recently, Hirano, Kawai, Hirose and Minematsu (2006) investigated the location of L2 Japanese learners' unfilled pauses in a read-aloud task. Participants were asked to read seven Japanese sentences within a paragraph, and their unfilled pauses were recorded for further analysis. The results demonstrated that while native Japanese speakers normally paused between the sentences with an awareness of syntactic boundaries, L2 learners' unfilled pauses were spread over different locations. The authors attributed this phenomenon to L2 learners' lack of syntactic planning.

3 The present study

So far, the existing literature has demonstrated that less proficient learners are more likely to have unfilled pauses in a sentence rather than between sentences, regardless of the task (Tavakoli, 2011). However, it is unclear whether advanced L2 learners also suffer from the same problem and what cause the occurrence of anomalous pauses in the articulatory process of speech production. In addition, more corpus-based studies are called for to gain a more comprehensive and clearer picture of where the L2 learners pause, why they pause, and how to solve this problem. As such, the current study attempts to fill in these research gaps by answering the following questions:

i. Do L2 English learners across different language proficiency levels have more anomalous pauses than native speakers?
ii. Where do anomalous pauses normally occur in English L2 speech?
iii. What are the phonological features of the words or phonemes that L2 learners have difficulty articulating?

3.1 Methodology

Two groups of participants took part in this study. The first group included 36 native speakers of English, all second-year students of Traditional Chinese and Chinese Literature at a university in Shandong Province of China. Of these, 19 were female and 17 were male, aged between 19 and 20, and spoke English as their first language. The second group included 36 L2 English speakers. They were from a six-semester academic English course specifically designed to prepare them for a four-year degree programme at a university in an English-speaking country. Their English language proficiency was

differentiated based on the Common European Framework of Reference. 18 of them were intermediate English learners (B1-B2) while the other 18 were advanced learners (C2). Of these 36 students, 25 were female and 11 were male, aged between 17 and 18. They spoke Mandarin as their native language.

All participants' oral performances were elicited through the use of a read-aloud task. Since this task only taps into participants' articulation, it does not require speakers' conceptualization and formulation processes because they do not need to plan what to say and encode the related information into specific words. Reading aloud is regarded as an effective pedagogical activity and is frequently used across different educational settings in China for English language teaching and assessment. All the reading materials in this study were selected from L2 learners' textbooks by a team of L2 English teachers and researchers, and in doing so, a range of different learner and task factors including familiarity, clarity, and cultural issues were considered. As the materials are all scripted texts that L2 learners had learned before, the problems of word frequency and semantic unfamiliarity were avoided. These texts were selected from *English Grammar in Use* (Murphy, 2015), *English Intensive Reading* (Fu & Chen, 2012), *Writing Academic English* (Oshima & Hogue, 2005), *Cambridge Academic English* (Hewings et al., 2012), Intermediate English Reading and Comprehension (Engelhardt, 2013) and *New Concept English* (Alexander, 1997). There were altogether 124 different texts (26081 words) for each participant to read. All the participants were reading the same texts as the study was progressing.

All the participants read aloud 7-8 texts every two days, and their reading was recorded for further analysis. All the L2 learners in this study had learned the grammar and vocabulary appearing in all the texts before (from one year to two days before the read-aloud task). To ensure that they were familiar with the vocabulary of each text, all participants were required to revise the word list of each text before reading it out aloud. In this way, any anomalous pauses could be attributed primarily to the articulatory factors rather than searching for unfamiliar words. In the progress of the read-aloud task, their recordings were collected and analyzed by two native English-speaking language instructors every two days to find out the silent intervals based on the operationalization of anomalous pauses discussed earlier.

After the read-aloud task was completed, Praat 5.7.08 (Boersma & David, 2017) was used to first measure all the unfilled pauses greater than one second occurring in and between sentence boundaries. The duration of each anomalous pause and the total number that each participant paused were also calculated. The data set was then inserted into SPSS 25.0 (IBM Corp, 2017) for statistical analyses, followed by a qualitative analysis to investigate the reasons for these pauses.

4 Results and discussions

As shown in Table 1, it is clear that L2 learners made significantly more anomalous pauses (t = 6.68, p = 0.001) and they were more likely to pause in a single sentence (t = 8.72, p = 0.001) than native speakers, which is in line with previous studies (Kowal et al., 1975; Hirano et al., 2006).

Table 1: The mean number and duration of anomalous pauses.

Read-aloud task	L2 learners	Native Speakers	*t*	*p*
Pauses in a single sentence	141. 67	38. 05	8.72*	0.001
Pauses between sentences	367. 31	339. 65	1.60	0.11
Pauses in formulaic sequences	68.93	17.32	13.17*	0.001
Pauses in a single word	72.74	20.73	9.62*	0.001
Total pauses	508.98	377.70	6.68*	0.001
Mean duration of pauses (s)	1.37	1.28	0.73	0.48

This finding answers the first question of this research. It is not surprising to gain this result because L2 learners are not able to articulate many words as fluently as native speakers. More importantly, L2 learners paused in the formulaic sequences, which contradicts previous findings in Tavakoli (2011) that formulaic sequences "contain very few pauses and therefore facilitates the learners' fluency" (p. 71). Because the word components within formulaic language are normally processed as an integral unit (Jiang, 2018) and L2 speakers in this study had learned all the formulaic sequences in the texts before, they should not have had too many pauses when producing these collocations, figurative expressions and core idioms.

A potential explanation for this finding is that the learners forgot the meaning of many multiword units because some texts were learned one year before the read-aloud task. It is noteworthy that this study did not test how well the participants knew these formulaic expressions, and revising the word list might not guarantee that they could retrieve the meaning of all formulaic sequences. According to Webb (2007) and Nation and Meara (2010), the unknown words are usually retained with repetitive encounters of the same word in different contexts. Compared with individual words, core idioms and figurative expressions are not frequently learned and used by the L2 speakers in this study, given that the learning objective of their course is to improve their academic English ability for studying abroad. Finally, no significant difference was found in the mean duration of pauses between L2 learners and native speakers, indicating that the length of anomalous pauses made by these two groups is similar to each other.

Of greater interest are the similarities and differences in anomalously occurring pauses between advanced and intermediate learners, as illustrated in the table below.

Table 2: The anomalous pauses made by learners across different levels

Read-aloud task	Intermediate	Advanced	*t*	*p*
Pauses in a single sentence	79.29	62.38	0.87	0.61
Pauses between sentences	173.28	194.03	0.76	0.70
Pauses in formulaic sequences	29.37	39.56	1.90	0.58
Pauses in a single word	24.52	48.22	4.52*	0.001
Total pauses	252.57	256.41	0.52	0.285
Mean duration of pauses (s)	1.47	1.27	0.90	0.38

A number of *t*-tests were conducted to find out whether there were any statistical differences between the intermediate and the advanced learners' performance with respect to the number of pauses in various conditions. Remarkably, no significant differences were found between advanced and intermediate learners' anomalous pauses ($t = 0.52$, $p = 0.285$) except for the pauses in a single word ($t = 4.52$, $p = 0.001$). This result contradicts previous research findings in Kowal and her colleagues (1975) as well as Suzuki and Kormos (2019) showing that higher language proficiency indicates more fluent L2 speech production. The reason for this phenomenon might be associated with the participants' different language backgrounds. While previous studies recruited English-speaking learners of German (Kowal ei al., 1975) and Japanese-speaking learners of English (Suzuki & Kormos, 2019) to participate in the read-aloud and story-telling tasks, the participants in the current study were Mandarin learners of English. Because English and Mandarin do not belong to the same language family and they do not even share a proto-language, it is more difficult for Mandarin learners to be fluent in reading. As such, it is important that learners of different second languages be further studied, especially languages that are not cognate to their first languages.

In order to answer the third research question, the problematic phonemes that L2 learners have difficulties articulating were selected and analyzed in terms of the phonological properties of learners' first and second language, as illustrated in Table 3:

Table 3: Phonological features of paused words

Phonemes	Features	Examples	Potential reason
/d/	voiced, alveolar	'difficult'	different VOT
/b/	voiced, bilabial	'suburb'	different VOT
/g/	voiced, velar	'glove'	different VOT
/dʒ/	voiced, palatal	'jet'	different places of articulation
/z/	voiced, alveolar	'zebra'	lack of L1 transfer
/ð/	voiced, dental	'that'	lack of L1 transfer

Although the voiced stops /d/, /b/ and /g/ exist in both Mandarin and English, the voice of onset time (VOT, the length of time between the release of air and onset of voicing) of these segments might be different between the two languages and thus causing difficulties in L2 speech production. The current study will use the Wavesurfer software

(Royal Institute of Technology, 2000) to investigate the VOT of these segments in both languages. Independent *t*-tests will be conducted to show whether the VOT for all three voiced stops in Mandarin is significantly different from that in English.

According to Flege and Eefting (1987), the different VOT of specific segments between learners' L1 and L2 might be causing the incorrect L2 pronunciation of these phonemes. If the present study can demonstrate the significant differences in the VOT of /d/, /b/ and /g/ between Mandarin and English, it would suggest that the different VOT probably lead to the anomalous pauses in L2 speech. This might be because that L2 learners struggle to articulate the problematic segments (whose VOT is different from their L1) in a correct and standard way, causing the long hesitation and pause.

In addition to /d/, /b/ and /g/, the English phoneme /dʒ/ in the fourth line of Table 3 sounds similar to Mandarin phoneme /zh/, but /zh/ is retroflex and voiceless. When it is articulated, the tip of tongue is curled back against the palate when articulating these sounds. Thus, the different phonation and places of articulation may also lead to anomalous pauses in learners' L2. Regarding the last two phonemes /z/ and /ð/ which do not exist in the Mandarin phonemic system, learners cannot transfer their L1 phonological representation to the articulation of these phonemes. As for the suprasegmental features of English (i.e., stress pattern, prosodic timing, peak alignment, intonation, etc.), no evidence was found that these factors cause Mandarin EFL learners' irregular pauses.

As discussed earlier, the formulaic sequences should have facilitated L2 learners' speech production (Tavakoli, 2011). Nevertheless, such evidence was not observed in this study. In addition, since all the L2 learners in this study do not have any speech problems in their native language (Mandarin), their L2 pauses are attributed primarily to different phonological properties between English and Mandarin. Drawing on Levelt's (1995) speech production model, this indicates that the processes of phonological encoding and phonetic planning are the main factors that cause Mandarin EFL learners' articulatory problems and speech pauses, regardless of their L2 proficiency levels. This may be due to the lack of English listening curriculums in China (Rost, 2014).

5 Pedagogical implications

Based on the findings of this research, instructors who teach English to native Mandarin speakers need to implement more phonemic discrimination tasks (e.g., minimal pairs) to strengthen students' understanding of the problematic phonemes in English. After students become aware of the different VOT of these segments and of the different places of articulation between Mandarin and English, teachers should implement more speaking activities (e.g., making sentences by using more words where the problematic phonemes are embedded:) to help learners practice the articulation of specific phonemes. Please see the following example (6):

6) I'm *reading* a *difficult book* / I'm *busy reading* a *book because* it is so *difficult.*

Additionally, the teachers can ask students to listen to and read aloud sentences where the problematic phonemes are embedded (as shown in the example below). It is self-evident that listening serves as the initial step of speaking, but the lack of English listening curriculums in China caused many learners' speaking problems (Rost, 2014). Please refer to the example given in (7):

7) I live in the su**b**ur**b** of Leth**b**ri**dg**e, which is not far from your **d**epartment.

Finally, students need to know some compensatory strategies to replace problematic words with other words (e.g., replace *difficult* with *not easy*). Although this method is not strongly recommended because different words can have subtle differences in meaning, it lets learners avoid strains caused by anomalous pauses when they need to read something out aloud in serious contexts, such as reading academic reports in a conference.

6 Conclusion

This study examined what anomalous unfilled pauses are usually associated with L2 English learners who speak Mandarin as their L1, providing explanations for the occurrence of these pauses from a phonological perspective. These research findings indicate that proficient L2 learners do not necessarily have fewer anomalous pauses and therefore the speaking activities for learners should be spread over all levels of English language courses. In addition, formulaic sequence does not necessarily facilitate learners' oral fluency, despite being processed as an integrated unit. Finally, English instructors should assign further emphasis on improving students' articulatory ability. As this study suggests, anomalous pauses occur even if learners skip the conceptualization and formulation of speech production. It should be noted that the participants in this study were engaged only in the read-aloud task, meaning that the results of this study do not necessarily apply to speaking contexts that are spontaneous and interactional in nature. Further research needs to investigate L2 learners' anomalous pauses during other speech acts such as the extemporaneous speech and everyday conversation.

References

Alexander, L. G. (1997). *New Concept English.* Beijing: Foreign Languages Teaching & Research Press.

Boersma, P., & Weenink, D. (2017). Praat: doing phonetics by computer [Computer program]. Version 5.7.08beta, retrieved 10 June 2017 from http://www.praat.org/.

Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second language fluency: Judgments on different tasks. *Language Learning 54*(4), 655-679.

Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics*, *30*(4), 474-509.

Engelhardt, D. (2013). *Intermediate English Reading and Comprehension*. Toronto: McGraw-Hill Education.

Fillmore, C. J. (1979). On fluency, In C. J. Fillmore, D. Kempler, & W. S. Wang (eds) *Individual differences in language ability and language behavior*, 85-101. Academic Press.

Flege, J. E., & Eefting, W. (1987). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics*, *15*(1), 67-83.

Fu, S. Y., & Chen, Y. (2012). *English Intensive Reading.* Beijing: Peking University Press.

Gan, Z. (2012). Understanding L2 speaking problems: Implications for ESL curriculum development in a teacher training institution in Hong Kong. *Australian Journal of Teacher Education*, *37*(1), 43-59.

Gilmore, A. (2004). A comparison of textbook and authentic interactions. *ELT journal*, *58*(4), 363-374.

Grant, L. E., & Nation, P. (2006). How many idioms are there in English?. *ITL-International Journal of Applied Linguistics*, *151*(1), 1-14.

Hewings, M., McCarthy, M., & Milanovic, M. (2012). *Cambridge Academic English: An Integrated Skills Course for EAP Student's Book*. Cambridge: Cambridge University Press.

Hirano, H., Kawai, G., Hirose, K., & Minematsu, N. (2006). Unfilled pauses in Japanese sentences read aloud by non-native learners. *INTERSPEECH* 2006: 725-728.

IBM Corp. (2017). IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.

Isaacs, T., & Trofimovich, P. (2012). Deconstructing comprehensibility: Identifying the linguistic influences on listeners' L2 comprehensibility ratings. *Studies in Second Language Acquisition*, *34*(3), 475-505.

Jiang, N. (2018). *Second language processing: An introduction*. New York: Routledge.

Kowal, S., O'Connell, D., O'Brien, E., & Bryant, E. (1975). Temporal aspects of reading aloud and speaking. *American Journal of Psychology, 88*(4), 549-569.

Levelt, W. J. (1995). The ability to speak: From intentions to spoken words. *European Review*, *3*(1), 13-23.

Murphy, R. (2015). *English grammar in use : A self-study reference and practice book for intermediate students of English.* Cambridge: Cambridge University Press.

Nation, P., & Meara, P. (2010). Vocabulary, In N. Schmitt (ed.) *An introduction to applied linguistics*, 34-52. London: Hodder Education.

Oshima, A., & Hogue, A. (2005). *Writing Academic English.* New York: Pearson Education ESL.

Riggenbach, H. (1991). Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes*, *14*(4), 423-441.

Rost, M. 2014. Developing listening fluency in Asian EFL settings, In Muller, T., Adamson, J., Brown, P., & Herder, S. (eds) *Exploring EFL fluency in Asia*, 281–296. Basingstoke: Palgrave Macmillian.

Royal Institute of Technology. (2000). WaveSurfer software for sound visualization and manipulation. Stockholm: Royal Institute of Technology.

Suzuki, S., & Kormos, J. (2020). Linguistic dimensions of comprehensibility and perceived fluency: An investigation of complexity, accuracy, and fluency in second language argumentative speech. *Studies in Second Language Acquisition*, *42*(1), 143-167.

Tavakoli, P. (2011). Pausing patterns: Differences between L2 learners and native speakers. *ELT Journal*, *65*(1), 71-79.

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, *28*(1), 46-65.

Yuan, J., Xu, X., Lai, W., & Liberman, M. (2016). Pauses and pause fillers in Mandarin monologue speech: The effects of sex and proficiency. *Proceedings of Speech Prosody* 2016: 1167-1170.

Contact Information

Shuo Kang

shuo.kang1@ucalgary.ca
School of Languages, Linguistics, Literatures, and Cultures
University of Calgary
Craigie Hall D310
2500 University Dr. N.W.
Calgary, AB, T2N 1N4
Canada

Plurality as a Phi-Feature in Non-Inflectional Plurals

Brittany McDonald

University of Calgary

Abstract

Some languages, such as Pirahã, express plurality through means other than plural inflectional morphology. Wiltschko (2008) calls these alternative plural marking strategies non-inflectional plurals *and develops several diagnostic criteria for determining whether or not a language is an inflectional plural-marking language, illustrated with examples from English (an inflectional plural-marking language) Halkomelem (a non-inflectional plural language). These criteria pertain to obligatoriness, agreement, compounding, and derivational morphology. This paper expands on these criteria, drawing two more from Greenberg's (1963) Universals, to answer the following research question: Do non-inflectional plurals possess a plural phi-feature? This paper explores this question by looking at certain properties of Khmer and Thai, two languages which appear to have non-inflectional plurals, to look for any evidence of the presence of phi-features in their respective plural-marking strategies.*

Key words: Syntax, morphology, inflection, derivation, plural, phi-features.

1 Introduction

Number is a feature which has been widely studied in language and, while it is generally thought to be well-understood, theories change and new discoveries can be made which force us to reconsider what we have learnt. For instance, it was once believed that all languages had grammatical number until the discovery of Pirahã provided evidence to the contrary (Corbett, 2000, p. 50). In this language, there are no plural forms for nominals, not even for pronouns (Corbett, 2000, p. 50). Instead, plural is expressed through conjunction as shown in (1), or the use of the comitative/associative postposition *xigí* (Corbett, 2000, p. 50).

1) **ti gíxai** pí-o ahá-p-i-í
   1 2 also-OBL go-IMP-PROX-COMPLETE.CERTAINTY
   (You and I will go)
   (Corbett, 2000, p. 51, Figure 35)

Wiltschko (2008) refers to these alternative means of expressing plurality as *non-inflectional plurals.* Knowing that there are languages that can express plurality by means other than grammatical number, the grammatical properties of these non-inflectional plurals come into question. This paper will attempt to answer the question in (2).

2) Do non-inflectional plurals possess a plural phi-feature?

In this paper, I conclude that it is entirely possible for non-inflectional plural forms to lack phi-features, though this may not be the case for all non-inflectional plural marking languages. This could have certain implications about the language's entire grammar. The most striking, perhaps, is the effect this would have on a language's agreement system. As there are other coding strategies besides agreement, such as word order and case-marking, it is plausible to consider that languages are able to express many of the same ideas, such as the idea of plurality, clearly without the use of phi-features or agreement.

In terms of markedness and Greenberg's Universal 34 in (3), there are certain expectations to be had regarding number marking systems in language. Essentially, that singular is the least marked number form, followed by plural, then by dual, etc.

3) No language has trial number unless it has dual. No language has dual unless it has plural.
   (Greenberg, 1963, p. 94, Universal 34)

Japanese has been described as a language with a plural form but no singular which violates (3) (Corbett, 2000, p. 38). It has been argued that since Japanese only optionally marks plural forms there is still a singular/plural distinction when plural is marked, therefore it does not violate Universal 34 (Corbett, 2000, p. 38). There are, however, several languages which behave like Japanese in that they optionally mark for plural, and sometimes have a single form for both singular and plural. Presumably, these languages (Japanese included) would have non-inflectional plural marking of some kind.

Section 2 of this paper will briefly introduce the idea of non-inflectional plural marking and some diagnostic criteria from Wiltschko (2008), as well as introduce new criteria for consideration. Section 3 will survey the plural marking systems of three languages; English (a review and continuation of the analysis performed in Witschko (2008)), Khmer, and Thai as well as apply each criterion to these languages to see which patterns emerge. Section 4 will discuss the findings in Section 3. Concluding remarks will follow in Section 5.

2 Non-inflectional plurals

Wiltschko (2008) explains the phenomenon of non-inflectional plurals in Halkomelem in great detail. She also provides four main diagnostic criteria for identifying whether a given language has an inflectional or non-inflectional plural marking system. Under this analysis, non-inflectional plurals are considered modifiers adjoined in the structure rather than heads of a NumP (#P) (Wiltschko, 2008, p. 646). Four main criteria are explored and are listed in (4). These criteria are used to differentiate between inflectional and non-inflectional plurals.

4) a. Is plural marking obligatory?
   b. Does plural marking trigger agreement?
   c. Can plural marking appear inside compounds?
   d. Can plural marking appear inside derivational morphology?
   (Wiltschko, 2008, p. 642)

I propose that these criteria do not only distinguish between inflectional and non-inflectional plural but are also relevant to the presence or absence of phi-features, in particular, criterion (4a). It is normally expected that a controller (in this case, a nominal) would invoke some sort of agreement on its target (verb, article/determiner, adjective/modifier) (Corbett, 2000, p. 178). This suggests a correspondence between the grammatical features of a controller and those valued on its target. If the features of a nominal are not phi-features, but rather some sort of semantic or pragmatic features, it follows that these same features cannot apply to the target in the grammar, not triggering agreement. Criteria (4a, c, and d) are also relevant to the presence or absence of phi-features, as the distribution of non-inflectional plurals does not seem to match that of their inflectional counterparts in the languages that have them. This could indicate that while inflectional plurals have one type of distribution dictated by phi-features, non-inflectional plurals have another based on whatever type of features they possess (i.e. semantic, pragmatic). In addition to the criteria in (4) I propose two more in (5). (5a) is motivated by Greenberg's (1963) Universal 36 shown in (6).

5) a. Does the language have grammatical gender?
   b. Do pronouns show a number distinction?

6) If a language has the category of gender, it always has the category of number.
   (Greenberg, 1963, p. 58, Universal 36)

While (5a) may not be a defining criterion for a language having plural number, it does reveal a general tendency of language which can be used in support of other criteria. (5b) is relevant because, as Corbett (2000) notes, languages are not always symmetric in how they mark plural in pronouns and other nominals, but if a language does not have a number distinction in its pronouns, it is unlikely that it would mark for plurality on other nominals (p. 27).

3 Analysis

Section 3.1 will apply the criteria in (4) and (5) to English, a language known to have inflectional plural, to set a baseline for one type of plural marking in order to make comparisons in the following sections. Section 3.2 will explore the different plural marking devices in Khmer and investigate the criteria proposed in the beginning of this section in relation to Khmer. Section 3.3 will do the same for Thai.

3.1 English

3.1.1 Inflectional Plurals in English

Wiltschko (2008) performs the following analysis of English plural marking in order to contrast it with the system found in Halkomelem. I will briefly show her findings in English and then apply the new criteria in (5). These will clearly show that English, an inflectional plural marking language, clearly follows the criteria in a particular pattern which will prove to be different than non-inflectional plural marking languages.

3.1.2 Criteria in English

Plural marking in English is obligatory. If a noun phrase has a cardinality greater than one, it generally must be marked with plural -/s/ or one of its allomorphs (Wiltschko, 2008, p. 642). This is seen in basic examples such as (7). For English, the answer to the question in (4a) is *yes*.

7) a. the **three** boy**s**
   b. *the **three** boy
   (Wiltschko, 2008, p. 642, figures (2a and b))

Plural marking in English triggers agreement within a DP, as seen in (8)a and b, and also triggers agreement on verbs if the subject is 3rd person singular as seen in (8)c and d. For English, the answer to the question in (4b) is *yes*.

8) a. **This boy** can sing
   b. ***These boy** can sing
   (Wiltschko, 2008, p. 643, figures (5c and d))
   c. The boy walk**s**.
   d. *The boy walk.

In English, plural marking is not permitted inside compounds. Even when a noun has a plural interpretation, such as in *toothbrush* since one would normally brush more than one tooth with it, the plural marking must appear on the entire compound and not on an element inside of it. This is shown in (9). The answer to the question in (4c) is *no*.

9) a. **tooth**-brush(**es**)
   b. *__teeth__-brush
   (Wiltschko, 2008, p. 644, figures (7a))

English plural marking cannot appear inside derivational morphology. Similar to compounding, the plural inflection must appear on the entire word rather than on the root before the derivational morphology. This is shown in (10). For English, the answer to the question in (4d) is *no*.

10) a. brother-hood(s)
    b. *brother**s-**hood
    (Wiltschko, 2008, p. 645, figure (10c))

English does have grammatical gender but almost exclusively for pronouns. Only the pronouns *he*/*him*/*his* (M) and *she*/*her*/*hers* (F) have a grammatical gender feature while the rest are neutral. Most regular nouns are neutral to gender (e.g. *chair*) but certain nouns which strongly refer to a male (e.g. *man*) or female (e.g. *woman*) trigger masculine or feminine pronouns when they co-refer as seen in (11). For English, the answer to the question in (5a) is *yes*.

11) a. **The woman**$_i$ ate **her**$_i$ lunch.
    b. **The man**$_j$ ate **his**$_j$ lunch.
    c. *__The woman__$_i$ ate **his**$_i$ lunch.
    d. **The woman**$_i$ ate **his**$_j$ lunch.

English pronouns distinguish for number. We see this in the corresponding 1st person forms *I* (sg) vs. *we* (pl) and the corresponding 3rd person forms *he/her/it* (sg) vs. *they* (pl). For English, the answer to the question in (5b) is *yes*. A summary of all the criteria as it applies to English is shown in Table 1 below.

Table 1. Comparative table of English

Criteria	**English**
4a. Obligatory	Yes
4b. Agreement	Yes
4c. Inside compounds	No
4d. Inside deriv. m'y	No
5a. Grammatical gender	Yes
5b. # distinction pronouns	Yes

3.2 Khmer

3.2.1 Non-inflectional plurals in Khmer

Khmer is the national language of Cambodia with approximately 14 million speakers (Haiman, 2011, p. xv). This language does not obligatorily mark for plural overtly, as seen in (12), but has certain devices to express plural meaning (Haiman 2011).

12) a. ***cru:k*** *muaj*
pig one
(a pig)

b. ***cru:k*** *pi*
pig two
(two pigs)
(Haiman, 2011, p. 75 figure (167a), p. 142 figure (5a))

Khmer also has a word *puak* ('group') which can be used with a noun or pronoun in order to express plurality, as seen in (13). This word is said to be undergoing a process of grammaticalization but the meaning is still transparent to speakers (Haiman, 2011, p. 44).

13) *puak neak Pnom Penj dael tveu: damnaeu tev twh kha:ng ceu:ng*
group person Phnom Penh who make trip go direction side north
(people from Phnom Penh who travelled north)
(Haiman, 2011, p. 165, figure (37a))

Plurality can also be expressed in Khmer with two modifying suffixes (or adjectives) - *niania* and *psee:ngpsee:ng*, meaning 'different' and 'various', respectively (Haiman, 2011, p. 50). These are said to be borrowings from Pali which motivates the suffix analysis, but also follow the Khmer word order of Head + Attribute (Haiman, 2011, p. 50). Examples of each are shown in (14).

14) a. *kamhoh-niania*
mistake-different
(mistakes)

b. *kamhoh-psee:ngpsee:ng*
mistake-various
(various mistakes)
(Haiman, 2011, p. 50, figures (21a) and(21b))

Finally, Khmer has a reduplication process which invokes a plural interpretation in DPs (Haiman, 2011, p. 164). In this process, the rightmost element of a DP is fully reduplicated, as seen in (15) (Haiman, 2011, p. 164). (15a) shows an example of adjectival reduplication, and (15b) shows adjectival reduplication for *cah* 'old', as well as noun reduplication for *srej* 'female'.

15) a. *koːn toːc toːc*
child small RED
(small children)

b. *srej srej nwng cah cah*
female RED and old RED
(women and old people)
(Haiman, 2011, p. 164, figures (36a) and (36b))

3.2.2 Criteria in Khmer

This section will apply Wiltschko's (2008) criteria to Khmer. In Section 3.2.1 we saw that plural marking is not obligatory in Khmer and is only used for disambiguation purposes. For Khmer, the answer to the question in (4a) is *no*.

Khmer is an isolating language and lacks inflectional agreement entirely[1] (see Haiman 2011). As such, there is no subject-verb agreement, nor is there adjectival or determiner agreement. The answer to the question in (4b) is *no*.

(4c) is not as easily answered, as Khmer forms extensive compounds of various types, where the reduplicative plural structure is often considered to be a type of compound in and of itself. Constructions with *puak* are also sometimes considered compounds as *puak* can be treated as either a head noun or as an agentive prefix (Haiman, 2011, p. 44). Khmer iconically-motivated compounding allows for plural marking within it, as seen in (16).

16) *cah* ***cah*** *piː daeum*
old **RED** from trunk
(the elders from the beginning/originally…)
(Haiman, 2011, p. 90, figure (6c))

Other types of coordinate compounds combine two uninflected words which derive a plural meaning, as seen in (17).

17) ʔəwpuk mdaːj
father mother
(parents)
(Huffman, 1970, p. 296-297)

It appears that certain types of compounds in Khmer allow for plural marking within them. However, compounds which are unmarked for number appear to be less marked. For Khmer, the answer to the question in (4c) appears to be *yes*.

Since Khmer has very little inflection, and what inflection it does have is prefixed or infixed, it is difficult to tell from Khmer sources if plural can be marked within derivational morphology since plural in this language is generally marked with a separate word. The data

[1] See Flores 2015 for discussion on formality agreement

analyzed for the purposes of this paper did not show cases of plural marking within derivational morphology, so for the purposes of this paper, the answer for the question in (4d) appears to be *no*.

Khmer does not have grammatical gender (see Haiman 2011). It does have two borrowed feminine suffixes from Pali (*-a* and *-ej*), but as inflection (suffixation especially) is foreign to Khmer it is not a productive process in the language and has no other implications in the grammar (Haiman, 2011, p. 36). Speakers are not consciously aware that these are considered gendered morphology (Haiman, 2011, p. 50-51). For Khmer, the answer for the question in (5a) is *no*.

Pronouns in Khmer do not distinguish for number (or person) with the exception of two of the 1st person pronouns *anj* (1st sg) and *jeu:ng* (1st pl). Since all other pronouns lack a person or number distinction, these two apparent exceptions may be specific speaker-oriented pronouns which instead of a 1st person phi-feature have a pragmatic [+Speaker] feature. The 1st person plural pronoun *jeu:ng* could quite plausibly be considered an associative (pragmatic) plural as opposed to an additive (grammatical) plural form since its semantics do not denote multiple speakers but rather one speaker and their associate(s). Therefore, the existence of distinct 1st person singular and plural pronouns in Khmer does not necessarily imply the presence of a plural phi-feature. The rest of the pronouns can be used to denote any person or number (Haiman, 2011, p. 185). Pronouns in Khmer can be made overtly plural with *puak* as seen in (18) for the 3rd person pronoun *koat*. For Khmer, the answer to the question in (5b) appears to be *no*.

18) *knjom cumriap sua* ***puak koat*** *ta:m lumdap lumdaoj*
1 greet group 3 follow gradual gradual
(I greeted them one by one)
(Haiman, 2011, p. 165, figure (37d))

A comparative table between English and Khmer is shown in Table 2 below.

Table 2. Comparative table of English and Khmer

Criteria	**English**	**Khmer**
4a. Obligatory	Yes	No
4b. Agreement	Yes	No
4c. Inside compounds	No	Yes
4d. Inside deriv. m'y	No	No*
5a. Grammatical gender	Yes	No
5b. # distinction pronouns	Yes	No

3.3 Thai

3.3.1 Plural marking in Thai

Thai, though it is from a different language family, has very similar plural marking strategies as Khmer. Thai also has one form for nouns for both singular and plural with no additional

morphology to signify plurality and relies on the discourse context to allow for the hearer to understand whether singular or plural number is intended (Smyth, 2000, p. 24). When number needs to be specified in Thai, numbers and quantifiers are used (Smyth, 2000, p. 25). A small number of nouns can undergo reduplication to invoke a plural interpretation as seen in (19) (Smyth, 2000, p. 25).

19) a. *phûan*
friend
(friend or friends)

b. *phûan phûan*
friend RED
(friends)
(Smyth, 2000, p. 25)

Similar to Khmer, Thai's personal pronouns do not distinguish for number. Like Khmer's plural word *puak* 'group', Thai has a very similar word *phûak* 'group' which can be used to disambiguate personal pronouns for singular or plural (Smyth, 2000, p. 40).

3.3.2 Criteria in Thai

Plural marking is not obligatory in Thai (Smyth, 2000, p. 24-25). This is demonstrated in (20) where there is no overt plural marking on the noun *dèk*, but it is interpreted as plural (Diller, 1988, p. 275). In Thai, the answer to the question in (4a) is *no*.

20) **dèk** hô'ng ní: ka:n-bâ:n mây khœ:y tham sàk thi:
child room this homework not ever do even time
(The **kids** in this room have never once done their homework!)
(Diller, 1988, p. 275, figure 3)

Just like Khmer and Japanese, Thai does not have any verb agreement. This is shown in (21) where the verb *kiàw* is uninflected for agreement in both sentences. Note that in example (21a) *khăw* is interpreted as 3rd person plural with no use of an overt plural marking device, but in (21b) it is interpreted as 3rd person singular. In Thai, the answer to the question in (4b) is *no*.

21) a. **khăw** càk kamlaŋ[2] **kiàw** khââw kan
3 will gather rice together
(**They** will be gathering rice together)

b. dɨannââ **khăw** kamlaŋ **kiàw** khââw lǽǽw
next.month 3 gather rice already
(Next month **he** will be gathering rice)
(Kanchanawan, 1978, p. 22, figures (20) and (21))

[2] Source provided no gloss for this word.

As in Khmer, Thai compounds do not offer such clear-cut answers for (4c). While the language does extensive compounding in nouns (22), adjectives (23), and serial verbs (24), since there is no obligatory plural marking it is difficult to find examples of data where there is plural marking within a compound.

22) phôɔ mêɛ
father mother
(parents)
(Smyth, 2000, p. 30)

23) cay yen
heart cool
(calm)
(Smyth, 2000, p. 84)

24) duu lɛɛ
see watch
(look after)
(Smyth, 2000, p. 60)

Since plurality is sometimes expressed through the use of quantifiers which appear before a noun, it is unlikely that these quantifiers or *phûak* would appear inside of a compound unless it was a phrasal type such as those found in Khmer. In Thai, the answer to the question in (4c) appears to be *no*.

While Thai lacks inflectional morphology, there is some derivational morphology (see Smyth 2000). Since all plural marking forms in Thai appear as separate words, it would not be expected that they could appear inside derivational morphology. In Thai, the answer to the question in (4d) appears to be *no*.

Thai has no grammatical gender (Smyth, 2000, p. 24-25), therefore the answer to the question in (5a) is *no*.

Thai has a larger set of personal pronouns than Khmer with slightly stricter distribution. Some pronouns can only be used by speakers of a certain gender, and the relative social status and age difference between the speaker and addressee strongly influences which forms are appropriate to use in which contexts. Kullavanijaya (2000) lists at least nine first person pronominal forms in Bangkok Thai; *chan, dichan, phom, raw, kha:p, pha, caw, ku:, phuak*, where *raw* can denote either singular or plural reference (p. 83). The rest of the personal pronouns can be used for singular or plural, and many of them can refer to different grammatical persons as well (Smyth, 2000, p. 41-42). In Thai, the answer to the question in (5b) is *no*. A comparison chart for English, Khmer, and Thai is shown in Table 3 below.

Table 3. Comparative table of English, Khmer, and Thai

Criteria	English	Khmer	Thai
4a. Obligatory	Yes	No	No
4b. Agreement	Yes	No	No
4c. Inside compounds	No	Yes	No*
4d. Inside deriv. m'y	No	No*	No*
5a. Grammatical gender	Yes	No	No
5b. # distinction pronouns	Yes	No	No

4 Discussion

The data from English, Khmer, and Thai in Section 3 clearly show certain patterns exhibited by inflectional (English) and non-inflectional (Khmer, Thai) plurals. The only "unexpected" results stemming from Wiltschko's (2008) criteria are that while Khmer and Thai appear to be non-inflectional plural marking languages, their plural forms do not (often) seem to appear inside of compounds or derivational morphology. This discrepancy is not unexpected, as Wiltschko (2008) developed these two criteria (4c) and (4d) in an analysis of Halkomelem, which is a polysynthetic language whose plural marking affixes are considered neither inflectional nor derivational (Wiltschko, 2008, p. 640). This language allows for plural marking affixes as well as derivational or inflectional ones to appear within a single word, whereas Khmer and Thai are highly isolating and would not permit the type of constructions described in the criteria for the most part.

As Khmer and Thai follow the rest of the criteria as if they are non-inflectional plural marking languages, I believe they still fall into this category based on the other characteristics of these languages. Criteria (4c) and (4d) appear to be more relevant in the diagnostic process of identifying inflectional plural marking languages but are not necessarily critical to the identification of non-inflectional plurals. In any case, these two criteria do not appear to be relevant for answering the question of whether non-inflectional plurals are comprised of phi-features because they are more descriptive regarding the type of plural marking a language has rather than that plural marker's implications on its syntax.

As for the two new criteria (5a) and (5b), English appears to pattern one way while Khmer and Thai pattern in another way. While criteria (5a) does not make any grand assumptions about the plural marking in these languages, it does support the tendency of languages with grammatical gender (e.g. English) to also have grammatical number, while Khmer and Thai lack both. Characteristics of the personal pronoun systems in these languages were revealed with criteria (5b). While, again, this criterion does not show any results that are surprising, it does support the idea that if a language's pronouns do not distinguish for number there will likely not be any grammatical number in the language. English personal pronouns do have certain number distinctions, and it follows that the rest of the nouns in the language might as well. As for Khmer and Thai, no (or very restricted) number distinctions are observed in the pronouns and this applies to their regular nominals as well.

5 Conclusion

As seen in the preceding sections of this paper, there appear to be distinct differences between inflectional and non-inflectional plural marking languages. There is no obviously apparent reason to suggest that the two non-inflectional plural marking languages discussed in this paper, Khmer and Thai, have a plural phi-feature in their grammars. Wiltschko's (2008) proposed criteria for distinguishing inflectional from non-inflectional plural marking proved to be quite revealing in terms of the presence or absence of phi-features, but to additional criteria were investigated in order to support the pre-existing ones.

The criterion identified in Section 3 as being potentially the most relevant to determining the presence or absence of phi-features (criterion 4b) remains so, as the languages with non-inflectional plural markers do not have phi-based agreement systems, while English has inflectional plural markers and does have phi-based agreement as well.

Further steps to be taken to determine whether non-inflectional plural markers contain a plural phi-feature would be to examine other phi-based characteristics of a language such as grammatical gender (this was examined elementarily in this paper), grammatical person, case, and definiteness. If languages that appear not to have phi-based number systems lack other phi-based characteristics in the language, if other indications point to a lack of phi-features it may lead to an interesting finding that there is the possibility of a language without any phi-features at all.

References

Corbett, G., (2000). *Number*. Cambridge, MA: Cambridge University Press.

Everett, D., (1986). Pirahã, In D.C. Derbyshire & J.K. Pullum (eds) *Handbook of Amazonian Languages*, 200-325, Mouton de Gruyter, Berlin.

Greenberg, J.H., (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J.H. Greenberg (ed) *Universals of Language*, 73-113, Cambridge, MA: MIT Press.

Haiman, J., (2011). *Cambodian: Khmer*, John Benjamins Publishing Company.

Huffman, F., (1970). *Modern Spoken Cambodian* (1998 ed.), Ithaca, NY: Southeast Asia Program Publications.

Kanchanawan, N., (1978). Expression for time in the Thai verb and its application to Thai-English machine translation. Austin, TX: ProQuest Dissertations Publishing.

Kullavanijaya, P., (2000). Power and Intimacy: A Contradiction in a Thai Personal Pronoun. *Oceanic Linguistics Special Publications* 29: 80-86.

Smyth, D., (2000). *Thai: An Essential Grammar*, Taylor & Francis Group.

Witschko, M. (2008). The syntax of non-inflectional plural marking. *Natural language & linguistic theory* 26: 639-694.

Contact Information

Brittany McDonald

brittany.mcdonald@ucalgary.ca
MA Student, Linguistics
School of Languages, Linguistics, Literatures & Cultures
University of Calgary
CHC 211
2500 University Drive NW
Calgary, AB T2N 1N4

Right-Edge Phonological Phenomena in Kaqchikel

Brett C. Nelson

University of Calgary

Abstract

This paper examines a slew of phonological phenomena that occur at the right edge of the prosodic word in Kaqchikel (cak), a Mayan language of Guatemala spoken by about 400,000 people (Heaton & Xoyón, 2016). Based on previous phonological work by Brown, Maxwell, & Little (2006) and Bennett (2018), I first introduce the phonemic inventory (22 consonants, 10 vowels) and prosodic structure of Kaqchikel, with the latter being composed of primarily stress-final, recursive prosodic words, and intonational prominence on the right-edge of the phrase. This is followed by a review of Bennett's (2016b) discussion of the Kaqchikel tense-lax distinction in vowels, which only surfaces in stressed (word-final) syllables. Thus, an underlying lax vowel { ɪ ɛ ə ɔ ʊ } surfaces as its corresponding tense vowel { i e a o u } in any unstressed syllable. I next discuss final aspiration of stops, and then spirantization of final sonorants, unifying them as a process of epenthesis of a [spread glottis] feature at the right edge of the word. Each phenomenon individually shows that the right-edge is a position of particular prominence in Kaqchikel; all together they demonstrate it is one ripe for future (and current) exploration into their acoustic correlates and their higher-level prosodic and morpho-syntactic implications.

Key words: phonology, right-edge, allophony, tense/lax, aspiration, spirantization, spread glottis

Languages: Kaqchikel (cak), English (eng), Spanish (spa)

Territorial Acknowledgement

This research was carried out upon traditional territories of several indigenous groups, upon which I was/am simply a guest. This includes Kaqchikel territory in southern Iximulew (Guatemala) between Armita (Guatemala City) and Lake Atitlán, most research of which took place in the Pan Choy/Pan Q'än valley where Antigua Guatemala now sits within the department of Sacatepéquez. Furthermore, due to the armed conflict (1960s-1990s) there are now thousands of displaced K'ichee' and Ixil Maya within those lands, which often bear Nahuatl-derived names. Additional analysis took place in Treaty 7 region in Southern Alberta, which includes the Blackfoot Confederacy (comprising the Siksika, Piikani, and Kainai First Nations), the Tsuut'ina First Nation, and the Stoney Nakoda (including the Chiniki, Bearspaw, and Wesley First Nations). Mohkinstsis (The City of Calgary) is also home to Métis Nation of Alberta, Region III. [1]

1 Introduction*

In describing the prosody of Kaqchikel, Brown, Maxwell, & Little (2006) noted that "word stress in Kaqchikel is generally on the final syllable of a word," and "in phrases, the primary stress falls on the last word" (p. 138). Additional prominences at the right edge are also discussed in work by Bennett (2016a; 2016b) with allophony in both vowels: "Lax vowels are restricted to the stressed syllable… which is almost always the ultimate syllable of the word" (Bennett, 2016b, p. 3), and consonants: "Plain stops are typically aspirated in word-final position" (Bennett, 2016a, p. 486) with a "parallel pattern of word-final sonorants devoicing" (p. 487).

Kaqchikel, a K'ichee'an language within the Mayan language family, is spoken by ~400,000 people in southern Guatemala, between Lake Atitlán and Guatemala City (Heaton & Xoyón, 2016). This area is shown in Figure 1. Most of these speakers are bilingual, also speaking Spanish (spa), the national language of Guatemala. Many are also able to speak other Mayan languages or English (eng), and the data analyzed in this paper were produced from wordlists and narrations by native speakers of Kaqchikel who were fluent in Spanish and had experience with English.

[1] Super special thanks to the native speakers of Kaqchikel who participated in this study, Aq'ab'al, B'alam, Ixnal, Kawoq, and Yab'un. Significant portions of this paper were developed after discussions with Dr. Darin Flynn, my supervisor, and Dr. Judith M. Maxwell, my former MA adviser. I thank them immensely. I also wish to thank the teachers, students, alumni, and associates of the Oxlajuj Aj Kaqchikel language and culture course. None of this would be possible without the initial and countless following steps I took into this language community with them there. I also appreciate the input from the attendees at NoWPhon 2019.
* The paper was presented at the *Northwest Phon{etics;ology} Conference* held on September 20th, 2019, at the University of Calgary.

Figure 1: Map of Guatemala with Kaqchikel area outlined. CC BY-SA 3.0 (https://commons.wikimedia.org/wiki/File:Idiomasmap_Guatemala.svg)

The body of this paper proceeds as follows: first, in section 2, I present the phonological inventory of Kaqchikel. This is followed by a discussion of the prosody of the language in section 3 with exemplifying spectrograms. Section 4 contains a discussion of the segmental processes that occur at the right edge of the word, and possibly elsewhere, in the language, again with accompanying spectrograms. This leads into a discussion of a previous analysis of Kaqchikel allophony by Nasukawa, et al. (2018), which was done under the framework of Element Theory. The ultimate section concludes by summarizing the findings of this paper.

2 Phonemic Inventory

Standard Kaqchikel has a phonological inventory of 32 phonemes, though, as discussed in this section, many speakers make fewer than this maximal number of distinctions. The consonantal inventory of the language is one typical of the Mayan language family and is presented as Table 1 (after Brown, Maxwell, & Little, 2006; Bennett, 2016a).

Table 1: Kaqchikel Consonantal Inventory

Place / Manner	Bilabial	Alveolar	Palatal	Velar	Uvular	Glottal
Plain Stop	p	t		k	q	ʔ
Glottalized Stop	ɓ	tʼ		kʼ	ɠ	
Plain Affricate		ts	tʃ			
Glottalized Affricate		tsʼ	tʃʼ			
Fricative		s	ʃ	x		
Nasal	m	n				
Lateral Approximant		l				
Approximant		ɾ				
Glide	w		j			

As shown, there are 22 consonantal phonemes spread across six places of articulation. The primary consonantal contrast among consonants is one of glottalization and is exhibited among the stops and affricates of the language. Thus, there are two series of four stop consonants each, and two series of two affricates each. Note that at the peripheral places of articulation (bilabial and uvular) the glottalized stops are realized as voiceless implosives, while all other glottalized consonants are ejectives. The glottal stop does not have any counterpart at its place of articulation. In addition to these 13 stop-like consonants, Kaqchikel has three (3) fricatives /s, ʃ, x/, two (2) nasal consonants /m, n/, and four (4) non-nasal sonorants /l, ɾ, w, j/. These three groupings become phonologically apparent after discussion of phonological processes in section 4.

To complete the segmental inventory, the ten (10) vowels of Kaqchikel are shown in Table 2. These ten vowels are distributed across five (5) general Places of Articulation, based on features of Height and Backness. Each of these Places has two (2) distinctive vowels: one Tense and one Lax. An interesting note about the Lax vowels is that they are reflexes of Proto-K'ichee'an Long vowels and are cognate with modern K'ichee' Long vowels (Campbell, 1977; Bennett, 2016a). Vowel length is therefore not distinctive in modern Kaqchikel, differing from most other Mayan languages.

Table 2: Kaqchikel Vocalic Inventory

	High, Front	**Mid, Front**	**Low, Front**	**Mid, Back**	**High, Back**
Tense	i	e	a	o	u
Lax	ɪ	ɛ	ɜ~ə~ɨ	ɔ	ʊ

Not all speakers exhibit the Lax distinction at every Place, with most speakers having between seven (7) and nine (9) phonemic vowels. The distinction is almost always made at the Low, Front Place of Articulation, and, as such, there is substantial inter- and intra-speaker variation in the phonetic realizations of the Lax counterpart at that Place, ranging from open [ɜ] through the most common realization of schwa [ə] to close [ɨ].

3 Prosody

With this knowledge of the segmental make-up of Kaqchikel, we can now move on to the prosodic suprasegmentals, namely word stress and phrasal intonation. One, and possibly more, of the segmental phenomena discussed in the subsequent section depends entirely upon the prosodic elements discussed here.

3.1 Word stress

As mentioned in the introduction, Brown, Maxwell, & Little (2006) described Kaqchikel as having word stress and that that word stress "is generally on the final syllable of a word" (p. 138). Bennett (2016a), reporting on Mayan languages more broadly, states that "final stress is the norm in K'iche[e']an languages" (p. 495). To verify these claims, this subsection examines words of various syllable lengths (one to four syllables), generally agreeing with the claims, although with a few exceptions.

Figure 2 shows the waveform and spectrogram for a production of monosyllabic *tuj* /tux/ 'sauna'. The single syllable in this noun trivially bears stress, and its vowel shows a rise and fall in intensity over its length.

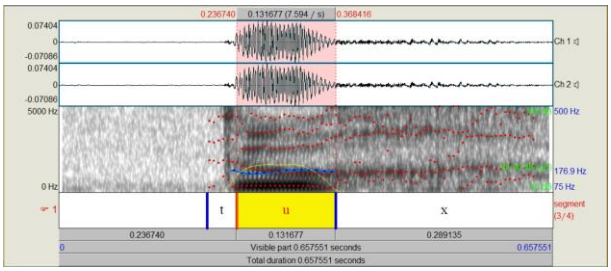


Figure 2: <tuj> /tux/ *produced by Aq'ab'al.*

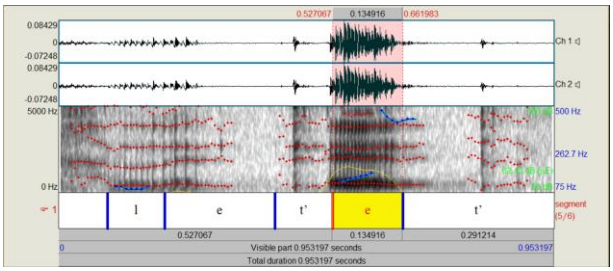


Figure 3: <let'et'> /leťeť/ *produced by Aq'ab'al.*

Moving up to a disyllabic example we have *let'et'* /leťeť/ 'bicycle' in Figure 3. Here we can see the first evidence in support of this final syllable analysis. This disyllabic word displays iambic stress, with stress on the second (i.e. final) syllable. Similar to the previous example, there is a marked rise and fall of intensity throughout the vowel of the final syllable, especially when contrasted with the first syllable's vowel.

Next we have the trisyllabic example in Figure 4, *chikopi'* /tʃikɔpiʔ/ 'animals'. This is an interesting case because the root of this word is *chiköp* 'animal', with //-iʔ// an inflectional (plural) suffix. Nevertheless, the stress appears on the rightmost syllable, that of the suffix. The acoustic realization of this stress is less clear than the previous examples, however. We do have the stressed syllable's vowel bearing a higher intensity, however the prior syllable also shows this. Both syllables are much more intense than the first syllable, which does appear to be more similar acoustically to the first, unstressed syllable observed in Figure 3. What that second syllable does not show, though, is a rise in pitch, which the final, stressed syllable does bear.

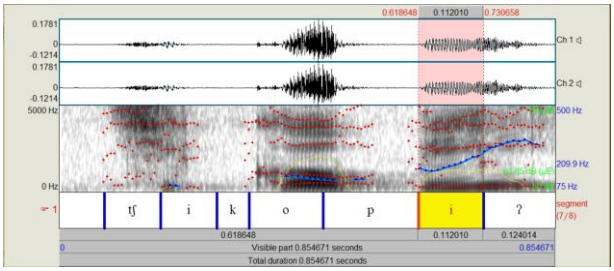


Figure 4: <chikopi'> /tʃikɔpiʔ/ *produced by Aq'ab'al.*

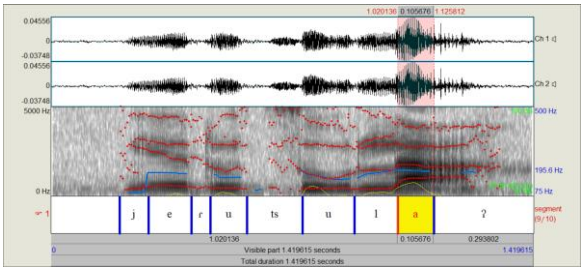


Figure 5: <yerutzula'> /jerutsulaʔ/ *by Speaker 1.*

Again, adding another syllable, we move on to the tetrasyllable *yerutzula'* /jerutsulaʔ/ 'he repeatedly looks at them' of Figure 5. This example shows another part of Kaqchikel phonology, with the verb root //tsuʔ// being inflected with the frequentative suffix and also with agreement prefixes. Again, however, the stress falls on the final syllable of the whole word, though the relative intensity measure of its vowel is a greater indicator in this example than the previous one. The pitch in this example, on the other hand, does not seem to be an

indicator of stress here, as it remains flat during the nucleus [a] after a slight rise during the Onset [l].

The preceding examples show that Kaqchikel does in fact place main word stress on the final syllable of the word. However, there are a few exceptions to this that must be mentioned. Word-final stress is by far the most common pattern for native vocabulary, yet there are examples of native vocabulary that do not match this pattern. Furthermore, non-native vocabulary is more variable, depending on the source language for its stress placement.

Figure 6 shows an example of native vocabulary, *janila* /xanila/ 'very', and how it does not have word-final stress. Instead, the higher relative intensity as well as the pitch rise are realized on the second (penultimate) syllable. This shows that stress in Kaqchikel is lexical, though the ultimate syllable is preferred. Note also that the standard orthography does not indicate stress in any way.

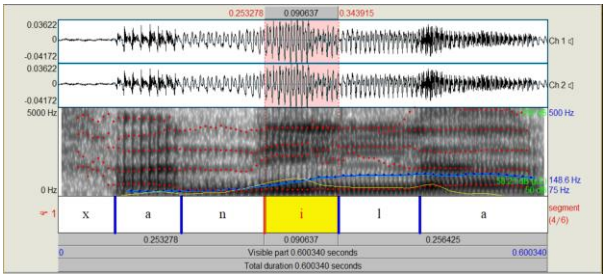


Figure 6: <janila> /xanila/ produced by Kawoq.

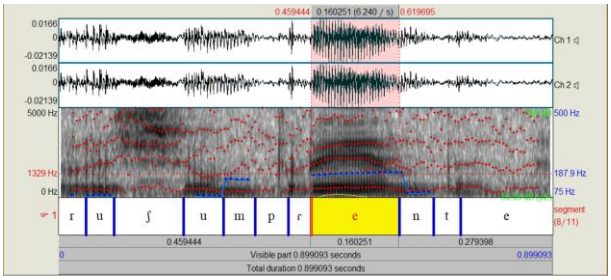


Figure 7: <ruxumprente>/ɾuʃumprente/; Speaker 1.

The final example in the subsection (Figure 7) shows a loanword (from Spanish), *ruxumprente* /ɾuʃumprente/ 'his (little) hat'. The Spanish source is *sombrerete* and bears typical Spanish penultimate stress. However, when adapted to Kaqchikel the antepenultimate and penultimate syllables contract into a single syllable, and that syllable bears stress, with its vowel being relatively more intense and longer.[2]

These examples show that, while Kaqchikel overwhelmingly prefers word-final stress, there is no absolute restriction against other patterns. Instead of stress being inherited from the prosodic hierarchy, it is lexically bound, so that lexical items determine the word-level prosody of the language.

3.2 Intonation

Moving up the prosodic hierarchy from the word to the phrase, this subsection discusses the phrase-level intonation claim of Brown, Maxwell, & Little (2006), that phrase-level prominence is realized on the last word of the phrase. The two examples shown here are drawn from a video-narration task (of the Pear Film (Chafe, 1980)) produced by the same

[2] This syllable also happens to be exceptionally complex for a Kaqchikel word-medial syllable, showing a complex onset and a coda, both of which are rare in Mayan languages which typically have /CVC/ roots with /CV-/ prefixes and /-VC/ suffixes (see Bennett, 2016a).

speaker. These data are very preliminary, and are the initial step in a process of an in-depth documentation and analysis of the intonation of Kaqchikel for the first time (Bennett, 2016a).

The first phrasal example is shown as Figure 8: *y k'a ri xuxïm chi ruqul.* 'and then he tied it to its neck.', a simple declarative with a discourse particle preceding it. At the right edge of this phrase there is a larger fall in pitch across the duration of the final syllable of the final word *ruqul* 'its neck'. Loudness modulates with each word, as expected with word stress, but at no other point other than the over the final syllable of the final word does the pitch change so drastically.

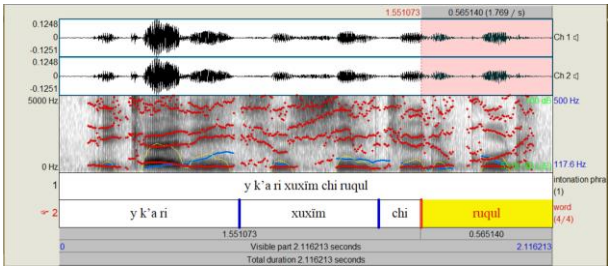


Figure 8: Phrase by Aq'ab'al: '... he tied it to its neck.'

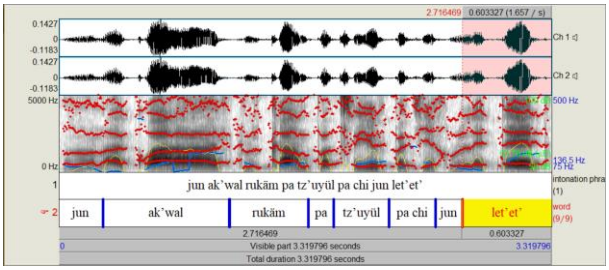


Figure 9: Iambic pentameter spoken by Aq'ab'al.

The next example comes from a more complex sentence featuring an intransitive verb modified by two locative/prepositional phrases: *jun ak'wal rukäm pa tz'uyül pa chi jun let'et'* 'a child is coming seated on a bike'. This is shown in Figure 9. Again the stress-indicating loudness is apparent in each word, but note again that at the right-most word of the utterance *let'et'* (and possibly of the initial intonational phrase *ak'wal*) features a sharp rise in pitch in the Onset of the final syllable followed by a slight fall in its nucleus. These two prominences create a salient boundary tone, which led Brown, Maxwell, & Little (2006) to describe declarative assertions as being marked with a falling tone on their respective final syllable. Brown, Maxwell, & Little (2006) further note that content (wh-) questions also bear a falling boundary tone, while polar (yes/no) questions carry a rising boundary tone. Due to the nature of the tasks, however, no questions appear in the current data set that confirm these claims.

4 Segmental processes at the right edge

Members of both segmental sets of vowels and consonants participate in particular phonological processes at the right edge of the word in Kaqchikel. This section discusses both of those sets. First, we consider the primary vocalic distinction of tense/lax and where it surfaces in stressed syllables. This is followed by a discussion of the consonants, particularly plain stops and sonorants.

4.1 Vowels

As introduced in section 2, the primary phonemic distinction among Kaqchikel vowels is one between a tense series and a lax series. However, this distinction does not surface in all positions of a word. In fact, lax vowels are restricted to stressed (word-final) syllables

(Bennett, 2016b). When affixation displaces an underlying lax vowel from the stressed syllable, it will always surface as its tense counterpart. Thus, the presence of a lax vowel serves as an indirect method to identify word stress, though the presence of a tense vowel does not serve as an indicator of non-stressed syllables, as they too may surface in stressed syllables.

Perhaps because of their limited surface distribution, there exists copious variation in the realization of the lax vowels cross-dialectically as well as within a single speaker. Indeed, not all dialects produce the tense-lax distinction for every vowel pair, with very few speakers having ten phonemic vowels (Patal Majzul, García Matzar, & Espantzay Serech, 2000). Furthermore, speakers of southern dialects of Kaqchikel, which provide the sample data for the current study, tend to have fewer distinctions than speakers of other dialects. Nevertheless, the next three examples strive to show some of those distinctions being produced by a speaker of one of those southern dialects.

In Figure 10 we have *tinamït* 'town; village', which has a high front vowel in both its first and last syllable. However, while the first syllable contains a tense vowel /i/, the final syllable's vowel is lax /ɪ/. As this syllable is stressed, the lax vowel is able to surface. This can be seen through the vowels' formant values. The unstressed tense vowel has midpoint F1 of 358hz and F2 of 2308hz, while those same values for the stressed lax vowel are 490hz and 1868hz, giving -132hz difference between their F1 and 440hz between their F2. Note also the increased length of the stressed lax vowel.

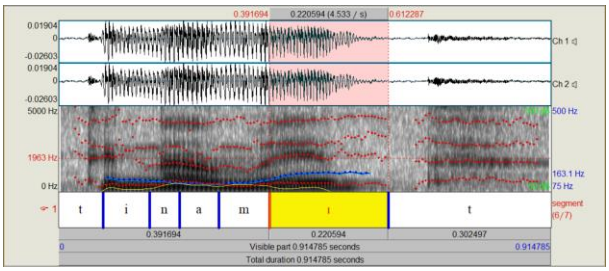


Figure 10: <tinamït> /tinamɪt/ *by Speaker 3.*

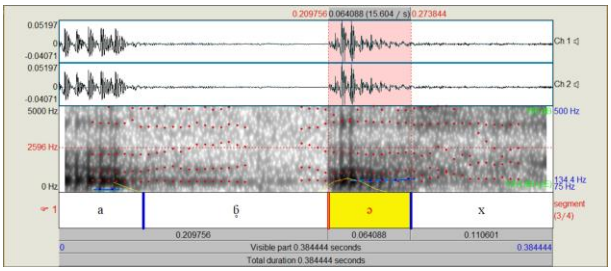


Figure 11: <ab'äj> /aɓ̥əx/ *produced by Speaker 3.*

The next pair of examples come from *ab'äj* 'stone' in Figure 11, which has both members of the low vowel pair /a-ə/. The lengths of these vowels are much more even, as are their formant values. The unstressed tense vowel's F1 is 650hz and its F2 is 1216hz, while the stressed lax vowel's F1 is 640hz and its F2 is 1079hz, for a difference between the vowels' F1 and F2 of 10hz and 137hz respectively, giving a slight distinction on F2.

The last pair of vowels examined in this paper are the mid back vowels /o/ and /ɔ/ in *jotöl* 'in a raised position' of Figure 12. The unstressed tense vowel is actually longer here, but the stressed lax vowel is more intense and has a higher pitch. The F1 values for these two vowels are 488hz and 503hz (-15hz difference), while the F2 values are 1115hz and 1299hz (-184hz difference). Thus, the lax vowels that surface only in stressed positions exhibit a more centralized realization, and that is particularly shown on F2.

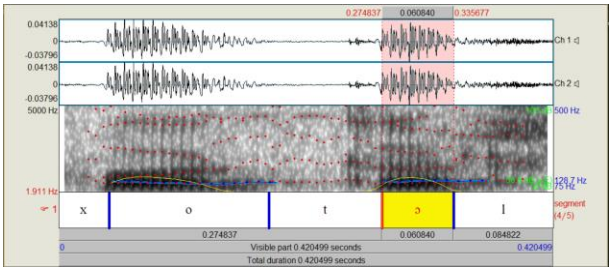


Figure 12: <jotöl> /xotɔl/ *produced by Speaker 3.*

4.2 Plain stops

The set of plain stops exhibit allophony of a different nature at the right edge. These four stops may each appear in both onset and word-final coda positions. This section shows examples of these consonants in those positions, first in Onset and then in word-final Coda. Note that all of these examples are CVC monosyllabic words, except for one. All syllables with the target stops contain the vowel [i].

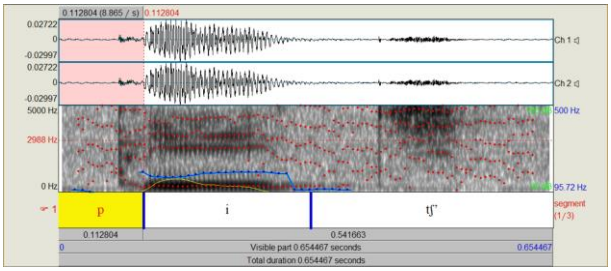


Figure 13: <pich'> /pitʃ'/ *produced by Yab'un.*

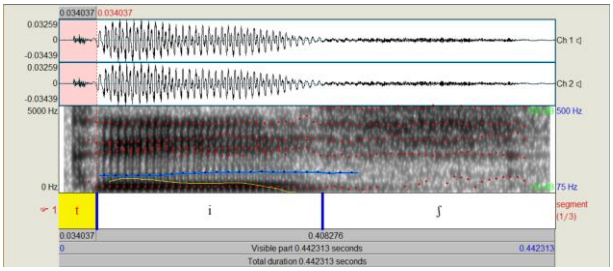


Figure 14: <tix> /tiʃ/ *produced by Yab'un.*

In Figure 13, *pich'* 'tender corn', we have the Onset example for the bilabial /p/. Following its initial burst, there is brief 32ms VOT, and only light frication during that time. A similar pattern is seen in the alveolar example in Figure 14 *tix* 'tapir', with a VOT of 23ms and similar levels of frication.

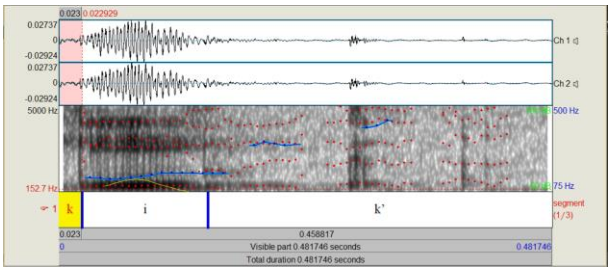


Figure 15: <kik'> /kik'/ produced by Yab'un.

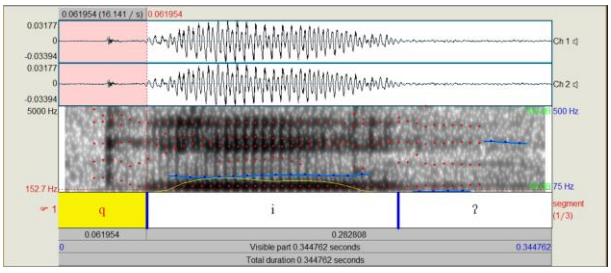


Figure 16: <qi'> */qiʔ/* produced by Yab'un.

The next two Onset examples are shown in Figure 15 *kik'* 'blood' and Figure 16 *qi'* 'ourselves', with the velar and uvular stops respectively. The velar stop has 20ms of VOT and minimal frication, while the uvular stop is produced with 29ms of VOT and little frication after the initial burst. These examples lead to the characterization of these stops in Onset as having

short lag voicing and little to no frication during that lag. They are plain, unaspirated voiceless stops.

The next four examples illustrate these same phonemes in word-final coda position. First, in Figure 17, we have the bilabial example *sip* 'tick', which is produced here with an extended release allowing for some frication across it. Though much less apparent than the following examples, it is transcribed as being an aspirate here due to its extended release.

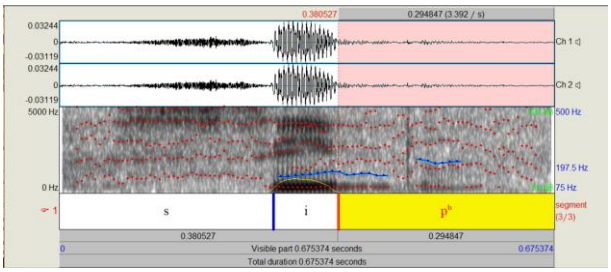


*Figure 17: <sip> /*sip*/ produced by Yab'un.*

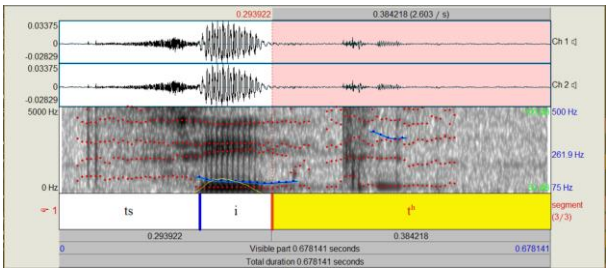


Figure 18: <tzit> /tsit/ produced by Yab'un.

The examples in Figure 18 *tzit* 'a little bit' and Figure 19 *jik* 'straight' display much more apparent aspiration/frication. The former, alveolar example has both an apparent period of silence prior to release of the stop, and 84ms of frication after the initial burst. The latter, velar example also has the offset after the vowel and 98ms of frication.

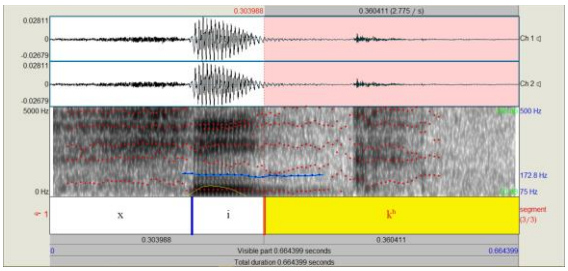


Figure 19: <jik> /xik/ produced by Yab'un.

The final stop example is shown in Figure 20 *nib'iq* 'it is degrained'. Although disyllabic, the target consonant remains as the Coda of the stressed syllable of the word. Again, we have a period of relative silence, followed by the initial burst and a period (111ms) of frication.

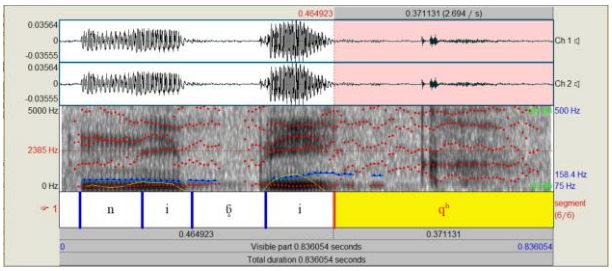


*Figure 20: <nib'iq> /*nɓiq*/ produced by Yab'un.*

These examples show that Kaqchikel plain stops /p t k q/ surface as aspirated stops /p^h t^h k^h q^h/ in word-final coda position, and that aspiration is characterized by a noisy release burst. Several possibilities exist for phonological processes that cause this allophony. Under the set of classical features of Featural Phonology, however, only the insertion of [continuant] and [spread glottis] are logical possibilities. At this point both are viable here. However further probing of Kaqchikel consonantal allophony changes that.

4.3 Non-nasal sonorants

The final set of Kaqchikel sounds analyzed in this paper are the non-nasal sonorants, of which there are four: two liquids /l ɾ/ and two glides /w j/. As with the previous section, in this section these four phonemes are exemplified in both Onset and Coda, in order to illustrate their positional allophony. We begin with these four in Onset.

The example shown in Figure 21 is *yesolon* 'they untie', with the segment of interest being the Onset of the final syllable, [l]. This liquid is produced with modal periodic voicing, with only slightly less intensity than the adjacent vowels. Thus, it is transcribed here as [l], a voiced sonorant.

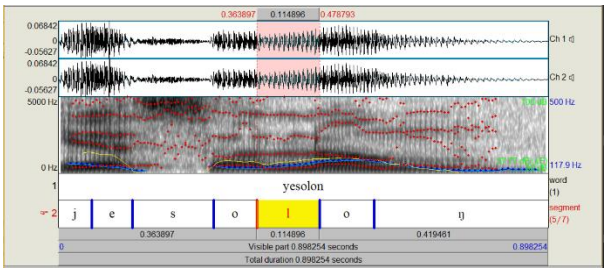


Figure 21: <yesolon> /jesɔlon/ *produced by Kawoq.*

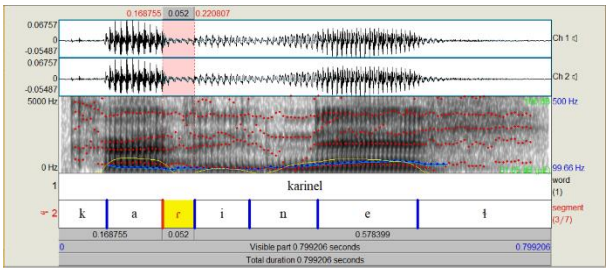


Figure 22: <karinel> /kəɾinel/ *produced by Kawoq.*

Moving on to Figure 22, we have the example for /ɾ/: *karinel* 'fisher'. Although this particular example removes the target segment from stressed position, we can see that it remains voiced and again, slightly less intense than the adjacent vowels, with no aperiodic frication. All of this leads to the transcription as [ɾ], the voiced sonorant.

Next we have the labiovelar glide /w/, shown by *82etwar* 'it gets cold' in Figure 23. Though the formants are distinct from the adjacent vowels, the segment retains semivocalic properties of those formants, modal voicing, and absence of aperiodic noise. In onset position, /w/ goes unchanged and surfaces as [w].

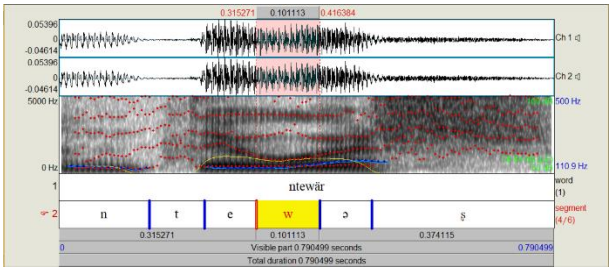


Figure 23: <ntewar> /ntewəɾ/ *produced by Kawoq.*

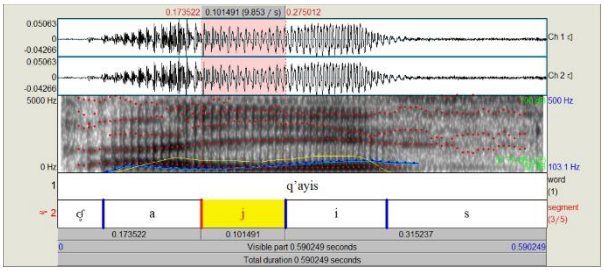


Figure 24: <q'ayis> /ɢˀəjis/ *produced by Kawoq.*

The final Onset analyzed in this paper is an example of the palatal glide /j/, shown by *q'ayis* 'weeds; trash' in Figure 24. As with the previous example, the semivocalic properties of this sound are apparent, especially its formants and lack of aperiodic noise. The voicing has a somewhat lower frequency than the following [i], and this, combined with the fact that adjacent vowels are disallowed in Kaqchikel leads us to the conclusion that this too surfaces unchanged as a voiced glide [j].

The preceding four examples show that the underlying forms of these sonorants do not change when they surface as Onsets. They display modal voicing and little aperiodic frication. The next four examples, all of which are based on the same roots as the previous four, show that this lack of allophony is again only found in onset position. Word-final coda position causes a particular allophonic pattern; one that will be explained in the following paragraphs.

The first Coda example of these sonorants shown in Figure 25 is *nkisöl* 'they untie it', the active voice form of *yesolon*. The target segment would be difficult to find were it not at the end of the word because its voicing has completely disappeared and instead has been replaced with aperiodic frication. The underlying /l/ surfaces in this word-final Coda as a lateral voiceless fricative [ɬ].

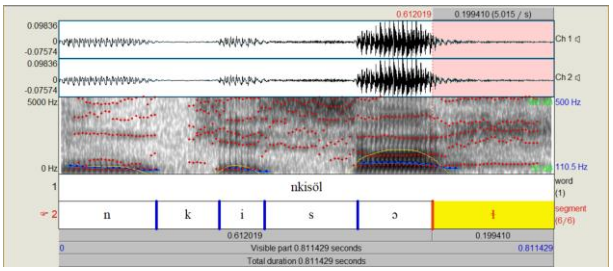


Figure 25: <nkisöl> /nkisɔl/ *produced by Kawoq.*

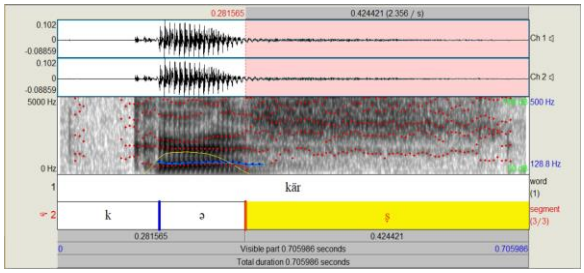


Figure 26: <kär> /kəɾ/ *produced by Kawoq.*

The example in Figure 26 shows the surface form of the underived root of *karinel*, *kär* 'fish', although the /ɾ/ in coda position (along with the preceding Lax vowel) obscure that relation at the surface. The /ɾ/ here is much longer than in *karinel*. It also has lost its voicing and is instead realized with voiceless frication; it surfaces as the retroflex fricative [ʂ].

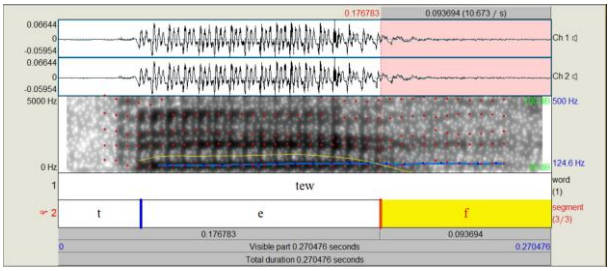


Figure 27: <tew> /tew/ *produced by Kawoq.*

The penultimate example of this paper shows /w/ in word-final Coda: *tew* 'cold' in Figure 27. As with the previous two examples, the segment in question here is not voiced, except perhaps briefly at its onset, is longer than its Onset counterpart, and is full of aperiodic frication. The underlying /w/ surfaces here as a voiceless labial fricative [f] or [ɸ] (the latter observed by Nasukawa et al., 2019).

Finally, we have Figure 28. Here, *nq'äy* 'it rots' reveals the Coda allophone of /j/. Again, there is the complete absence of the low-frequency voicing band, but extreme prevalence of high-frequency, aperiodic noise. This voiceless fricative retains the palatal place of articulation of its underlying sonorant, and surfaces as the palatal fricative [ç].

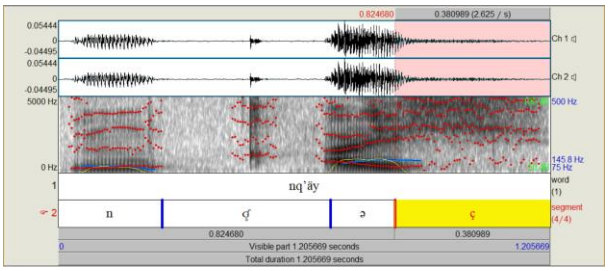


Figure 28: <nq'äy> /nɰ̊əj/ *produced by Kawoq.*

These last eight examples have demonstrated that there is a positionally motivated allophony for sonorants, just as there was for the voiceless stops in their eight examples. While Onsets retain their underlying form of (voiced) sonorants, the Coda allophones are all voiceless fricatives at identical or nearby places of articulation. Brown et al. (2006) describe these allophones as voiceless varieties of those sonorants, however Lombardi (1991) and Clements (1985) and Mester and Ito (1989) before her, argue that these are fundamentally, and featurally, equivalent: they are both underlyingly aspirates.

The positional allophony exhibited by Kaqchikel non-nasal sonorants closely mirrors the stop allophony described in section 4.2, and, following Vaux (1998) there is one phonological feature that unites these two processes of spirantization and aspiration, [spread glottis]. Vaux (1998), in examining data from numerous languages, found that they suggest [+ spread glottis] is present in unmarked voiceless fricatives, in other words there is an inherent connection between voiceless fricatives and [spread glottis], and in systems that do not contrast fricatives laryngeally, these fricatives nonetheless adopt [spread glottis]. Thus adding [spread glottis] to the non-nasal sonorants, in the absence of a [voice] feature, creates the voiceless fricatives observed in Kaqchikel word-final Codas.

Insertion of [spread glottis] also achieves the aspiration exhibited by Kaqchikel plain stops in word-final Coda. Lombardi (1991) among others, have shown that aspirated stops occur due to an [aspiration] feature, however, admits that this, again, is fundamentally equivalent to [spread glottis]. Thus, the same feature that can account for the spirantization of sonorants can account for the aspiration of stops. The right edge of the word, and possibly every syllable, is marked by the insertion of [spread glottis].

5 Nasukawa, et al. (2019): Element Theory

The current paper is not the first examination and unification of these two processes as a single edge-marking process in Kaqchikel. Nasukawa et al. (2019) examine similar data to come to a similar conclusion: there is a single rule of insertion that can account for both sets of allophony. They use this to argue that Kaqchikel has preference for Coda rather than the near-universal preference for Onset.

However, Nasukawa et al (2019) argue for this under the framework of Element Theory. This theory is similar to traditional featural theories of phonology in that every phonological unit can be accounted for with smaller atomic units. However, under Element Theory, there are only six of these atomic Elements: three vocalic in nature and three consonantal in nature (Backley, 2011). The critical element for Nasukawa et al.'s (2019) analysis is the [H] element, which they name 'Edge'.

This Edge element, Nasukawa et al. (2019), claim, is added at the right edge of the prosodic unit of the syllable to mark its prominence. This is realized among the voiceless stops as aspirated stops, and among the four non-nasal sonorants as voiceless fricatives. However, adding [H] to sonorants only produces voiced fricatives, therefore another [H] must be added to achieve the desired outcome of voiceless fricatives. Thus, their unified analysis is not unified: Stops receive one [H] element in Coda, while sonorants receive two.

Furthermore, Nasukawa et al. (2019) do not account for why only these classes of consonants exhibit this kind of positional allophony, only mentioning the other classes of consonants in passing while describing the whole inventory. Indeed, only these two classes of plain stops and non-nasal sonorants. However, under the [spread glottis] view argued for here, their application to the other classes can be argued away in a series of ways. Most simply for the three fricatives /s, ʃ, x/, these are already voiceless fricatives, so inherently have [spread glottis]. Inserting that feature again does not change anything, therefore they do not change. The glottalized consonants of Kaqchikel have [constricted glottis] underlyingly. This feature is antithetical to the newly inserted [spread glottis], so nothing happens.

Nasals, of which there are two phonemes in Kaqchikel, also do not exhibit any analogous allophony.[3] A possible explanation for this may lie in the particular feature geometry (Clements, 1985) underlying these segments that prevents [nasal] and [spread

[3] There is, however, an ongoing process of velarization of nasals in word-final position in some dialects of Kaqchikel not analyzed here:

/m/ > [n] / _#

/n/ > [ŋ] / _#

glottis] from co-existing. Nonetheless, Nasukawa et al. (2019) do not provide any explanation for these consonants' lack of allophony, and it remains unclear how their implementation of Element Theory could account for those cases.

6 Conclusion

This paper has surveyed the various phonetic, phonological, and prosodic phenomena that occur at the right edge of prosodic units in Kaqchikel. First, Kaqchikel stress was shown to be overwhelmingly bound to the right edge of the word, no matter the length of the word. There are, however, few examples of both native and loan words that do not bear this pattern. With these word stresses, phrase level intonational prominences also appear at the right edge of the phrase, with the final syllable bearing a phrasal boundary tone.

Allophonic variation is pervasive at the right edge of the word in Kaqchikel. Vocalically, this is seen in the variation between Tense and Lax vowels. Lax vowels, which are cognate with long vowels elsewhere in the K'ichee'an branch of Mayan, only surface as Lax in stressed, word final position. Outside of these positions, these vowels surface as their Tense counterpart.

Consonantally, there exist two seemingly parallel patterns of allophony among plain stops and non-nasal sonorants when found at the right edge of the word. Plain stops surface as aspirated stops, while the sonorants spirantize to voiceless fricatives. The current proposal argued for here is that this is due to the marking of the right edge of the word via a [spread glottis] feature. All of the examples analyzed here show this to be occurring at the right edge of the word, though it may in fact be the case that this process applies at the right edge of every syllable. Further analysis of data should shed light on the precise locus of this process. Additional statistical analyses of the acoustics and frequency distributions of these processes may also provide critical insight into the phonological possibilities of not only Kaqchikel, but other languages manipulating these and other features.

These findings provide valuable insight to positions of interest in the study of the acquisition of Kaqchikel. Languages differ in the features that underly their segments, but they may also differ in how they manipulate the features they may otherwise share. How these types of differences may impact acquisition is a major question of future research. At the very least, these findings show that the right edge must not be ignored in such studies of acquisition.

References

Backley, P. (2011). *An introduction to Element Theory.* Edinburgh: Edinburgh University Press.

Bennett, R. (2016a). Mayan phonology. *Linguistics and Language Compass, 10*(10), 469-514.

Bennett, R. (2016b). La tension vocálica en el Kaqchikel de Sololá, Guatemala: un estudio preliminar. In *Las actas de la cuarta mesa de trabajo del Seminario PHONOLOGICA*. Mexico City: Colégio de México. Retrieved March 25, 2019, from https://www.academia.edu/29152041/La_tensi%C3%B3n_voc%C3%A1lica_en_el_kaqchikel_de_Solol%C3%A1_Guatemala_un_estudio_preliminar

Bennett, R. (2018). Recursive prosodic words in Kaqchikel (Mayan). *Glossa: A Journal of General Linguistics, 3*(1), 1–33.

Campbell, L. (1977). *Quichean linguistic prehistory.* University of California Publications in Linguistics, *81*. Berkeley, CA: University of California Press.

Chafe, W. (Ed.) (1980). *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production.* Norwood, NJ: Ablex.

Clements, G. N. (1985). The geometry of phonological features. *Phonology Yearbook, 2*, 225-252.

Brown, R. M., Maxwell, J. M., & Little, W. E. (2006). *¿La ütz awäch? Introduction to Kaqchikel Maya Language*. Austin, TX: University of Texas Press.

Heaton, R., & Xoyón, I. (2016). An Assessment of Linguistic Development in a Kaqchikel Immersion School, *Language Documentation & Conservation, 10*(April), 497–521.

Lombardi, L. (1991). *Laryngeal features and laryngeal neutralization.* (Doctoral dissertation). University of Massachusetts, Amherst. Retrieved from ProQuest Dissertations Publishing, (9132881).

Nasukawa, K., Backley, P., Yasugi, Y., & Koizumi, M. (2018). Challenging cross-linguistic typology: Right-Edge consonantal prominince in Kaqchikel. *Journal of Linguistics, 55*(3), 611-641.

Patal Majzul, F., García Matzar, L. P. O., & Espantzay Serech, I. C. (2000). *Rujunamaxik ri Kaqchikel chi' : Variación dialectal en Kaqchikel*. Guatemala, Guatemala: Cholsamaj.

Rill, J. (2013) *Vowel Tensing in Kaqchikel Maya.* Poster presented at the 37th Penn Linguistics Colloquium, Philadelphia, PA.

Vaux, B. (1998). The laryngeal specifications of fricatives. *Linguistic Inquiry, 29*(3), 497-511.

Contact Information

Brett C. Nelson

brett.nelson@ucalgary.ca
School of Languages, Linguistics, Literatures and Cultures
University of Calgary
Craigie Hall D 310
2500 University Drive NW
Calgary, AB, T2N 1N4
Canada

Distribution of PRO in Serbian Subjunctives

Dušan Nikolić

University of Calgary

Abstract

The paper investigates distribution of PRO in Serbian subjunctives by following the tenets of Landau's (2000, 2004) theory of control. The goal of the paper is to present theoretically-based and empirical evidence for two types of Serbian subjunctives: Type I and Type II, and to argue that PRO is generated and motivated independently of Case or Binding Theory. I make three main claims in this study. First, PRO in Serbian subjunctives depends on the interaction of Tense [T], Agree [Agr], and Reference [R]. Second, Type I subjunctives allow only PRO, which is anaphoric, or [-R], while Type II subjunctives allow a lexical DP or pro, which is [+R], or an independent reference. In addition, I argue that Type I subjunctives are untensed, whereas Type II subjunctives are tensed. Third, Landau (2004) assumes that [-R] prohibits PRO from being dispatched to the spell-out immediately, and that PRO enters Agreement with the matrix functional head. I argue that his theory has fallen short of providing conceptually strong arguments for such a stipulation. Instead, I propose Harwood's (2015) variable phase approach that heeds to more dynamic phase structure. By 'shifting' the phase to a higher level domain, that is to the matrix clause vP, PRO becomes accessible to the probe, and the derivation of subjunctives converges.

Key words: PRO, Serbian, subjunctives, tense, agreement

1 Introduction

Standard grammars (Stanojčić & Popović, 1992; Piper et al., 2005) group Serbian subjunctives among *da* (complementizer) + present constructions. This classification is mainly based on their structure, which is illustrated in (1).

1) a. Ona voli da radi.
She like.PRES.3sg that work.PRES.3sg
(She likes to work.)

b. Ona nastoji da bude poznati maneken.
She intend.PRES.3sg that be.PRES.3sg famous model
(She intends to work.)

Latest work on Serbian syntax maintains, though, that Serbian subjunctives should be treated as a separate syntactic unit analogous to subjunctives in other languages (Janić, 2012, Miškeljin, 2012; Todorović, 2012). Yet, irrespective of some recent analyses, descriptions of Serbian subjunctives are generally sparse in literature on Serbian syntax, and theoretical considerations that underlie the exploration of Serbian subjunctives are mainly absent from the research.

In this paper, I present a detailed exploration of Serbian subjunctives by making three major assertions. First, there are two types of Serbian subjunctives: Type I and Type II, both of which are complement clauses. Subjunctive types are determined based on the selecting predicates and the subjects generated in the complement clause. Type I subjunctives bear PRO, while Type II subjunctives are specified for a lexical DP or *pro*. This is shown in (2)a and (2)b, respectively.

2) a. Ona počinje da PRO broji.
She start.PRES.3sg that PRO count.PRES.3sg
(She is starting to count.)

b. Ona je odlučila da *pro* učestvuje.
She AUX decide.PAST.3sg that *pro* participate.PRES.3sg
(She decided to participate.)

Second, to classify Serbian subjunctives and systematically account for the distribution of PRO (and other empty categories), I adhere to Landau's (2000, 2004) theory of control whose fundamental claims are grounded in Chomsky's (2000, 2001) Minimalist approach. I, therefore, provide an analysis of Type I and Type II Serbian subjunctives. The analysis is shown in (3)a and (3)b.[1]

[1] I assume that there is a [T] feature that can be assigned to the head INFL. Therefore, I will use [T] to indicate the feature Tense and I^0 and IP to indicate the head INFL and the INFL phrase, respectively.

3) a. [$_{CP}$… DP… *v* [$_{CP}$ C^0 $_{[-T]}$ [$_{IP}$ PRO$_{[-R]}$ [$_{I'}$ I^0 $_{[-T, +Agr, -R]}$ [$_{VP}$ t$_{PRO\ [-R]}$]]]]]

Agree Agree [-T] Agree [+Agr, -R]

Agree [+Agr]

b. [$_{CP}$… DP… *v* [$_{CP}$ C^0 $_{[+T, +Agr, + R]}$ [$_{IP}$ [$_{I'}$ I^0 $_{[+T, +Agr, +R]}$ [$_{VP}$ DP/*pro* $_{[+R]}$]]]]]

Agree Agree [+T, +Agr, + R] Agree [+Agr, +R]

Third, according to Chomsky's (2001) phase approach (which constitutes an important part of Landau's (2000, 2004) theory), PRO finds itself within the IP of the embedded clause wherefrom it is sent to the spell-out. This renders it inaccessible for further syntactic operations. Thus, Landau's (2004) argument that PRO enters Agreement with the functional head of the matrix clause cannot hold. To disentangle this conundrum, I maintain Harwood's (2015) concept of the variable phase approach. In this approach, sub-numerations constitute a phase, but are not contingent on any particular head. I claim that the numeration has not been exhausted and that the CP does not need to be a phase. Instead, the vP of the matrix clause is the phase where PRO may freely enter Agreement with the functional head.

In the following section, I introduce Serbian subjunctives and theoretical assumptions that underlie the present analysis. The theoretical concepts comprise the summary of approaches and descriptions of PRO, and Landau's (2000, 2004) theory of control. In Section 3, I identify the main issues that have provided an impetus for the study, and I outline the objectives of the study. A comprehensive analysis of Serbian subjunctives and their interpretation is presented in Section 4. In Section 5, I summarize the main points of the study.

1.1 Serbian Subjunctives

Subjunctive structures in Serbian are embedded complement clauses preceded by the complementizer *da* (Janić, 2012). Their content is licensed by the predicate of the matrix clause that determines whether subjunctive clauses will allow tense variations or the tense will be restricted to the present tense.[2] For example, in (4)a and (4)b, the embedded clause is specified for the present and past tense, respectively, while the subjunctive clause in (4)c only permits the present tense.

4) a. On se nada da Marko dolazi.
He self hope.PRES.3g that Marko come.PRES.3sg
(He hopes that Marko is coming.)

[2] The tense that is being used in these subjunctive clauses is purely morphological, that is, it has agreement morphology.

b. On se nada da je Marko došao.
He self hope.PRES.3sg that AUX Marko come.PAST.3sg
(He hopes that Marko came.)

c. On uspeva da postigne pogodak
He manage.PRES.3sg that score.PRES.3sg goal
(He manages to score the goal.)

d. On uspeva da *je postigao pogodak
He manage.PRES.3sg that *AUX score.PAST.3sg. goal
(He manages to *scored the goal.)

(4)d is ungrammatical because the matrix verb cannot select the subjunctive structure that bears the tense other than the present tense.[3]

While standard grammars do not allocate too much of their resources on subjunctives in Serbian, some of more recent works have explored the status of Serbian subjunctives (Belić, 2005; Kim, 201; Janić, 2012; Miškeljin, 2012; Todorović, 2012). Janić (2012) claims that representations of subjunctives are complementary to those of indicatives in that the main distinction is made between indicative and subjunctive selecting predicates. On one hand, indicative verbs are semantically permeated with the content that is part of a real-life, factual activity or state. On the other hand, subjunctive selecting verbs are usually the verbs that refer to wishes, possibilities, desires, or the so-called irrealis state. Serbian subjunctives are not substantially different in this respect. Following Krapova's (1998, 2001) assumption of distinct types of subjunctives, Janić (2012) distinguishes between weak and strong subjunctives. The division is mainly based on the properties of the selecting matrix verb and the subjects of the complement clause they allow (PRO or DP/*pro*). Unlike Janić (2012), who does not adhere to any particular syntactic framework to account for Serbian subjunctives, Krapova (1998, 2001) offers a classification of Bulgarian subjunctives that is predicated on the Minimalist assumptions (Chomsky, 2000, 2001). The author (1998) contrasts Type I and Type II Bulgarian subjunctives. In addition, she assumes that there is a V-to-T-to-C movement within the CP phase and that PRO, in Type I subjunctives, receives the Null Case (Martin 2001). Miškeljin's (2012) paper follows, not only Krapova's (1998) categorization, but also Landau's (2000, 2004) theory of control. Her study is one of the few (if not the only one) that provides some evidence for Landau's (2000, 2004) system based on subjunctive structure in Serbian. The author (2012) explores untensed and tensed infinitivals and subjunctives, whereby the latter are divided into Type I and Type II. Miškeljin (2012) unorthodoxly, and contrary to Landau's (2004) assumptions, argues that both types allow PRO as their subject. In addition, the author divides Type II subjunctives into non-obviative and obviative. The former allow only a lexical DP or *pro*, whereas the latter license only PRO.[4]

[3] The tense of the subjunctive selected by the verb 'hope' is an exception. The majority of Type II selecting predicates can only license the present tense morphology, which I indicate in the analysis. I also show that both types of Serbian subjunctives do not differ in terms of the tense they permit, but only in the selection of subjects.

[4] This is incorrect as the only difference between obviative and non-obviative subjunctives is in [Agr] on C^0 (Landau, 2004, p. 857).

Governed by the previous research on Serbian subjunctives, I provide evidence for two types of subjunctives. They are selected based on the semantic content of the matrix verb and the interaction of tense and agreement (which are syntactically motivated). Contrary to Miškeljin (2012), I do not maintain that both types allow PRO. Instead, I claim that Type I Serbian subjunctives allow only PRO, while Type II can license either a DP or *pro*.

2 Theoretical Considerations

In the following two sections, I present relevant background theoretical assumptions. First, I describe the null category, PRO. Then, I elaborate on the fundamentals of the theory that underlies the entire present paper analysis – Landau's (2000, 2004, 2006) theory of control.

2.1 PRO

PRO is considered to be a null element specified and licensed by the verb of the matrix clause. It thus receives its theta-role from the matrix verb and assumes the position of the subject of the embedded non-finite clause (Adger, 2003). The main motivation behind stipulating PRO stems from the Theta Criterion (Chomsky, 1981) which specifies that each argument of the verb must be assigned a theta-role. Therefore, sentences such as (5) require an empty category to behave as a subject that is phonetically null. PRO in (5) is co-indexed with John and, thereby, anaphoric.

5) John$_i$ managed to PRO$_i$ pay for the tickets.

Chomsky (1981) defines PRO as a pronominal anaphor, which indicates that it is both [+pronominal] and [+anaphoric]. According to the Binding Theory, PRO can be both bound (Principle A) and free (Principle B) (Chomsky, 1981). The distribution of PRO as both an anaphor and a pronominal has been temporarily solved by the PRO Theorem, which argues that PRO cannot be governed (Adger, 2003; Carnie, 2007). If PRO is ungoverned, then it cannot be case-marked. However, evidence from Icelandic (Sigurðsson, 2008), Russian (Moore & Perlmutter, 2000; Landau, 2008), and Hungarian (Tóth, 2000) reveals that PRO can actually bear case like any other lexical DP (quirky case in Icelandic, structural in Russian and Hungarian). Chomsky and Lasnik (1993) have invested a lot of effort into solving this issue by stipulating the Null Case that would be assigned to PRO to escape the violation of the Visibility Condition. Null Case can be assigned only to PRO, that is, no other lexical or empty element is permitted to receive this case. Such a stipulation and the fact that, to account for PRO, one needs both the PRO Theorem and Case theory, is in a direct opposition to the Minimalist Program that adheres to the idea of reducing theoretical stipulations to minimum (Hornstein, 1999).

Therefore, a number of scholars have investigated PRO and sought to develop a theory of control that pertains to the principles of Minimalism.[5] One of these theories is Landau's (2000, 2004) theory of control. In the following section, I will describe basic mechanisms that govern this theory.

2.2 Theory of Control

Landau's (2000, 2004, 2006) theory of control is an attempt to find a unifying theoretical framework that would define control cross-linguistically. The author (2000, 2004, 2006) attempts to develop a system that "generates a typology of clausal complementation with minimal recourse to external stipulations" (Landau, 2000, p. 814). The 'minimal recourse' is constituted of certain parameters and operations. I will now describe the theoretical vocabulary indispensable for understanding the theory.

There are two types of syntactic objects necessary to understand any control theory: nominal phrases and complement clauses. Landau (2000, 2004) categorizes the latter into infinitives, subjunctives and indicatives. These are, in turn, either tensed or untensed. As regards nominal phrases, i.e. subjects of the embedded clauses, their properties ought to indicate a referential contrast between lexical DPs and *pro*, on one hand, and PRO, on the other hand. Landau (2000, 2004) uses a rich scope of privative features specified on functional heads and DPs to show the distribution of subjects. Functional heads I^0 and C^0 are specified for [T] and [Agr], while DPs and null elements (*pro* and PRO) bear a reference feature [R] (Reuland & Reinhart, 1995). [T] can occur on C^0 if, and only if, the tense of the embedded complement is selected by the matrix predicate. If the tense is free, i.e. independent, C^0 does not bear [T]. Anaphoric tense, [-T], is assigned to the two functional heads if these are specified for the tense identical to the one in the matrix clause. The dependent tense [+T] is assigned elsewhere. The configuration of [T] is summarized in (6).

6) a. Anaphoric tense = [-T] on I^0/C^0
   b. Dependent tense = [+T] on I^0/C^0
   c. Independent tense = [+T] on I^0, Ø on C^0
   Landau (2004, p. 839)

Tense is semantic, which indicates that a complement clause may bear [+T], i.e. it can be tensed, but carry no tense morphology, or it can be [-T] and possess the tense morphology. What regulates the morphology of the embedded clause is the feature [Agr]. This feature is

[5] Hornstein (1999) lays out the movement theory of control in which he dispenses with theta-roles on lexical DPs and treats theta-roles as features. PRO is forgone as its place is taken by a lexical DP or *pro* that emerges in the vP, and moves in consecutive steps to the subject position of the matrix clause. Culicover and Jackendoff (2001) claim that this is not the way control works. They raise the questions of control with verbs such as 'promise' and 'vow', and extend their argument to a number of examples stemming from nominal structures. The authors (2006) argue that control ought to be completely handed over to semantics. I follow Landau's (2000, 2004) theory of control as it offers a middle ground between these theories. It balances syntactic, morphological and semantic descriptions in order to provide motivation for PRO, and it does this by relying mostly on features and operations that are already present in the theory.

strictly morphological for it stands for a bundle of phi-features. Thus, if there is agreement morphology, [+Agr] is present on I^0, while its presence on C^0 is parasitic on [+T].

Another feature mentioned above that is quite important for the distribution of lexical DPs and null elements is [R], or the reference feature. It was first stipulated by Reuland and Reinhart (1995), who postulated it for the purposes of explaining Dutch and German *sich*. It is specified on PRO as anaphoric, that is, [-R], while it is [+R] on *pro* or a lexical DP. Landau (2004) considers PRO to be an SE-like anaphor, which is "not subject to condition A (since they lack a reflexivizing function) but are subject to Chain Condition, which forbids them from heading chains" (p. 841). The feature is assigned a plus or minus value on the nominal phrase under specific configurations. For example, "whenever I^0 or C^0 are specified for [+T, +Agr], then they automatically come to bear [+R]; any other feature constitution – [+T, -Agr], [-T, +Agr], [-T, -Agr] – is associated with [-R]" (ibid., p. 842). Albeit a straightforward stipulation, the R-assignment rule serves to distinguish PRO from lexical DPs or *pro* on independent interpretive grounds.

Lastly, in order for the clausal and nominal syntactic objects to interact, as is the goal of every control theory, there needs to be a connecting element. Syntactic operations (as per Chomsky 2000) provide the 'machinery' for this. The operation Agree, specifically, ensures that all the objects enter Agreement and match and check their features. According to Landau (2004), the probe is the matrix functional head (either T or v) whose potential goal is PRO due to its anaphoric [-R] (p. 843). The probe, an essential bolt in the Agree operation, is a way of coindexing PRO's phi-features with those of antecedent. When two syntactic objects enter the agreement relationship, their features are checked against each other.

Therefore, feature configurations and the interaction of operations[6] are sufficient for a derivation of infinitives, indicatives, and subjunctives. (7)a below shows Landau's (2004) 'calculus' of an indicative, no-control clause, and no-control F(ree)-subjunctive clause is displayed in (7)b.

7) Indicative clause

a. [$_{CP}$… DP… V [$_{CP}$ C^0 [$_{IP}$ [$_{I'}$ I^0 $_{[-T, +Agr, +R]}$ [$_{VP}$ DP/*pro* $_{[+R]}$]]]]]

Agree Agree $_{[+Agr, +R]}$

F-subjunctive clause

b. [$_{CP}$… DP… V [$_{CP}$ C^0 $_{[+T, +Agr, +R]}$ [$_{IP}$ [$_{I'}$ I^0 $_{[+T, +Agr, +R]}$ [$_{VP}$ DP/*pro* $_{[+R]}$]]]]]

Agree Agree $_{[+T, +Agr, +R]}$ Agree $_{[+Agr, +R]}$

Since indicative complements, (7)a, have independent tense, the embedded C^0 is assigned no [T, Agr] features, which further implies that R-assignment rule cannot be applied to C^0. (7)b, on the other hand, indicates that Balkan F-subjunctives bear dependent tense and agreement. All the features are present on the two functional heads, C^0 and I^0, and they

[6] Landau (2004) refers to this interaction as the 'calculus of agreement and tense'.

cancel each other by checking and matching under Agree. Landau (2004) argues that "checked features persist to the end of their phase, the controller DP does not "use up" the features of that functional head, which can enter another checking relation before they are erased" (p. 843). Therefore, [-R] features on PRO are interpretable and visible to Agree from the outside of the CP phase, while [+R] features check off each other on I^0 and C^0 without triggering Agree between the matrix functional head and the lower-level DP or *pro*.

3 Objectives

There are in particular three main issues that I attempt to address in this study. First, previous research has shown that Serbian subjunctives have not been explored in a principled way within more recent theories, such as Landau's (2000, 2004) theory. Second, the classification of Serbian subjunctives lacks solid empirical and conceptual support. Third, mechanisms that trigger the emergence of PRO have frequently been linked to Case and Binding, which requires a lot of additional structure. I believe that syntactic 'instruments' independent from Case or Binding can accommodate the distribution of PRO.

To provide answers to the raised issues, I adhere to the basic tenets of Landau's (2000, 2004) theory of control. My objective is to present a typology of Serbian subjunctives with a complementizer *da*, and to argue that the distribution of PRO is derived on the grounds of tense and agreement interaction rather than the intrinsic properties of PRO, Binding or Case theory. Therefore, the main research question of the present study is: *How is PRO distributed in Serbian subjunctives?*

4 Analysis

In the analysis, I first provide a categorization of Serbian subjunctives and their feature specification. The categorization is contingent on semantic properties of selecting predicates, whereas the feature configuration development is predicated on Landau's (2000, 2004) theory. Next, I show the interaction of mechanisms behind the distribution of PRO in Serbian subjunctive clauses, which presents the longest part of the analysis. Lastly, I draw parallels of PRO in Serbian subjunctives with other languages (Balkan languages).

4.1 Categorization

Maintaining Krapova's (1998) and Janić's (2012) categorization of Bulgarian and Serbian subjunctives, I classify Serbian subjunctives into Type I and Type II subjunctives. To account for the categorization, though, governed by the fact that semantic properties of the matrix verb guide the selection of subjunctives, I have developed a classification of Serbian subjunctive selecting predicates. These can be viewed in the list below.[7]

[7] The list is tentative and based on other authors' works (Giannakidou 2009, Landau 2004 (EC & PC selecting predicates), Todorović 2012, Janić 2012).

Serbian subjunctive selecting predicates

Type I:

Aspectual (početi, stati, nastaviti...) ('begin, start, stop, continue…')

Modal (moći, morati, trebati, biti sposoban da…) ('can, must, need, be able to…')

Implicative (usuditi se, uspeti) ('dare, manage…')

Factive (mrzeti, ne voleti…) ('hate, dislike…')

Type II:

Desiderative (želeti, preferirati, odlučiti, zahtevati, nadati se, planirati...)

('want, prefer, decide, demand, hope, desire, plan…')

Commissive (ponuditi, obećati, zavetovati se…) ('offer, promise, vow…')

Directive: (narediti, predložiti…) ('order, suggest, recommend…')

According to some accounts (Belić, 2005; Janić, 2012; Miškeljin, 2012), Type I predicates select complement clauses that restrict the tense specification to the present tense. Type II subjunctives are slightly more flexible with respect to their morphological tense specifications as they allow more variable tense selection. However, I claim that this view is incorrect because both subjunctive types allow only the present tense morphology. This is shown in examples in (8).[8]

8) a. Type I
Ne volim dugo da pešačim.
pro [I] not like.PRES.3sg long that walk.PRES.3sg
(I don't like to walk for a long period of time.)

b. Type I
Ne volim dugo da *sam pešačio.
pro [I] not like.PRES.3sg long that AUX walk.PAST.3sg
(I don't like to *walked for a long period of time.)

c. Type II
Ponudio sam Marku da dodje.
pro$_i$ [I] offer.PAST.3sg AUX Mark that come.PRES.3sg
(I offered Mark to come.)

[8] I have not specified whether the matrix clauses select PRO or *pro* as this will be shown in the subsequent section.

d. Type II

	Ponudio	sam	Marku	da	*je došao
pro$_i$ [I]	offer.PAST.3sg	AUX	Mark	that	AUX come.PAST.3sg

Type II selecting predicates, akin to Type I predicates, prohibit any other tense but the present tense morphology. Therefore, the difference between Type I and Type II subjunctives is only visible through the selection of their subjects (PRO or DP/*pro*).

4.2 Feature Specification

Whether PRO or DP/pro will be the subjects of the embedded clause is predicated on the argument structure that matrix verbs license (Janić, 2012). Hence, the semantic classification of subjunctive selecting predicates is most certainly quite useful. Still, in order to precisely determine the distribution of subjunctive subjects, I put forward a feature configuration of the two types of Serbian subjunctives. I assume that Type I subjunctives are analogous to Landau's (2004) Balkan C(ontrolled)-subjunctives, whereas Type II subjunctives are equated with Balkan F(ree)-subjunctives. The configuration of features on Serbian subjunctive functional heads is presented in Table 1.

Table 1: Feature configuration of I^0 and C^0 of Serbian subjunctives

	Type I	**Type II**
I^0	[-T, +Agr]	[+T, +Agr]
C^0	[-T]	[+T, +Agr]
Subject	[+R]	[+R]

By definition (R-assignment rule), Type I subjunctives ought to select PRO, as it is [-R], while the subject of Type II subjunctives is [+R], that is, either a lexical DP or *pro*. The full feature specification is shown in Table 2.

Table 2: Complete feature configuration of Serbian subjunctives

	Type I PRO	**Type II DP/*pro***
I^0	[-T, +Agr]	[+T, +Agr]
C^0	[-T]	[+T, +Agr]

4.3 Interpretation

Such a specification should thus ensure a full and systematic interpretation of Serbian subjunctives. I present the analysis of Serbian subjunctives in (9), whereby PRO is the subject of Type I (9)a, and *pro* is the subject of Type II subjunctives (9)b.

9) a. [$_{CP}$… $_{DP}$ Marko… *v* pokušava [$_{CP}$ C^0 da $_{[-T]}$ [$_{IP}$ PRO [$_{I'}$ I^0 $_{[-T, +Agr, -R]}$ [$_{VP}$ t_{PRO} $_{[-R]}$ glumi]]]]]

Agree Agree $_{[+Agr]}$ Agree $_{[-T]}$ Agree $_{[+Agr, -R]}$

Marko*i* try.PRES.3sg that PRO*i* act.PRES.3sg
(Marko is trying to act.)

b. [$_{CP}$…$_{DP}$ Marko… *v* želi [$_{CP}$ C^0 da $_{[+T, + Agr, +R]}$ [$_{IP}$ [$_{I'}$ I^0 $_{[+T, +Agr, +R]}$ [$_{VP}$ DP/*pro* $_{[+R]}$ glumi]]]]]

Agree Agree $_{[+T, +Agr, +R]}$ Agree $_{[+Agr, +R]}$

Marko$_i$ want.PRES.3sg that *pro$_i$* act.PRES.3sg
(Marko wants to act.)

(9)a does not allow a lexical DP or *pro* to assume the position of the subject of a subjunctive clause because of the anaphoric [-R]. [-R] can be specified only on anaphoric PRO which enters the Agreement relationship with the functional head of the matrix clause. [-T] on I^0 is checked against [-T] on C^0, while [-R] is checked on I^0. PRO moves to the [Spec, IP] as it is required to find the closest slot to the functional head of the matrix verb that licenses it through Agree. Although PRO is not at the edge of the phase, it is visible to the probe because of its anaphoric feature. In (9)b, the functional head of the matrix clause does not enter any relationship with a lexical DP or *pro* because the subject of the subjunctive clause bears [+R] feature. [+R] is first checked off at I^0, and then at C^0, at the end of the phase. Therefore, the predicate of the matrix verb in (9)b specifies both the tense and agreement configuration, which are [+T] and [+Agr], and the reference, which is independent of the subject of the matrix clause, i.e. it is non-anaphoric [+R]. This configuration and the interaction of features provide syntactic evidence for what appears to be a semantic and morphological issue. In addition, tense and agree feature specifications and their interaction indicate that Type I Serbian subjunctives are untensed ([-T]) with a full morphological structure [+Agr], whereas Type II Serbian subjunctives are both tensed and morphologically inflected. This is shown in (10).[9]

10)a. [$_{CP}$… $_{DP}$ Marko… *v* uspeva [$_{CP}$ C^0 da $_{[-T]}$ [$_{IP}$ PRO [$_{I'}$ I^0 $_{[-T, +Agr, -R]}$ [$_{VP}$ t_{PRO} $_{[-R]}$ glumi]]]]]

Agree Agree $_{[+Agr]}$ Agree $_{[-T]}$ Agree $_{[+Agr, -R]}$

Marko*i* manages.PRES.3sg that PRO*i* act.PRES.3sg
(Marko manages to act.)

[9] For more examples, view Appendix A and B.

b. [$_{CP}$… $_{DP}$ Marko… *v* preferira [$_{CP}$ C^0 da $_{[+T, + Agr, +R]}$ [$_{IP}$ [$_{I'}$ I^0 $_{[+T, +Agr, +R]}$ [$_{VP}$ DP/*pro* $_{[+R]}$ glumi]]]]]
Agree Agree $_{[+T, +Agr, +R]}$ Agree $_{[+Agr, +R]}$

Marko$_i$ prefer.PRES.3sg that *pro$_i$* act.PRES.3sg
(Marko prefers to act.)

Landau's (2000, 2004) theory assumes that PRO is accessible for Agree despite its position within the IP, that is, it enters the Agreement relationship with the functional head of the matrix clause by crossing the phase. Chomsky's phase theory (2000, 2001) argues that there are two major phases, CP and vP. They are determined in the numeration that comprises sub-numerations. These in turn constitute phases. CP would be a higher phase, while vP is the clause-internal phase. Still, phases are not sent to the spell-out in their entirety, but

> "only the complement of the phase head, the spell-out domain, is sent to spell-out, whilst the phase head and its specifier, the phase edge, remain in the syntax and are only spelt out with the higher phase. A side effect of all this is that any material contained inside the completed spell-out domain would be unavailable for further syntactic computations" (Harwood, 2015, p. 526).

Phases and spell-out domains are shown in (11).

11) [$_{CP}$… $_{DP}$ Marko… *v* uspeva [$_{CP}$ C^0 da $_{[-T]}$ [$_{IP}$ PRO [$_{I'}$ I^0 $_{[-T, +Agr, -R]}$ [*v*P *v* [$_{VP}$ t$_{PRO\ [-R]}$ glumi]]]]]
phase spell-out domain phase spell-out domain

Hence, PRO emerges in the spell-out domain and should not be further accessible to syntactic computations, that is, "if an item has been shipped off, along with the rest of the spell-out domain, to PF and LF, then it is no longer visible to the syntactic component, and so cannot enter into any further syntactic operations with elements outside of the phase" (Harwood, 2015, p. 825).

Following Landau (2004), a way to solve this hindrance is to make another stipulation in claiming that PRO is visible to the probe due to its [-R] feature, which is anaphoric, and it needs to be coindexed with the subject of the matrix clause. This, however, seems to be a stipulation that does not completely solve the issue. Therefore, I maintain Harwood's (2015) proposal of dynamic approach to phases.[10] The author's (2015) main argument is that "the phase is not complete until the last item in sub-numeration has been merged into the workspace, irrespective of what the last item is" (p. 557). Such a state of affairs permits the last item in the sub-numeration to project to vP,

[10] This approach is based on some previous phase structure approaches such as Bobaljik and Wurmbrand (2005), Wurmbrand (2014), and Bošković (2013, 2014).

which would make up a phase consisting of [v, CP, C, IP, I], that is, instead of being the phase head, C^0 is merely the head of the internal CP.[11] This is shown in Figure 1.

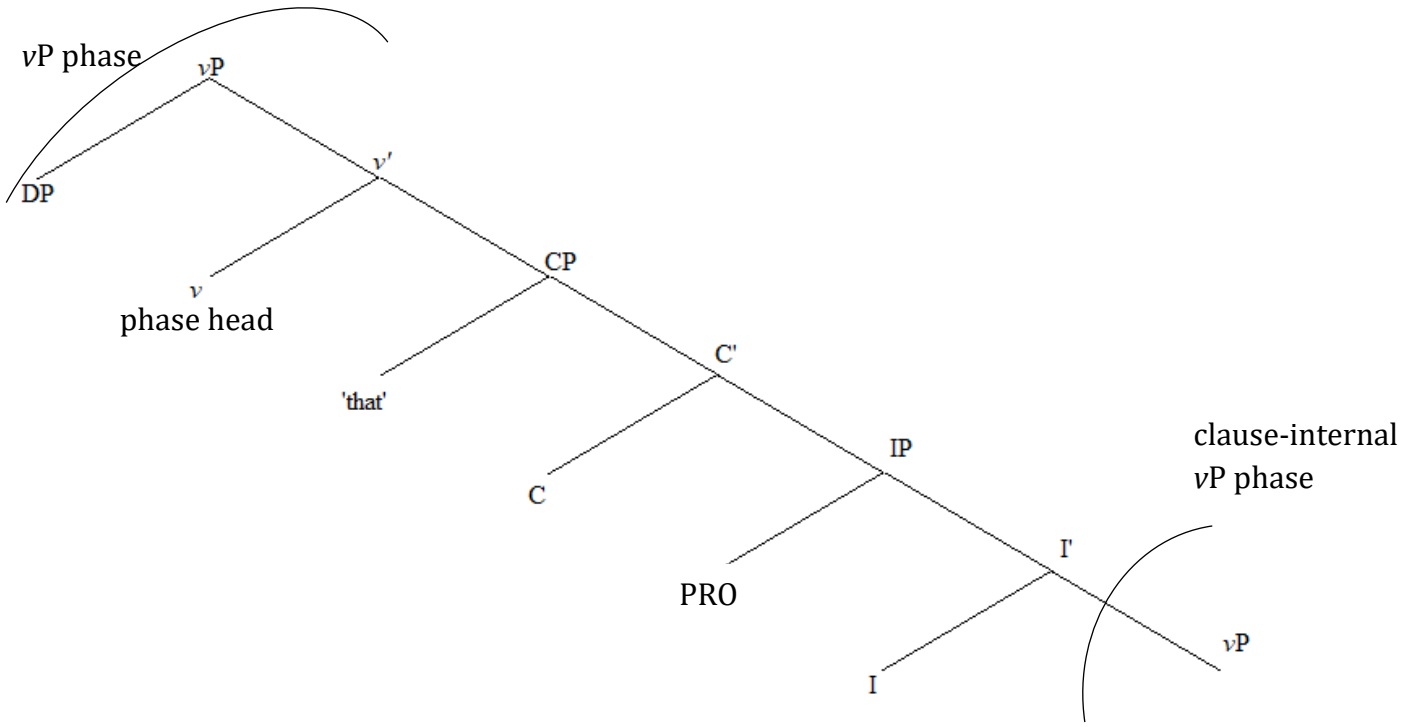


Figure 1: Variable phase structure – Type I Serbian subjunctives

In this way, the distribution of PRO is thoroughly and efficiently motivated without resorting to redundant rules or stipulations. PRO is visible to the functional head, and, thereby, accessible for further syntactic computations when it moves to the spec IP position. Contrary to PRO, a lexical DP or *pro* of Type II subjunctives in Serbian does not enter agreement with the functional head of the matrix clause due to their [+R]. Therefore, the internal CP phase is retained (Fig. 2).

[11] Type I subjunctives are usually selected by predicates that are referred to 'complex predicates', especially aspectual and modal verbs. This is another reason why it can be assumed that a phase is not complete until all the elements in the numeration merge.

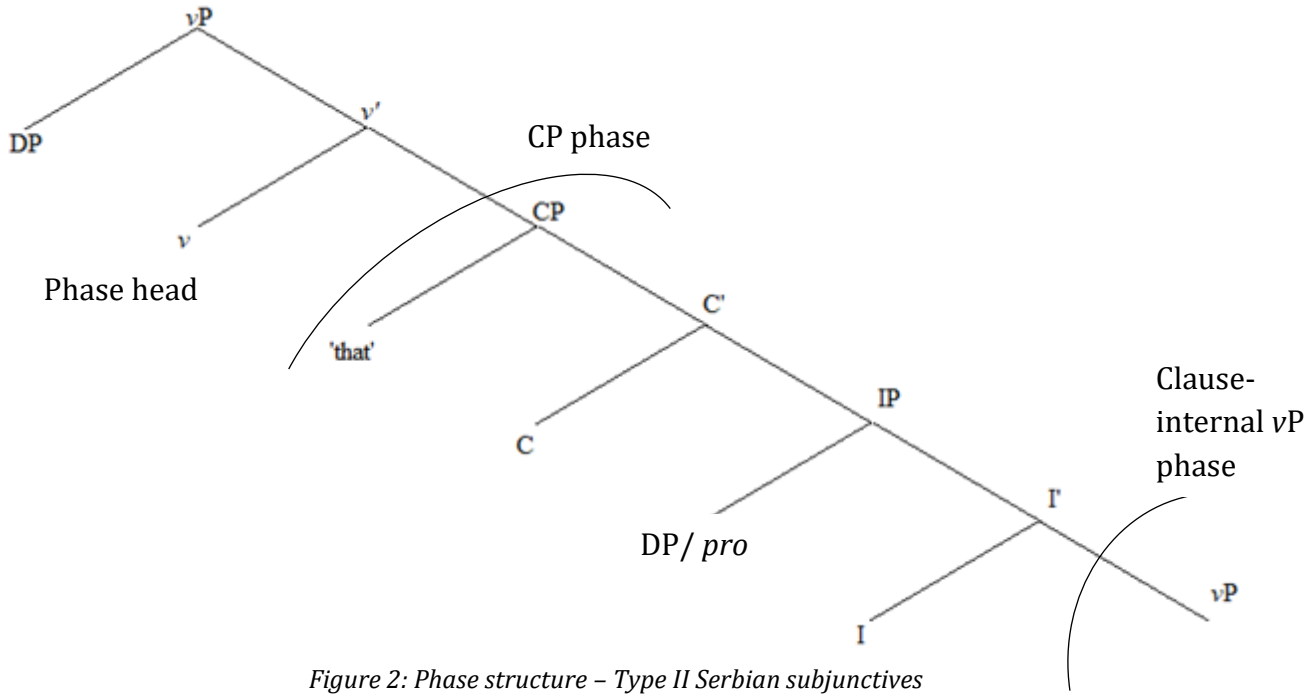


Figure 2: Phase structure – Type II Serbian subjunctives

4.3.1 Balkan Languages

Landau's (2004) calculus and Harwood's (2015) phase structure approach can in a unifying and systematic way explain the motivation for PRO in Type I Serbian subjunctives. I argue that the analysis can be applied to subjunctives in other Balkan languages such as Bulgarian (Krapova, 1998), Romanian (Dobrovie-Sorin, 1993), Greek (Terzi, 1997), and Albanian (Dobrovie-Sorin, 2001). Although almost all these languages belong to different language families, they are spoken mainly in the Balkans, which renders them Balkan languages. I show in (12) that PRO is generated in Bulgarian, Romanian, Greek, and Albanian subjunctives in the same manner as in Serbian Type I subjunctives.[12]

12) a. Bulgarian
Ivan$_1$ se opita da PRO$_1$ razbere vprosa
Ivan$_1$ try.PAST.3sg that PRO$_1$ understand.PRES.3sg question
(Ivan tried to understand the question.) (Krapova, 1998, p.74)

b. Romanian
Maria$_1$ a încercat să PRO$_1$ plece.
Maria$_1$ try.PAST.3sg that PRO$_1$ leave.PRES.3sg
(Maria tried to leave.) (as per Dobrovie-Sorin, 1993)

[12] For a more detailed analysis of these examples see Appendix C.

c. Greek
I Maria$_1$ prospathise na PRO$_1$ divasi.
The Maria$_1$ try.PAST.3sg that PRO$_1$ read.PRES.3sg
(Maria tried to read.) (Terzi, 1997, in Landau 2004)

d. Albanian
Beni$_1$ filloi të PRO$_1$ lexonte Tre Muskëtjerët.
Beni$_1$ start.PAST that PRO$_1$ read.PRES.3sg Three Musketeersë
(Beni started to read Three Musketeers.) (as per Dobrovie-Sorin, 1993)

PRO is generated as the subject of the subjunctive clause within the vP phase because it is anaphoric [-R], and enters Agreement with the matrix functional head. Checking and deletion of [T] and [Agr] on I^0 and C^0 control the derivation of the entire sentence, wherein the subjunctive is untensed, but bears the agreement features.

5 Conclusion

In this paper, I have managed to provide an answer to the research question which explores how PRO is distributed in Serbian subjunctive clauses.

First, I have shown that there are two subjunctive types, Type I and Type II. The former selects PRO, while the latter licenses the DP or *pro*. Type I subjunctives are [-T], but [+Agr], which means that they are untensed despite agreement morphology on the embedded verb, while Type II Serbian subjunctives are both tensed and permit tense morphology in the embedded clause. Type I subjunctives are, therefore, analogous to Landau's (2004) C-subjunctives, and Type II subjunctives share the identical configuration to that of F-subjunctives in Balkan languages (p. 869). Furthermore, if one is to maintain Landau's (2004) control typology, it can be claimed that Type I subjunctives exhibit Obligatory Control, whereas Type II Serbian subjunctives pertain to No Control (p. 869).

In addition to categorizing Serbian subjunctives and describing how their subjects are generated the analysis reveals that Landau's (2000, 2004) theory requires a different phase structure approach, for Chomsky's (2000, 2001) phase theory is limiting with regard to the distribution of PRO within Serbian subjunctives. The approach I assume is Harwood's (2015) phase structure approach. The author (2015) puts forward an idea of a derivation system that "continues to merge items from the sub-numeration until there is no more material left to (externally) merge" (p. 558). Under these circumstances, one can assume that, with Type I Serbian subjunctives, CP is not a phase, but a complementizer of the vP. PRO is thus freely allowed to enter the Agree relationship with its licensor under the same phase. By assuming the variable phase approach, and by following Landau's (2000, 2004) concepts of control, I have illustrated that the distribution of PRO in Serbian subjunctives is detached from either Case or Binding. In addition, it can be motivated not only in Serbian subjunctives within this theoretical framework, but also in subjunctives

cross-linguistically. I have demonstrated that the distribution of PRO can be accounted for in other Balkan languages such as Bulgarian, Romanian, Greek and Albanian.

References

Adger, D. (2003). *Core syntax: A minimalist approach* (Vol. 20). Oxford: Oxford University Press.

Belic, B. (2005). Complement verb variation in present-day Serbian. *Unpublished doctoral dissertation*. The Ohio State University.

Bobaljik, J. D., & Wurmbrand, S. (2005). The domain of agreement. *Natural Language & Linguistic Theory*, 23(4), 809-865. Netherlands: Springer.

Bošković, Ž. (2013). Phases beyond Clauses. In L. Schürcks, A. Giannakidou, & U. Etxeberria (Eds.) *The nominal structure in Slavic and beyond*, 116, (75-127). Boston: DeGruyter.

Bošković, Ž. (2014). Now I'm a phase, now I'm not a phase: On the variability of phases with extraction and ellipsis. Linguistic inquiry, 45(1), 27-89. Cambridge, MA: MIT Press.

Chomsky, N. (1981). *Lectures on Government and Binding*. Foris, Dordrecht.

Chomsky, N., & Lasnik, H. (1993). Principles and parameters theory in syntax: *An international handbook of contemporary research*. Mouton De Gruyter, 506-570.

Chomsky, N. (2000). Minimalist inquiries: The framework. In R. Martin, D. Michaels & J. Uriagereka (Eds.) *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89-155. Cambridge, MA: MIT Press.

Chomsky, Noam. (2001). Derivation by phase. In M. Kenstowicz (Eds.) *Ken Hale: A Life in Language*, 1-52. Cambridge, MA: MIT Press.

Culicover, P. W., & Jackendoff, R. (2001). Control is not movement. *Linguistic inquiry*, 32(3), 493-512. Cambridge, MA: MIT Press.

Culicover, P. W., & Jackendoff, R. (2006). Turn over control to the semantics!. Syntax, 9(2), 131-152.

Dobrovie-Sorin, C. (1993). *The syntax of Romanian: Comparative studies in romance*. Retrieved from https://ebookcentral-proquest-com.ezproxy.lib.ucalgary.ca

Dobrovie-Sorin, C. (2001). Head-to-head merge in Balkan subjunctives and locality. In M. L. Rivero, & A. Ralli (Eds.). *Comparative syntax of Balkan languages*, 44-73. New York, NY: Oxford University Press.

Giannakidou, A. (2009). The dependency of the subjunctive revisited: Temporal semantics and polarity. *Lingua*, 119(12), 1883-1908.

Harwood, W. (2015). Being progressive is just a phase: celebrating the uniqueness of progressive aspect under a phase-based analysis. *Natural Language & Linguistic Theory*, 33(2), 523-573.

Hornstein, N. (1999). Movement and control. *Linguistic inquiry*, 30(1), 69-96.

Janić, A. (2012). Kategorija subjunktiva u srpskom i bugarskom jeziku – uporedno proucavanje. *Unpublished MA Thesis*. Niš, Serbia: University of Niš.

Kim, B. R. (2010). *Non-Finite Complements in Russian, Serbian/Croatian, and Macedonian*. ProQuest LLC. 789 East Eisenhower Parkway, PO Box 1346, Ann Arbor, MI 48106.

Krapova, I. (1998). Subjunctive Complements, Null Subjects and Case Checking in Bulgarian. *Working Papers in Linguistics*, 8(2). Venice, Italy: University of Venice.

Krapova, I. (2001). Subjunctives in Bulgarian and modern Greek. In M. Rivero & A. Ralli (Eds.), *Comparative syntax of Balkan languages* (105-126). New York: Oxford University Press Inc.

Landau, I. (2000). *Elements of control: Structure and meaning in infinitival constructions* (Vol. 51). Netherlands: Kluwer Academic Publishers.

Landau, I. (2004). The scale of finiteness and the calculus of control. *Natural Language & Linguistic Theory*, 22(4), 811-877.

Landau, I. (2006). Severing the distribution of PRO from case. *Syntax*, 9(2), 153-170.

Landau, I. (2008). Two routes of control: Evidence from case transmission in Russian. *Natural Language & Linguistic Theory*, 26(4), 877-924.

Martin, R. (2001). Null case and the distribution of PRO. *Linguistic inquiry*, 32(1), 141-166.

Miškeljin, I. (2012). Infinitival complements in Serbian-A generative approach. *Филолог–часопис за језик, књижевност и културу*, (5). Banja Luka, Bosnia: University of Banja Luka.

Moore, J., & Perlmutter, D. M. (2000). What does it take to be a dative subject. *Natural Language & Linguistic Theory*, 18(2), 373-416.

Piper, P., Antonić, I., Ružić, V., Tanasić, S., Popović, L., & Tošović, B. (2005). *Sintaksa savremenog srpskog jezika* – Prosta rečenica. Belgrade, Serbia: Institut za srpski jezik SANU.

Reuland, E., & Reinhart, T. (1995). Pronouns, anaphors and case. In *Studies in comparative Germanic syntax* (pp. 241-268). Dordrecht: Springer.

Sigurðsson, H. Á. (2008). The case of PRO. *Natural Language & Linguistic Theory*, 26(2), 403-450. Netherlands: Springer.

Stanojčić, Ž., & Popović, L. (1992). *Gramatika srpskoga jezika: udžbenik za I, II, III i IV razred srednje škole*. Belgrade, Serbia: Zavod za udžbenike i nastavna sredstva.

Terzi, A. (1997). PRO and null case in finite clauses. *The Linguistic Review*, 14(4), 335-360. Netherlands: Mouton de Gruyter.

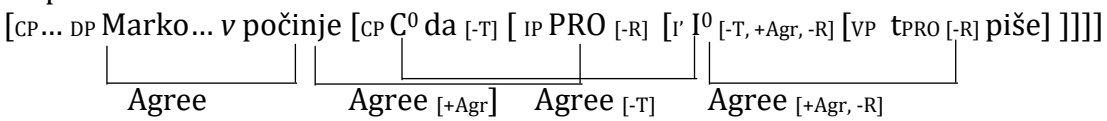
Todorović, N. (2012). The Indicative and Subjunctive" da"-complements in Serbian A Syntactic-Semantic Approach. *Unpublished doctoral dissertation*. Chicago: University of Illinois.

Tóth, I. (2000). *Inflected infinitives in Hungarian*. Tilburg, Netherlands: Tilburg University, Faculty of Arts.

Wurmbrand, S. (2014). Tense and aspect in English infinitives. *Linguistic Inquiry*, 45(3), 403-447. Cambridge, MA: MIT Press.

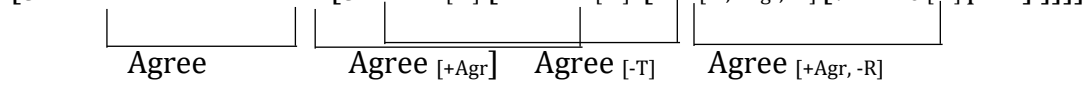
Appendix A: Type I Serbian Subjunctives

13)a. Aspectual



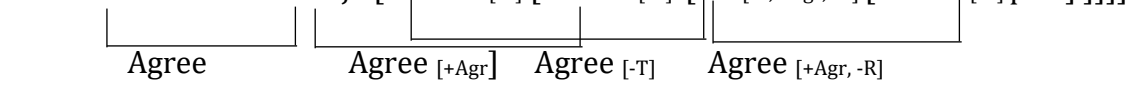
Marko*i* start.PRES.3sg that PRO*i* write.PRES.3sg
(Marko is starting to write.)

b. Modal



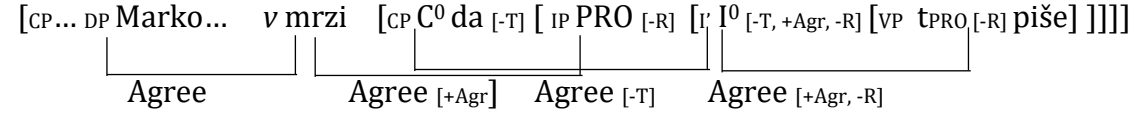
Marko*i* can.PRES.3sg that PRO*i* write.PRES.3sg
(Marko can write.)

c. Implicative



Marko*i* self.dare.PRES.3sg that PRO*i* write.PRES.3sg
(Marko dares to write.)

d. Factive



Marko*i* hate.PRES.3sg that PRO*i* write.PRES.3sg
(Marko hates to write.)

Appendix B: Type II Serbian Subjunctives

14)a. Desiderative

[$_{CP}$… $_{DP}$Marko…*v* zahteva [$_{CP}$ C^0 da $_{[+T, + Agr, +R]}$ [$_{IP}$ [$_{I'}$ I^0 $_{[+T, +Agr, +R]}$ [$_{VP}$ DP/*pro* $_{[+R]}$ piše]]]]]

Agree Agree $_{[+T, +Agr, +R]}$ Agree $_{[+Agr, +R]}$

Marko$_i$ demand.PRES.3sg that *pro$_i$* write.PRES.3sg
(Marko demands to write.)

b. Commissive

[$_{CP}$… $_{DP}$Marko…*v* se nudi [$_{CP}$ C^0 da $_{[+T, + Agr, +R]}$ [$_{IP}$ [$_{I'}$ I^0 $_{[+T, +Agr, +R]}$ [$_{VP}$ DP/*pro* $_{[+R]}$ piše]]]]]

Agree Agree $_{[+T, +Agr, +R]}$ Agree $_{[+Agr, +R]}$

Marko$_i$ self.offer.PRES.3sg that *pro$_i$* write.PRES.3sg
(Marko offers to write.)

c. Implicative

[$_{CP}$…$_{DP}$Marko…*v* predlaže [$_{CP}$ C^0 da $_{[+T, + Agr, +R]}$ [$_{IP}$ [$_{I'}$ I^0 $_{[+T, +Agr, +R]}$ [$_{VP}$ DP/*pro* $_{[+R]}$ piše]]]]]

Agree Agree $_{[+T, +Agr, +R]}$ Agree $_{[+Agr, +R]}$

Marko$_i$ suggest.PRES.3sg that *pro$_i$* write.PRES.3sg
(Marko suggests that he should write.)

Appendix C: PRO in Balkan Languages

15)a. Bulgarian

[$_{CP}$… $_{DP}$Marko…*v* se opita [$_{CP}$ C^0 da $_{[+T, + Agr, +R]}$ [$_{IP}$ [$_{I'}$ I^0 $_{[+T, +Agr, +R]}$ [$_{VP}$ DP/*pro* $_{[+R]}$ razbere v prosa]]]]]

Agree Agree $_{[+T, +Agr, +R]}$ Agree $_{[+Agr, -R]}$

(Ivan tried to understand the question.) (Krapova, 1998, p. 74)

b. Romanian

[$_{CP}$…$_{DP}$Maria…*v* a încercat [$_{CP}$ C^0 să$_{[+T, + Agr, +R]}$ [$_{IP}$ [$_{I'}$ I^0 $_{[+T, +Agr, +R]}$ [$_{VP}$ DP/*pro* $_{[+R]}$ plece]]]]]

Agree Agree $_{[+T, +Agr, +R]}$ Agree $_{[+Agr, -R]}$

(Maria tried to leave.) (as per Dobrovie-Sorin, 1993)

c. Greek

[$_{CP}$…$_{DP}$ I Maria…*v* prospathise [$_{CP}$ C^0 na$_{[+T, + Agr, +R]}$ [$_{IP}$ [$_{I'}$ I^0 $_{[+T, +Agr, +R]}$ [$_{VP}$ DP/*pro* $_{[+R]}$ divasi]]]]]

Agree Agree $_{[+T, +Agr, +R]}$ Agree $_{[+Agr, -R]}$

(Maria tried to leave.) (Terzi, 1997, in Landau, 2004)

d. Albanian

[$_{CP}$…$_{DP}$ Beni…*v* filloi [$_{CP}$ C^0 të$_{[+T, + Agr, +R]}$ [$_{IP}$ [$_{I'}$ I^0 $_{[+T, +Agr, +R]}$ [$_{VP}$ DP/*pro* $_{[+R]}$ lexonte Tre Muskëtjerët]]]]]

Agree Agree $_{[+T, +Agr, +R]}$ Agree $_{[+Agr, -R]}$

(Beni started to read Three Musketeers.) (as per Dobrovie-Sorin, 2001)

Contact Information

Dušan Nikolić

dusan.nikolic1@ucalgary.ca
University of Calgary
School of Languages, Linguistics, Literatures & Cultures
University of Calgary, CHC 211
2500 University Drive NW
Calgary, Alberta, T2N 1N4
Canada

Case Assignment on Bare Direct Objects in Turkish

Metehan Oguz

University of Calgary

Abstract

Bare direct objects (BDOs) in Turkish have attracted the interest of many researchers (Aydemir, 2004; Enç, 1991; Kamali, 2015; Kornfilt, 1994, 2003; Öztürk, 2005, 2009; Travis, 2010; among others). Studies so far mostly agree that BDOs in Turkish get into a relationship with the verb, however, the hypotheses for the type of the relationship and the data provided are inconsistent. Kornfilt (1994) suggests that head of the BDO incorporates into the verb and creates a new verb head, following Baker's (1988) proposal. Aydemir (2004) proposes that BDOs in Turkish are not syntactic arguments and that they do not occupy object position, while Öztürk (2005) proposes that BDOs pseudo-incorporate with the verb, following Massam's (2001) proposal. Kamali (2015) disagrees with some Turkish data provided by Öztürk (2005), but still agrees with her pseudo-incorporation proposal for Turkish. Kamali (2015) concludes that lack of overt case in BDOs is due to a weak accusative feature, which leads to either pseudo-incorporation or an indefinite reading of the BDO. This paper suggests that the evidence put forward to support both incorporation and pseudo-incorporation analyses could be explained by adopting Karimi's (2005) Two Object Position Hypothesis (TOPH) to Turkish.

Key words: syntax, bare objects, Turkish, Case

1 Introduction

This paper discusses whether bare direct objects in Turkish are assigned case or not. A bare object is a noun functioning as the grammatical object in a sentence, but unlike an accusative-marked object, it is not marked with any morphological case affixes. An example of a case marked direct object is given in (1a), and a bare direct object in (1b).

1) a. Ali kitab-ı oku- du.
Ali book-ACC read-PAST
(Ali read the book.)

b. Ali kitap oku- du.
Ali book read-PAST
(Ali read a book/ books.)

Bare direct objects in Turkish have attracted the interest of many researchers (Aydemir, 2004; Enç, 1991; Kamali, 2015; Kornfilt, 1994, 2003; Öztürk, 2005, 2009; Travis, 2010; among others). Enç (1991) observed that objects with overt accusative case (like 1a) have a "specific", or "presuppositional" reading, but that this specificity disappears in caseless (bare) objects (like 1b). Studies so far mostly agree on the fact that the bare object (henceforth, BDO) in Turkish enters into a relationship with the verb, however, the hypotheses for the type of the relationship and the data provided are inconsistent. Kornfilt (1994) claims that head of the BDO incorporates into the verb and creates a new verb head, following Baker's (1988) proposal. Further, adopting Kornfilt (1994), Aydemir (2004) claims that BDOs in Turkish are not syntactic arguments and that they do not occupy an object position. In contrast, Öztürk (2005) argues that head incorporation in Baker's sense is impossible in Turkish and proposes that BDOs pseudo-incorporate with the verb, following Massam's (2001) proposal. Kamali (2015) disagrees with some Turkish data provided by Öztürk (2005), but still agrees with her pseudo-incorporation proposal for Turkish. Kamali (2015) concludes that the lack of overt case in BDOs is due to a weak accusative feature, which leads to either pseudo-incorporation (formation of a complex predicate) or an indefinite reading of the BDO.

In this paper, I analyze arguments of both incorporation and pseudo-incorporation approaches, demonstrate their weaknesses and strengths, and provide new Turkish data related to the issue. I show that the evidence in support of an incorporation analysis is not convincing, and I argue that Turkish BDOs do not incorporate at all. In sections 2 and 3, I introduce the existing approaches to BDOs in Turkish. In section 4, I show that except for one piece of it, namely, behaviour of non-derived modifiers in Turkish, the evidence put forward for incorporation is problematic. In section 5, I propose a new explanation for the behaviour of non-derived modifiers. In section 6, I conclude that the issue, which has been analyzed as (a kind of) incorporation so far, is indeed unrelated to incorporation and is a result of strong and weak case features on direct objects in Turkish.

2 Head incorporation of BDOs in Turkish

Baker (1988) proposes that direct objects of transitive verbs sometimes undergo head-to-head movement and incorporate with the verb head, resulting a new verb head. Turkish has been cited as a language in which incorporation in Baker's (1988) sense is observed (cf. Mithun, 1984; Knecth, 1986; Kornfilt, 1994, 2003; Aydemir, 2004; among others). Kornfilt (1994) claims that BDOs in Turkish incorporate into the verb by a head-to-head movement and create a novel verb head. Aydemir (2004) takes this suggestion further and claims that BDOs in Turkish do not occupy the object position. Instead she proposes that BDOs function as simple noun heads (Ns) while case marked nouns, or nouns following determiners like *bir* "a/an" function as noun phrases (NPs). Aydemir (2004) presents differences in the behaviour of BDOs and case assigned NPs as evidence for the incorporation analysis of BDOs. She claims that ellipsis of a BDO is not possible (2a), while ellipsis of an NP is possible (2b).

2) a.*Bütün gün **kitap** oku- du-m, san-a da oku- ma-n-ı
All day book read-PAST-1SG, you-DA also read-NOM-AGR.2SG-ACC
tavsiye ed-er-im.
recommend-AOR-1SG
(I read books/did book reading all day, I recommend you to read (it) too.)

b. Dün **bir kitap** oku-du-m, san-a da oku-ma-n-ı
Yesterday a book read-PAST-1SG, you-DAT also read-NOM-AGR.2SG-ACC
tavsiye ed-er-im.
recommend-AOR-1SG
(I read a book yesterday, I recommend you to read (it) too.)
(from Aydemir, 2004, p. 468)

It can be seen in (2a) that a BDO noun cannot be deleted in the following clause; it has to be repeated in the VP. However, (2b) shows that an NP object can be deleted in the following clause. The contrast in (2) suggests that BDO Ns are different from NPs in that they cannot be omitted. Aydemir (2004) suggests that BDOs do not form an independent constituent within in the VP, and so they cannot be separated from the verb. She supports her claim by stating that the BDOs must always be immediately preverbal, while scrambling of NP objects is possible. Aydemir (2004) also claims that BDOs in Turkish cannot be referentially active. In her own words, it is impossible to refer back to bare N (3a) whereas NPs (3b) are referentially active.

3) a. ***Dün** **film**$_i$ seyret-ti-m, **onu**$_i$ sen de seyret-meli-sin.
Yesterday movie watch-PST-1S it-ACC you too watch-MOD-2S
(I did movie-watching yesterday, you should also watch it.)

b. Dün **bir film**$_i$ seyret-ti-m, **onu**$_i$ sen de seyret-meli-sin.
Yesterday a movie watch-PST-1S it-ACC you too watch-MOD-2S
(I watched a movie yesterday, you should also watch it.)
(from Aydemir, 2004, p. 468)

Another diagnostic that Aydemir (2004) uses — the most important in my view — is the meaning change in use of modifiers like *hızlı* "fast", *kötü* "bad", and some others. Underived modifiers like *hızlı* "fast", and *kötü* "bad" function as adjectives when they precede a noun, and function as adverbs when they precede a verb (Taylan, 1984). Aydemir (2004) provides (4) to demonstrate this difference.

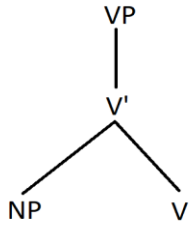
4) a. Mehmet [**kötü** [araba kullanıyor]].
Mehmet bad car use
(Mehmet drives badly.)

b. Mehmet [[**kötü** bir araba] kullanıyor].
Mehmet bad a car use
(Mehmet drives a bad car.)

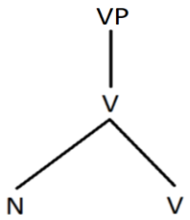
(from Aydemir, 2004, p. 467)

The modifier *kötü* "bad" functions as an adverb in (4a), while it functions as an adjective in (4b). Aydemir (2004) suggests that since BDOs are not syntactic arguments, they are not barriers between the modifier and the verb, which causes the modifier to modify the verb. In fact, she argues that the BDO is a part of the verb, so technically the modifier immediately precedes the verb in (4a), and that is why it functions as an adverb. In contrast, in (4b) the modifier *kötü* "bad" precedes an N and functions as an adjective modifying this N. Based on the evidence she provides, Aydemir (2004) hypothesizes that VPs that contain BDOs in Turkish form the structure in (5b), rather than the one in (5a).

5) a.



5) b.



It is worth noting here that many native Turkish speakers (including me) accept the sentences (2a) and (3a) as grammatical. These sentences are the main evidence Aydemir (2004) builds her arguments on, so I believe that further discussion is needed, which Öztürk (2005) and Kamali (2015) point out (see section. 3). It is also worth noting here that the sentence (4a) is ambiguous; it has the meanings given in both (4a) and (4b). I will get into details about this issue in sections 4 and 5.

3 Pseudo-incorporation of nouns

Massam (2001) proposes that the non-referential nominal objects in Niuean are phrasal categories. In other words, that they are NPs. However, the object NP and the verb act as a single unit in the Syntax. (6) is an example for the type of nominal object Massam (2001) is concerned with.

6) Ne inu **kofe kono** a Mele.
PST drink coffee bitter ABS Mele
(Mary drank bitter coffee.)

(from Massam, 2001, p. 158)

Massam (2001) proposes that in (6), the direct object *kofe kono* "bitter coffee" and the verb *inu* "drink" act as a single unit- as if the NP were incorporated. She suggests that what is observed in (6) is *Pseudo Noun Incorporation (PNI)*, because phrasal categories cannot incorporate by means of the head-incorporation operation of Baker (1988). Massam (2001) argues that PNI and Baker's NI are two different types of incorporation one can observe in the world's languages.

3.1 Pseudo Noun Incorporation in Turkish

Following Massam's (2001) analysis, Öztürk (2005) applies PNI analysis to Turkish BDO data. To begin with, Öztürk (2005) disagrees that BDOs in Turkish are N heads, rather than NPs. She shows Taylan's (1986) data (7) to show that focus particles like *bile* "even" can intervene between the BDO and the verb.

7) Ali kitap *bile* okudu.
Ali book even read
(Ali did even book reading.)

(from Öztürk, 2005, p. 215)

Öztürk (2005) points out that (7) is evidence for the fact that the BDO and the verb do not form a morphologically complex verb head. This follows from the assumption that adverbs, such as *bile* "even" cannot occur internal to the verb. She also shows that the BDO can be modified in (8).

8) Ali ekşi elma ye-di.
Ali sour apple eat-PAST
(Ali did sour apple eating.)

(from Öztürk, 2005, p. 216)

Considering the evidence in (6) and (7) (and some others), Öztürk (2005) suggests that BDOs in Turkish are not incorporated Ns, but rather that they are independent NPs. For this reason, she claims that head incorporation analysis is not compatible with the facts of Turkish.

Even though Öztürk (2005) disagrees that BDOs in Turkish are Ns, she agrees with Aydemir (2004) that BDOs cannot be referential, and that they must be in pre-verbal position. She draws a parallel between Turkish BDOs and Niuean incorporated NPs and suggest that BDOs in Turkish pseudo-incorporate into the verb to create complex predicates. Öztürk (2005) supports her claim by two different sets of facts: impersonal passives (9-10), and non-derived modifiers (11). The latter were also used by Aydemir (2004).

9) a. Ali *kitab-ı* oda-da oku-du. (active)
Ali book-ACC room-LOC read-PAST
(Ali read the book in the room.)

b. Kitap oda-da oku-n-du. (personal passive)
book room-LOC read-PASS-PAST
(The book was read in the room.)

c. Ali oda-da *kitap* oku-du. (active)
Ali room-LOC book-ACC read-PAST
(Ali did book-reading in the room.)

d. Oda-da kitap oku-n-du. (impersonal passive)
room-LOC book read-pass-past
= (Book-reading was done in the room.)
≠ (The book was read in the room.)

(from Öztürk, 2005, p. 217-218)

Öztürk (2005) points that accusative marked theme in (9a) is visible to operations in the Syntax such as passivization, as shown in (9b). However, pseudo-incorporation of themes (seen in (9c)) does not allow a personal passive like (9b) but an impersonal passive shown in (9d). She claims that verbs with pseudo-incorporated themes act alike unergatives (10).

10) a. İnsanlar koş-tu. (unergative)
people run-PAST
(People ran.)
b. Koş-ul-du. (impersonal passive)
run-PASS-PAST
(Running was done.)

Öztürk (2005) claims that (10) indicates that BDOs are not in true object position in the Syntax. When passivized, the sentences show the same behaviour as unergatives. She also provides similar examples to data in (4), repeated here as (11), to claim that the BDO incorporates into the verb and creates a complex predicate with it.

11) a. Mehmet [kötü [araba kullanı-yor]].
Mehmet bad car use-PRES
(Mehmet drives badly.)

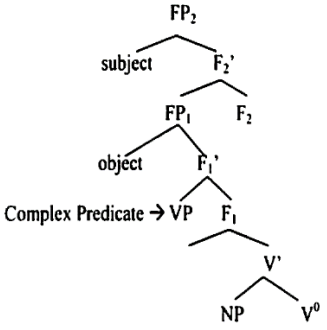
b. Mehmet [[kötü bir araba] kullanı-yor].
Mehmet bad a car use-PRES
(Mehmet drives a bad car.)

(from Aydemir 2004, p. 467)

Based on the evidence in (9), (10), and (11), Öztürk (2005) suggests that pseudo-incorporated nouns are not syntactic objects, they function as part of the complex predicate

verb. She argues that the complement NP and the lexical verb form a complex predicate at the V' level, and proposes the syntactic structure in (12).

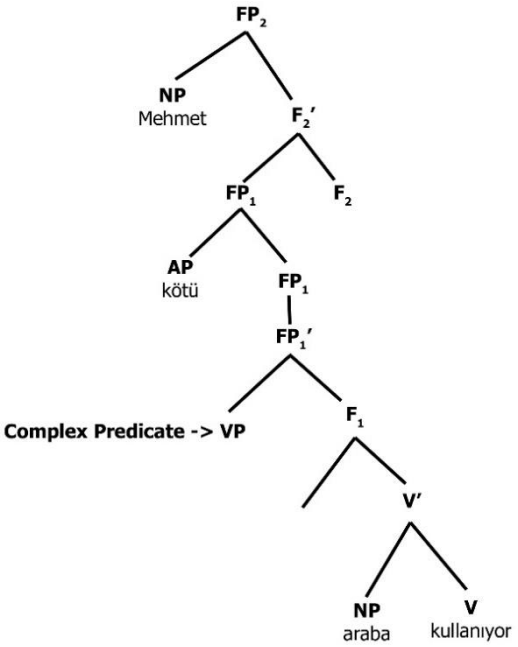
12)[1]



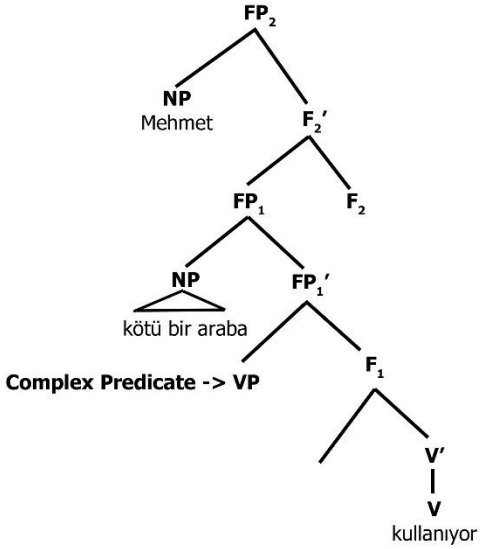
(from Öztürk, 2005, p. 222)

Öztürk (2005) states that in the structure (12a), any NP that is immediate sister to the lexical verb will be interpreted as part of the complex predicate. In order to be interpreted as referential direct objects, NPs need to occur in the Spec of higher functional projections. Assuming her proposal, the structures for sentences (11a) and (11b) looks like (13a) and (13b), respectively.

13) a.



13) b.



The object, *araba* "car", is a part of the complex predicate in (13a), while it stays outside of the complex predicate and functions as a (regular) object in (13b).

[1] Öztürk (2005) assumes a full-blown Neo-Davidsonian model, where all of the verb arguments are introduced via separate functional projections. For details, see Öztürk (2005).

3.1.1 Theta-role assignment in pseudo-incorporation

Öztürk (2005) proposes that even though they are not fully independent syntactic elements, incorporated BDOs are associated with specific theta-roles. She suggests this is possible due to the fact that pseudo-incorporated NPs can check Case features against functional heads. She provides evidence for Case checking in causative structures in Turkish. Taylan (1984) shows that when an unergative is causativized (14a), the agent of the unergative always gets marked for accusative case as shown in (14b). In contrast, when a transitive structure with an already existing accusative case marked object (15a) is causativized, the agent of the transitive verb is dative marked, but not accusative (15b). Taylan (1984) claims that this is due to the double-case constraint in Turkish, which does not allow one case to be realized more than once in the same clause.

14) a. Ali koştu.
Ali ran
(Ali ran.)

b. Ayşe Ali-**yi** koşturdu.
Ayşe Ali-ACC run-CAUS-PAST
(Ayşe made Ali run.)

15) a. Ali balığ-ı tuttu.
Ali fish-ACC caught
(Ali caught the fish.)

b. Ayşe Ali'**ye** balık tutturdu.
Ayşe Ali-DAT/*ACC fish catch-CAUS-PAST
(Ayşe made Ali catch (a) fish.)

(from Öztürk, 2005, p. 223)

Based on (14) and (15), Öztürk (2005) proposes that pseudo-incorporated nouns are associated with accusative case, even though it is not morphologically realized, and therefore it is not allowed to have another accusative case marked element in the structure.

Considering the evidence shown in sections 3.1 and 3.1.1, Öztürk (2005) concludes that (overt) caselessness of objects is as result of pseudo-incorporation, not head-incorporation in Baker's (1988) sense, in contrast to Kornfilt (2003), Aydemir (2004), and others.

3.2 Bare Objects and weak accusative Case

Kamali (2015) agrees with Öztürk's (2005) pseudo-incorporation analysis on Turkish BDOs, but she points that she disagrees with judgements on the data provided by Aydemir (2004) and Öztürk (2005). Kamali (2015) claims that BDOs are not necessarily much different from overtly Case-marked NPs in the Syntax, but they are different in terms of Semantics (definiteness and specificity). Kamali (2015) introduces evidence that BDOs in

Turkish can be referentially active, in contrast to what Aydemir (2004) and Öztürk (2005) suggest.

16) Bir saattir oğlan-lar-ı izli-yorum.
one hour boy-s-ACC watch-PRES.1SG[2]
(I've been watching the boys for the last hour.)

a. Emre portakal getiri-yor.
Emre orange bring-PRES
(Emre does orange-bringing.)

b. Ali de onu soyu-yor.
Ali COMM it-ACC peel-PRES
(And Ali peels it.)

c. Ama sonra Ø ye-mi-yorlar. Ø Biriktir-iyorlar.
but then eat-NEG-PRES.3PL save-PRES.3PL
(But after that they don't eat. They save.)
(from Kamali, 2015, p. 120, re-glossed)

Kamali (2015) claims that in the discourse (16), the BDO in (16a) is referred to in (16b) by a pronoun, and a dropped pronoun in (16c). She also provides evidence that BDOs do not have to be pre-verbal because they can scramble (17).

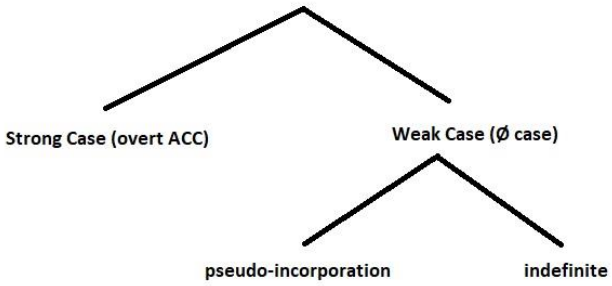
17) A: I wonder how tall lions are.[3]
B: Ben t_i gör-dü-m aslan$_i$. 2 metre var.
I see-PAST-1S lion 2 meter exists
(The lion, I've seen it. About 2 meters.)
(from Kamali, 2015, p. 118, re-glossed)

The data Kamali (2015) provides in (16) and (17) are actually very critical in that we observe the BDO *aslan* "lion" moves scrambles into SVO order. Kamali (2015) proposes that apparent caselessness is in fact the realization of weak accusative case in Turkish. She claims that strong accusative case has to be morphologically overt, while weak accusative case could be morphologically empty. Kamali (2015) claims that not all objects without overt Case-marking are the same and makes a distinction between objects with morphologically null case and bare objects (BDOs). She claims that objects with morphologically null case (different from BDOs) are referential (16) and able to refer back to elements (17), the only difference between objects with null case and objects with overt Case-marking is that an object with null case is a result of weak case, leading to non-specific/indefinite reading.[4] On the other hand, BDOs pseudo-incorporate and become a

[2] There is no gloss here in the original paper and it was added by me for this paper.
[3] It is worth noting here that the setup sentence A is necessary for the actual sentence B.
[4] Kamali (2015) provides more semantic differences between case-marked and caseless objects in Turkish, but I will not go into details of her analysis for purposes of this paper.

part of the complex predicate. Kamali (2015) claims what ties caseless objects and BDOs is the fact that weak accusative case is realized as Ø. She proposes the taxonomy in (18).

18)



(from Kamali, 2015, p. 121)

It is seen in (18) that when the object checks strong case, it must carry overt accusative case and when it checks weak case, it lacks overt case marker and results with either an indefinite reading, or incorporation. When they incorporate, they must be pre-verbal and non-referential.

4 Evaluating previous analyses of Turkish Bare Objects

In this section, I will review the arguments introduced earlier in this paper, and provide new data where it is necessary. I will point out weaknesses of each hypothesis and suggest that a new proposal for case assignment of bare objects in Turkish is needed.

4.1 The Head-incorporation analysis

Baker's (1988) head-incorporation proposal has had a significant effect on the literature, especially on the studies on compounds. Even though Turkish BDOs do not show morphological compounding, the topic has been cited as head-incorporation many times (Mithun, 1984; Nilsson, 1986; Schroeder, 1999; Kornfilt, 1994, 2003; Aydemir, 2004; among others). Among these studies, the most extreme claim came from Aydemir (2004), who claimed that indeed pre-verbal BDOs were not visible to the Syntax. Aydemir (2004) provided some evidence for her claims as seen in (2), (3), and (4) above. In (2) and (3), she showed that the bare objects are not phrases and as a result for that, ellipsis of them is not possible and they cannot be referential. She shows these by the ungrammaticality of (2a) and (3a). However, both (2a) and (3a) are acceptable to many Turkish native speakers (including me). This indeed indicates that BDOs can be deleted, or referentially active — exact opposite of what Aydemir (2004) claims. Later, Öztürk (2005) demonstrates that BDOs can be deleted, and Kamali (2015) demonstrates that BDOs can be referentially active (16). However, the judgments on very similar structures in Turkish are inconsistent. Moreover, Öztürk (2005) provides very good evidence for the phrasal status of BDOs. In (7), it is seen that focus particles like *bile* "even" can intervene between the BDO and the verb. This suggests that the verb and the BDO can be separated from each other. Also, it is shown in (8) that the BDO can be modified by adjectives. If BDOs were simply bare Ns,

there would be no position for an adjunct Adjective Phrase. Both (7) and (8) are strong evidence that BDOs are not simply bare N elements but rather that they are NPs. I also add another piece of evidence for this claim in (19).

19) a. [Çay ve sigara] iç-tim.
tea and cigarette drink-PAST.1S
(I drank tea and smoked.)
Lit. Translation: "I drank tea and cigarette"

b. [Çorba ve makarna] pişir-dim ve ye-dim.
soup and pasta cook-PAST.1S and eat-PAST.1SG
(I cooked soup and pasta and ate them.).

It is seen in both (19a) and (19b) that the BDO can indeed be a Conjunction structure. Following Kayne (1994), I argue that the BDOs in (19) are Conjunction Phrases (&P). This shows that the BDOs are not bare Ns , but phrases (NP or &P). The &P is also referred to as the object of *yedim* 'eat' in conjoint sentence.

Head status of the objects is the key feature in Baker's (1988) incorporation. Considering the evidence shown in (7), (8), and (19), I agree with Öztürk (2005) and Kamali (2015) that BDOs in Turkish are phrasal structures. Thus Baker's (1988) head-incorporation analysis is not applicable to Turkish BDOs.

4.2 The Pseudo-incorporation analysis

Massam's (2001) pseudo-incorporation analysis offers a strong alternative for some languages which do not fit Baker's (1988) analysis. As shown in section 3, Massam (2001) develops a pseudo-incorporation analysis of Niuean phrasal objects which act as a single unit with the verb in the syntax. Considering behaviour of modifiers like *kötü* "bad" shown in (4)-repeated here as (20) , Öztürk (2005) suggests that Turkish BDOs are also pseudo-incorporated and together with the verb form a complex predicate. She claims that the reading we get in (20a) shows that BDO *araba* "car" incorporates into the verb, and acts as a part of the complex predicate (verb) after the incorporation. As a result, the modifier is preceding the complex predicate, and functions as an adverb modifying this predicate.

20) a. Mehmet [kötü [araba kullanı-yor]].
Mehmet bad car use-PRES
(Mehmet drives badly.)

b. Mehmet [[kötü bir araba] kullanı-yor].
Mehmet bad a car use-PRES
(Mehmet drives a bad car.)

(from Aydemir, 2004, p. 467)

Some other behaviours that would be expected from pseudo-incorporation of Massam (2001) are non-referentiality, and stableness in pre-verbal position. Öztürk (2005) agrees with claims of Aydemir (2004) that Turkish BDOs are not referential (shown in (3)) and

that they must be pre-verbal. Under these assumptions, Öztürk's (2005) proposal for pseudo-incorporation works perfectly. However, as mentioned earlier in this paper, many native Turkish speakers accept sentences in which the BDOs are referential. Furthermore, as Kamali (2015) shows, many Turkish speakers accept movement of BDOs, under specific contexts. Massam's (2001) proposal emphasizes many times that the pseudo-incorporated object must be non-referential. The fact that BDOs in Turkish can be referential strongly contradicts the pseudo-incorporation analysis Massam (2001) develops. In addition to that, BDOs' ability to move is also not expected in Massam's (2001) proposal. Considering these facts, Kamali (2015) argues persuasively against a pseudo-incorporation analysis, pointing that BDOs can be referential and they can move around.

Even though she provides strong counter evidence for pseudo-incorporation analysis, Kamali (2015) still claims that some BDOs in Turkish are pseudo-incorporated. I believe the reason why she wants to keep pseudo-incorporation is to be able to account for data like (20), in which a modifier like *kötü* "bad" preceding the object functions as an adverb. Even though it seems to cover all possible structures, I want to point that Kamali's (2015) theory is too strong. She first shows that some BDOs (which she analyses as caseless indefinite objects, not BDOs) can show behaviours unattested in pseudo-incorporation analysis, then creating a different class of caseless objects, she accounts for data like (20). I find the motivation for separating indefinite caseless objects from bare objects (BDO) is not satisfying. The assumptions are not generalizable, and are instead too specific. She concludes that there are three types of objects; overt case-marked objects, bare objects (which incorporate), and indefinite caseless objects. Aside from the behaviour in (20), there is no evidence for two different overt caseless objects. Kamali (2015) also states that by assuming only bare objects pseudo-incorporate, we lose the generalization on caseless objects.

To sum up, the head-incorporation analysis is inapplicable for Turkish BDOs in that the BDOs are phrasal elements, not heads. Pseudo-incorporation analysis suggested by Öztürk (2005) does not cover all behaviours of BDOs in Turkish and makes overgeneralization, claiming that BDOs must be pre-verbal and non-referential. Pseudo-incorporation in Kamali's (2015) theory seems to predict correct results, however, it is very specific and sounds very ad hoc. For these reasons, a novel analysis on bare objects in Turkish is necessary.

5 A novel proposal

Considering the data Kamali (2015) provided, the BDOs are not that different from overtly case-marked objects. They both can move, be deleted (ellipsis), and be referential. All that remains to be accounted for is to explain the behaviour of non-derived modifiers like *kötü* "bad" in sentences like (20). If one can explain the behaviour in (20a), the fact that the modifier precedes a noun but functions as an adverb, any other assumptions like (any kind of) incorporation might be unnecessary. To account for the behaviour in (20), I am going to adapt to Karimi's (2005) analysis of Persian objects.

Karimi (2005) proposes that there is a special relation between the verb and its non-specific object in Persian. She adopts a version of Distributed Morphology (Halle & Marantz, 1993)

suggested by Marantz (1997). Karimi (2005) claims that the verb and the non-specific object provide a special meaning that cannot be produced directly by the lexicon. This operation occurs in the Syntax and Karimi (2005) names it *Syntactic Word Formation* (SWF). She claims that the V' is the domain level for SWF. The meaning is formed within the V' level in the Syntax. Considering the semantic, syntactic and morphological behaviours of Persian specific and non-specific objects, Karimi (2005) claims that specific and non-specific objects in Persian take different positions in the Syntax. She proposes "Two Object Position Hypothesis" (TOPH) shown in (21).

21) Two Object Position Hypothesis (TOPH)
a. [$_{VP}$ DP$_{[+specific]}$ [$_{v'}$ PP V]]
b. [$_{VP}$ [$_{v'}$ PP [$_{v'}$ DP$_{[-specific]}$ V]]

(from Karimi, 2005, p. 105)

It can be seen that a specific object takes a specifier position in the verb phrase (VP) (21a), while a non-specific object takes the complement position (21b). Karimi (2005) suggests that in (21a), the specific object originates in the complement position, and moves higher to check its "specific" feature (for details, see Karimi, 2005).

I suggest that Turkish also exhibits TOPH. At this point, recall that according to Kamali's (2015) proposal, the appearance of overt case is dependent on a strong accusative Case feature. Kamali (2015) also shows that objects with a strong Case feature have a specific reading, while objects with a weak Case feature are non-specific. Let's assume that movement in TOPH is motivated by strong accusative Case feature in Turkish. If the object has strong accusative Case feature, it must move higher than the V' level, namely to Spec, vP (22).

22) a. Ben [$_{VP}$ çok [çocuk sev-iyor-um]].
I much child$_{[non-specific]}$ love-PRES-1S
(I like children a lot.)

b. [$_{vP}$Ben [$_{v'}$ çocuk-u$_i$ [$_{VP}$ çok [$_{V'}$ t$_i$ sev-iyor-um]]].
I child$_{[specific]}$-ACC much love-PRES-1S
(I like the child (a specific child) a lot.)

c. *Ben [$_{VP}$ çok [$_{V'}$ çocuk-u sev-iyor-um]]].
I much child$_{[specific]}$-ACC love-PRES-1S

d. *Ben [$_{VP}$ çok [sen-(i) sev-iyor-um]].
I much you$_{[specific]}$-ACC love-PRES-1S

e. [$_{vP}$Ben [$_{v'}$ sen-i$_i$ [$_{VP}$ çok [$_{V'}$ t$_i$ sev-iyor-um]]].
I you$_{[specific]}$-ACC much love-PRES-1S
(I love you so much.)

It is shown in (22) that an overt case-marked object (object with strong Case feature) moves to specifier position in the vP (22b), while the object without an overt case marker can stay in its original position (22a). The obligatoriness of the movement for specific objects is observed with the ungrammaticality of (22c). (22d) has a second person pronoun, which is always specific definite, in its object position. It can also be seen in (22d) and (22e) that without movement of the pronoun to Spec vP, the sentence is ungrammatical. The contrast shown in (22) indicates that original position for adverbs in Turkish is pre-nominal. Originally, the adverbs precede the objects, however, the strong Case feature motivates the object to move higher than the adverb. Considering this, if an object does not have a strong Case feature and does not move higher than the adverb, it is very natural that the adverb precedes the object in the surface structure. The next step is to determine how strong and weak Case features are introduced.

Bliss (2004) does a syntax-semantics analysis on BDOs in Turkish. She argues that BDOs lack the semantic properties of number and specificity, which are present on other NPs, while NPs with determiners like *bir* "a/an" or *bazı* "some" and case-marked NPs are always number specific. She suggests that maximum level of BDOs is NP level, which causes the lack of number and specificity, while maximum level of cased objects and [Det + BDO] structures is Determiner Phrase (DP) level, which enables the object to have number and specificity features. I assume her proposal and claim that strong case feature is introduced by Ds, while weak case feature is introduced by Ns. This implies that no case marked object can originate between the verb and the adverb in Turkish because that position is occupied by the (BDO) NP. Hence, I modify Karimi's (2005) TOPH as in (23). In Karimi's (2005) original proposal, non-specific determiner *bir* "a/an" must be lower than the Adv. (which is not the case in Turkish) because it lacks *specific* feature. This modification allows non-specific determiner *bir* "a/an" to hold a position higher than VP and motivate the NP to move there to check strong Case feature.

23) Two Object Position Hypothesis (TOPH) (modified for Turkish)
a. [$_{vP}$ DP$_{[+ACC*]}$ [$_{VP}$ AdvP V]] (* implies strong feature.)
b. [$_{VP}$ [$_{v'}$ AP [$_{v'}$ **NP**$_{[+ACC]}$ V]]

(23) implies that DPs are higher than NPs and AdvPs in the structure. When the object is specific, it needs to move to Spec, vP in order to satisfy strong accusative feature. This also explains the difference between (24a) and (24b).

24) a. Ben **hızlı** araba sür-dü-m.
I fast car drive-PAST-1S
Meaning 1: (I drove fast)
Meaning 2: (I drove (a) fast car(s) (I did fast car-driving.))

b. Ben **hızlı** bir araba sür-dü-m.
I fast a car drive-PAST-1S
(I drove a fast car (non-specific).)

(24a) can have two different structures, as illustrated in (24a') and (24a''). However, (24b) can only have the structure in (24b').

24) a'. Ben [$_{VP}$ [$_{V'}$ **hızlı** [$_{V'}$ [$_{NP}$ araba] sür-dü-m]].
I fast car drive-PAST-1S
(I drove fast.)

a''. Ben [$_{VP}$ [$_{V'}$ [$_{NP}$ **hızlı** araba] sür-dü-m]].
I fast car drive-PAST-1S
(I drove a (non-specific) fast car.)

b'. Ben [$_{VP}$ [$_{V'}$ [$_{DP}$ **hızlı** bir [$_{NP}$ araba$_i$] [$_{V'}$ t$_i$ sür-dü-m]].[5]
I fast a car drive-PAST-1S
(I drove a (non-specific) fast car.)

The determiner *bir* "a/an" in (24b) motivates the NP *araba* "car" to go higher than the adverb position in VP, thus, there is no way of interpreting *hızlı* "fast" in the structure as an adverb. The ambiguity in (24a) is a result of the position *hızlı* "fast" originally merges. It merges into the VP in (24a'), where it functions as an adverb. However, it merges into the NP in (24a''), where it functions as an adjective. (24b) shows that the modification I made in Karimi's (2005) original suggestion is necessary. If we assume that the movement in TOPH is derived by specificity in Turkish, we would expect *hızlı* "fast" in (24b) to have an adverbial reading as well because the object is non-specific.

Öztürk (2005) proposed pseudo-incorporation based on the assumptions that BDOs are non-referential, that they must be preverbal, and on the structural ambiguity of modifiers in (20). Kamali (2015) provides evidence against the first two assumptions. This left only the structural ambiguity of modifiers in (20) as evidence for a pseudo-incorporation analysis. However, it is shown in this section that the behaviour in (20) is not related to any kind of incorporation. Indeed, it is a result of strong/weak Case features of the objects.

Lastly, the necessity of weak Case for NPs can be questioned. Kamali (2015) accepts Öztürk's (2005) proposal that bare NPs check accusative case. The evidence for this comes from causative structures in Turkish, shown in (14-15). (14) shows that the agent of an unergative is accusative case marked when causativized, however, the agent of a transitive cannot get accusative case, (15b), due to double case constraint in Turkish (Taylan, 1984). This suggests that even though it is not morphologically realized, bare NPs in Turkish also check accusative case.

6 Conclusion

To conclude, I have proposed that in contrast to what has been claimed so far, a bare object is not the result of (any kind of) incorporation but is rather the result of strong/weak Case feature. It is shown that Turkish bare objects show distinct behaviours from both the incorporation analysis of Baker (1988) and the pseudo-incorporation analysis of Massam

[5] This structure is considered problematic in that the adjective *hızlı* is higher than the indefinite article *bir*. The status of indefinite article *bir* can be further discussed. The fact that *bir* is also used as numeral "one" in Turkish, it might function as Numeral, rather than a Determiner. Leaving this discussion for later, I point here that the important issue is *bir* is higher than the NP and AdvP, within the DP.

(2001). Pointing out that behaviour of non-derived modifiers like *kötü* "bad" are expected in the Two Object Position Hypothesis, I assert that this alternative approach to BDOs has greater empirical coverage than either the incorporation or pseudo-incorporation analysis, which leads to questioning the conceptual necessity of incorporation of BDOs. Under these circumstances, I suggest that BDOs do not incorporate. However, different from case-marked objects, BDOs check weak accusative Case feature.

References

Aydemir, Y. (2004). Are Turkish preverbal bare nouns syntactic arguments? *Linguistic Inquiry* 35: 465-474.

Baker, M. C. (1988). *Incorporation*. Chicago: University of Chicago Press.

Bliss, H. (2004). The semantics of the bare noun in Turkish. In I. Mezhevich and M. B. Dobrovolsky (Eds.), *Calgary Papers in Linguistics*, 25: 1–65.

Enç, M. (1991). The semantics of specificity. *Linguistic Inquiry* 22: 1–27.

Kamali, B. (2015). Caseless direct objects in Turkish revisited. *ZAS Papers in Linguistics* 58: 107-123.

Karimi, S. (2005). *A minimalist approach to scrambling: Evidence from Persian*. Berlin/Boston: De Gruyter, Inc.

Kayne, R. S. (1994). *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Kornfilt, J. (1984). Case Marking, Agreement, and Empty Categories in Turkish. *Doctoral dissertation*: Harvard University.

Kornfilt, J. (1994). Türkçe'de geçişim ve sözcük dizimine etkisi. *Dilbilim araştırmaları*, 42-53. Ankara: Hitit Yayinevi.

Kornfilt, J. (2003). Scrambling, subscrambling and case in Turkish, word order and scrambling. In Karimi, S. (Ed.), Word order and scrambling. Blackwell, Oxford, 125–155.

Massam, D. (2001). Pseudo noun incorporation in Niuean. *Natural Language and Linguistic Theory* 19: 153–197.

Öztürk, B. (2005). *Case, referentiality and phrase structure*. Amsterdam, Philadelphia: John Benjamins.

Öztürk, B. (2009). Incorporating agents. *Lingua* 119: 334–358.

Taylan, E. (1984). *The function of word order in Turkish grammar*. Berkeley, CA. University of California Press.

Travis, L. D. (2010). *Inner aspect: The articulation of VP.* Montreal: Springer.

Contact Information

Metehan Oguz

metehan.oguz@ucalgary.ca
University of Calgary
School of Languages, Linguistics, Literatures & Cultures
University of Calgary
CHD 514
2500 University Drive NW
Calgary, AB T2N 1N4
Canada

The Lives of Coordinate Structures: Evidence from Distribution and Cases of Three (Or More) Conjuncts

Brooklyn Sheppard

University of Calgary

Abstract

Coordination has long been ignored in much of the previous syntactic literature. The goal of this article is to begin the task of explaining a variety of phenomena involving coordinate structures, such as their distribution and cases where there are three (or more) conjuncts. I begin by comparing two previous accounts of coordinate structures and investigate if and how each approach could account for the phenomena of the distribution of coordinate structures and cases where there are three (or more) conjuncts. I then conclude that these structures are best analyzed as being headed by the conjunction itself. Finally, I propose a set of minimalist features on the conjunction itself that can help us account for both the distribution and selectional properties of coordinate phrases.

Key words: coordination, conjunction, minimalism

1 Introduction

This paper critically evaluates two approaches to coordinate structures through a direct comparison of how these structures account for certain data. Specifically, this paper looks at the issue of how each theory may explain the distribution of such constructions, as well as the lesser-discussed issue of how these theories could account for cases in which three or more conjuncts are being coordinated. The first approach considered is *The Adjunct Approach* by Munn (1993). According to this approach, the conjunction is merged with the second conjunct, creating a Boolean phrase (BP; referred to here as &P), and that this phrase is then adjoined to the first conjunct (Figure 1a). The second approach considered is *The Conjunction Phrase Approach* (Kayne, 1994; Johannessen, 1998; Shepherd, 2014). This approach states that the conjunction is the head of the &P and that this head selects the second conjunct as its complement and the first conjunct as its specifier (Figure 1b).

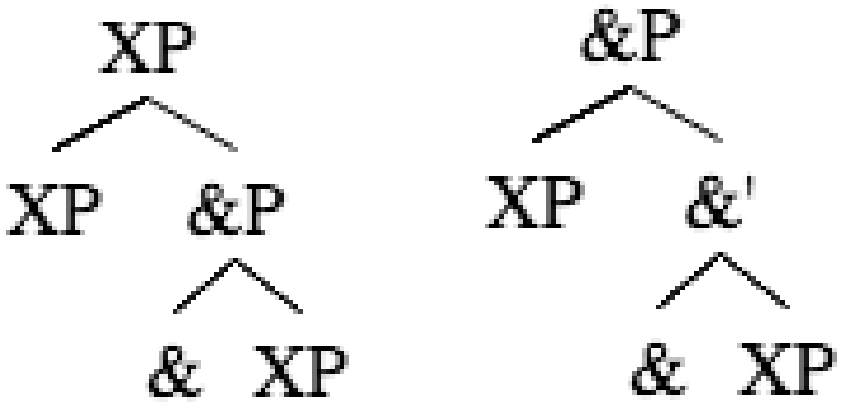


Figure 1. a. The Adjunct Approach *b. The Conjunction Phrase Approach*

For consistency, all phrases headed by the conjunction *and* (&) are labeled as &Ps, although these were originally labeled BPs in the Adjunct Approach.

Each approach is analyzed through the investigation of how the approach is able to account for both the distribution of coordinate structures as well as cases involving three (or more) conjuncts. From this analysis of the two approaches, I conclude that coordinate structures are best analyzed as being headed by the conjunction itself i.e. The Conjunction Phrase Approach. I further argue that both the issues of selectional properties of conjunctions as well as the distribution of these phrases can be accounted for by positing minimalist features on the conjunction itself.[1]

In Section 2, I outline the Adjunct Approach and showcases evidence in favour of this approach as well as outline its issues. In Section 3, I analyze the Conjunction Phrase Approach through a discussion of evidence in favour of and against this approach. Section 4

[1] That is, this paper assumes the minimalist syntactic framework first proposed by Chomsky (1995), and later presented by Adger (2003), Hornstein (2005), etc

then discusses whether each approach could account for cases involving three or more conjuncts and if so, how. Section 5 discusses if and how each approach could account for the distribution of coordinate phrases. In Section 6, I propose features on the conjunction *and*. Following these features, I account for both the selectional properties of these items and the distribution of conjunction phrases. This section also includes a brief discussion of the consequences of such a proposal. Finally, Section 7 summarizes the arguments made and provides direction for future work.

2 The Adjunct Approach

The main claim of this approach is that the structure of a coordinate phrase involves the conjunction, *and,* (labeled here as &) merging with the second conjunct (the XP on the right) to form a &P, followed by this &P being adjoined to the first conjunct (the XP on the left). This is demonstrated in Figure 1a, repeated below in Figure 2.

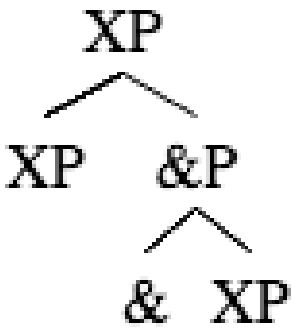


Figure 2. The general underlying structure of coordination under the Adjunct Approach.

To begin, Munn (1993) states that coordinate structures can be used to conjoin two phrases of different categories and that this is an argument in favour of the Adjunct Approach. This phenomenon is demonstrated in (1) below.

1) *Perot expects to run and that he'll win.*
(Munn 1993, p. 70)

In 1) above, we can see that the conjunction *and* conjoins the two phrases [*to run*] and [*that he'll win*]. From the presence of the overt complementizer *that* in the second conjunct, we can confidently say that this phrase is likely a CP. Although one could potentially analyze the phrase [*to run*] as being headed by a null C, this is not what is typically accepted in the literature (e.g. Adger, 2003; Hornstein, 2005). The reason for this is that *expect* is considered an ECM (exceptional case-marking) verb, where the subject of the embedded clause receives accusative case from the matrix verb (e.g. *Perot expects him to win*). Assuming that case checking cannot occur across a CP boundary, we would expect no null C in this case and that the phrase [*to run*] is indeed a TP. Therefore, Munn argues that the example in 1) appears to be a case of conjoining two unlike phrases – a CP with a TP.

It is not clear, however, that these are indeed two different categories. In 1), it appears that the first conjunct is instead a case of control. That is, *Perot* is both the agent of *run* and the experiencer of *expect*. Under the standard theory of control (Chomsky, 1981; Adger, 2003), we would then analyze this first conjunct [*to run*] as a CP with a controlled PRO. Following Chomsky & Lasnik (1993), PRO has a *u*Case feature that must be checked by a [*u*Case: NULL]. Such a feature is assumed to be found on a null C. Thus, in order for the PRO's case feature to be checked, this conjunct must be headed by a null C and therefore be a CP.

The second conjunct is clearly a CP with the presence of the overt complementizer *that*, and it could very well be the case that this is simply a case of two coordinated CPs. Another issue with this example lies in what happens if the example no longer contains the complementizer *that* as shown 2) below.

2) *Perot expects to run and he'll win.*

Here, we see that without the complementizer *that*, we cannot conjoin *to run* and *he'll win*, but instead can only get the reading of *expects to run* being conjoined with *he'll win*.

If, as Munn proposes, *to run* is indeed a TP, we should expect the elimination of the C *that* to make the coordination more acceptable. This is not the case as we can see in 2). This example then suggests that perhaps there is more to this example than Munn may have initially thought.

If, instead, it is the case that conjunctions can combine different types of phrases, why should this be evidence that the &P and the first conjunct are related by adjunction? It could be argued that if the two conjuncts were the complement and specifier of a single phrase, we would expect them to be more closely related (e.g. agreement) than if one were an adjunct above the phrase of the other. From this closer relationship, however, it does not necessarily follow that the category of the first conjunct should be identical to that of the second conjunct. In fact, it is quite typical that a certain head will accept specifier whose category is distinct from its complement. For example, the ditransitive verb *give* can take the DP argument [*a book*] as well as the PP argument [*to Mary*]. That is, even if we do see the coordination of two different categories, it does not rule out the Conjunction Phrase Approach as a viable option. More importantly, it is noted by Munn that these cases of distinct categories being coordinated are rare and only occur with very specific combinations and orderings. It is not clear, then, how adjunction would be able to account for such constraints on these categories and ordering.

A second argument in favour of the Adjunct Approach comes from what pieces we are able to extract from this structure, as seen in 3).

3) a. *John bought a book and a newspaper yesterday.*
   b. *John bought a book yesterday, and a newspaper.*
   c. **John bought a newspaper yesterday a book and.*
   d. **John bought a book and yesterday, a newspaper.*
   (Munn 1993, p.15)

In 3), in the conjunction [*a book and a newspaper*] (as seen in 3a)), it is possible to move the second conjunct with the conjunction out to the right edge 3b). In 3c), we cannot, however, move only the first conjunct and the conjunction, while leaving the second conjunct behind. Munn further argues that 3d) is also ungrammatical as the second conjunct has been moved without the conjunction itself. While this may be grammatical under a certain interpretation (i.e. with focus on *yesterday*), it is ungrammatical under the interpretation that both the book and the newspaper were bought yesterday. The facts in 3) suggest that the conjunction and the second conjunct form a maximal projection, and therefore form a constituent.

It is important to note that these facts have also been used in previous work as an argument for the Conjunction Phrase Approach (Shepherd, 2014). In their dissertation, Shepherd points out that these facts can be interpreted as evidence that the conjunction and second conjunct hold a Head-Complement relation, and thus also form a constituent. While this is true, this is also the case under the Adjunct Approach. In the Adjunct Approach, the conjunction and the second conjunct still hold a Head-Complement relationship, just as in the Conjunction Phrase Approach. The difference between the two approaches instead lies in whether this Head-Complement construction is a maximal projection and how the first conjunct is related to these two elements. This argument of whether the second conjunct and the conjunction form a constituent can therefore not help us decide between the Adjunct Approach and the Conjunction Phrase Approach.

The third argument of the Adjunct Approach is, I would argue, the strongest. This argument notes the fact that the distribution of these coordinate structures is clearly determined by what elements are being conjoined. This is demonstrated in 4) below.

4) a. [*Bill and Dave*] *played catch* → [*They*] *played catch*
   b. *Bill* [*ran a marathon and ate some cake*] → *Bill* [*did so*]

In 4a), we see a DP coordinated with a DP. This entire coordinate structure can easily be replaced by a simple DP pronoun. 4b) shows another example of this where our two conjuncts are instead VPs, and that this coordination can be replaced by the single VP *did so*. Here it can be seen that the distribution of the coordinate structure is determined not by the presence of a conjunction, but by the category of the conjuncts. This argument is discussed in more detail in Section 5, where I discuss how each theory of coordinate structure can account for distribution.

One issue with the Adjunct Approach that has not been brought up yet in the literature cited is that of agreement. Let us examine 5) below.

5) a. *John and Mary is*/are hiking.*
   b. *John is/are* hiking.*

In 5a), we see that in the case of conjoined DPs, we see plural agreement on the verb. It is clear from the contrast with 5b) that it is the presence of two conjuncts that triggers this agreement. If the first conjunct is simply an adjunct, we would not expect it to play any role in agreement. It is then not clear, under this approach, why we would see plural agreement in the presence of two conjuncts if one is simply an adjunct.

Although there have been several arguments made in favour of the Adjunct Approach, all have been questioned here. It is true that this approach can easily account for the distribution of coordinate phrases, however, the variety of other arguments put forth have some major flaws. The following section next outlines a different approach to such structures and discusses the arguments both in favour and against such an approach.

3 The Conjunction Phrase Approach

As mentioned in Section 1, the main premise of the Conjunction Phrase Approach is that a conjunction acts as the head of a conjunction phrase (&P), with the first conjunct being in the specifier position and the second conjunct in the position of the complement. This is illustrated in Figure 1b, repeated here as Figure 3.

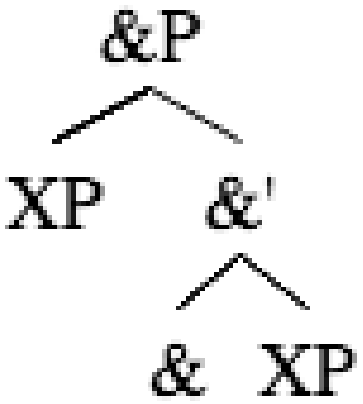


Figure 3. The general underlying structure of coordination under the Conjunction Phrase Approach

Several syntacticians have argued for such an approach (Kayne, 1994; Johannessen, 1998; Shepherd, 2014) for various reasons.

A first argument in favour of the Conjunction Phrase Approach is that provided by Kayne (1994) as a consequence of his theory of linearization. This theory states that the linear ordering of syntactic elements is directly determined by their c-command relationship. That is, any syntactic element which c-commands another, must precede it as well in linear order. In discussing the consequences of such a powerful claim, Kayne states that this theory of linearization can also account for the ordering of conjuncts in a coordinate structure. For example, in the flat structure hypothesis (Figure 4), the conjunction *and* c-commands both the first and the second conjunct which, according to Kayne's theory, would predict that the conjunction also precedes both conjuncts in terms of linear order. This is clearly not the case, and so Kayne argues that the flat structure hypothesis must be incorrect.

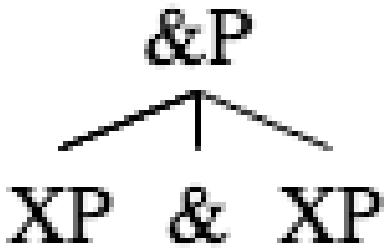


Figure 4. Coordination under the Flat Structure Hypothesis.

Kayne uses this comparison as evidence for the Conjunction Phrase Approach. It must be noted, however, that Kayne's theory of linearization does not consider the Adjunct Approach as it is stated here. Thus, although this theory has been used to argue for the Conjunction Phrase Approach in subsequent work (e.g. Shepherd, 2014), this work has not considered how Kayne's theory would rule out the Adjunct Approach. In the Adjunct Approach discussed in this paper, it is still the case that the first conjunct c-commands the conjunction, and the conjunction c-commands the second conjunct. Thus, it does not appear that the Adjunct Approach necessarily violates Kayne's theory. To conclude on this matter, although previous work has used Kayne's theory to argue in favour of the Conjunction Phrase Approach when compared to the flat structure hypothesis, it appears that this theory cannot help us decide between the Conjunction Phrase Approach and the Adjunct Approach.

The second argument that has been used in favour of the Conjunction Phrase Approach follows the arguments made by Shepherd in their (2014) dissertation. Shepherd explored the odd case marking seen in English DPs in coordinate structures and collected grammaticality judgements from native speakers. Several examples of this are given in (6).[2]

6) a. Me and him went shopping.
   b. She and he went shopping.
   c. She and him went shopping.
   d. *Him and she went shopping.
   e. She saw me and him yesterday.
   f. *I saw she and he yesterday.

From the data collected, Shepherd is able to make several generalizations. First, it seems that the first person singular DP *I* behaves differently from all other DP pronouns (see Shepherd, 2014 for arguments). After assuming that *I* should be analyzed separately, Shepherd makes the following generalizations. In a subject position, NOM+NOM,

[2] These are not exact examples from Shepherd, but are instead used here to illustrate different possibilities of case markings in English according to Shepherd's study. Although Shepherd decided to make these generalizations, the data seems quite complex, and much of these acceptability judgements are subject to inter-speaker variation. Regardless of the exact judgements and generalizations made by Shepherd and their participants, the data does indeed show clear ordering constraints that may be difficult for the Adjunct Approach to explain.

NOM+ACC, and ACC+ACC are all possible case combinations (6b, 6c, and 6a, respectively), while ACC+NOM is never possible (6d). In an object position, however, the only acceptable case combination is ACC+ACC (6e, 6f). Given this data, it seems that the case marking of the second conjunct is somewhat dependent on the case of the first conjunct – that is, in a subject position, we may only see NOM case on the second conjunct when the first conjunct also has NOM case. Shepherd argues that this is due to the fact that the first and second conjuncts form a chain and enter into an agreement relationship. Specifically, it is argued that the first conjunct's case feature is checked by the [*u*Case: Nom] on T (Chomsky, 1995) and that this NOM case feature is then passed on from the first conjunct to the second one through a Specifier-Complement relationship. This explanation for the case marking seen in English DP conjunctions is evidence that the first and second conjuncts do indeed have a much closer relationship than Munn (1993) originally proposed. On the Adjunct Approach, it is not clear why such case marking patterns seen in coordinate structures are attested. In a Conjunction Phrase Approach, however, it is quite conceivable how the Specifier and Complement of a single phrase could share case features. Thus, under the Conjunction Phrase Approach, we are able to explain the odd case marking patterns seen in coordinated DPs. It is unclear how the Adjunct Approach would account for the data in 6).

There are also several other arguments for such a structure that are not specifically outlined in the previous literature. First, as was stated in Section 2, Munn (1993) uses the fact that there are certain cases in which we see coordination between two phrases of different types as evidence in favour of the Adjunct Approach. An example of this is given in 1), repeated here as 7).

7) *Perot expects to run and that he'll win.*

Even if we agree that this is indeed a case of coordination of differing phrase types (see Section 2 for a detailed discussion), Munn (1993) himself states that these cases are exceptional and subject to ordering constraints. The fact is that the vast majority of cases (if not all) require that the two conjuncts are matching in category. A process of adjunction (as proposed by the Adjunct Approach) would not be constrained enough to predict this. Under the conjunction phrase approach, however, it is possible that this requirement for category matching and/or ordering constraints is related to some aspect of the Spec-Complement relationship in &Ps.

Another argument in favour of the Conjunction Phrase Approach is that a conjunct *must* attach two elements. Under the Adjunct Approach, the conjunction and the second conjunct together form a maximal projection, and the &P is an adjunct to the first conjunct. It is thus not clear under such an approach why it is required that we have two conjuncts and not simply one. This requirement is better explained by the Conjunction Phrase Approach, in that the head of a phrase may require both a complement and a specifier. Specifically, under the feature checking approach of minimalist syntax, we could assume that the head & has two features that must be checked or valued by different phrases (see Section 6 for a detailed discussion).

In the following sections (Sections 4 and 5), I investigate how each of these theories would hold up when trying to account for the structure of coordinate phrases with three or more conjuncts and, importantly, how each of these theories can account for the distribution of the coordinate structures discussed here.

4 Three (Or More) Conjuncts

This section discusses the consequences of cases with three or more conjuncts for each theory outlined in Sections 2 and 3. We will first consider the consequences for the Adjunct Approach in Section 4.1, followed by the consequences for the Conjunction Phrase Approach in 4.2.

4.1 Three (Or More) Conjuncts Under the Adjunct Approach

If we assume the Adjunct Approach, we could very well assume that a third conjunct simply means that we can optionally adjoin more conjuncts. Such a construction is depicted in Figure 5.

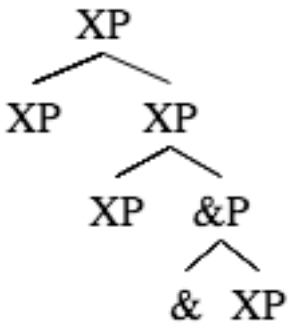


Figure 5. The general structure of a coordination of three conjuncts under the Adjunct Approach

Here, we see that the third conjunct is the complement of the conjunction and that the first and second conjuncts are adjuncts added on top of the &P. Such a structure is proposed for sentences like that in 8) below.

8) *John, Bill and Mary play soccer together*.

It is easy to see how this theory would then be able to account for the fact that we can add an infinite number of conjuncts, as they could just continue to be adjoined. This is exactly what is proposed briefly by Munn. It is also important to consider the fact that these cases may also contain an optional *and* between the first and second conjunct, as can be seen in 9).

9) *John and Bill and Mary play soccer together.*

The case shown in 9) can also potentially be explained by the Adjunct Approach. In this case, we would assume that the *and* between *John* and *Bill* selects the DP [*Bill and Mary*] as its

complement, forming the &P, and the DP John is adjoined to this &P. Such a solution, however, requires us to agree that a structure like that in 8) is underlyingly different than the one in 9). Each of these different structures are shown in Figures 6a and 6b below.

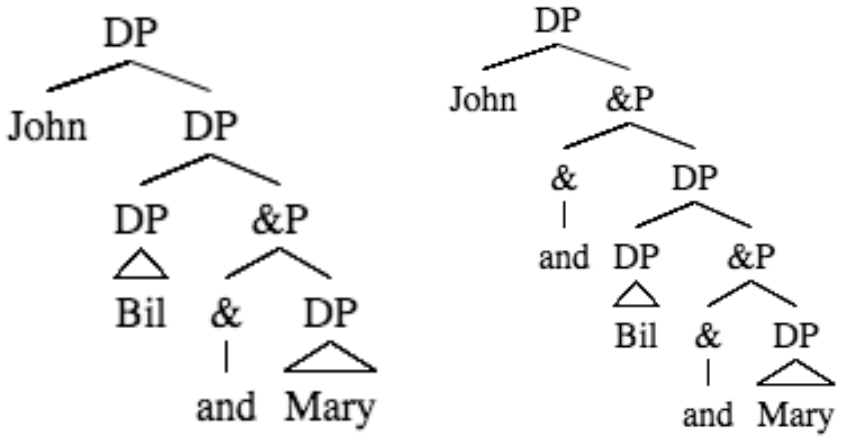


Figure 6. a. No optionally overt & *b. Presence of an optionally overt &*

It seems odd to assume that these are entirely different structures rather than simply a difference in whether or not one chooses to overtly pronounce the higher conjunctions.

Another issue with the Adjunct Approach in light of three conjuncts is the fact that it assumes DPs are adjoined to other DPs in the case of three or more conjuncts. This is not typically a process in English, and therefore begs the question of when is such adjunction possible? It seems that a DP may only be adjoined to another DP when there is some DP lower that is adjoined to a &P. Such a constraint is very arbitrary and, therefore, does not seem to be a plausible explanation.

4.2 Three (Or More) Conjuncts Under the Conjunction Phrase Approach

If we apply the Conjunction Phrase approach to a case with three or more conjuncts, a null element is necessary. If coordinate structures are headed by the conjunction, each level of the coordinate structure must contain a conjunction. In a case of three conjuncts, only one overt conjunction is required. To account for these structures under this theory, we would then need to posit that an overt conjunction can project an infinite number of optionally null conjunctions. An example of such a derivation is given in Figure 7.

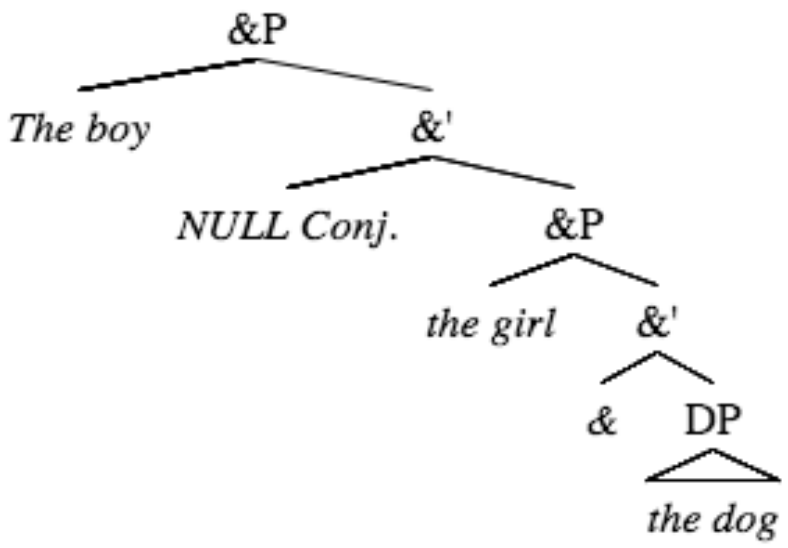


Figure 7. Derivation of a &P assuming null conjunctions.

In Figure 7, we see that the optionally null conjunction takes the lower &P as its complement and the highest DP as its specifier. Given that we already have a & head, it is simply an option of whether or not the speaker chooses to pronounce the higher conjunctions or not.

This is not the first time something like this has been proposed in the grammar. For example, it is commonly agreed that verbs show some sort of VP shell, or little *v* layer. In this case, we also see an element (the V) with overt content projecting an outer layer that is null (the little v; or VP shell analysis in earlier work (e.g. Adger, 2003; Hornstein, 2005)). A similar analysis has also been proposed for nouns, in that we have a little *n* shell above the overt N. The proposal for what we see in coordinate structure with three or more conjuncts then mirrors these phenomena. That is, we see an overt conjunction lowest in the tree and higher optionally null &-shells.

While this solution may seem elegant, it comes with major consequence. It was stated in Section 3 that in the vast majority of (if not all) cases, conjunctions take two elements of identical category. As was also expressed in Section 3, it seems that these cases where there is a category mismatch are highly exceptional and highly constrained. In the case of Figure 7 above, we see a conjunction selecting both a &P and a DP. These are clearly of different categories and this solution would therefore require us to accept yet another exceptional property of conjunctions – that they may conjoin two unlike elements in the case of three or more conjuncts. At face value, this seems like a major flaw in the current approach. This problem, however, stems from the issue of the distribution of such phrases. This is discussed in detail in the following section and a solution to this type of problem is proposed in Section 6 of this paper.

5 Explaining the Distribution

It is well known that the category of any syntactic element can be determined by its distribution. For example, any DP can be substituted by any other element of the same category – that is, a DP. This is demonstrated in 10) below.

10) $_{DP}$[Bill] can do math.
$_{DP}$ [The boy] can do math.
$_{DP}$[He] can do math.

In this section, we explore how the distribution of coordinate phrases can be accounted for under the two theories discussed in this paper. Section 5.1 looks at distribution under the Adjunct Approach, while 5.2 focuses on distribution under the Conjunction Phrase Approach.

5.1 Distribution Under the Adjunct Approach

As was noted briefly in Section 2, one argument in favour of the Adjunct Approach is the fact that it can account for the distribution of coordinated phrases. The reason for this is that this approach assumes that the head of the first conjunct is the head of the entire coordinate structure, and will thus pattern just as the first conjunct would on its own. Thus, if we have two DPs conjoined, the overall structure is headed by the higher DP, and will show the distribution of a DP. It can be seen that the distribution of coordinate structures then follows naturally from the Adjunct Approach.

5.2 Distribution Under the Conjunction Phrase Approach

Distribution of coordinate structures under the Conjunction Phrase Approach does not come as naturally as it does to the Adjunct Approach. If the head of the phrase is the conjunction itself, it is not clear why the conjunction phrase shows the distribution of whatever categories it immediately dominates. An example of this was given in 4), repeated here as 11).

11) a. [*Bill and Dave*] *played catch* → [*They*] *played catch*
b. *Bill* [*ran a marathon and ate some cake*] → *Bill* [*did so*]

Here, we see that the &P in 11a) shows the distribution of a DP, while the &P in 11b) shows the distribution of a VP. Hornstein et al. (2005) argue that selectional properties should be blind to any features within the phrase that are not directly under the head. It seems as though the distribution of these phrases is instead depended on the category features of the complement and the specifier – not the head. At first glance, it seems that the specifier and the complement are being selected for here, making conjunction phrases different from all of other phrases we know. Thus, under this approach, the distribution of conjunction phrases is not so easily explained. In the following section, I provide an overall evaluation of both approaches to coordinate structure and a possible solution to the issue of distribution and selectional properties of conjunctions is proposed.

6 A Possible Solution

6.1 Evaluation and a Possible Solution

To begin this section, I provide a brief summary of the arguments discussed in the previous sections as well as a general evaluation of each at this point in time. I will first note that neither of the theories discussed here can successfully explain the selectional properties of conjunctions, which should be considered as a serious flaw.

It is also clear that each of these theories have some serious issues that must be explained before either can be considered satisfactory. The Adjunct Approach can easily account for the distribution of coordinate phrases. Aside from this, however, it seems that there are some serious problems. As discussed in previous sections, this approach fails to constrain the structure of coordinate phrases in any way. That is, it is not evident from this approach why a conjunction *must* have two conjuncts or why these two conjuncts must match in category. There are additionally issues of agreement in this approach.

The Conjunction Phrase Approach, on the other hand, can account nicely for the majority of the patterns we see in conjunctions. This approach can account for all of the issues discussed in this paper other than the distribution. The fact that the Conjunction Phrase Approach can account for everything, but the distribution of these phrases makes it quite appealing. This approach, however, also has the issue of explaining why higher &s (in cases with three or more conjuncts) connect &Ps with other XPs, seemingly violating the fact that conjunctions only conjoin elements of identical category.

As was stated in Section 4, this issue is directly linked to the issue of distribution. That is, to explain the distribution of &Ps, we must realize that they show properties of whatever it is they are conjoining. Since these phrases clearly behave as the items they conjoin in terms of distribution, it is not surprising that they seem to satisfy the selectional restrictions of higher &P shells (see Figure 7 from Section 4).

If there is then a way for this approach to explain the distribution of coordinate structures, it would fix a major flaw in the theory so far. Although several previous theories have attempted to predict the structures of coordinate phrases, exactly what features a conjunction may contain has yet to be proposed.

First, in order to account for the distribution of &Ps, I propose that there is some interpretable feature, [X:], on & which must be valued by the category, X, of its complement through a c-command relationship. In essence, such a proposal assumes that conjunctions are without a category, and that this feature must instead be valued by one of its arguments. The purpose of using the variable X is to express the fact that & is impartial to which category its complement is.

As was discussed in Section 3, it is also the case that a conjunction requires two arguments. This being the case, I propose that there is a second feature on & that is similar to an uninterpretable selectional feature. Such a feature, having the shape of *u*X, must be checked by the specifier of &. The X on this feature now tells us that the category of the specifier must be identical to the category of the complement of & which is also now the

category of & (as it has been valued by the complement). An example derivation involving these two features is given in Figure 8.

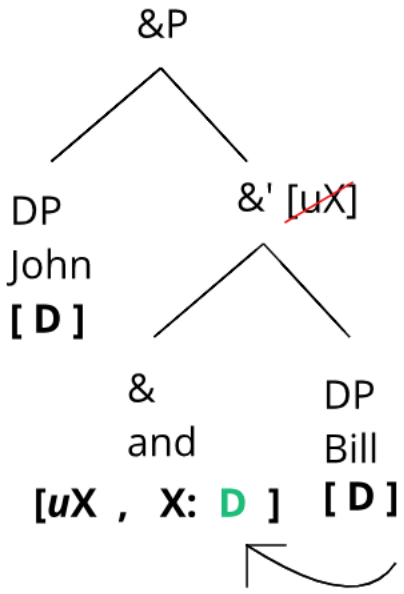


Figure 8. The derivation of a &P, assuming the [uX, X:] features on &.

In Figure 8, we see that the & first merges with the DP Bill, creating &'. During this process, the D feature on Bill values the category feature [X:] on &. The uninterpretable *u*X feature must then be brought to the &' level as it has yet to be checked. This uninterpretable feature is then checked by the merging of the specifier. Since the &'s category has already been valued, there must be some constraint on *u*X which states that the category it selects must match with the category of & and of the complement of &.

Both uninterpretable selectional features and interpretable valuable features are already present in minimalist syntax. The only additional element that coordination brings us is how the features of this conjunction are somehow able to communicate with one another. In order for the *u*X to select the correct item in the specifier, it must know what category the [X:] has been valued with. A detailed account of the exact mechanisms responsible for such communication is beyond the scope of this paper, but is vital to a complete theory of coordination. Finally, the following section outlines some consequences of the present proposal.

6.2 Consequences

The goal of this section is to briefly outline some consequences of the proposal made in the previous section. One consequence that was mentioned in Section 6 is that we can now explain the selectional properties of higher conjunctions. That is, if a higher &-shell is selecting the lower &P, this &P must have already had their category valued and is therefore

masquerading as an XP. Thus, it seems that higher &s are indeed correctly selecting for the correct features on the head of the &P.

Another interesting consequence comes with the coordination of DPs specifically. In a case where we have coordinated DPs in the subject position, it is clear that the entire &P moves up to satisfy the EPP and not simply the higher DP within the &P. This fact is exemplified in 12) below.

12) [TP Bill & John vP[<Bill & John> ate a sandwich]

If the head of the &P, &, has its category valued by its complement (a DP), it is then behaving just as a DP would. This &P is then the highest "DP" and we would expect the EPP to target it. It is clear that the consequences outlined here are far from complete, and must be further considered in future work. From this brief section, however, we can see how this proposal brings with it some clear benefits.

7 Conclusions & Future Directions

This paper has reviewed two different approaches to the structure of coordinates phrases. Specifically, I have argued that the Conjunction Phrase Approach, could account for the majority of phenomena discussed, but lacked the crucial ability to explain the distribution of such structures.

Given the fact that the Conjunction Phrase Approach can account for the majority of coordinate phenomena, while the Adjunct Approach cannot, it seems that the Conjunction Phrase Approach is a more viable option for a theory of coordination. I additionally argue that the Conjunction Phrase Approach can indeed account for both the distribution and the selectional features of coordinate structures if we posit certain minimalist features on the conjunction itself. Namely, we must posit a [X:] feature in order to account for the distribution, and a *u*X feature in order to account for the selectional properties of conjunctions.

Our work on conjunctions is far from complete. As stated in Section 6, a clear next step is to provide a more complete discussion of the consequences of the proposal made here. It was also stated in Section 6 that the exact mechanism of ensuring the complement and the specifier of & be of the same category is yet to be explored.

There are also several other issues that were not touched on at all in this paper. For example, this paper has (as have many others) only looked at the case of the conjunction *and*. The conjunctions *or* and *but* are not considered at all. In order for a theory of coordination to be complete, we must have an analysis that can account for the differences seen between different types of conjunctions as well.

Another limitation of the present paper is that it only considers English. It would be extremely beneficial to a theory of coordination for a future project to focus on documenting the behaviour of coordinated structures cross-linguistically. There are yet even more interesting questions we have left to solve when it comes to coordination. For

example, there seems to be some constraints on binding in coordination. In 13) below, we see that *himself* can only be co-indexed with *Tom*, but not with *Bill*.

13) *Tom*$_{i}$ *likes Bill*$_{j}$ *and himself*$_{i/*j}$.

The world of coordinate structures is one that has largely been ignored in the syntactic literature. As I hope is clear from the present paper, these structures are immensely complex. It can often be the case that such structures cause problems for theories of syntax (e.g. binding, agreement, case, etc.). Rather than ignore such structures, the present paper attempts to begin the difficult task of finding a place for these structures in a theory of syntax. Although these constructions are difficult to explain syntactically, they are not at all rare in the actual productions of speakers, and are therefore an important aspect of any syntactic theory. In order for a theory of syntax to be complete, we must be able to account for the strange lives of coordinate structures – the present paper aims to act as a first step in this pursuit.

References

Adger, D. (2003). *Core syntax: A minimalist approach.* Oxford: Oxford University Press.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, N., & Lasnik, H. (1993). The theory of principles and parameters. In *Syntax: An international handbook of contemporary research*, Vol. 1, ed. Jacobs, J., von Stechow, A., Sternefeld, W. & Vennemann, T. 506-569. Berlin: Walter de Gruyter.

Hornstein, N., Nunes, J., & Grohmann, K. (2005). *Understanding Minimalism*. Cambridge/New York: Cambridge University Pres

Johannessen, J. B. (1998). *Coordination*. Oxford: Oxford University Press.

Kayne, R. S. (1994). *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Munn, A. B. (1993). *Topics in the syntax and semantics of coordinate structures* (Doctoral dissertation, University of Maryland at College Park).

Shepherd, A. (2014). *" Him and me" or" he and I": a minimalist analysis of case variation in English conjunction* (Doctoral dissertation, University of Southampton).

Contact Information

Brooklyn Sheppard

brooklyn.sheppard1@ucalgary.ca
University of Calgary
19 Cawder Drive, N.W.
Calgary, AB, T2L 0L8
Canada

Expletive Pronoun Deletion Elicitation

Jesse Weir

University of Calgary

Abstract

Under the current minimalist program, most spoken English sentences are required to have an overt subject by virtue of the strong EPP feature at T. In informal spoken English however, it is possible to omit an expletive or referential pronoun before the raising verb seems. In sentences with a referential pronoun subject, participants will take notice of referential pronoun constructions and assign stress to the subject pronoun, even if it is absent in what they are reading aloud (Weir, 2019). When listening to degraded audio, interspeaker phonological variation plays a significant role in determining the likelihood of reproducing a sentence with the subject pronoun absent. Previous research on the topic of subject pronoun deletion in spoken English has been approached from both a pragmatic approach (Mack et al. 2012) and a phonological approach. The results of the present study suggest a combination of these two approaches which explains the interspeaker variation in the audio recreation data of the present study. This paper also argues that the phonetic patterns that appeared in Weir (2019) are the result of participants adjusting based on the syntactic differences in sentences based on the referentiality of the pronoun subject.

Key words: Pronoun deletion, Phonology, Syntax

1 Introduction

Sentences containing a raising verb like *seems* can either employ an expletive pronoun such as *it* or *there* in English, or raise the embedded clause subject to a higher position in order to satisfy the EPP feature present on the tense head associated with *seems.* Examples of these sentences can be seen in (1).

1) a. (It) seems he likes baking
   b. (There) seems to be smoke coming off the engine
   c. (It) seems to be broken

The pronouns in (1a) and (1b) are examples of expletive pronouns which, under current minimalist theories, are brought into the derivation to satisfy the strong EPP feature, and do not receive a theta role in the process. The pronoun in (1c) is a referential pronoun, which was present in the derivation before *seems* merged in, and received a theta role from its initial position in the embedded clause. All three of these types of pronouns can be deleted in informal spoken English, although, context is required to set up a sentence like (1c). Deletions like those in (1) seem to be acceptable to the majority of native English speakers in a spoken context, but this acceptability decreases when they are presented in writing[1]. The results of previous studies have indicated that participants will make phonetic adjustments to their speech to account for the referentiality of a subject pronoun, but this study finds that it is unlikely that this may be a unique result of the cross-modal delivery of previous research. This paper argues that the phonetic variation in Weir (2019) is influenced by syntax due to the method of delivery of the stimuli, while the results of this study and Mack et al. (2012) are likely influenced by phonology and pragmatics.

Section 2 of this paper will consist of a review of several works concerning both the deletion of words in spoken English, and the use of speech rate modulations to change what participants perceive. Section 3 will outline the present study methodology and results. Section 4 will discuss the results in detail while providing possible explanations and implications for them. Section 5 will provide future directions to both improve the methodology of the present study and conclude the paper.

2 Background Literature

The literature discussed in this section concerns experimental views on the deletion of left edge material, specifically subject pronouns, in spoken and written English (Mack et al, 2012; Weir, 2019).

[1] When corresponding with my supervisor about this topic over email, the Grammarly™ extension found sentences with the expletive subject pronoun deleted to be incorrect and recommended I change it before sending my email.

2.1 Mack et al. (2012)

The research conducted by Mack et al. (2012) uses a novel variant of the speech restoration method to test participants' usage preferences in sentences with expletive *it* as the subject. "Speech restoration occurs when listeners hear, or think they hear, segments of speech that are either absent or highly degraded in the acoustic input." (Mack et al., 2012, p. 3) By using the speech restoration method, and presenting listeners with degraded and ambiguous audio, Mack et al. (2012) aimed to examine listeners' usage preferences based on what they think they may have heard. When looking at expletive subjects in English, three observations are noted. First, initial deletion can apply to many different linguistic elements which suggests a prosodic phenomenon as opposed to a syntactic one. This is due to the fact that a syntactic explanation may require the creation of many ad hoc rules to explain all of the potential deletions, while a prosodic explanation would be much simpler. Second, only sentence initial subjects can be deleted in English (shown in (2)). Finally, unstressed constituents cannot be left behind and become the start of the sentence after subject deletion (shown in (3)), which suggests a phonological constraint that prevents English sentences from beginning with weak prosodic material.

2) a. (It) seems like Tom is happier than usual today.
   b. *Today (it) seems like Tom is happier than usual.
   (From Mack et al., 2012: 4)

3) a. (It's) pretty nice out today.
   b. *(It) is pretty nice out today.

While these examples do suggest a phonological analysis is sufficient to explain subject deletion in English, Mack et al. (2012) point out that there are pragmatic constraints and differences in sentences with deleted subjects. The deletion of a subject makes the sentence more 'subjective' than sentences with an overt subject (Mack et al., 2012). This difference is most obviously seen when comparing a first-person experiencer and a third-person experiencer as seen in (4).

4) a. It looks to me/Sam like things are going pretty well.
   b. (It) looks to me/?Sam like things are going pretty well.
   (From Mack et al., 2012: 5)

The questionable acceptability of a deleted subject when using a third person experiencer suggests that sentences with deleted subjects like (4b) express some sort of immediate judgement information, that only clearly allows the deletion of the subject pronoun when dealing with a first-person experiencer. The idea of immediate judgement in this case refers to a judgement formed by the speaker just before the sentence was uttered (Mack et al., 2012). With a third person experiencer, it is no longer an idea realized by the speaker and cannot be classified as an immediate judgement. Attempting to delete the third person subject pronoun makes the sentence ungrammatical. By expanding this immediacy to both personal and temporal immediacy, Mack et al. (2012) predict that participants will be more

likely to produce sentences with deleted subjects that have some sort of immediacy, in the sense that they have recently formed their own judgements on them.

The study conducted by Mack et al. (2012) had participants come into the lab and listen to short pre-recorded conversations consisting of three sentences. The final sentence of each conversation was a sentence that would permit the deletion of an expletive subject. An example of one of the conversations used can be seen in (5).

5) A: Joe and I are studying like crazy for the German exam.
   B: Oh yeah?
   A: It seems/seemed to me/him like there are thousands of words to learn.
   (Mack et al., 2012)

The conversations that were created controlled for both temporal immediacy by alternating between *seems* and *seemed*, as well as personal immediacy by alternating between a first-person and third-person experiencer. The 24 stimuli conversations also covered the three main verbs that were examined in a corpus search they had conducted (*seem, look,* and *sound*) (Mack et al., 2012). The audio of the conversations was presented to participants with the subject position distorted by a white noise filter to make it ambiguous whether the subject was present or absent (Mack et al., 2012). No manipulation was done to the audio to change the rate of speech or pitch. Participants were asked to listen carefully to the audio and, when prompted, repeat the last sentence that they heard out loud before then being asked to manually type the first five or six words of that same sentence (Mack et al., 2012). A strong emphasis was placed on ensuring that participants repeated exactly what they heard, regardless of ungrammaticality or any other perceived 'mistakes' (some of which intentionally occurred in the filler items) (Mack et al., 2012).

An analysis of the spoken data found that participants would produce an expletive subject 68.3% of the time, with participants being more likely to produce the subject pronoun in both past tense situations, and situations where the experiencer was third-person (Mack et al., 2012). The results of this support their findings from a previous corpus analysis and provide further evidence that both temporal immediacy and personal immediacy have a pragmatic effect on the likelihood of an expletive subject pronoun being deleted. When comparing the spoken results to the written results that participants provided, Mack et al. (2012) found that participants were more likely to produce the expletive subject pronoun in writing compared to their spoken responses. This difference is likely "due to effects of register or prescriptive norms" (Mack et al, 2012, p. 11).

By showing the effects of temporal and personal immediacy on the rate of pronoun deletion, the data in this study suggest that a purely phonological approach to deleted subjects is not sufficient. Mack et al. (2012) conclude their article by stating that usage preferences potentially alter the grammatical structure that participants assign to ambiguous input they are hearing. The present research aims to look beyond these usage preferences, and instead consider the effect that the phonetic qualities of the input might have in the rate of deletion of the subject pronoun.

2.2 Weir (2019)

Research conducted at the University of Calgary investigated the phonetic changes that participants made when asked to read aloud sentences that contain a subject pronoun, versus sentences that had the subject pronoun omitted. The purpose of the research was to look at the adjustments that occurred on the immediately adjacent raising verb *seems*, by having participants read sentences like (6) presented on a computer monitor.

6) The toaster did not pop on schedule.
   (It) seems to be broken.
   I guess it's time for a new one.

The key sentence was always the second sentence, ensuring that participants had sufficient context for all items. Each set of sentences like (6) would have two versions, one that contained the subject pronoun in the second sentence, and one where the subject pronoun was absent. Each participant would only see one version of each sentence, and would see an equal amount of each type of sentence. The stimulus lists were counterbalanced to control for any possible effects of item ordering. The research explored sentences that used expletive *it*, expletive *there* and referential *it* as subject pronouns, and only looked at the present tense verb *seems* (Weir, 2019). Participants were presented with a blank screen, and upon advancing in the experiment, would be presented with a set of sentences like (6) and immediately begin reading. Unlike Mack et al. (2012), the study by Weir (2019) avoided the potential confound of recovering a non-expletive pronoun by not using a copy-raising construction with any experiential verbs.

Initial measurements taken concerned the length of the raising verb *seems* (measured in milliseconds) in isolation, as well as the fundamental frequency (F0) at the onset of the vowel in *seems* in an attempt to observe any significant changes in between the two conditions. The research found that there was no significant difference in either length or frequency between conditions where the subject pronoun was present, versus conditions where it was absent. Instead, the research found that participants would make distinction of both length and F0 of the raising verb based on the type of pronoun that belonged in the subject position (Weir, 2019). Further to this point, Participants would say *seems* differently after an expletive *it* pronoun, compared to any other type of pronoun. This pattern occurred independently of pronoun presence in what participants were asked to read aloud (Weir, 2019). This suggests that participants were processing the type of pronoun that belonged in the subject position, and adjusting the way that they were saying the word *seems* to conform to this underlying pattern, even when the pronoun was absent in what they were reading (Weir, 2019). It was also found that, on average, participants would speak the raising verb 0.140 ms faster if the subject pronoun was an expletive *there* compared to both expletive and referential *it*. Further to this, participants would speak the word *seems* 0.220 ms faster when the pronoun was referential *it* compared to expletive *it*. Both of these rate phenomena occurred independently of pronoun presence (Weir, 2019). When looking at the F0 values of *seems* after the referential *it* pronoun, Weir (2019) found that participants would significantly lower their F0 value 6.75 Hz on average when the pronoun in the subject position was a referential *it* (Weir, 2019). This occurred regardless

of the pronoun presence, suggesting that even when the pronoun was omitted from the sentence that participants were reading, they would still have a lower F0 value on the raising verb *seems* on average (Weir, 2019). The explanation offered by Weir (2019) is that the referential *it*, due to its semantic content and theta role, will receive some amount of lexical stress which results in a decreased F0 on the word that immediately follows it. The fact that this occurs independently of the pronoun presence might indicate that participants are processing the sentence as though the pronoun were present in the subject position, which would align with a PF deletion hypothesis.

When first designing the study that was conducted in 2019, there was an underlying hypothesis that speed of delivery of the sentence was affected in these sentences somehow. Whether it was a result of sentences having a slower rate of speech due to the pronoun being absent and more emphasis being placed on the sentence to denote that, or if it would result in sentences without the subject pronoun being spoken faster due to being associated with quick speech, there was no definitive evidence either way. The goal of the study in this paper was to ultimately try and determine whether such associations could be drawn between speech rate and deletion of subject pronouns. The hypothesis behind this study is that if there is a meaningful distinction between sentences that contain the subject pronoun versus those that do not, we might be able to use artificially altered audio to try and recover those judgements from participants when they are asked to recreate them. The previous research on this topic did not produce any significant patterns to suggest this, but it may have been affected by the cross-modal delivery of written stimulus and audio responses. The present study uses a unimodal delivery and response system which eliminates a potentially confounding variable. Previous research has looked into the idea of participants either adding in or deleting words that they feel they may have heard and this will be covered in the next section.

3 Current Study

The study conducted at the University of Calgary is based on a modified version of Mack et al. (2012). Using their method of speech restoration combined with the claim that deletion of subject pronouns may affect speech rate, this study aimed to measure participants ability to recognize speech rate changes and produce sentences with either overt or deleted subject pronouns. This study avoids the use of cross modal response methods that were observed in all previous research on this topic. Participants heard only audio with no visual prompts, and were asked to reply with only spoken responses.

3.1 Method

3.1.1 Participants

33 undergraduate students from the University of Calgary participated in the experiment in exchange for course credit[2]. All participants were native English speakers with normal

[2] The original design of the experiment called for 40 participants, but the data collection was abruptly halted by the COVID-19 pandemic which ceased in person interactions at the University of Calgary.

hearing. The study was approved by the Research Ethics Board at the University of Calgary. Demographic information such as age, gender, and languages spoken was also collected. Participant ages ranged from 18 to 32 with a median range of 19. The majority of participants were from Alberta with a total of 28 participants listing their hometown as an Alberta city or town. The number of languages spoken is summarized in Table 1, sorted by gender identity reported.

Table 1: Participant language demographics

Gender Identity	One Language Spoken	Two Languages Spoken	Three Languages Spoken	Total Participants
Female	12	14	3	29
Male	3	1	0	4

3.1.2 Stimuli

The experimental stimuli were created in three groups similar to Weir (2019): expletive *it*, expletive *there*, and referential *it*. The conversations were designed similar to the style of Mack et al. (2012), consisting a short dialogue between two speakers of different genders. Each stimulus group consisted of 10 short conversations for a total of 30 unique stimulus conversations. The rate of speech of the recorded target sentences was calculated, and the averages of the speech rates is presented in Table 2.

Table2: Average speech rate (syllables per second) per condition

Pronoun Type	Male Speaker	Female Speaker	Combined Average
Expletive *it*	5.13	4.99	5.06
Expletive *there*	4.58	4.87	4.73
Referential *it*	4.95	5.01	4.98
Total Average	4.89	4.95	4.92

Each sound file had the final sentence speech rate artificially increased and decreased to create five total variants of each file. The five conditions were as follows:

- Speech rate decreased by 20%
- Speech rate decreased by 10%
- Unaltered speech rate
- Speech rate increased by 10%
- Speech rate increased by 20%

After the rate of speech modifications, a white noise filter with a duration of 1.60 ms and an amplitude of 0.2 was generated in Audacity version 2.1.3 . This filter was applied at the onset of the final sentence to make the presence of the pronoun ambiguous in all conditions, and at a second location near the main verb of the sentence. The same process for applying the white noise filters was applied to the filler items, with random starting locations for the filter.

A total of 10 lists were created, each consisting of 30 experimental trials and 60 filler items. The stimuli conversations created consisted of short, three-line dialogues similar to Mack et al. (2012), without the use of evidential markers such as *like* or *as.* The complete list of stimuli conversations used in this study is provided in the Appendix of this paper. The filler items consisted of three sentence conversations like the stimuli sentences, but there was no raising verbs or expletive pronouns of any kind included.

3.1.3 Experiment Design

Participants were randomly assigned to one of 10 separate lists. The hypothesis put forward is that participants will be able to notice the speech rate changes, and be more likely to recover the subject pronoun in sentences that were sped up. All stimulus sentences were produced with the subject pronoun with no additional intonational cues. The independent variables considered were the gender of the speaker in the key sentence, and the rate of speech of each sentence. The dependent variable considered was the number of tokens each participant produced in which they recovered the subject pronoun. If the speed of the audio that participants hear is a phonological factor that cues the presence or absence of an overt subject, we would expect to see participants producing the subject pronoun significantly more frequently when the audio is sped up, and less frequently when the audio is slowed down. If no significant effect is observed as a function of the speech rate changes, this would suggest that these phonetic adjustments are not meaningful phonological changes that a listener would be able to interpret. The independent variable of speaker gender is being considered to control for the sociolinguistic idea that young female speakers are most often the driving force of linguistic change. This is not a main focus of the study and no significant difference is being anticipated for this, this factor was added for the purposes of balancing the number of items delivered across all lists.

3.1.4 Procedure

The experiment was created in PsychoPy Version 1.8.4 (Pierce et al., 2020). Participants were situated in a sound attenuated booth, with a computer monitor, keyboard, and microphone in front of them. Participants were instructed to listen to each conversation fully, before being prompted to begin recording their response. Participants were in complete control of when the recording began, and were provided with four seconds to record their response. A pure tone would play after each recording time elapsed to let the participants know that they could advance to the next conversation. This was done both to allow the participants to move at their own pace through the experiment, and to ensure that the experiment would not crash due to the memory overload of audio playing and recording occurring simultaneously. Unlike Mack et al. (2012), participants were not asked to provide a second written or typed response to the prompts.

3.2 Results

A total of 990 recordings were assessed and coded based on the presence or absence of a subject pronoun in the initial position. Tokens were coded as “Pronoun Produced” if

participants produced any pronoun before the word *seems,* even if it was not the intended pronoun. These included sentences meant to begin with referential *it* that instead start with a *he* or *she,* or sentences where participants produced an expletive *it* instead of an expletive *there.* Of these 990, five tokens were discarded from the analysis due to inaudible or entirely missing audio. Across all productions, participants produced utterances that contained a subject pronoun 83% of the time (815 total tokens), which is much greater than the results provided by Mack et al. (2012). There appeared to be an underlying trend in the participants that divided them into three broad categories:

- 23 Participants that deleted less than 5 subject tokens
- 7 Participants that deleted between 5 and 25 subject tokens
- 3 Participants that deleted more than 25 subject tokens

The results presented by Mack et al. (2012) do not mention any breakdown by participants so there is no indication that they encountered a similar patterning of results. This does indicate a strong interspeaker variation within this data with the majority of participants being resistant to any subject pronoun deletion. The participants who were inclined to delete any of the subject pronouns would either delete a small number of the possible subject pronouns, or would prefer to delete nearly all of the subject pronouns with few random exceptions. While several participants deleted a total of zero subjects (17 total), no participant in this study deleted all 30 subject pronouns (the maximum was 29 of 30).

All statistical analyses were conducted in R version 3.6.1 (R Core Team, 2019). Considering the initial hypothesis, no significant effect of speech rate was observed. Table 3 presents the data from this initial analysis which shows that increasing or decreasing the speech rate artificially did not affect the rate of subject pronoun deletion in participants ($\chi^2 = 6.8837$, df = 4, p = 0.142).

Table 3: Pronoun production tokens by rate of speech alteration

Rate of Speech	Pronoun Produced	Pronoun Deleted	Deletion Percentage
Speech rate decreased by 20%	172	26	13.13%
Speech rate decreased by 10%	166	32	16.16%
Unaltered speech rate	152	45	22.84%
Speech rate increased by 10%	164	33	16.75%
Speech rate increased by 20%	161	35	17.86%
Total	815	171	17.34%

Ignoring any effects of rate of speech, the three pronoun types were split and a Chi squared analysis was run on the data. The results of this test were not significant, but still worthy of some discussion ($\chi^2 = 5.612$, df = 2, p = 0.060). The data is presented in Table 4.

Table 4: Pronoun production tokens by pronoun type

Pronoun Type	Pronoun Produced	Pronoun Deleted	Deletion Percentage
Expletive *it*	259	70	21.28%
Expletive *there*	281	48	14.59%
Referential *it*	275	53	16.16%

There appears to be a trend emerging in this data suggesting that participants were more likely to delete an expletive *it* pronoun more than any other pronoun type. Although the statistical analysis was not significant for this data, this is still being mentioned due to the fact that data collection was ended prematurely for this study, so it is possible that if data collection had completed, there may have been a significant result here.

When examining the effect of the number of languages spoken on their deletion rates, a somewhat surprising result was observed. Participants who spoke two languages fluently were significantly more likely more likely to delete subject pronouns of all types ($\chi^2 = 19.75$, df = 2, $p < 0.001$). These differences can be seen in Figure 1 and Figure 2 below.

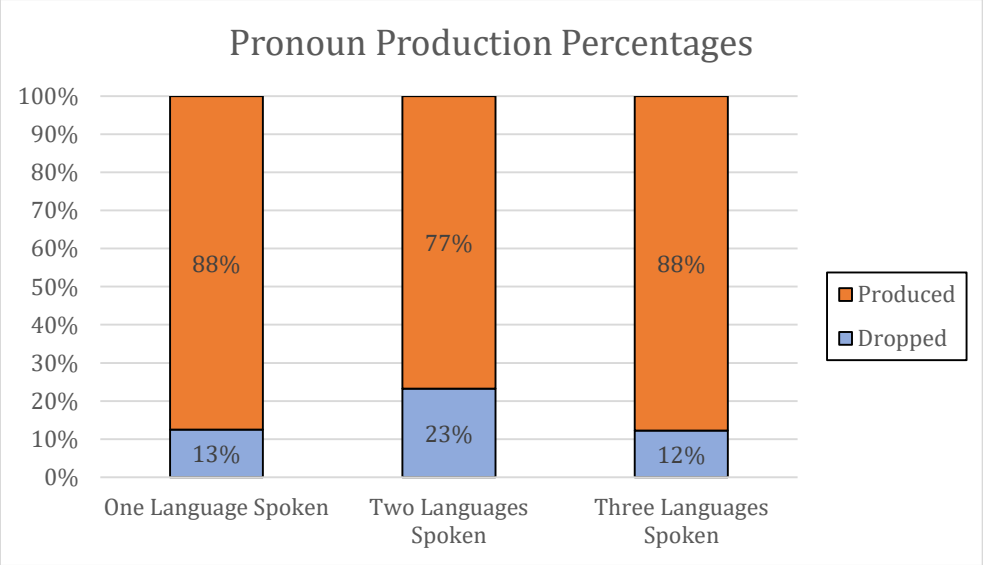


Figure 1: Pronoun production by number of languages spoken

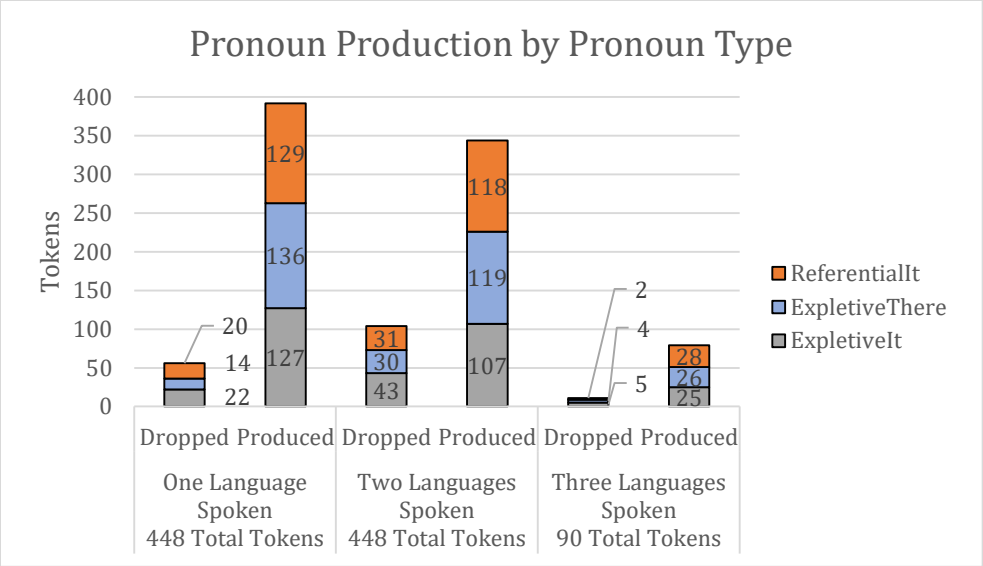


Figure 2: Pronoun production by type and number of languages spoken

Looking specifically at bilingual participants, there was also an effect based on what their second language was. Participants who spoke Mandarin and Cantonese appeared to be more likely to delete subject pronouns in this study ($\chi^2 = 149.59$, df = 5, $p < 0.001$). Table 5 presents a quantitative breakdown of the bilingual participants based on their L2.

Table 5: Bilingual speaker pronoun productions

Language	Number of Speakers	Pronoun Produced	Pronoun Deleted	Deletion Percentage
Arabic	1	28	2	6.67%
Cantonese	3	32	58	64.44%
French	7	195	13	6.25%
Malayalam	1	27	3	10.00%
Mandarin	2	33	27	47.37%
Urdu	1	29	1	3.33%

There were no significant effects in the multilingual speakers, and none of the groups when separated by languages spoken had any significant effects on pronoun type or rate of speech.

Finally, participant F0 on the raising verb *seems* was calculated for all tokens. This was done to observe if the pattern observed in Weir (2019) where the F0 in the referential *it* sentences was significantly lower would appear in this strictly auditory context. The speakers who were asked to create the stimuli sentences were instructed to deliver them as neutrally as possible, so the pattern did not exist in the audio that participants were hearing. If this pattern emerged in Weir (2019) as a result of participants noticing the syntactic differences in sentences with referential subjects, we would not expect participants to be able to recreate this pattern if they are only hearing the audio files, especially if that audio does not also contain this pattern. The mean F0 values by pronoun type are presented in Table 6.

Table 6: Mean F0 observed at raising verb by pronoun type

Pronoun Type	Mean F0 (Hz)
Expletive *it*	251.30 Hz
Expletive *there*	252.32 Hz
Referential *it*	253.09 Hz

A statistical analysis in R version 3.6.1 (Kuznetsova et al., 2017; R Core Team, 2019) using a linear mixed effects model showed no significant effects. There was also no significant interaction between the F0 on the raising verb, and the presence of a subject pronoun in the sentences that participants were producing, or the type of pronoun in the sentence. The initial model of this test is shown in (7)

7) F0 ~ PronounType * PronounProduced + (1|Participant) + (1|ItemID)

When comparing models, it was determined that removing the PronounType effect made for a better and simpler model. This implies that the type of pronoun had no significant effects on the F0 of the raising verb, which contrasts with the results from Weir (2019). The simplest and best model is expressed in (8) and summarized in Table 7.

8) F0 ~ PronounProduced + (1|Participant) + (1|ItemID)

Table 7: Linear mixed effect model analysis of F0 values

	Estimate	Std. Error	DF	t value	Pr (>|t|)
(Intercept)	249.5650	8.760	34.792	28.430	<2e-16***
PronounProduced: TRUE	3.467	2.126	890.872	1.63	0.103

--- Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

The random effects in (8) of participant and item variance yield a significantly better model when compared to models where either of those effects are removed. The reason for these random effects improving the model is likely related to the effects of number of languages spoken, and the variability in some items which allow different combinations of possible subjects to be recovered.

4 Discussion

In terms of the initial hypothesis, there was no significant effect of speed observed, meaning that participants were not influenced by the extralinguistic information that may be conveyed by someone deleting a subject pronoun. There are two possible reasons for why no effect was seen here. The first is that the difference may have been too subtle for participants to pick up on and make adjustments for. Participants were not given specific instructions on what they should be looking for before the experiment began, and were simply told to fill in any missing information they may have heard while being as accurate as possible. No mention was made of the changing speeds before the study began. A second possible reason that they were reluctant to delete subject pronouns is due to the fact that

their personal grammar was influencing what they were saying regardless of what they were hearing. Even though they were hearing audio which may have sounded like there was no subject pronoun, they would still have to make a choice on what they would say when repeating the sentence. Based on their own internal grammar, they may want to produce a sentence with a subject pronoun. This lack of an effect when asking participants to recreate audio like this suggests that subject pronoun deletion in English is most likely a phonological phenomenon. The pragmatic effects noted in Mack et al. (2012) likely still affect this so it is not solely on the phonology to determine what should be deleted, as much as it determines what is able to be deleted.

When splitting the pronoun types, there seems to be a potential for an effect to emerge surrounding expletive *it* being more likely to be deleted, and expletive *there* being less likely to be deleted. We can look at research done by Dilley and Pitt (2010) to examine a potential explanation on why *there* is appearing to behave differently than both expletive and referential *it*. They found that function words were likely to appear or disappear as a function of speed changes when they were able to coarticulate with some surrounding. Example (9) shows how the function word might be able to blend into the previous word if there was no discernible boundary between them.

9) Claire said that sour and bitter (*are*) both…
[klɜɻ sɜd ðɑy sow and **bɪɾ̩ ɾ̩ boθ**]
(Dilley and Pitt, 2010)

With the speed increased, one could imagine how [bɪɾ̩ ɾ̩ boθ] might become [bɪɾ̩ boθ] With the sentences created for this study, sentences with an expletive or referential *it* were able to blend into the following word in a similar way, not through coarticulation, but through the minimization of weak prosodic material. In fact, some of the participants in this study did not produce a complete [ɪt sims], and would sometimes minimize their production to [tsims]. This occurred a in total of four tokens. These occurrences were counted as “Pronoun Produced” when coding data, as they still made some effort to include some phonetic material before the *seems*, and there were not enough instances of this to warrant the creation of its own data code. There is no possibility of coarticulation or minimization of expletive *there* however. The rhotic ending in [ðɜɻ] does not coarticulate easily with the alveolar fricative beginning of [sims], and it is much harder to minimize it, possibly due to its higher sonority. With these differences in mind, it would make sense that we see a difference between the pronoun types. In order to look further into this, another set of sentences would need to be crafted explicitly with expletive *there* in mind by using a different raising verb. An example of an alternative raising verb is shown in (10).

10) There appears to be...
[ðɜɻ əpiɻz tu bi]

By using the raising verb *appears*, we are providing a better environment for the ending of *there* to coarticulate to a degree. It is at least possible with a sentence like this to imagine participants producing something like [ɻəpiɻz tu bi], which is closer to the minimization we were seeing with [tsims], and would potentially make it easier for participants to delete.

Considering the differences between expletive and referential *it* in Table 2, there may also be an effect emerging suggesting that participants would be less likely to delete a referential subject pronoun. This would be a logical result as the referential subject pronoun contains semantic information and would be slightly more difficult to recover, even with context. Looking back at the results of Weir (2019), we do see results of referential *it* being treated differently than expletive *it*, with participants still having a lower F0 on the word *seems* in the referential sentences, even with the pronoun being deleted. From this we can conclude that referential *it* has some sort of lexical stress on it due to the semantic content that it has, which causes it to have phonetic effects on its surrounding environment that we do not observe in any of the expletive pronoun variants. Combining with this idea with the F0 results that we observe in the present study, we do not see the same effect that was observed in Weir (2019). This lack of an effect when using audio stimulus delivery and audio responses means that there may be something unique about the cross-modal methodology of Mack et al. (2012) and Weir (2019). If participants are able to take notice of the sentences with referential pronoun subjects, and produce them in a different way with regards to the lexical stress that they put on the subject when reading them aloud, but do not when hearing them, suggests that there is something else afoot.

The effect that occurred when comparing bilingual participants and monolingual participants appears to occur because of the particular second language that these participants have. Participants who speak either Mandarin or Cantonese as their L2 appear to be significantly more likely to incorporate subject pronoun deletion in their English L1. While there is a record of their languages they claim to speak, no questions were asked regarding claimed level of fluency, or family history of language use. The idea that an L2 may impact the structure of an L1 is one that is not entirely unfounded, especially if these participants were bilingual learners from a young age. Sato (2014) looks at data from Colloquial Singapore English (CSE) concerning strict/sloppy readings with null subjects and null objects. He describes CSE as "an English-lexified contact variety which has evolved with a constant Sinitic/Malay substratum in the multilingual endogenous contact ecology in Singapore" (Sato, 2014, p. 3). While still being considered a contact language, CSE has reached a point in its growth where speakers are learning it as their native language despite its low prestige in Singapore (Sato, 2014). The grammar of CSE exhibits influences from multiple sources including English, Malay, Cantonese, Mandarin and several other native languages, but recent language policy changes have made the influences from Mandarin more prominent. Sato (2014) presents data showing subject/object asymmetry in interpreting sloppy/quantificational sentences containing empty arguments. This asymmetry would need to come from one of the substrate languages such as Mandarin, or Cantonese rather than the lexifier language English. By attempting to draw a parallel from this to the present study, we can claim that the particular variety of English spoken by the bilingual participants in this study may have been influenced by their L2. Because Mandarin and Cantonese are both *pro*-drop languages, it is possible that the increased frequency of subject pronoun deletion we see in these English sentences comes from this influence. This hypothesis is not airtight, however, as Arabic is also a *pro*-drop language, and there is no significant increase in that particular participant's pronoun deletion. Because no data was collected regarding family histories or degree of fluency, it is difficult to make any strong claims on this particular data, however, the data does align with this possible explanation.

5 Conclusion

The results of this study were substantially different from what was seen in Mack et al. (2012). This may be due to regional dialectical differences, as well as other underlying factors that could not be compared. The lack of a positive result regarding the main hypothesis is informative for how participants are processing these types of sentences and recovering them. The strict divide between participants' willingness to delete pronouns suggests that there is certainly a strong inter-speaker phonological difference. A useful direction that I would like to look at in the future would involve recreating both the study from Weir (2019) and the present study as a two-part single study rather than two separate studies. This would allow us to draw a link between participants who are more likely to naturally speak sentences without a subject pronoun and observe whether they make significantly different phonetic adjustments when reading these sentences aloud compared to participants who resist deleting the subject pronouns.

Although there are many theories at play which describe the constraints and limitations of phenomenon such as subject deletion, there does not appear to be a unifying explanation that can account for all the data that we are observing at this point. Based on the data from Mack et al. (2012) and the study in this paper, the explanation appears to be mostly phonological. A phonological explanation does not account for the pattern observed in Weir (2019), which suggests that participants may be processing the syntactic structure of what they are reading, and making unconscious adjustments to the adjacent word to distinguish sentences which have a referential subject as opposed to an expletive subject. The results of Weir (2019) and their contrast with the results in this paper serve as evidence that deletion of material in spoken English may not solely be the responsibility of phonology. I argue that the context of participants' interaction with deleted material dictates what is 'responsible' for the deletion of phonetic information. If written deletion is indeed syntactic, and we are seeing consistent patterns emerging from it, one could argue that it is a syntactic rule that is causing this pattern. If spoken deletion is indeed phonological, and we observe that participants behave at highly variable rates, then one could argue that this is due to their own phonological system. There is no reason to state that it needs to be strictly phonology, or strictly syntax that is causing this. Both explanations are within the realm of possibility depending on the context. Although this may not be the simplest explanation of the data, I believe it captures all of the results we see in the written context of Weir (2019), and the audio contexts of the present study.

References

Dilley, L. C., & Pitt, M. A. (2010). Altering context speech rate can cause words to appear or disappear. Psychological Science, 21(11), 1664-1670.

Kuznetsova, Alexandra, Per B. Brockhoff, and Rune H. B. Christensen. (2017). lmerTest package: Tests in linear mixed effects models. Journal of Statistical Software 82:1–26.

Mack, J. E., Clifton Jr, C., Frazier, L., and Taylor, P. V. (2012). (not) hearing optional subjects: The effects of pragmatic usage preferences. Journal of memory and language, 67(1):211–223.

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J. (2020). PsychoPy2: experiments in behavior made easy. Behavior Research Methods. 10.3758/s13428-018-01193-y

R Core Team. (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org.

Sato, Y. (2014). Argument ellipsis in Colloquial Singapore English and the anti-agreement hypothesis. Journal of Linguistics, 50(2), 365-401.

Weir, J. (September 21, 2019). Expletive Pronoun Deletion. Retrieved from: http://hdl.handle.net/1880/111084.

Appendix – Stimuli Sentences

Expletive *it*

1. Employment
Speaker A: I heard that Adam got a new job.
Speaker B: How is that going?
Speaker A: It seems he is much happier.

2. Food
Speaker A: I think I got food poisoning from the restaurant.
Speaker B: How did that happen?
Speaker A: It seems the meat wasn't cooked through.

3. Baking
Speaker A: My mother made these cookies for us.
Speaker B: Why did she do that?
Speaker A: It seems she likes treating us well.

4. Doctor
Speaker A: James visited the doctor yesterday.
Speaker B: Is everything okay?
Speaker A: It seems he has an awful cold.

5. Cattle
Speaker A: Farmer Ted sold all of his cows.
Speaker B: Why did he do that?
Speaker A: It seems they were costing him lots.

6. Shoes
Speaker A: My shoes have a hole in the bottom.
Speaker B: Will they last through the summer?
Speaker A: It seems that I need new ones soon.

7. Dating
Speaker A: Andrew is trying online dating.
Speaker B: Is that working for him?
Speaker A: It seems he is having success.

8. Computer
Speaker A: Amy bought a new computer yesterday.
Speaker B: Why did she need to do that?
Speaker A: It seems her old one stopped working.

9. Friends
Speaker A: Hank looks unhappy.
Speaker B: Is everything okay with him?
Speaker A: It seems he needs to be cheered up.

10. Heat
Speaker A: It is very cold in my house.
Speaker B: Did you turn up the heat?
Speaker A: It seems that doesn't help at all.

Expletive *there*

1. Car Problems
Speaker A: My car broke down yesterday.
Speaker B: Do you know why?
Speaker A: There seems to be a fuse that blew.

2. Pests
Speaker A: You should call an exterminator.
Speaker B: Why is that?
Speaker A: There seems to be a rat issue.

3. Draft
Speaker A: I think I need to replace my windows.
Speaker B: What's the problem?
Speaker A: There seems to be a leaky seal.

4. Hair
Speaker A: Can you call the waiter.
Speaker B: Why do you need the waiter?
Speaker A: There seems to be no forks here.

5. Reading
Speaker A: Jake dropped his university classes.
Speaker B: Why did he do that?
Speaker A: There seems to be lots of reading.

6. Printing
Speaker A: My document will not print.
Speaker B: Do you know why?
Speaker A: There seems to be a paper jam.

7. Fire
Speaker A: I think there is a fire in that building.
Speaker B: How can you tell?
Speaker A: There seems to be smoke billowing.

8. President
Speaker A: The debate last night was interesting.
Speaker B: Were all the candidates equal?
Speaker A: There seems to be a clear favorite.

9. Choices
Speaker A: I think I need to break up with my partner.
Speaker B: What makes you say that?
Speaker A: There seems to be no better way.

10. Leftovers
Speaker A: Carol is trying to give me some turkey.
Speaker B: Why is she doing that?
Speaker A: There seems to be lots leftover.

Referential *it*

1. Cat
Speaker A: Look at my cat lying in the sunbeam.
Speaker B: What is happening?
Speaker A: It seems to be sleeping soundly.

2. Horse
Speaker A: The horse is running wild in the field.
Speaker B: Did something happen?
Speaker A: It seems to be spooked by something.

3. Toaster
Speaker A: My roommate wants to throw out our toaster.
Speaker B: Why is that?
Speaker A: It seems to always burn the bread.

4. Laptop
Speaker A: My laptop won't turn on anymore.
Speaker B: What the problem?
Speaker A: It seems to have faulty hardware.

5. Rock
Speaker A: That rock is falling down the mountain.
Speaker B: Should we back up?
Speaker A: It seems to be picking up speed.

6. Vulture
Speaker A: That vulture has been circling us for a while.
Speaker B: Do we look that far gone?
Speaker A: It seems hungry this afternoon.

7. Fire
Speaker A: I am getting more wood for the fire.
Speaker B: Do you think you need to?
Speaker A: It seems to be going out quick.

8. Plants
Speaker A: My plant is looking brown.
Speaker B: Why not get some water?
Speaker A: It seems to be dead already.

9. Water Bottle
Speaker A: I need to replace my water bottle.
Speaker B: What is wrong?
Speaker A: It seems to be leaking on top.

10. Phone
Speaker A: I think I lost my phone.
Speaker B: How can you be sure?
Speaker A: It seems to have grown legs and left.

Contact Information

Jesse Weir

jesseweir908@gmail.com
University of Calgary
2500 University Dr NW,
Calgary, AB
T2N 1N4
Canada