Prosodically Conditioned Realization of Voiced Stops and Vowels in Yucatecan Spanish

Inaugural-Dissertation

zur Erlangung des Doktorgrades

der Philosophischen Fakultät

der Universität zu Köln

im Fach Phonetik

vorgelegt von

María Nuria Martínez García

Köln, September 2020

Table of Contents

III Read speech studies

Appendices

ABSTRACT

This dissertation investigates the acoustic nature and distribution of prosodic strengthening in relation to the Prosodic Word domain and prosodic prominence in Yucatecan Spanish. In order to do so, phonologically voiced stops and word-initial vowels are examined in a corpus of sociolinguistic interviews and a read speech task with 16–21 speakers of the variety. Previous accounts of this Spanish variety have reported that phonologically voiced stops are mainly produced as voiced stops and that word-initial vowels are frequently glottalized, unlike common descriptions of many other Spanish varieties, and have also tried to relate these realizations to several speaker-specific characteristics. Based on these accounts, this dissertation also examines two speaker-specific factors in connection to prosodic strengthening: Yucatec Maya–Yucatecan Spanish bilingualism in terms of language dominance, and gender.

The first part of the dissertation surveys prosodic constituency and prosodic strengthening crosslinguistically, and specially in Spanish. The phonetic characteristics of voiced stops and word-initial vowels are reviewed in Spanish in general and in Yucatecan Spanish in particular. Language dominance and gender are examined as possible sources of speaker-specific variation.

The second part of the dissertation examines prosodic strengthening by means of acoustic recordings of sociolinguistic interviews of 20 speakers (for the voiced stop study) and 16 speakers (for the glottalization of word-initial vowels). The acoustic manifestations of prosodic strengthening are longer duration and presence of a release burst (voiced stops), and presence of glottalization (vowels). The results of several Bayesian models for the voiced stops /b/ and /d/ show that prosodic position in the Prosodic Word domain has an effect on prosodic strengthening, with more strengthened realizations at the beginnings of Prosodic Words and lexical words; they also show an effect of word class, with more strengthened realizations for lexical than function words. The effect of lexical stress is mixed, as are the effects of language dominance and gender. For vowels, lack of resyllabification and glottalization are examined. The results of several Bayesian models show that lack of resyllabification and glottalization show similar patterns, with more instances of the two phenomena across Prosodic Words than within them, in stressed syllables, and in the case of Yucatec Maya dominant speakers. The similar pattern observed for lack of resyllabification and glottalization indicates that glottalization is a cue to lack of resyllabification. Further random forest analyses suggest that many other factors may play a role in the distribution of prosodic strengthening.

The third part of the dissertation examines prosodic strengthening with acoustic recordings of a read speech task by 21 speakers. A new factor, repetition, is introduced. The acoustic manifestations of prosodic strengthening are longer duration, greater change in intensity in the consonant and presence of a release burst (voiced stops), and presence of glottalization (vowels). The results of several Bayesian models for the voiced stops show that lexical stress has an effect on prosodic strengthening, as does prosodic position, although the distribution of realizations differs from the results obtained in the second part of the dissertation. Moreover, there is some evidence for more strengthening for lexical than for function words. Language dominance and gender do not lead to more strengthened realizations, while there is evidence for more strengthening in first than in second mentions. The results for the vowels show that prosodic position and lexical stress have an effect on strengthening, with more instances of strengthened realizations at the left edge of the Prosodic Word and under lexical stress. There is little to no evidence for an effect of word class, language dominance, gender, or repetition. Furthermore, the phonetic nature of glottalization is examined qualitatively and quantitatively. The results show that, in Yucatecan Spanish, glottalization is produced primarily as prototypical creaky voice, and also that H1*–H3* and cepstral peak prominence are two acoustic correlates to this glottalization.

The fourth part of the dissertation investigates the possible role of accent on prosodic strengthening. A preliminary analysis of the data examined in the previous chapters strongly suggests that accentuation may have played a role in the results. Finally, a proposal is made for strengthening as the marking of the left edges of a recursive Prosodic Word in Yucatecan Spanish.

Dedication

Para mi madre y mi hermana. Para mi padre, *in memoriam.*

ACKNOWLEDGMENTS

I cannot express enough thanks to Martine Grice, Melanie Uth, and Francesco Cangemi for their continued support and encouragement. I would like to thank my first supervisor, Martine Grice, for her guidance and knowledge, both of which have helped me to shape and reshape this dissertation on many occasions. It still amazes me how an apparently simple comment can be the seed of many fruitful hours of thinking and writing. I owe to my second supervisor, Melanie Uth, a special debt of gratitude for choosing me as *wissenschaftliche Mitarbeiterin* in Project A05 *Prominence marking and language contact in Spanish* more than three years ago, and for the opportunities that have come with it. I would also like to thank her for her invaluable suggestions for revisions, which on many occasions have been food for thought outside the phonetic box. I am also grateful to Francesco Cangemi, who has been kind enough to offer me his help whenever I needed it. I hope I have learned from his structured creativity, and that some sparks of it show on these pages.

The completion of this dissertation could not have been accomplished without the support of several members of Project A05. I would like to thank Leonard Rick, Anna Wördehoff, Lena Pillkahn, and Verena Schlemmer. Extra thanks go to Patrick Auhagen, who was part of Project A05 since the beginning and who helped me with the annotation in Part III and the preparation of Appendix C.

I have learned a great deal in conversation with many people in the SFB and the University of Cologne Phonetics Lab. Thanks go to Aviad Albert, Christine Röhr, Constantijn Kaland, Diego Romero Heredero, Julio Manero González, Kurt Malcher, Doris Mücke, Lukas Müller, Maria Bardají i Farré, Max Bonke, Simona Sbranna, and Theo Klinker. Extra heartfelt thanks go to Maximilian Hörl, whose patience and dedication have helped me on many occasions. I would also like to thank the lunch and brunch bunches for all the good and sometimes bizarre conversations we have shared over the past three years.

I would like to thank Thomas Kisler and Florian Schiel for providing the original G2P Spanish set of rules and for their technical support with it. I would also like to thank them, as well as Uwe Reichel, for kindly answering BAS-related questions. Thanks also to Erich Round, who helped me with some issues related to the script he wrote, on which I've built the study on voiced stops of Part III. Thanks to Bodo Winter and Bruno Nicenboim for their useful workshops. Many thanks to Marianela Fernández Trinidad, who helped me with the coding in Chapter 6. I would also like to thank Joaquim Llisterri and Juana Gil Fernández for their support over the years.

I would like to thank my sister and my friends for their long-distance support during the writing of this dissertation. My heartfelt thanks to my boyfriend, who has put up with me during the pandemic months while we have had to work at the same desk.

Me gustaría darles las gracias a todos los que participaron en las grabaciones de esta tesis, no solo por ello, sino también por acogerme con tantísimo cariño en Carrillo Puerto. El trabajo se convierte en placer estando en la mejor de las compañías.

Jach ya'abkach níibóolal ti' Ama yéetel u láak'tsilo'ob uchik u k'amikeno'ob yéetel ya'abkach yaabilaj tu yotocho'ob. Mix bik'in in tu'ubsik le kmáano'on náaysaj k- óol yéetel coochée, le u yich paak'alo'ob ich kool jach máan ki'tako'ob u máak'a'ntal u ki'iwajil, mix xan uchik u k'amikeno'ob bey juntúul u láak'o'obe'.

The writing of this dissertation was funded by the *Deutsche Forschungsgemeinschaft* (DFG, German Research Foundation), Project-ID 281511265 – SFB 1252 *Prominence in Language*, in the project A05: *Prominence marking and language contact in Spanish*, at the University of Cologne. This dissertation was accepted as a dissertation thesis by the Faculty of Arts and Humanities of the University of Cologne.

List of Figures

List of Tables

Chapter 1
Introduction

1.1 Overview and research questions

This dissertation presents an acoustic study on the nature of prosodic strengthening in voiced stops and vowels in Yucatecan Spanish with a focus on the Prosodic Word domain and prosodic prominence, taking into account speaker-specific factors that may have an effect on prosodic strengthening.

Yucatecan Spanish is spoken in the three Mexican states of the Yucatán Peninsula (Yucatán, Campeche, and Quintana Roo), where Yucatec Maya is also spoken (see Uth, 2018, for a study on language contact). Because of the historic and present-day language contact between the two languages and the fact that a sizeable number of speakers are bilingual, several authors have attributed some linguistic particularities of Yucatecan Spanish to the influence of Yucatec Maya. These are, for example, its rhythm (Michnowicz & Hyler, 2020), the strengthened realization of voiced stops (i.e., as stops rather than approximants; Lope Blanch, 1987), and the presence of vowel glottalization (Barrera Vásquez, 1946/1977; Lope Blanch, 1987; Suárez Molina, 1945/1996). This influence has often been discussed without adequate reference to the complexity of language contact or to bilingualism.

Most studies on Yucatecan Spanish have been conducted in the state of Yucatán, especially in its largest cities, Mérida (e.g., Michnowicz, 2009; Michnowicz & Kagan, 2016; Rosado Robledo, 2011) and Valladolid (e.g., García Fajardo, 1984). A few studies have investigated the variety in Chetumal (e.g., Pérez Aguilar, 2002), and an increasing body of research has been concerned with Yucatecan Spanish as it is spoken in Quintana Roo (e.g., Uth, 2018). Since there are no significant dialectal differences within the Yucatán Peninsula according to what we know from previous research (e.g., Lope Blanch, 1987; Suárez Molina, 1945/1996), it is common practice to consider that the Spanish spoken in the three states corresponds to the same variety. This dissertation examines Yucatecan Spanish as it is spoken in Felipe Carrillo Puerto, a relatively large municipality in the middle of the so-called *Mayan area* in Quintana Roo. Approximately 92 % of the population of Felipe Carrillo Puerto characterize themselves as indigenous, and two thirds of them speak Yucatec Maya (Instituto Nacional de Estadística y Geografía-INEGI, 2015).

Yucatecan Spanish shares the same five-vowel system (/i e a o u/) as most Spanish varieties, and its consonants do not seem to differ greatly from those of Mexico City Spanish. Table 1.1, which is an adaptation of the one provided in Avelino (2018) for Mexico City Spanish,[1] shows that there are six plosives, a characteristic also shared with other Spanish varieties. In this dissertation, the phonological term *(oral) stop* is used for these Spanish consonants. The sounds that are mostly used in loanwords of Nahuatl origin have been excluded from the table. There are Nahuatl loanwords in Yucatecan Spanish, but none seem to keep sounds alien to Spanish (see Suárez Molina, 1945/1996, p. 115, for a list of Nahuatl loanwords in Yucatecan Spanish).

Some Yucatec Maya words have been incorporated into Yucatecan Spanish (e.g., *k'ol* 'sauce with tomato and achiote', *pib* 'underground baking', *xix* 'remainder', *xux* 'wasp'; Suárez Molina, 1945/1996, pp. 97–113). Whether these loanwords keep their original pronunciation or not is unclear, since there are currently no studies that examine whether Mayan sounds are used in Yucatecan Spanish.

Table 1.1. Yucatecan Spanish consonants. Table adapted from the one for Mexico City Spanish consonants in Avelino (2018, p. 223). See the text for an explanation of the modifications made to the original table.

	Bilabial	Labio-dental	Denti-alveolar	Post-alveolar	Velar
Plosive	p b		t d		k g
Affricate				tʃ	
Nasal	m		n	ɲ	
Trill			r		
Flap			ɾ		
Fricative		f	s	ʝ	x
Approximant				j	w
Lateral approximant			l		

Yucatec Maya is the only Mayan language that has developed lexical tones (Kügler et al., 2007). In terms of quality, there are five Yucatec Maya vowels, which can be of four types in terms of length, phonation type, and tonal characteristics: short /i e a o u/, long-low /ìi èe àa òo ùu/, long-high /íi ée óo úu/, or long-high and glottalized /íḭ éḛ óo̰ úṵ/

[1] In Avelino (2018), the postalveolar affricate appears in parentheses, which are used there to indicate that the consonant is used mainly in loanwords. This is undoubtedly an error and it has been corrected here. Also, the inclusion of the sounds [j] and [w] as consonants is not supported by all researchers of Spanish. Their consonantal or vocalic character has been and is still debated in Spanish phonology. The interested reader is referred to Colina (in press) for an overview.

(Frazier, 2009). The consonants of Yucatec Maya are shown in Table 1.2, which is based on the table in Frazier (2009, p. 18), since there is no International Phonetic Association illustration for Yucatec Maya. Two airstream mechanisms are used: pulmonic (most consonants) and glottalic (ejective and implosive consonants). Plosives and ejectives are voiceless in Yucatec Maya; the only voiced stop is implosive.[2]

Table 1.2. Yucatec Maya consonants. Adapted from Frazier (2009, p. 18).

		Labial	Alveolar	Post-alveolar	Palatal	Velar	Glottal
Stop	Plosive	p	t			k	ʔ
	Ejective	p’	t’			k’	
	Implosive	ɓ					
Affricate	Plosive		ts	tʃ			
	Ejective		ts’	tʃ’			
	Nasal	m	n				
	Fricative		s	ʃ			h
	Approximant				j	w	
	Lateral approximant		l				

Several studies on Yucatecan Spanish claim that the realization of voiced stops and vowels in word-initial position differs from that of most Spanish varieties. It is claimed that underlying voiced stops in intervocalic context are produced frequently as voiced stops (e.g., García Fajardo, 1984; Lope Blanch, 1987; Michnowicz, 2009), whereas the classic description of allophonic variation in Spanish in general indicates that voiced stops in intervocalic context are produced as approximants (e.g., Alarcos Llorach, 1965/1991; Navarro Tomás, 1918/1996; Quilis, 1999). Moreover, it has been claimed that, in Yucatecan Spanish, the stop realization of voiced stops is due to Yucatec Maya influence (Lope Blanch, 1987, p. 34) and that stop realizations in intervocalic context are more frequently found in word-initial position (versus word-medial position) and in stressed syllables (versus unstressed syllables; Michnowicz, 2011). However, these and other accounts provide little to no acoustic evidence. Moreover, in contrast to this allophonic perspective, articulatory and acoustic studies on several varieties of Spanish have shown that there is a continuum of realizations (e.g., Eddington, 2011; Parrell, 2011; Simonet et al., 2012; Soler & Romero, 1999). Finally, some of these studies have also shown that, even in varieties that presumably have more weakened (lenited) realizations of voiced stops in intervocalic context than

[2] However, the existence of an implosive sound in this language is not supported by some authors (e.g., Colazo-Simón, 2007).

Yucatecan Spanish, prosodic position, lexical stress, as well as other linguistic and nonlinguistic factors may influence the realization of voiced stops.

Several accounts claim that a glottal stop may be inserted before vowels in word-initial position in Yucatecan Spanish (Barrera Vásquez, 1946/1977; Lope Blanch, 1987; Michnowicz & Kagan, 2016), especially when the vowels are in stressed syllables (Barrera Vásquez, 1946/1977, p. 342). Some authors claim that the phenomenon is due to Yucatec Maya influence (Lope Blanch, 1987, pp. 35–36). Importantly, it has been suggested that the insertion of a glottal stop (i) blocks resyllabification (resyllabification being a pervasive phenomenon in Spanish) and (ii) helps mark word boundaries (Michnowicz & Kagan, 2016). Similarly, Martín Butragueño (2014) points out that the glottal stop usually appears between two words, which suggests that it could be related to the formation of the Prosodic Word.[3] Nevertheless, *glottal stop insertion* may not be the most appropriate label for the phenomenon. Instead, *glottalization of word-initial vowels*, that is, the presence of glottal stops and perceptually equivalent realizations (which are usually produced with creaky voice) at the beginning of word-initial vowels, seems a better characterization for two reasons: (i) this phenomenon has been found to signal prosodic boundaries in some languages, such as English (e.g., Dilley et al., 1996), and (ii) the study on Yucatecan Spanish by Michnowicz and Kagan (2016) indicates that not only glottal stops but also creaky voice are used.

Taken together, the abovementioned accounts suggest that the realization of voiced stops and vowels in Yucatecan Spanish may differ depending on their position in the word and on whether they appear in stressed syllables or not; additionally, their realization appears to be influenced by Yucatec Maya. Other factors, such as gender, have also been investigated in several sociolinguistic studies (e.g., García Fajardo, 1984; Pérez Aguilar, 2002).

In this dissertation, I investigate whether the variation observed in the realization of the bilabial and dentialveolar voiced stops and vowels in word-initial position can be best characterized in terms of *prosodic strengthening*, that is, as “the interaction between prosody and the segmental realization of sounds, and/or . . . the specific phonetic manifestations of the phenomenon” (Georgeton & Fougeron, 2014, p. 83). Prosodic strengthening involves inherently strong positions, which are the edge positions of prosodic constituents and syllables under prominence. In the case of edge positions, strengthening may serve to mark

[3] “Tal propiedad sugeriría explotar el papel de la clausura [glotálica] en la constitución de la palabra fonológica” (Martín Butragueño, 2014, p. 251).

prosodic boundaries (Keating, 2003, p. 120). This is referred to as *domain-initial strengthening*. In articulatory terms, strengthening means a more extreme articulation, which may be correlated with a longer duration of the articulatory gesture. In acoustic terms, strengthening can be indicated by several measures that reflect the articulatory strengthening, such as longer acoustic duration, a greater change in acoustic intensity, or differences in spectral tilt. For Yucatecan Spanish, I will investigate acoustic strengthening in relation to the Prosodic Word domain, specifically whether there is more strengthening at the left edge of the Prosodic Word domain, and to prosodic prominence, especially under lexical stress.[4] It is generally accepted that the Spanish Prosodic Word has primary stress (see, e.g., Elordieta, 2014; Harris, 1983). However, its internal structure has not been the focus of much attention (notable exceptions are Elordieta, 2014, and Hualde, 2009). Consequently, I will also take into account its internal structure by considering the lexical and function words that may form it. Furthermore, I will also examine the role of Yucatec Maya–Yucatecan Spanish bilingualism, in terms of language dominance, and gender.
In sum, the main research questions to be addressed are:

1) Is there phonetic evidence for prosodic strengthening in Yucatecan Spanish?
    a. For voiced stops?
    b. For vowels?
2) If so, where does prosodic strengthening occur?
    a. Initially in the Prosodic Word?
    b. In stressed syllables?
3) What are the sources of speaker-specific variation in the extent and distribution of prosodic strengthening?
    a. Language dominance?
    b. Gender?

1.2 Structure of the dissertation

In the first part of the dissertation (Part I, Chapters 2 to 4), I discuss the theoretical issues on which the subsequent studies are founded. Chapter 2 presents the main characteristics of prosodic constituents in Spanish and in Yucatec Maya, as well as of prosodic strengthening. In this chapter, the relationship in Spanish between lexical stress and accent, the internal

[4] In this dissertation, the terms *Prosodic Word* and *word* will not be used as synonyms. While *Prosodic Word* refers to the prosodic constituent, which will be characterized in detail later in the dissertation, *word* refers to the orthographic word, unless otherwise noted. The role of accent will be considered in relation to lexical stress.

structure of the Prosodic Word, and resyllabification are discussed in detail. Chapter 3 discusses language dominance and introduces gender in phonetic studies, both crosslinguistically and for Yucatecan Spanish. Finally, Chapter 4 discusses prosodic strengthening in Spanish and Yucatecan Spanish. In this chapter, the articulatory and acoustic characteristics of voiced stops and the glottalization of word-initial vowels are presented. Although the focus is on Spanish and Yucatecan Spanish, reference is also made to Yucatec Maya and other languages.

In the second part of the dissertation (Part II, Chapters 5 and 6), I examine prosodic strengthening for voiced stops (Chapter 5) and the lack of resyllabification and glottalization of word-initial vowels (Chapter 6) in Yucatecan Spanish based on a corpus of sociolinguistic interviews that were recorded in Felipe Carrillo Puerto in 2017. There are 20 participants in the study in Chapter 5, and 16 in the study in Chapter 6. The acoustic cues to prosodic strengthening that are considered for the voiced stops are duration and the presence of a release burst. The analysis of word-initial vowels in Chapter 6 investigates lack of resyllabification and glottalization based on an auditory analysis (accompanied by the visual inspection of spectrograms and waveforms), with a post hoc random forest analysis.

In the third part of the dissertation (Part III, Chapters 7 to 9), I present a read speech task that was conducted in Felipe Carrillo Puerto in 2019 (Chapter 7) in order to examine once again the prosodic strengthening of voiced stops (Chapter 8) and vowels in word-initial position (Chapter 9). The number of participants in both studies is 21. The acoustic parameters investigated in the study of voiced stops are duration, change in intensity, and the presence of a release burst. The glottalization of word-initial vowels is examined both qualitatively, by means of a classification of creaky voice types, and quantitatively, by means of an analysis of spectral measures and periodicity/noise. Moreover, repeated mentions are also examined. At the time of the design of the read speech task, most of the findings presented in Chapters 5 and 6 were already known; consequently, those findings are taken into account in the read speech studies. Moreover, Parts II and III are complementary in that they investigate different speech styles, spontaneous and read speech, respectively.

In the fourth part of the dissertation (Part IV, Chapters 10 and 11), the relationship between lexical stress and accent in the studies of Parts II and III is discussed (Chapter 10), followed by a summary and discussion of the main findings and a conclusion to the dissertation (Chapter 11). Overall, the results indicate that there is evidence for prosodic strengthening in the Prosodic Word domain, both for consonants and vowels. Strengthening occurs most frequently under lexical stress and in Prosodic Word initial position. A preliminary analysis also suggests that accent may have an effect on strengthening. Finally, the results suggest that neither language dominance nor gender has an effect on prosodic strengthening.

Part I

Theory

Chapter 2

Prosodic constituents and prosodic strengthening

This chapter provides an overview of prosodic constituency in Spanish and Yucatec Maya, followed by an introduction to prosodic strengthening. In Section 2.1, the prosodic hierarchy is introduced. In Section 2.2, the prosodic constituents of Spanish are introduced. The constituents here presented are tentatively thought to apply to Yucatecan Spanish as well, since prosodic constituency in this variety appears not to have been discussed. In Section 2.2.1, a discussion of several issues related to the Prosodic Word in Spanish is provided. More specifically, the relationship between lexical stress and accent is discussed (Section 2.2.1.1), followed by an overview of prosodization of function words (Section 2.2.1.2), an overview of resyllabification (Section 2.2.1.3), and a summary (Section 2.2.1.4). In Section 2.3, the prosodic constituents of Yucatec Maya are presented. In Section 2.4, an introduction to prosodic strengthening is provided, which includes a description of its main characteristics (Section 2.4.1), an analysis of domain-initial strengthening in articulatory studies (Section 2.4.2), a review of acoustic studies on prosodic strengthening (Section 2.4.3), and a summary (Section 2.4.4).

2.1 The prosodic hierarchy

The stream of speech can be divided up into smaller parts. These parts constitute the prosodic structure, which is considered to be hierarchical. Consequently, much of the literature on prosody refers to a prosodic hierarchy and its constituents. In the current dissertation, the definition of prosody proposed by Shattuck-Huffnagel and Turk (1996) is followed. Unlike definitions that take into account either the acoustic parameters that signal boundaries and prominence or a higher phonological constituent organization, their definition

> merges the phonetic and phonological aspects of prosody, including both the higher level organization, with its constituent boundaries and prominences, and the phonetic reflexes of this organization in the pattern of F0, duration, amplitude and segment quality/reduction within an utterance (p. 196).

The Strict Layer Hypothesis (SLH) proposed by Selkirk (1984, p. 26) formalizes the idea of prosodic structure as a hierarchical organization (Shattuck-Huffnagel & Turk, 1996, p. 207). The SLH can be worded as follows: “A single constraint requiring that a prosodic constituent of level C^i immediately dominate only constituents of the next level down in the prosodic hierarchy, C^{i-1}" (Selkirk, 1996/2004, p. 467). Figure 2.1 shows a schematic example of a prosodic tree. In the example, the names of prosodic constituents are substituted for numbers. A number at a higher level dominates the one immediately below it in the hierarchy, that is, 1 dominates 2, 2 dominates 3, and so on. The figure depicts an understanding of the hierarchy in which a constituent can dominate more than two constituents at the immediate lower level, that is, it is *n-ary branching* (e.g., Nespor & Vogel, 1986/2007).[1]

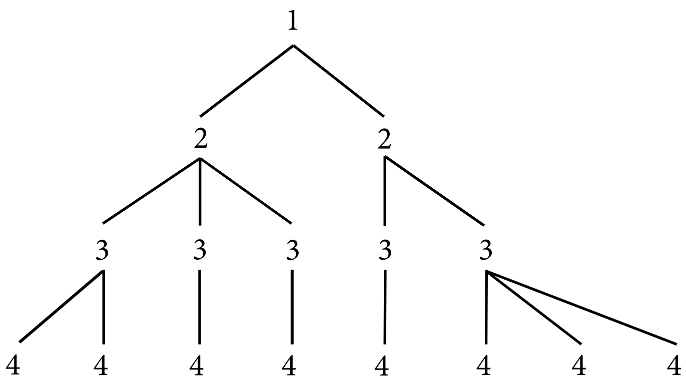

Figure 2.1. Schematic representation of the prosodic structure of a fictional utterance.

The SLH comprises four principles: *layeredness* (a constituent cannot dominate another of a higher level), *headedness* (a constituent has to dominate at least one other of the level immediately below it, if there is a level immediately below), *exhaustivity* (if one constituent dominates another, the latter must be in the level immediately below the former), and *nonrecursivity* (a constituent cannot dominate another of the same type). According to this conception of the hierarchy, the beginning of a given constituent coincides with the beginning of a constituent of the level immediately below it. For example, in Figure 2.1 this means that the beginning of constituent 1 coincides with the beginning of the first constituent 2 (and also with the beginning of the first constituents 3 and 4).

Layeredness and headedness are generally accepted by researchers who adhere to the idea of a prosodic hierarchy, whereas exhaustivity and nonrecursivity have been widely disputed. Instead of exhaustivity, *skipping of levels* is proposed. Instead of nonrecursivity, *recursivity* (also called *recursion*) is proposed. Supporting the skipping of levels does not imply that nonrecursivity is accepted as well, and vice versa. Furthermore, which principles of the SLH are adopted and which are dispensed with has consequences for the number of

[1] A tree in which no more than two constituents are dominated is *binary branching* (e.g., Selkirk, 1980).

constituents in the prosodic hierarchy (see Ito & Mester, 2009, for discussion). As Vigário (2010, p. 487) summarizes, exactly which constituents belong to the prosodic hierarchy is still a controversial issue.

2.2 Prosodic constituents in Spanish

In this section, the main prosodic constituents that have been put forward for Spanish are briefly examined. These constituents are the Prosodic Word (PW), the Phonological Phrase (PhP), the intermediate phrase (ip), and the Intonational Phrase (IP). The syllable and the foot will be discussed in relation to the PW.

The framework that has been commonly used in these studies is the Autosegmental-Metrical (AM) model of intonation and, in particular for Spanish, the Autosegmental-Metrical/Spanish Tones and Breaks Indices (Sp_ToBI) framework. Within the AM model, the two most important parameters of prosody are phrasing and prominence. In terms of *phrasing*, several prosodic constituents are considered, which may or may not be the same for all the languages studied. For example, Spanish does not have an Accentual Phrase, or AP, while French and Korean do.[2] *Prominence* is studied at both the lexical and postlexical levels. Postlexical prominence can be found at the head of the constituent (e.g., by means of a pitch accent), at the edge of the constituent (e.g., by means of a boundary tone), or at both locations (Jun, 2005b).

The Prosodic Word (PW) is the lowest constituent in the hierarchy that makes use of nonprosodic (or, specifically, morphological) information (Hildebrandt, 2015).[3] What constitutes a PW in one language may not apply to others; that is, the definition of a PW is to a great extent language-specific (Hildebrandt, 2015; Jun & Fletcher, 2014; Nespor & Vogel, 1986/2007, p. 109). According to proponents of the AM framework, all languages have PWs and the PW is "the lowest unit that can be defined by intonation" (Jun & Fletcher, 2014, p. 501).

[2] The AP is larger than a Prosodic Word, is marked by an edge tone or a tonal melody, and can have a maximum of one pitch accent (if the language under consideration has pitch accents; Jun & Fletcher, 2014, pp. 501–502). For example, for French, Delais-Roussarie et al. (2015) argue that the AP has phrasal stress that marks its right edge, and that it is composed of at least one lexical word plus all the function words that are governed by it, although other factors, such as rhythmic constraints, may also play a role (p. 67). The AP has also been proposed for Tokyo Japanese (e.g., Beckman & Pierrehumbert, 1986) and for Korean (e.g., Jun, 2005a; Keating et al., 2004).

[3] The PW has also been referred to as a *phonological word* in several works (e.g., Hall, 1999; Nespor & Vogel, 1986/2007). The two terms can be considered synonyms in many studies, although not in all (e.g., Pentland & Laughren, 2005).

The presence of prominence at the lexical level may differ among languages: thus, prominence may be indicated by means of lexical stress (e.g., Greek), lexical pitch accent (e.g., Japanese), tone (e.g., Mandarin Chinese), or none (e.g., Korean; Jun 2005b, p. 431). It is also possible for a language to have lexical stress and lexical pitch accent at the same time (e.g., Swedish), or to have lexical stress and tone (e.g., Cantonese; Jun 2005b, p. 431).

The Spanish PW has primary stress (Aguilar et al., 2009; Elordieta, 2014; Harris, 1983; Hualde, 2007, 2009). In Spanish, stress is assigned lexically. All lexical (or content) words (LWs) are stressed; as such, they can constitute PWs on their own. However, function words (FWs) may be stressed or unstressed (see Section 2.2.1.1).[4] In some languages, such as English, the foot plays a role in stress assignment; conversely, the foot is rarely taken to be a prosodic constituent in Spanish. Although some authors argue that (primary) lexical stress in Spanish is assigned on the basis of moraic trochee feet (Hayes, 1995; Lipski, 1997), these accounts are full of exceptions (see, for example, the analysis in Lipski, 1997). Furthermore, Hualde (in press) argues that there is no evidence in support of the foot in Spanish because there is no secondary stress at the lexical level, whereas, in English, words need to be parsed into feet in order to account for secondary stress. Thus, in Spanish, the syllable (σ) is the immediate lower level of prosodic constituency to the PW.

The Phonological Phrase (PhP) is one of the prosodic constituents above the PW and below the intermediate phrase (ip) proposed for Spanish. The PhP is included in the AM analysis of some languages, such as for example European Portuguese (Frota, 2014). In Frota's proposal, the PW, the PhP, and the Intonational Phrase (see below) coexist. Jun and Fletcher (2014), however, point out that some authors consider that the PhP may correspond to an AP or to an intermediate phrase (p. 30; see below). The existence of the PhP as a different, lower constituent to the ip in Spanish is a matter of debate. Aguilar et al. (2009) leave open the possibility of a PhP in Spanish, arguing that the edge of the PhP may be indicated by the percept of a break, but that this break has no tonal manifestation. Consequently, the PhP as a constituent different from the ip is usually not included in AM studies of Spanish.

The ip is a constituent larger than the PW. In the Spanish ip, continuation rises (which in Sp_ToBI normally correspond to a boundary tone H-) may appear at the end of the ip. For example, Prieto (2006) analyzed a corpus of read speech in European Peninsular Spanish and concluded that the ip has a perceived prominent accent on its last stressed

[4] In the current dissertation, *lexical word* and *function word* are abbreviated as *LW* and *FW*, respectively, for ease of presentation.

syllable as well as a phrase break, with an optional continuation rise at its right boundary.[5] Although some proposals include phrase accents, which “mark the presence of an intermediate phrase” (Prieto & Roseano, 2010, p. 4), other works exclude them in unequivocal terms: “In Sp_ToBI, phrase accents are not used” (Aguilar et al., 2009; see Hualde, 2003, and Prieto & Roseano, 2010, for discussion).

The Intonational Phrase (IP) is the domain of the minimal tune and phrase-final lengthening. It consists of at least one pitch accent and one boundary tone, and it also sets the position for pauses (Prieto & Roseano, 2010, p. 3). The main difference between the ip and the IP in Spanish is of a perceptual nature: the degree of disjuncture between two ips is smaller than between two IPs (Aguilar et al., 2009). For Castilian Spanish, Estebas-Vilaplana and Prieto (2010) claim that the same boundary tones may appear at the end of the ip and at the end of the IP (p. 20), although this may be different for other varieties. In fact, it has been argued that there is only a partial overlap between the boundary tone inventory at the ip and IP edges (e.g., Aguilar et al., 2009; Prieto & Roseano, 2010).

According to the AM framework, all languages have IPs, which are the highest-level prosodic constituent. Nevertheless, some authors have proposed a larger unit than the IP, the Utterance (U; e.g., Beckman & Pierrehumbert, 1986, for Japanese). Keating et al. (2004, p. 150) point out that the difference between the IP and the U is controversial, and that from an intonational point of view they may not be different.

2.2.1 The Prosodic Word

2.2.1.1 Lexical stress and accent

Hualde and Prieto (2015) point out that “the practice among analysts of Spanish intonation has been to consider that essentially every content word, with few exceptions, is accented” (p. 358; see also Hualde, 2007). However, mismatches between lexical stress and accent can take place in the form of postlexical secondary stress and deaccenting.

Postlexical secondary stress, which is a phenomenon common in public discourse, refers to “prominence on a syllable other than the one that is lexically specified to carry (primary) word-level stress” (Hualde, 2007, p. 79; see also Hualde, in press; Llisterri, in press, for literature review and discussion). Further evidence for Hualde’s analysis of postlexical secondary stress comes from the analysis of a corpus of European Peninsular Spanish speech read by two news broadcasters (Aguilar & Gutiérrez-González, 2018). In the corpus,

[5] Prieto uses the term *Phonological Phrase*, not *intermediate phrase*. Nevertheless, it is clear that her definition corresponds to the intermediate phrase (cf. Toledo, 2007).

13 % of the total number of pitch accents corresponded to instances of postlexical secondary stress.

Deaccenting refers to the phenomenon in which lexically stressed syllables, which are usually accented in Spanish, are not. Several works have shown that the extent of deaccenting varies as a function of speech style, being most frequent in spontaneous speech (Face, 2003; Rao, 2005, 2007), although it can also be found in map tasks (Rao, 2007), public discourse (Aguilar & Gutiérrez-González, 2018), and read speech in experimental settings (Ortega-Llebaria & Prieto, 2007). These studies are surveyed below.

Face (2003) conducted a preliminary phonetic study of declarative sentences produced in read speech and spontaneous speech in central European Peninsular Spanish. One of the several intonational differences between the two speech styles was deaccenting (manifested by means of a lack of rise in F0 accompanying stressed syllables). Face argued that deaccenting was uncommon in read speech, but rather common in spontaneous speech. He noted that 30 % of all "accentable words" (p. 122) in the spontaneous speech sentences he examined were deaccented. Moreover, most instances corresponded to verbs (especially *ser* 'to be', *estar* 'to be', and *haber* 'to be/to have'), adverbs, and stressed function words (see Section 2.2.1.2).

Following Face (2003), Rao (2005) conducted a preliminary study using declarative sentences, classified according to their different pragmatic meanings. These sentences were read by one speaker of Mexico City Spanish. Deaccented instances ranged between 7 and 18 % of the total, and were more common in verbs (*ir* 'to go', *estar* + adjective or gerund) and in phrase-final position. In another study, Rao (2007) examined deaccenting in declarative sentences in spontaneous speech and by means of a map task. In this study, "a stressed lexical item is considered as deaccented when any type of F0 movement is absent from its stressed syllable" (p. 203). Seventeen speakers of Barcelona Spanish participated in the acoustic recordings. For spontaneous speech, 23 % of words were deaccented. Five variables favored deaccentuation: (i) high lexical frequency (versus low frequency), (ii) lower number of syllables in the word, (iii), grammatical category (verb > adverb > adjective > noun, although no effect was found for stressed pronouns and stressed conjunctions, the other two grammatical categories included in the study), (iv) recent repetition in discourse, and (v) position in the ip (medial > initial > final > single).[6]

[6] Rao (2007) used the term *Phonological Phrase*, not *intermediate phrase*. Nevertheless, it is clear from the examples provided in the article that his definition corresponds to the intermediate phrase. *Single* refers to PWs that constituted an ip on their own.

Moreover, there were more instances of deaccenting for (i) verbs and adverbs that had recently been repeated in discourse and (ii) verbs, adverbs, and adjectives (but not nouns) made up of less syllables. The three variables studied that had no effect on deaccentuation were: (i) repetition in discourse, (ii) stress pattern (oxytone, paroxytone, and proparoxytone words), and (iii) position in the IP (initial, medial, and final). For the map task, 24 % of PWs were deaccented. The variables that had an effect on deaccenting were fewer than in spontaneous speech: only (i) lower number of syllables of the word and (ii) position in the ip (medial > initial + final > single).

Deaccented syllables are also present in read speech, although to a lesser degree. In the news corpus of European Peninsular Spanish studied in Aguilar and Gutiérrez-González (2018), around 5 % of the total number of accentable instances were deaccented. In the study by Ortega-Llebaria and Prieto (2007; see below), who examined read speech in an experimental setting, deaccented syllables appeared in parenthetical utterances.[7]

In several studies on Spanish, the acoustic correlates of lexical stress are often taken to be longer duration, greater amplitude (intensity), and higher F0 values (pitch; see Llisterri, in press, for a thorough survey). However, the reason for taking F0 as a correlate of lexical stress is that many studies have conflated lexical stress and accent, not only in Spanish (see discussion in Llisterri, in press; Ortega-Llebaria, 2006) but in other languages as well (see Gordon 2014; Roettger & Gordon, 2017, for crosslinguistic examples and discussion). There is evidence in a number of languages (e.g., Arabic, Dutch, European Portuguese, English, Finnish) that shows that pitch accented syllables are longer than unaccented stressed and unstressed ones (Fletcher, 2010, p. 531, and references therein).[8] However, as will be argued below, Spanish may not be one of them. As Fletcher (2010) remarks, "not all languages show equal degrees of stress- or accent-related lengthening, nor do they have a three-way duration contrast between unstressed, stressed unaccented, or stressed accented syllables" (p. 532).

There are some studies on Spanish that have controlled for the effect of accent when examining the acoustic correlates of lexical stress. Their results indicate that duration and probably intensity are acoustic correlates of stress. In a study of Barcelona Spanish, Ortega-Llebaria and Prieto (2007) investigated whether there were acoustic differences between stressed syllables that also bore a pitch accent, produced in declarative utterances,

[7] "In parenthetical intonation, F0 is flat across the utterance and shows no pitch accent" (Ortega-Llebaria & Prieto, 2007, p. 158).

[8] See also Section 2.4.3, which overviews studies on Dutch and English that show that there is an effect of accent on duration (as well as on VOT).

and stressed syllables that were deaccented, produced in parenthetical utterances. They also examined unstressed syllables in both declarative and parenthetical utterances. Results showed that duration was the most important acoustic correlate of stress: stressed syllables (whether accented or not) were significantly longer than unstressed ones; however, accented syllables were not longer than deaccented ones. In sum, "the presence of an accent does not obligatorily trigger lengthening on the stressed syllable" (p. 164). Interestingly, results for intensity showed (i) that there were differences between accented and unstressed syllables, but with contradictory patterns among the vowels studied, and (ii) that there were no differences between deaccented and unstressed syllables. Ortega-Llebaria and Prieto suggested that the differences in intensity may be due to an effect of accent, but not of stress. However, in a study of European Spanish with experimental data, Torreira et al. (2014) examined the role of duration and intensity in phrase-medial unaccented tokens (both stressed and unstressed). The results showed that lexical stress was cued by longer duration and also by greater intensity, duration being the most important cue of the two. Thus, the results for intensity are at odds with the results from Ortega-Llebaria and Prieto (2007), who found no differences in terms of intensity between stressed and unstressed syllables. The results of these studies suggest that duration is an acoustic correlate of stress, but not of accent, and that intensity is also probably a correlate of stress, but again, not of accent. In sum, further research is needed to elucidate the weight of these acoustic correlates for lexical stress and pitch accent in Spanish.

2.2.1.2 Prosodization of function words

Crosslinguistically, the relationship between the PW and the grammatical word has been discussed widely (e.g., Beckman, 1996; Fletcher, 2010; Hall & Kleinhenz, 1999; Nespor & Vogel, 1986/2007). Researchers usually agree on the idea that there is nonisomorphism, in most instances, between the PW and the grammatical word, and also on the idea that the edges of PWs align with morphosyntactic edges. Clitics, which are "a hybrid and widely defined class of small, function (vs. content), or 'weak' words" (Hildebrandt, 2015, p. 238), play a fundamental role in the nonisomorphism between the PW and the grammatical word. Consequently, the PW has been defined in different ways depending on the role given to clitics. Moreover, the principles of the SLH may not necessarily be followed.

In Spanish, LWs comprise nouns, adjectives, verbs, and adverbs; FWs, unlike LWs, lack descriptive content, constitute closed paradigms, and are usually phonologically or morphologically dependent on a LW (Escandell Vidal & Leonetti, 2000). Function words include, among others, articles, pronouns, demonstratives, prepositions, and conjunctions. However, this primarily semantic characterization is not clear-cut. For example, some adverbs (*aquí* 'here', *así* 'this way', *allí* 'there') and some verbs (*ser* 'to be', *haber* 'to have') can be considered FWs (Real Academia Española & Asociación de Academias de la Lengua Española, 2009, pp. 43–44).

While all LWs are stressed in Spanish, which means that they can constitute PWs on their own, FWs may have stress (e.g., *un* 'a-M.SG', *una* 'a-F.SG', *nuestro* 'ours-M.SG') or not (e.g., *el* 'the-M.SG', *la* 'the-F.SG', *a* 'to', *de* 'of', *para* 'for/to').[9] Stressed FWs have not received much attention in the study of stress in Spanish; Hualde (in press) analyzes them as constituting PWs in themselves. Moreover, any unstressed FW can become stressed when it is cited, nominalized, or contrastively focused (Elordieta, 2014; Hualde, 2007, 2009), thus also being able to constitute a PW on its own (see below). For example, in the sequence *el artículo* la 'the *la* article', both *artículo* and ***la*** would have a stressed syllable (Hualde, 2009, p. 202).

The few studies that have focused on the PW in Spanish have analyzed unstressed FWs, but not stressed FWs. Hualde (2009) analyzed unstressed FWs in standard Peninsular Spanish within a metrical framework and put forth two possible analyses. In the first one, all unstressed FWs are in fact lexically stressed, but they surface as unstressed because "they are subject to a rule of prosodic merger with following elements within the syntactic phrase, creating a single prosodic word domain" (p. 199). Hualde further argued that this is due to a rule in Spanish according to which, within the PW domain, only the rightmost stress is kept. In the second analysis, unstressed FWs are simply unstressed. Elordieta (2014) presented an analysis of some types of unstressed FWs (determiners, preverbal object pronouns, and possessive pronouns). His analysis closely follows Selkirk's (1996/2004) analysis; consequently, Selkirk's proposal is presented first.

Selkirk (1996/2004) considered FWs in relation to the PW. Whereas LWs can be PWs on their own, FWs may be prosodized in two ways: either as PWs on their own or as *prosodic clitics*, which can be of three types:

- *free clitic*: the FW attaches directly to the prosodic constituent higher than the PW, that is, the PhP. This type violates the SLH's exhaustivity principle. Thus, the clitic could be represented in a PhP as $(\mathbf{FW}\,(\mathrm{LW})_{\mathrm{PW}})_{\mathbf{PhP}}$.

[9] Quilis (1999) provides a thorough classification of stressed and unstressed words in standard Peninsular Spanish, which is generally agreed upon by researchers (e.g., Hualde, 2007; Rao, 2007; Sosa, 1999).

- *affixal clitic*: the FW is within a recursive PW. This type violates nonrecursivity. It can be represented as $((\mathbf{FW}\ (\mathrm{LW})_{\mathrm{PW}})_{\mathbf{PW}})_{\mathrm{PhP.}}$
- *internal clitic*: the FW forms, together with the LW, a (nonrecursive) PW. It can be represented as $((\mathbf{FW}\ \mathrm{LW})_{\mathbf{PW}})_{\mathrm{PhP.}}$

The prosodization of FWs varies not only according to the language in which they appear, but may also vary in particular instances. For English, Selkirk (1996/2004) noted that a FW such as *at* is prosodized as a free clitic, but when it is uttered in isolation, focused, or in phrase-final position, then it constitutes a PW on its own. Although analyzing those free clitics as affixal clitics would also be possible, Selkirk argued that they must be considered free clitics because of the presence of other cues that mark the beginning of the PW. Selkirk's example is that voiceless stops in English are aspirated in PW-initial position, for example at the beginning of the LW *tomatoes* in *grow tomatoes* (p. 473). This aspiration is a cue to the beginning of the PW. However, in the case of FWs that appear in a possible PW-initial position, such as, for example, *to* in *They grow to the sky*, the voiceless stop in *to* is not aspirated (p. 474).

Elordieta (2014) analyzed proclitics in Spanish as either affixal clitics or as free clitics, that is, as part of a recursive PW or as clitics attached directly to the PhP, respectively. Figure 2.2 exemplifies the two analyses for the proclitic *el* 'the' in the sequence *el árbol* 'the tree'.

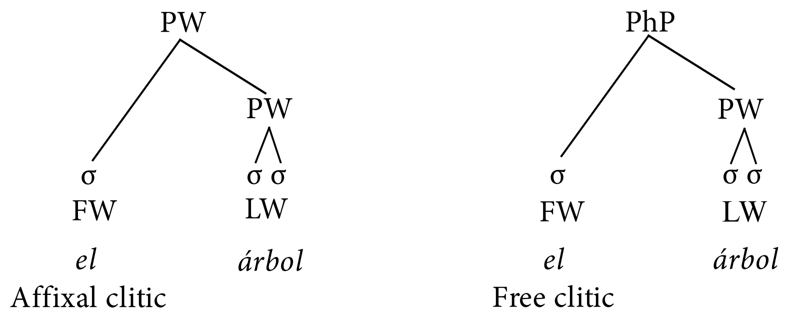

Figure 2.2. Two possible prosodic analyses of the proclitic *el* 'the' in *el árbol* 'the tree', after Elordieta (2014).

Elordieta (2014) argued that proclitics in Spanish cannot be considered internal clitics because of two segmental phenomena that apply at the beginning of the LW, namely the realization of rhotics and *e*-epenthesis. With respect to rhotics, Elordieta (after Harris, 1983) argued that root-initial rhotics are pronounced as a flap [ɾ] in word-internal position, but as a trill [r] in word-initial position. Elordieta claimed that this may indicate the existence of a "word boundary at the left edge of the lexical words" (p. 30). Thus, based on the root /ɾupt-/, the word *erupción* 'eruption' is pronounced with a flap, but *ruptura* 'rupture' is pronounced with a trill [rup.ˈtu.ɾa], even when a proclitic is attached to it (e.g., *tu ruptura* 'your rupture' [tu.rup.ˈtu.ɾa]) (p. 30). It follows that, if *tu* were an internal clitic, the rhotic in *ruptura* would be a flap. Furthermore, *e*-epenthesis before a sequence of /s/ + consonant only takes place at the left edge of LWs, but not within them. For example, *esfera* 'sphere' is possible, but **estratoesfera* is not, the correct word being *estratosfera* 'stratosphere' (p. 31). There is also *e*-epenthesis with a proclitic (e.g., *la estratosfera* 'the stratosphere'), which shows a similar pattern to that of the realization of rhotics.

In sum, the prosodization of FWs in Spanish, whether stressed or unstressed, is still in need of further research. In particular, it remains to be clarified whether proclitics should be analyzed as affixal or as free clitics in Spanish.

2.2.1.3 Resyllabification

Resyllabification can be defined as the process by which the syllabic affiliation of a sound of a syllable changes. In Spanish, resyllabification across word boundaries (as well as *within* word boundaries) is widespread (Obediente Sosa & Méndez Seijas, in press), although several phonetic studies have shown that resyllabification does not always take place (see below). Resyllabification being widespread in Spanish, lack thereof may help to indicate some kind of "break", that is, it may serve to signal boundaries between prosodic constituents.[10]

Nespor and Vogel (1986/2007, pp. 72–76) examined a group of phonological rules for Spanish put forward by Harris (1983) and concluded that the domain of application of resyllabification is either the IP or the U. Consequently, resyllabification should take place not only across the PWs that may form an IP or an U, but within them as well. For example, if a PW is made up of a FW that clitizes to the following LW, we would expect resyllabification also to take place between the FW and the LW. However, it should be noted that Harris (1983) and Nespor and Vogel (1986/2007) examined only *consonantal resyllabification*, that is, cases in which the syllable-final sound is a consonant (e.g., *e**n** agua* [e.ˈna.ɣwa] 'in water', *todo**s** andaban* [ˈto.ðo.san.ˈda.βan] 'all walked'), and not *vocalic resyllabification*, that is, cases in which the syllable-final sound is a vowel (e.g., ***tu** amigo* [twa.ˈmi.ɣo] 'your friend'). Nevertheless, in the Hispanic phonological literature, it is well established that Spanish disfavors the "adjacency of heterosyllabic vowels" (Aguilar, 2003,

[10] This possibility will be discussed for Yucatecan Spanish in Section 4.3.4.

p. 2111), which is usually known in Spanish as *tendencia antihiática* ('antihiatic tendency'). Consequently, Nespor and Vogel's claim could also be applied to vocalic resyllabification.

For consonantal resyllabification, two studies on the production of /s/ in European Peninsular Spanish show that position in the word modulates the realization of the fricative. Hualde and Prieto (2014) examined speech obtained through a map task. Results showed that the fricative was more voiced in orthographic word-final position than in word-medial position, and also shorter in word-final position than in word-initial and word-medial positions. Strycharczuk and Kohlberger (2016) built on this study and measured the duration of the fricative in similar positions. They found that word-final /s/ in a resyllabification context was shorter than word-initial /s/. Strycharczuk and Kohlberger (p. 12) conclude that this result does not support a strong resyllabification hypothesis, that is, one in which the duration of the fricative in both positions (word-initial and word-final) would be the same. However, in a perception study that investigated word-initial and word-final /s/, /n/, and /l/ in European Peninsular Spanish (Lahoz-Bengoechea, 2018), while a short realization of /s/ was interpreted as word-final, which would support the results of Hualde and Prieto (2014) and Strycharczuk and Kohlberger (2016), the opposite effect was found for /l/, with shorter /l/ being identified as word-initial, whereas there was no effect of position on the duration of /n/.

In the case of vowels across words, they can be resyllabified or not resyllabified (in which case there is a hiatus). Resyllabification can take place either through the creation of a diphthong (*fuego y* [ˈfwe.ɣoi̯] 'fire and') or through the reduction of the vowels. Reduction includes instances both of *elision* of one of the two vowels (e.g., *se hace* [ˈsa.θe] 'it is done') and of *coalescence* between the two (e.g., *nombre exacto* [ˈnom.bɾe.ˈsak.to] 'precise name'), which may result in the vowel having longer duration or even a different vowel quality from either of the two original vowels.

There are some studies that have shown that, while vocalic resyllabification across word boundaries is indeed frequent, several factors play a role in the lack of resyllabification (Alba, 2006, for New Mexican Spanish; Aguilar, 2003, 2005; Hualde et al., 2008, for European Peninsular Spanish). These studies are surveyed below.[11]

Alba (2006) auditorily inspected sequences of word-final /a/ followed by any of the five Spanish vowels in word-initial position to investigate whether hiatuses across word

[11] Phonological studies provide a similar analysis. For example, Hualde (1989) argues that vocalic resyllabification depends on multiple factors, which are speech rate, stress, and vocalic quality (after Navarro Tomás, 1918/1996, pp. 148–149). In particular, fast speech rate, unstressed syllables, and two adjacent vowels with the same quality would favor resyllabification.

boundaries were maintained or resolved. The data set used was a large corpus of sociolinguistic interviews with 36 speakers of New Mexican Spanish.[12] The results showed that there was lack of resyllabification in 26 % of all cases. The main factors that favored lack of resyllabification were lexical stress (the word-final and/or the word-initial vowels were in a stressed syllable) and the presence of high vowels (/i u/). To a lesser extent, lack of resyllabification was also favored by the vowels being across LWs (versus FW + LW), by the fact that they belonged to words that had not been mentioned previously, and by having a lower string frequency (i.e., "the token frequency of a multiword sequence"; p. 275).

The results of three studies on European Peninsular Spanish (Aguilar, 2003, 2005; Hualde et al., 2008) agree in part with those in Alba (2006), with presence of lexical stress on the syllable that contains the word-initial vowel and high vowels favoring lack of resyllabification. The discrepancies between Alba (2006) and the studies on European Peninsular Spanish may be due to the fact that they are examining different varieties, but also to the fact that the latter studies used read speech, examined vowel sequences across noun + adjective constructions only, and had a smaller number of participants (4–5 participants). Even for such relatively small data sets, there was great variation in how the vowel sequences were resolved. In particular, Aguilar (2003) studied combinations of word-final /a/ followed by /a e i/ in all possible combinations of lexical stress patterns, in three positions in the IP (initial, medial, and final). Aguilar reported a 45 % of lack of resyllabification, which was independent of vowel identity and position in the IP. On the other hand, the presence of lexical stress on the syllable of the word-initial vowel favored lack of resyllabification, especially if the vowel in the preceding syllable (i.e., the word-final one) was also stressed. In Aguilar (2005), the five vowels appeared in word-final or word-initial position. The results showed that 14 % of all cases were not resyllabified, which is a much lower percentage than that obtained in Aguilar (2003). Similarly to Alba (2006) and Aguilar (2003), the presence of lexical stress on the syllable of the word-initial vowel favored lack of resyllabification. Also in agreement with Alba (2006), the presence of /i/ and /u/ in word-final position favored lack of resyllabification. Aguilar (2005) argued that the

[12] Alba (2006) discusses the limitations of distinguishing between hiatuses and diphthongs in some doubtful cases that seem to fall in-between: "resolution versus maintenance of hiatus is not a binary distinction, but is instead gradient. Acknowledging this, we based our coding of each token on the *degree* of reduction or syllable merger that had occurred; in other words, whether the hiatus seemed 'more maintained' or 'more resolved'" (p. 284).

presence of the two words in the same Clitic Group[13] favored resyllabification, which is in line with Alba's finding of more resyllabification for FW + LW sequences than for LW + LW sequences. Hualde et al. (2008) took into account only nonhigh vowels (/e a o/) across noun + adjective sequences. An acoustic analysis of duration and F1–F2 showed that the hiatuses were resolved either by means of reduction (e.g., [e.a] > [a]) or diphthongization (e.g., [e.a] > [ja]). Interestingly, they did not report any instances of hiatus maintenance, that is, of lack of resyllabification.

2.2.2 Summary

This section provided an overview of prosodic constituents proposed for Spanish (the PW, the PhP, the ip, and the IP), as well as a discussion of some constituent-related issues. Below the PW, the syllable is a constituent, whereas the foot is not. The Prosodic Word was explained in more detail, with reference to prominence at the lexical and postlexical levels, the prosodization of function words, and resyllabification. On the PW level, a PW has primary stress. It is unclear whether duration and intensity are acoustic correlates of lexical stress exclusively or whether they could also be a correlate of accent. Unaccented FWs (proclitics) merge with the following LW. Nevertheless, there is still uncertainty regarding how stressed FWs and proclitics should be analyzed. Resyllabification in Spanish is a common phenomenon across varieties, having a phonetic realization that seems to be conditioned by a variety of factors, being lexical stress the most important.

2.3 Prosodic constituents in Yucatec Maya

The literature on Yucatec Maya, although scarce, provides some phonological information on several prosodic constituents, such as the syllable or the IP. This information could potentially be of help in explaining the distribution of the realization of voiced stops and glottalization of word-initial vowels in Yucatecan Spanish.

In Yucatec Maya, syllables may be of different types, namely CV(V)C, CV(V), CCV(V)C, and CCV (Sobrino Gómez, 2018). All syllables are characterized as having an obligatory onset, which can be a glottal stop, which is inserted in the surface representation of an underlying VC syllable (Sobrino Gómez, 2007, 2010, 2018). Morphological roots and words (including loanwords from Spanish) start and end with a consonant (Frazier, 2009).

[13] Aguilar (2005) follows Nespor and Vogel's (1987) proposal of prosodic constituency. The *Clitic Group* is composed of a Phonological (Prosodic) Word that contains a nonclitic word plus a clitic or clitics (Nespor & Vogel, 1986/2007, p. 162).

Hiatuses within a word are not attested (Orie & Bricker, 2000), while diphthongs can only be postlexical, arising from the common postlexical deletion of the glottal consonants between vowels (Gussenhoven & Teeuw, 2008).

The foot is usually taken into consideration in stress assignment (e.g., Krämer, 2001). It has been suggested that lexical stress is noncontrastive (Frazier, 2009, p. 22), and that heavy syllables attract stress (e.g., England & Baird, 2017; Krämer, 2001). Furthermore, an acoustic and perceptual study by Kidder (2013) showed that duration is not an acoustic correlate of lexical stress (pp. 168–170).

The IP in Yucatec Maya is discussed in a study on clitics (Skopeteas, 2010) and in another on focused constructions (Verhoeven and Skopeteas, 2015). The IP is the domain of downstep, its right edge is associated with a tonal target, and it is often followed by a prosodic break.[14] Moreover, Skopeteas (2010) indicated that there are some enclitics that are associated with the right edges of IPs, thus also providing morphological information to mark the right edge of the prosodic domain.

2.4 Prosodic strengthening

2.4.1 Main characteristics

Prosodic strengthening refers to "the interaction between prosody and the segmental realization of sounds, and/or to the specific phonetic manifestations of the phenomenon" (Georgeton & Fougeron, 2014, p. 83). Prosodic strengthening involves inherently strong positions, which are the edges of constituents and syllables under prominence. The phonetic manifestation of these positions is one of strength: a segment in a strong position is produced with more strength than another in a weaker position; at the same time, that strength is what cues the position as strong (Keating, 2003, p. 120).

The obvious conceptual opposite of prosodic strengthening is *prosodic weakening*. Prosodic strengthening/weakening is considered equivalent to *fortition/lenition* in several works on prosodic strengthening (e.g., Fougeron & Keating, 1997; Keating, 2003) and also in works on historical linguistics (e.g., Bauer, 2008; Bye & de Lacy, 2008). Thus, some studies on fortition/lenition can be informative in terms of prosodic strengthening/weakening (e.g., Soler & Romero, 1999, for Spanish; see Section 4.2). While

[14] "Phonetic evidence for intonational phrases in Yucatec Maya is provided by the following phenomena: (*a*) an intonational phrase determines the phonological domain within which the tonal events are downstepped; and (*b*) the right edge of an intonational phrase is associated with a tonal target (which is high for non-final intonational phrases) and is frequently accompanied by a prosodic break" (Verhoeven & Skopeteas, 2015, p. 32).

prosodic strengthening studies are synchronic, works nominally on fortition/lenition tend to be diachronic in nature (Ségéral & Scheer, 2008, p. 134), although synchronic studies also exist (e.g., Kingston; 2008; Lavoie, 2001). A more fundamental difference between the two pairs of terms is their scope of application: whereas prosodic strengthening/weakening has been studied in connection to constituents at all levels of the prosodic hierarchy, studies on fortition/lenition are mostly limited to the (lexical) word and its subcomponents, most notably the syllable.

In terms of the "specific phonetic manifestations of the phenomenon" (Georgeton & Fougeron, 2014, p. 83), strengthening can be defined articulatorily: articulators adopt a more extreme articulation in strong positions, that is, the articulatory effort is greater and the articulatory target may be achieved in full. This is the *spatial dimension* of strengthening. Also, a strong articulation may be correlated with longer articulatory (and acoustic) duration; this is the *temporal dimension* of strengthening (Cho, 2016, p. 126; Fougeron & Keating, 1997, p. 3737). It follows, then, that many works on prosodic strengthening are articulatory in nature, but its acoustic manifestation has also been examined, either in connection to articulation (e.g., Fougeron & Keating, 1997; Keating et al., 2004; Onaka, 2003) or in itself (e.g., Cho & McQueen, 2005, and references therein; Cole et al., 2007).

2.4.2 Domain-initial strengthening in articulatory studies

Domain-initial strengthening (DIS), also called *postboundary lengthening*, is the prosodic strengthening that takes place at the beginning of a constituent, thus contributing to the marking of its left edge.[15] Keating (2003) suggests that "a strengthened segment indicates a break, the start of a new domain, while domain-internal spans of segments are not interrupted by strengthening" (p. 120). Studies on DIS have taken articulatory and acoustic measures mainly of consonants, although glottalization of vowels at the left edge has also been interpreted as prosodic strengthening (Dilley et al., 1996; Keating, 2003). Thus, Keating (2003) argues that this glottalization of vowels can be considered strengthening because it "gives them a more consonantal quality" (p. 120).

The linguistic function of prosodic strengthening at the edges has been hypothesized to be one of phonological *contrast enhancement*: strengthened segments maximize the contrast with others both at the syntagmatic and the paradigmatic levels (Cho, 2016; Cole et al.,

[15] *Phrase-final lengthening* (also called *preboundary lengthening*) is a boundary marking phenomenon that cues the right edge of constituents by means of longer duration.

2007; Georgeton & Fougeron, 2014). At the syntagmatic level, the strengthened segment contrasts with its neighbors. A strengthened consonant becomes more consonant-like and a vowel becomes more vowel-like. In the case of consonants, for example, this might be reflected by means of longer Voice Onset Time (VOT) and/or longer closure duration (Cho, 2016, p. 133), although it is also possible that shorter VOT values may reflect strengthening (see Cole et al., 2007, for English, and Cho and McQueen, 2005, for Dutch, both discussed in Section 2.4.3). Thus, prosodic strengthening at the syntagmatic level may help to mark the disjuncture between two constituents. At the paradigmatic level, the strengthening may result in the enhancement of phonological features that make a segment more distinct from others of the same class (Cho, 2016, pp. 133–134; Cole et al., 2007, p. 182). For example, Georgeton and Fougeron (2014) studied French rounded and unrounded vowels. They reported that both classes of vowels were articulatorily strengthened at the beginning of an IP (when compared to word-initial position within an IP), but that the strengthening (by means of lip aperture and lip width) was greater for unrounded vowels. They interpreted this result as the enhancement of the phonological paradigmatic contrast between unrounded and rounded vowels (pp. 93–94). However, it is not always possible to tease apart syntagmatic and paradigmatic strengthening because they often co-occur (Cho, 2016, p. 134).

The precise extent of application of DIS seems to be subject to crosslinguistic differences. One important question is whether strengthening applies only to the very beginning of the constituent (whether it is conceived of as a segment or as the start of the articulatory gesture) or if it can extend further into the constituent. For example, in English, the extent of the strengthening at the left edge is mostly limited to the first syllable and especially to its first segment, which in a CV syllable would mean the consonant. Incidentally, the influence of prominence seems to be centered on the vowel, and its effect on the preceding consonant seems to gradually diminish leftward (Cho, 2016, pp. 128–129).

Fougeron and Keating (1995, 1997) put forth the hypothesis that DIS is *cumulative*, namely that strengthening at the left edge of a constituent is more pronounced the higher in the prosodic hierarchy that constituent is (see also Cho et al., 2007, and references therein, as well as Keating, 2003). This is related to *prosodic boundary strength*, which is "the degree of prosodic disjunction between abutting prosodic units, roughly in proportion to the level of the units in the constituent hierarchy" (Cho, 2016, p. 123). Thus, the strength of a boundary is proportional to that of the constituent it is related to, meaning that boundaries marking higher prosodic constituents are considered to be stronger than those for lower prosodic constituents.

There is some evidence in support of a cumulative effect of DIS in several languages (Fougeron & Keating, 1997, for English; Keating et al., 2004, for English, French, Korean, and Taiwanese; Onaka, 2003, for Japanese), although this cumulative effect is better thought of in terms of an overall tendency (Keating, 2003). For example, Keating et al. (2004) examined /n/ and /t/ at the beginning of several prosodic constituents in English, French, Korean, and Taiwanese. They measured articulatory seal duration, that is, the duration between the first and the last frames in an electropalatography (EPG) that show complete closure of the oral cavity, acoustic closure duration, and VOT for /t/. The consonants were inserted into meaningful sentences in the desired prosodic positions: the left edge of the U (which they defined as being set off by pauses; all languages), the IP (all languages), the AP (French and Korean), the PW (all languages), the small phrase (Taiwanese), and the syllable (that is, PW-medial; all languages). What was considered a PW necessarily differed across languages, but all of them were instances of LWs. Results showed that there was a similar pattern of cumulative DIS (measured in terms of articulation duration and acoustic duration) pointing to the existence of different levels of phrasing, although measures varied not only by language, but also by speaker and consonant. In more concrete terms, all speakers made a distinction between two levels of phrasing, one higher and one lower, but it was not the same levels of phrasing for all speakers (even of the same language). Overall, measures in PW-initial position did not differ from those in PW-medial position. Utterance and IP were only distinguished from each other in Korean; it was also only in Korean that VOT could differentiate between U, IP, AP, and PW levels.

Fougeron and Keating (1997, p. 3737) argued that increased duration of the articulatory gesture may help to explain the nature of articulatory strengthening at the edges. In principle, a more extreme articulation may be achieved if there is more time to achieve the articulatory target. Conversely, a shorter duration of the gesture may result in *articulatory undershoot* (see also Cho, 2016, p. 127; Kingston, 2008; and Parrell & Narayanan, 2018, for a similar approach). Fougeron and Keating (1997) hypothesized that a positive correlation between consonantal duration and linguopalatal contact would support the hypothesis that both are due to the same mechanism, that is, articulatory undershoot for nonstrengthened realizations.

Keating et al. (2004; see above) is also an example of a study that examines such a correlation, specifically between articulatory seal duration and linguopalatal peak contact.

In English, the correlation was marginal, which led them to conclude that the evidence did not support the articulatory undershoot hypothesis. On the other hand, articulatory seal duration was well correlated with linguopalatal peak contact in Korean (and, to a lesser degree, also in French); moreover, acoustic duration was also cumulative. Keating et al. (2004) suggest that, in Korean, unlike in English, there may be some evidence in support of the articulatory undershoot account, specifically that "there may well be a special pairing of temporal and spatial properties in domain-initial position in Korean compared to other languages" (p. 161). Further studies have yielded more evidence of a strong correlation between linguopalatal contact and duration (both articulatory and acoustic) for Korean (Cho & Keating, 2001), and also for Japanese (Onaka, 2003).

Fougeron and Keating (1997, p. 3737) suggest three possible ways in which DIS could help in perception, provided that the DIS has acoustic, perceivable features (for a similar observation, see Fletcher, 2010, p. 544, and references therein). First, it may help in the chunking of the stream of speech by marking the prosodic boundaries. Second, it may provide some information about the strength of those boundaries, at least by distinguishing between two levels of prosodic constituents (e.g., it could help to distinguish between the IP and the PW levels). Third, it may help in lexical access by enhancing what characterizes a particular segment, especially since it may enhance the word-initial segment, which seems to be important in word recognition. Nevertheless, there is still much research needed to clarify the precise nature of prosodic strengthening (Georgeton & Fougeron, 2014, p. 84).

Great interspeaker (and intraspeaker) variation has been reported in several DIS studies, even when the number of participants is limited to a handful because of the practical constraints, such as having custom-made pseudopalates for EPG (e.g., Fougeron & Keating, 1997; Georgeton & Fougeron, 2014; Keating et al., 2001; Keating et al., 2004). Other variables that may play a role in speaker variation, such gender, age, socioeconomic background, or language knowledge, are not included in the above cited research.

2.4.3 Acoustic studies on prosodic strengthening

Prosodic strengthening has also been examined in acoustic studies. In this section, two studies are presented, one on Dutch (Cho & McQueen, 2005) and another on American English (Cole et al., 2007). These studies examined the role of several factors on prosodic strengthening. In comparison to the articulatory studies reviewed above, they also investigated the role of accent and lexical stress, thus enriching our understanding of prosodic strengthening.

In a read speech study with 11 participants, Cho and McQueen (2005) examined the effect of prosodic boundaries, phrasal accent, and lexical stress on the phonetic realization of several Dutch consonants (/t/, /d/, /s/, and /z/). More specifically, they examined acoustic duration (closure duration in the case of the stops), VOT (for /t/), and voicing (of the closure for /d/, of the whole fricative for /s/ and /z/), among other parameters. The strong prosodic positions were word-initial position (in PW-initial, Small-Phrase-initial, and Big-Phrase-initial positions),[16] accented syllables (versus unaccented), and stressed syllables (versus unstressed). Overall, the tokens were produced with longer duration in strong positions, thus providing evidence for a prosodic strengthening effect on the phonetic realization of the consonants (manifested through lengthening). There was an interaction between accent and stress for several of the measures, indicating that the effect of accent was greater on stressed than on unstressed syllables.[17] The consonants for which the prosodic effects were more evident were /d/ and /t/. For the closure and voicing duration of /d/, accented tokens were longer than unaccented but stressed ones, although unaccented but stressed tokens were not longer than unaccented but unstressed ones. For the VOT of /t/, the interaction between accent and stress showed that "there was a cumulative effect of Stress and Accent on VOT" (p. 134). Overall, VOT values were shorter when in a strong position, in contrast with English, arguably because Dutch does not have aspirated voiceless stops (but see Cole et al., 2007, below).

In a study of American English, Cole et al. (2007) conducted an acoustic study on the effects of prosodic position (IP-initial versus IP-medial) and accent (accented versus unaccented; unaccented could be stressed or unstressed) on the phonetic manifestations of voicing and place of articulation of stops (/p/, /t/, /k/, /b/, /d/, /g/). In order to do so, they analyzed the realizations of stops in the recordings of four professional news announcers from the *Boston University Radio News* corpus (Ostendorf et al., 1995). Overall, in IP-medial position, there was a significant effect of accent on VOT, closure duration, and F0, with tokens in accented syllables having greater values for these parameters. The data for the comparison between IP-initial and IP-medial were drawn from the tokens for /t/ and /d/. Contrary to their expectations of an enhancement of the consonants in IP-initial

[16] Cho and McQueen (2005) used these groupings, which are based on prosodic boundary strength, instead of commonly used prosodic constituents, because of uncertainty about the number and characteristics of phrase-level constituents in Dutch.

[17] "Therefore, when it is said in the present study that test consonants in unstressed syllables are accented, it means that the whole target-bearing words are accented with the test consonants being in pre-accented syllables" (Cho & McQueen, 2005, p. 127).

position, there were no significant effects for any of the acoustic parameters, neither in accented nor in unaccented tokens. Further examination of the standard deviations of the parameters showed lower values in IP-initial position than in IP-medial position. Cole et al. interpreted this result as the consonant in IP-initial position being "produced with greater precision in their laryngeal and supralaryngeal articulations" (p. 202), which they argued could constitute "a different kind of strengthening, if the result is greater perceptual salience of those contrasts for the listener" (p. 202). This is in line with the proposal by Fougeron and Keating (1997) that prosodic strengthening may help to mark prosodic boundaries (see Section 2.4.2).

In sum, not only prosodic position but also lexical and postlexical prominence may have an effect on prosodic strengthening. The two studies also show that the effect of lexical stress and accent may be language dependent. In both Dutch and English, there was more strengthening in accented syllables (vs. unaccented syllables), while only in Dutch was there a cumulative effect of stress and accent.

2.4.4 Summary

This section provided an overview of the main characteristics of prosodic strengthening, which included a detailed characterization of domain-initial strengthening (with examples of articulatory studies) and two acoustic studies that investigated prominence at the postlexical and lexical levels. These studies show that several factors, such as prosodic position or lexical prominence, can have an effect on prosodic strengthening.

Chapter 3

Speaker-specific variation: Language dominance and gender

This chapter presents two variables, language dominance and gender, as possible sources of interspeaker variation. Speaker-specific variation in phonetic research has started to gain attention only recently. Foulkes et al. (2010) argue that, in phonetic studies, interspeaker differences have been "frequently treated as undesirable noise in the data" (p. 716), whereas in the field of sociolinguistics, an assessment of interspeaker variation has been lost due to the averaging of data points for each speaker group.[1]

The remainder of this chapter is organized as follows. In Section 3.1.1, bilingualism is defined and the Speech Learning Model (SLM), which is a model of second language (L2) phonological development, is introduced. This model has been used in previous studies to address the perception and production of some sounds (e.g., McKinnon, 2020), and it will be used in Part II of the present dissertation to model the possible influence of bilingualism on prosodic strengthening. In Section 3.1.2, language dominance and the Bilingual Language Profile are introduced. Language dominance, which is one way of operationalizing bilingualism, will also be employed in Parts II and III of the present dissertation by means of the Bilingual Language Profile, which is a tool for assessing language dominance. In Section 3.1.3, an overview of the treatment of bilingualism in studies on Yucatecan Spanish is provided. In Section 3.2, an introduction to phonetic variation as an index of gender is provided, as well as an overview of how gender has been investigated in studies of Yucatecan Spanish. This section provides the theoretical basis for the study of gender in Parts II and III of the present dissertation.

3.1 Bilingualism and language dominance

3.1.1 Bilingualism and the Speech Learning Model

Bilingualism can be defined as "the use of two or more languages on a regular basis, irrespective of proficiency and age of acquisition as a bilingual practice" (Piller & Pavlenko, 2006, p. 489). Studies have stressed that a bilingual speaker is not two monolingual speakers in one (Birdsong, 2015, p. 104; Grosjean, 1989; 2008; Silva-Corvalán & Treffers-Daller, 2015a, p. 2), and that bilinguals are usually not equally fluent in both languages (Grosjean, 2008, p. 243). Rather, each bilingual speaker has a particular language configuration

[1] There are exceptions to these trends. See Foulkes et al. (2010) and references therein.

(Grosjean, 1989, p. 3). Moreover, in the processing, production, and perception of sounds, there is evidence that supports an integrated view of the two languages, with dynamic crosslinguistic interactions taking place (Simonet, 2016).

Three main models of L2 phonological development have been proposed. The *Perceptual Assimilation Model* (PAM-L2; Best & Tyler, 2007) and the *Second Language Linguistic Perception* model (L2LP; Escudero, 2005) deal with the process of learning an L2 from its earliest to latest stages. Both models account for perception; the L2LP also accounts for production. The third model is the *Speech Learning Model* (SLM; Flege, 1995, 1999, 2002).[2] This model addresses the production and perception of sounds by bilingual speakers, especially by those who have used their L2 for many years but who still retain some foreign accent. Because the SLM offers hypotheses that can be applied to the production of adult bilingual populations (such as the one whose speech will be examined in this dissertation), its main characteristics are presented below.

The SLM aims to explain why bilinguals may not be able to attain native-like pronunciation, in particular of *phonetic categories*, that is, "language-specific aspects of speech sounds . . . specified in long-term memory representations" (Flege, 1995, p. 239). Age of acquisition of the L2 and the perceived distance between the sound categories of the first language (L1) and the L2 are important variables in the model. Overall, the model proposes that an early age of acquisition of the L2 will facilitate more native-like production of sounds. However, while age of acquisition may be easy to determine (e.g., from biographical information provided by the participants), the perceived distance between sound categories is problematic because it is unclear how to quantify it: "An obstacle to testing hypotheses such as these is the lack of an objective means for gauging degree of perceived cross-language phonetic distance. It is uncertain, also, what metric bilinguals use in doing so" (Flege, 1995, p. 264).

The SLM is a dynamic model, not only because it takes into account that the perception and production of a phonetic category may change over time due to factors such as an increase in the use of either of the two languages, but also because the categories of the L1 and the L2 interact with one another in the same "phonological space" (Flege, 1995, p. 239). The model makes predictions about how the L1–L2 interaction may take place by proposing two mechanisms: category assimilation and category dissimilation. By means of *category assimilation*, bilinguals interpret phonetic categories of the L2 in terms of the L1. There are two possible results: (i) the L2 sound may be produced as identical to the L1 sound if the

[2] See Elvin et al. (2019) for a comparison of the three models.

bilingual fails to perceive the difference between the L1 and L2 categories, or (ii) a merged L1–L2 category is created, which usually takes place when the L1 and L2 categories are phonetically similar (Flege, 2002, pp. 225–226).[3] Flege (2002) exemplifies this second result of assimilation with a previous study (Flege, 1987) of the voice onset time (VOT) of /t/ in French–English bilinguals. The bilinguals were grouped into those who had French as their L1 and those who had English as their L1. Results for both groups showed that the VOT values of French and English productions did not match those of monolingual speakers of either French or English; that is, neither group had created two different categories, one for French /t/ and another for English /t/. In fact, the VOT values of both groups were in between the values that would be expected for French (i.e., lower VOT values) and those that would be expected for English (i.e., higher VOT values).

By means of *category dissimilation*, bilinguals may create new L2 phonetic categories in their phonetic inventories. The possibility of this happening increases (i) if the difference between the L1 and the L2 categories is perceived as high, and (ii) the sooner the learning of the L2 starts. The new L2 phonetic category (which may not be identical to the corresponding monolingual category) and the L1 category become *dissimilated*, that is, the difference between them is increased so that the two categories (the L1 category and the L2 category) are more distinct from each other phonetically. For example, in two studies of VOT of /p t k/ with Spanish–English bilinguals (Flege & Eefting, 1986, 1987), the VOT values for both languages were shorter than in the productions of monolingual speakers, but bilinguals still seemed to create categories for the L1 and the L2 (cf. the French–English bilinguals example above). Importantly, the VOT values for Spanish were even shorter than those of Spanish monolinguals; that is, the difference between the L1 (Spanish) and L2 (English) categories was increased so that they were more distinct phonetically.

The examples provided so far refer to voiceless stops, which exist in the three languages examined (English, French, and Spanish). This raises the question of what the result would be if one of the two languages compared lacked the phonetic category under study. McKinnon (2020) examined the production of Spanish voiced stops /b d ɡ/ by Kaqchikel–Spanish bilinguals. Kaqchikel is a Mayan language spoken in Guatemala, which, like Yucatec Maya, does not have voiced stops. On the basis of the SLM model, McKinnon (2020) hypothesized that bilingual speakers "would form a new L2 category for the Spanish

[3] "One would expect phonetic category assimilation to operate when the L1 and the L2 possess speech sounds that are close to one another in phonetic space, but are not physically identical to one another" (Flege, 2002, p. 226).

voiced stops; however, the category could be different from monolingual Spanish speakers" (p. 162). The comparison of the relative intensity of voiced stops produced by bilingual Kaqchikel–Spanish speakers to that reported for other varieties of Spanish (not to that of monolingual Spanish speakers from Guatemala) led McKinnon to conclude that the lower intensity values obtained in his study (vs. the values reported for other varieties) support the hypothesis that the bilingual speakers had created a new L2 category.

The interaction between the phonetic categories of the two languages of a bilingual appears to be even more dynamic than the SLM posits. Simonet (2016) provides an overview of empirical studies that show that bilingual speakers can modify the phonetic categories and the connection between them to an even greater extent than Flege's model suggests. For example, Simonet argues that exposure to more variable speech may lead to a more native-like categorization of the sounds of the L2, at least when this is tested in an experimental setting. Moreover, the communicative context may play a role as well. To prove this point, Simonet (2016) reviews two studies on stops, namely Olson (2013) for English–Spanish bilinguals and Simonet (2014) for Catalan–Spanish bilinguals. In these studies, L1 and L2 stops presented different acoustic characteristics as a function of whether the session in which they were produced was unilingual or bilingual: L1 and L2 stops produced in unilingual sessions were acoustically more different from each other than L1 and L2 stops elicited in bilingual sessions.

3.1.2 Language dominance and the Bilingual Language Profile

Language dominance is a complex construct that has been used "to capture the bilingual experience" (Silva-Corvalán & Treffers-Daller, 2015a, p. 1). The dominant language is "that in which a bilingual has attained an overall higher level of proficiency at a given age, and/or the language that s/he uses more frequently, and across a wider range of domains" (Silva-Corvalán & Treffers-Daller, 2015a, p. 4).[4]

Birdsong (2015, p. 86) makes a distinction between dimensions and domains of language dominance. The *dimensions* of language dominance refer to all the features that have been associated with being bilingual, whether linguistic (e.g., fluency of speech) or not (e.g., degree of identification with the cultures of each language). Language dominance is

[4] See Silva-Corvalán and Treffers-Daller (2015b) for a survey of how several authors have defined language dominance.

thus multidimensional. Furthermore, it is noncategorical,[5] because the dimensions that are considered for language dominance (e.g., fluency of speech) are not categorical. It is also relativistic because it is necessary to compare the two languages in order to assess language dominance. The *domains* of language dominance refer to the contexts in which the languages are used (e.g., in an academic lecture). Because the speaker may be able to choose which language to use, domains can be volitional (e.g., the speaker can choose what language to use in inner speech).

Language dominance is closely associated with the concepts of balanced bilingualism and language proficiency. *Balanced bilinguals* are those who have similar levels of performance in both languages (Birdsong, 2015, p. 95); that is, they are not dominant in neither language. *Language proficiency* refers to linguistic ability and is commonly assessed by means of examinations that test the four classic skills (speaking, listening, writing, and reading; Birdsong, 2015; Edwards, 2006).[6] Language dominance, however, is assessed by comparing the knowledge about two languages that one speaker has (Birdsong, 2015, pp. 91–92). If the speaker has a similar degree of dominance in both languages, it is possible to talk about balanced bilingualism. However, being dominant in one language does not imply being proficient in that language (Birdsong, 2015, p. 92), even though language proficiency can be considered "one component of dominance" (Gertken et al., 2014, p. 208; see also below).

Language dominance has been assessed by means of several methods (see Solís-Barroso & Stefanich, 2019, for an empirical comparison of language dominance assessment methods). The *Bilingual Language Profile*, or *BLP* (Birdsong et al., 2012; Gertken et al., 2014) assesses language dominance in bilinguals. The definition of bilingualism that is considered in the BLP is speaking two (or more) languages. This broad definition makes the BLP a useful tool for assessing language dominance in many bilingual contexts (Gertken et al., 2015, p. 221). Bilingual Language Profile scores are obtained by means of a questionnaire. In it, participants self-report on their knowledge in four areas: language history, language use, language proficiency, and language attitudes. The four modules, which receive equal weighting, include several questions that reflect both dimensions and domains of language dominance. The resulting BLP score is a continuous measure of

[5] *Noncategorical* is used here for simplicity. Birdsong (2015) characterizes the dimensions as gradient or continuous: "As these examples suggest, the dimensions along which the construct of dominance is operationalized are inherently *gradient* or *continuous,* not categorical" (p. 86).

[6] The concept of language proficiency has also received multiple definitions. See Austin et al. (2019, pp. 1–5) and Gertken et al. (2014, pp. 211–212) for an overview.

language dominance. Each end of the continuum is indicative of monolingualism (Amengual & Simonet, 2019, p. 8).

There are several advantages of a continuous measure over a categorical grouping of participants according to their dominance in one language or the other. First, by using a continuous variable, the researcher is not confronted with having to decide where to set the cutoff point between groups of bilinguals, such as dominant and balanced bilinguals (see Treffers-Daller, 2015, p. 268, for further discussion). Nevertheless, the BLP score can still be used as a tool for creating categorical variables. For example, Amengual Watson (2013) and Amengual and Simonet (2019) grouped the Catalan–Spanish bilinguals in their studies into Catalan-dominant and Spanish-dominant by setting the cutoff point at 0. Second, a continuous measure reflects interspeaker differences that would be lost if the participants were grouped. Moreover, no information is lost in terms of statistical power (Birdsong, 2015, p. 92). Third, treating language dominance as a categorical or continuous variable has consequences in terms of the predictions it makes about linguistic behavior. This point is exemplified by Solís-Barroso and Stefanich (2019, p. 15). If two speakers who show great interspeaker differences between themselves are categorized as dominant in language A, the prediction would be that their linguistic performance should be similar. The same can be said about two very dissimilar speakers categorized as dominant in language B. However, such a categorization would overlook the possibility that the less "extreme" language-dominant speakers of languages A and B may be similar to each other. In other words, "the more balanced a bilingual is, the greater chance s/he has of being inconsistently classified" (p. 16).

3.1.3 Bilingualism in Yucatecan Spanish studies

Bilingualism in studies on Yucatecan Spanish has usually been regarded as a sociolinguistic variable for which speakers are categorized as either monolingual or bilingual.[7] Speakers are usually grouped on the basis of self-reporting of language knowledge, although in some instances the researchers do not provide any information about how the categorization was applied (e.g., Michnowicz & Carpenter, 2013). Thus, speakers of Yucatecan Spanish are categorized as monolinguals if they do not speak Yucatec Maya,[8] while bilinguals are those

[7] There are, however, some works that use the BLP score to measure language dominance, such as Martínez García (2017) and Martínez García and Uth (2019).

[8] Note that this is not the same as saying that the speakers categorized as monolinguals are those who *only* speak Spanish. The studies reviewed here do not report knowledge (or absence thereof) of languages other than Yucatecan Spanish and Yucatec Maya.

who may speak and/or understand Yucatec Maya. In Rosado Robledo (2011) and Michnowicz (2011), bilingual speakers are those who can speak Maya, whereas in Michnowicz and Kagan (2016), bilingual speakers are those who speak it fluently. Michnowicz (2009) describes bilinguals as "active" if they speak Maya "in everyday interactions with their families" (p. 73) and as "passive" if they can understand it (p. 73). In Uth (2016), both Spanish-dominant and balanced bilinguals understand Yucatec Maya, but whereas the former speak it "to a certain degree" and use it rarely (p. 259), the latter consider themselves native speakers of both languages and use them regularly (p. 259).

3.2 Gender

A large body of research has shown that the speech of female and male speakers differs in several regards (see Simpson, 2009, for an overview). Importantly, whereas there are differences due to biological sex (e.g., differences in F0), there are also differences due to gender, that is, to how female and male identities are socially constructed. In fact, not only gender but also sexual identity may be indexed by means of phonetic features (Podesva & Kajino, 2014). However, it is not necessarily clear whether differences between females and males have a biological basis or a social one (see Munson & Babel, 2019, for an overview). For example, Oh (2011) examined the VOT of Korean stops and found an effect of gender on long-lag stops, with male speakers displaying longer VOT values. This result contrasts with studies on English, which have consistently indicated that the female speakers, not the male ones, are the ones displaying longer VOT values for long-lag stops (Oh 2011, p. 60). Oh concluded that the VOT differences found in Korean and in English may be sociophonetic and not physiologically grounded (p. 66).

Because gender is socially constructed, the phonetic differences that index gender may index other social factors as well, such as social class. Moreover, the way that phonetic features are used to index gender and other social categories can evolve over time. This is the case of the realization of /s/ in Glaswegian. Stuart-Smith (2007) found differences as a function of gender in terms of Center of Gravity and peak frequency of /s/ in a corpus of spontaneous speech acoustic recordings from 1997. The study also showed that the results for working-class girls (13–14 years of age) did not pattern with those of the other female speakers, but that they seemed to pattern with those of male speakers instead. However, the 1997 recordings were reanalyzed by Stuart-Smith (2020) for Center of Gravity and spectral slope and compared to acoustic recordings of Glaswegian that predated 1997. The results suggested that the working-class girls recorded in 1997 were reverting to older female vernacular norms, which were more similar to the lower spectral frequency used by males, but also that the /s/ of working-class boys was changing towards /ʃ/, shifting away from that of their female counterparts.

Gender has been included as a sociolinguistic variable in several studies of Yucatecan Spanish. Most studies have found no differences between female and male speech, but there are some exceptions. For voiced stops, one study (Pérez Aguilar, 2002) reported a gender difference, with female speakers producing more stop realizations than male ones. For the glottalization of word-initial vowels, Michnowicz and Kagan (2016) reported no differences as a function of gender.[9] Finally, for the labialization of word-final nasals (e.g., *Yucatán* produced as *Yucatám*), Yager (1989) and Michnowicz (2008) indicated that it is more frequent among female speakers, whereas García Fajardo (1984) found no gender differences. In sum, there is a need for further studies to assess the role of gender-related variation in Yucatecan Spanish.

[9] The gender-related findings about voiced stops and the glottalization of vowels will be discussed further in Chapter 4.

Chapter 4

Prosodic strengthening in Spanish and in Yucatecan Spanish

4.1 Overview

Voiced stops and vowels in Yucatecan Spanish differ from those in other Spanish varieties in that they present higher rates of strengthened realizations, that is, more stop-like realizations of voiced stops (vs. more approximant-like realizations) and glottalization of word-initial vowels. The studies that will be reviewed in this chapter provide information about the acoustic characteristics and distribution of strengthened realizations of these sounds in Spanish and in other languages, as well as in Yucatecan Spanish. However, most studies on voiced stops and glottalization in Yucatecan Spanish are sociolinguistic in nature; moreover, they base their conclusions on impressionistic observations or on auditory impressions without providing phonetic information. This dissertation aims to provide an acoustic analysis of Yucatecan Spanish voiced stops and glottalization of word-initial vowels in terms of prosodic strengthening.

In Yucatecan Spanish, voiced stops are frequently strengthened intervocalically, especially when they occur in word-initial position. However, several phonetic studies on other Spanish varieties have shown that intervocalic stops are weakened, even if they appear in word-initial position (e.g., Carrasco et al., 2012; Colantoni & Marinescu, 2010; Eddington, 2011; Ortega-Llebaria, 2004). Moreover, glottalization of vowels is also frequent in word-initial position in Yucatecan Spanish (e.g., Lope Blanch, 1987) and rare in other varieties of Spanish (e.g., Bissiri et al., 2011). Thus, according to previous accounts, strengthened realizations of voiced stops and glottalization of vowels occur word initially in Yucatecan Spanish. This dissertation aims to investigate whether the distribution of strengthened realizations in Yucatecan Spanish is amenable to an explanation in terms of prosodic structure, and more specifically, in relation to the Prosodic Word (PW). It will investigate whether strengthened realizations serve to mark the left edge of the PW domain, that is, if there is domain-initial strengthening (see Section 2.4.2), and also if lexical stress has an effect, that is, if realizations are strengthened to a greater extent in stressed syllables (see Section 2.4.3).

The studies on Yucatecan Spanish that will be reviewed in this chapter have also investigated the role of speaker-specific variation. Thus, there are several sociolinguistic studies that have considered factors such as knowledge of Yucatec Maya, gender, age, or sociocultural background as sources of variation for strengthened realizations (e.g., Michnowicz & Kagan, 2016; Rosado Robledo, 2011). In this dissertation, two factors that contribute to speaker variation will be investigated, namely language dominance (Yucatec Maya–Yucatecan Spanish language dominance) and gender.

The current chapter presents previous research about Yucatecan Spanish voiced stops within the context of Spanish phonetics and phonology (Section 4.2) and research on nonmodal phonation and glottalization of vowels in several languages (including Yucatec Maya), as well as in Yucatecan Spanish (Section 4.3), followed by a brief summary of the chapter (Section 4.4).

4.2 Voiced stops in Spanish

In this section, a phonological characterization of the voiced stops in Spanish is presented (Section 4.2.1). This is followed by an overview of articulatory and acoustic studies on Spanish voiced stops, as well as of sources of variation in their phonetic realization (Section 4.2.2). The characteristics of voiced stops in Yucatecan Spanish are then presented (Section 4.2.3).

4.2.1 Phonological characterization

Spanish has three voiced oral stops (/b d g/) and three voiceless ones (/p t k/). In Mexican Spanish, as well as in other varieties of Spanish, they are weakened intervocalically (Avelino, 2018; see discussion below). Voiceless stops in Spanish are usually described as unaspirated (e.g., Martínez-Celdrán et al., 2003, for Castilian Spanish; Avelino, 2018, for Mexico City Spanish; Cho & Ladefoged, 1999, for Puerto Rican Spanish). The study of Spanish voiced stops has been undertaken on the theoretical assumption that the phonemes /b d g/ can be manifested phonetically either as stop allophones [b d g] or as approximant allophones [β ð ɣ]. In the earliest works on Spanish (e.g., Alarcos Llorach, 1965/1991; Navarro Tomás, 1918/1996; Quilis, 1999), the approximant realizations of voiced stops were taken to be fricatives. However, experimental studies have shown that such realizations are in fact approximants (e.g., Eddington, 2011; Parrell, 2011).

Stop and approximant allophones are taken to be in complementary distribution (e.g., Carrasco et al., 2012; Hualde, 2013). The realizations are determined by the preceding sound: stop realizations appear after a pause and after a nasal (e.g., *mambo* [ˈmambo] ‘mambo’, *mundo* [ˈmundo] ‘world’, *mango* [ˈmango] ‘mango’); for /d/, they also appear after /l/ (e.g., *caldo* [ˈkaldo] ‘broth’). In all other postconsonantal contexts and in all postvocalic contexts, approximant realizations are produced. In intervocalic contexts, Hualde, Simonet, and Nadeu (2011) argue that voiced stops are produced as approximants in all speech styles, even in careful read speech. According to these authors, "this lenition process has no lexical exceptions and it applies both inside words and across word boundaries" (p. 304); furthermore, it takes place regardless "of the presence of morphological or syntactic boundaries" (p. 305). Nevertheless, they point out that there is variation in the degree of constriction of these approximant sounds, which can go from very close to very open, and that the latter may result in deletion (p. 304).

The issue of whether the phonological representation of these sounds should be /b d g/, which is the most extended practice, or /β ð ɣ/, has been discussed widely in Spanish phonology. For example, Veiga and Arias (in press) argue that their representation as stops is not justified in terms of frequency because approximant realizations are more frequent. More importantly, the issue of phonological representation has also been understood in terms of strengthening/fortition (see Section 2.4.1), according to which [b d g] are the strengthened realizations of /β ð ɣ/, or in terms of weakening/lenition, according to which [β ð ɣ] are the weakened realizations of /b d g/. Lavoie (2001) points out that either approach may be equally justifiable (p. 6). Thus, there are some proponents of the strengthening account (Barlow, 2003; Eddington, 2011; Lavoie, 2001), but the weakening approach is arguably the most common, maybe owing to the parallelism that can be drawn with diachronic analyses of Latin intervocalic /p t k/, which lenited into Spanish /b d g/. For example, Harris (1969) proposed a "Rule of Spirantization" (pp. 37–40) to reflect the stop/approximant alternation shown above. For this rule, Harris assumed that stops may become "continuant" (i.e., approximants), and not the other way around, an assumption that Harris admits may be incorrect (p. 38). Nevertheless, there have also been phonological accounts that propose instead that voiced obstruents in Spanish are underspecified for the feature [continuant] (Colina, 2016, pp. 116–129; Lozano, 1979, in Veiga & Arias, in press). In the present dissertation, the most extended practice (i.e., /b d g/) is followed for convenience. Importantly, a large body of phonetic research has shown that a binary allophonic categorization is a simplification because there is a continuum of realizations (e.g., Eddington, 2011; Hualde, 2013; Hualde, Shosted, & Scarpace, 2011; Hualde, Simonet, & Nadeu, 2011; Martínez Celdrán & Fernández Planas, 2007; Simonet et al., 2012; Soler & Romero, 1999). These phonetic studies will be discussed in the following sections.

4.2.2 Phonetic studies

Phonetic studies on Spanish voiced stops have been conducted primarily from the perspective of weakening/lenition. Whereas some works focus on voiced stops in intervocalic contexts (e.g., Colantoni & Marinescu, 2010; Martínez Celdrán, 2013), others aim to establish whether there are also approximant realizations in contexts for which strengthened realizations with an oral closure are expected, following the categorical distinction presented in Section 4.2.1 (e.g., Hualde, Shosted, & Scarpace, 2011; Simonet et al., 2012). Moreover, some studies aim to determine what differentiates Spanish voiceless stops from voiced ones (e.g., Hualde, Simonet, & Nadeu, 2011; Parrell, 2011), whereas others compare voiced stops in English and Spanish (e.g., Lavoie, 2001; Ortega-Llebaria, 2004; Parrell & Narayanan, 2018) and also study other consonants from the perspective of weakening/strengthening (e.g., Lavoie, 2001; Lahoz Bengoechea, 2015). In this section, I focus on the research regarding Spanish voiced stops mainly in intervocalic context.

4.2.2.1 Articulatory studies

The articulatory realizations of voiced stops can be placed along a continuum of realizations that go from complete constriction of the supraglottal articulators to an absence of such a constriction. In acoustic terms, the realizations go from full stop realizations to vowel-like approximant realizations. Articulatory and acoustic studies of Spanish voiced stops are consistent in their description of the continuum of phonetic realizations. Thus, strengthened realizations have longer articulatory and acoustic closure duration, decreased acoustic intensity, and, at the far end of the continuum, they present complete closure of the articulators. Conversely, weakened realizations have shorter articulatory and acoustic closure duration, increased acoustic intensity, and, at the far end of the continuum, they present the complete opening of the articulators.

Articulatory studies show that the more constricted (i.e., strengthened) realizations appear in the vicinity of prosodic boundaries, especially those corresponding to the Intonational Phrase (IP). For example, Lavoie (2001) measured linguopalatal contact data by means of electropalatography (EPG) for /d/ and /g/. Two Mexican participants read reiterant speech made up of disyllabic words (voiced stop + /o/) in different stress patterns. There were no substantial differences as a function of lexical stress (stressed vs. unstressed). Results showed that there was an increase in linguopalatal constriction that resulted in closure in IP-initial position. By contrast, in intervocalic contexts (i.e., all other instances) the stops were commonly articulated as approximants. Similarly, Hualde, Shosted, and

Scarpace (2011) conducted an EPG study of /d/ in word-medial position with three speakers of European Peninsular Spanish. The results showed that only 12 % of realizations were produced with full closure of the articulators.

Parrell (2011) conducted an electromagnetic articulometer (EMA) study of /b/ with two European Peninsular Spanish speakers. The consonants were placed in three prosodic positions: phrase-initial, (phrase-medial) word-initial, and word-medial. The examples of carrier sentences provided in the article show that the phrase-initial condition can be equated to IP-initial (e.g., *La chica juega.* ***V****aga también* 'The girl plays. She wanders as well', p. 429).[1] Results show an effect of prosodic position on the realization of /b/, with results for word-initial and word-medial positions (i.e., phrase-medial) patterning together for both the articulatory measurements of duration of the gesture (total duration, constriction duration, movement duration) and degree of constriction (lip aperture). The duration of the gesture and the degree of constriction were larger in IP-initial position, where /b/ had a full articulatory closure. The acoustic data also showed that they were produced as full stops with a period of prevoicing of up to 103 ms. In phrase-medial position, constriction duration and constriction degree measures were positively correlated. The acoustic data showed that tokens of phrase-medial /b/ were all produced as approximants.

Parrell and Narayanan (2018) obtained similar results for /d/ to those of Parrell (2011). This was a real-time Magnetic Resonance Imaging (MRI) study of European Peninsular Spanish and American English /d/, /t/, and /n/. The three prosodic positions were phrase-initial, word-initial, and word-medial. Similarly to Parrell (2011), phrase-initial position can be equated with IP-initial position. Unlike in Parrell (2011), the stimuli for the word-initial condition were designed to elicit a prosodic boundary, although smaller than that of IP-initial position. The results of the three measurements, namely tongue tip, tongue body, and jaw movement were rather similar for the two languages and for all consonants. For /d/, results for word-initial position patterned with IP-initial position for tongue-tip and tongue-body movement, although for jaw movement the three conditions did not pattern together. Additionally, Parrell and Narayanan indicated that in 7 % of the tokens of /d/ in IP-initial and word-initial positions full closure was not achieved since palatal contact was absent; for word-medial position, 50 % of tokens showed no contact of the tongue with the palate. Thus, results for the voiced stops in (phrase-medial) word-initial position differ between Parrell (2011) and Parrell and Narayanan (2018) since in Parrell (2011) the word-initial condition patterned with the word-medial condition. I suggest that

[1] The bilabial stop can be represented in Spanish orthography by both <b> and v>.

this could be explained by the insertion of a prosodic boundary before the target sounds in word-initial condition in Parrell and Narayanan (2018).

Parrell (2011) and Parrell and Narayanan (2018) reject a view of voiced stop variation in terms of allophonic categories because the data they obtained is not amenable to a categorical analysis. In particular, Parrell and Narayan (2018) show that the results for the three prosodic conditions overlap to a great extent. Parrell (2011) and Parrell and Narayanan (2018) offer an articulatory undershoot account for the issue of the underlying realization of Spanish voiced stops, which is in line with the argumentation in connection to domain-initial strengthening in Fougeron and Keating (1997; see Section 2.4.2). Parrell (2011) argues that these sounds should be viewed as stops underlyingly, and not as approximants. The crucial idea is that these consonants have an articulatory spatial target. Whether that target is reached or not depends on the duration of the articulatory gesture. Parrell and Narayanan (2018) argue that the spatial magnitude of the gesture is conditioned by its duration (p. 165).[2] If the duration of the gesture is long enough, the spatial target is reached, which results in full closure of the articulators. This is the case in IP-initial position, as the measurements of gesture duration and constriction in IP-initial position showed. On the other hand, if the duration is shorter, the spatial target is not reached; that is, there is gestural undershoot, which results in reduced (weakened) realizations. This is the case in word-medial position, where the gesture may be undershot, with a shorter duration and lesser constriction degree. The results in Parrell (2011) and Parrell and Narayanan (2018) are similar to those obtained for Korean (Cho & Keating, 2001) and Japanese (Onaka, 2003; see Section 2.4.2), languages in which a positive correlation between the temporal and the spatial dimensions of prosodic strengthening is observed. In sum, the different articulatory productions of voiced stops can be explained as "the dynamic consequences of interactions of a single, invariant spatial target and variable gesture duration" (Parrell 2011, p. 427).

4.2.2.2 Acoustic studies

Acoustic studies of Spanish voiced stops have investigated acoustic correlates of articulatory constriction degree (intensity, duration, spectral tilt, and presence of a release burst), in many cases in relation to linguistic factors such as prosodic position or lexical stress (see Section 4.2.2.3). With a few exceptions that have analyzed spontaneous speech (e.g.,

[2] The duration of the gesture is modulated by prosodic structure, as these studies have shown, and it may also be modulated by other segmental and suprasegmental factors (Parrell & Narayanan, 2018, p. 156).

Colantoni & Marinescu, 2010; Eddington, 2011), the data have been obtained by means of scripted speech (in the form of games or narration tasks) or read speech tasks. Table 4.1 provides an overview of these studies, which will be discussed in this and the next section (Section 4.2.2.3).

Table 4.1. Overview of acoustic studies on Spanish voiced stops. Variables discussed in the main text are included in the table. Read speech was used in these studies, unless indicated otherwise. 'European Pen.' – European Peninsular; 'spont.' – *spontaneous*; 'bil.' – *bilinguals*; 'mon.' – *monolinguals*; 'f' – *female*; 'm' – *male.*

Study	**Variety and speech style**	**Number of participants**	**Stops**	**Evidence for strengthening**	**Linguistic variables**
Carrasco et al. (2012)	Costa Rican, European Pen., *scripted & read*	Costa Rica: 10 (5 f); Spain: 3 (2 f)	/b d g/	· intensity	· prosodic position · lexical stress
Colantoni & Marinescu (2010)	Argentinian *spontaneous*	6 (6 m)	/b d g/	· intensity · duration	· lexical stress · flanking vowels · lexical frequency
Cole et al. (1999)	European Pen.	3 (1 f)	/g/	· intensity · duration	· prosodic position · flanking vowels
Eddington (2011)	7 varieties *spontaneous*	8 (5 f)	/b d g/	· intensity	· prosodic position · lexical stress · lexical frequency
Hualde, Shosted, & Scarpace (2011)	European Pen.	3 (—)	/d/	· intensity	—
Hualde, Simonet, & Nadeu (2011)	Majorcan bil. *scripted & spont.*	20 (10 f)	/b d g/	· intensity	—
Kingston (2008)	Ecuadorian, Peruvian	2 (2 f)	/b d g/	· intensity	· lexical frequency
Lahoz Bengoechea (2015)	Mexican	6 (3 f)	/b d/	· intensity · duration	· prosodic position
Lavoie (2001)	Northern Mexican	4 (4 m)	/b g/	· duration	· prosodic position
Martínez Celdrán (2013)	—	3 (3 f)	/b d g/	· intensity · duration · release burst	—
Martínez García (2017)	Yucatecan bil. & mon. *spontaneous*	12 (6 f)	/b d g/	· release burst	· prosodic position · lexical stress
McKinnon (2020)	Guatemalan bil. *scripted & spont.*	36 (18 f)	/b d g/	· intensity	· prosodic position · lexical stress · flanking vowels
Ortega-Llebaria (2004)	Caribbean varieties *scripted*	10 (10 f)	/b g/	· intensity	· lexical stress · flanking vowels
Polo Cano & Elordieta (2016)	European Pen.	4 (2 f)	/b d g/	· intensity · duration · release burst	· prosodic position
Soler & Romero (1999)	European Pen.	4 (m)	/b/	· intensity · duration	· speech rate
Simonet et al. (2012)	Majorcan bil. *scripted*	40 (20 f)	/d/	· intensity · spectral tilt	· lexical stress · flanking vowels

In terms of intensity, Cole et al. (1999) compared the intensity of the voiced stop to that of the whole word; Ortega-Llebaria (2004) compared it to the VCV context in which the voiced stop appeared. The most frequent measure is the comparison to the following vowel (Hualde, Shosted, & Scarpace, 2011; Martínez Celdrán, 2013; Soler & Romero, 1999). The difference in acoustic intensity between the lowest value in the stop and the highest value in the following vowel has been measured either as a difference (Colantoni & Marinescu, 2010; Eddington, 2011; Simonet et al., 2012) or as a ratio (Carrasco et al., 2012; Hualde, Simonet, & Nadeu, 2011; Lahoz Bengoechea, 2015; McKinnon, 2020).

Several measures of the abruptness of the transition between the consonant and the following vowel have been used. The maximum rising velocity between the consonant and the following vowel (both measured at their midpoints) was used by Hualde, Shosted, and Scarpace (2011). Kingston (2008), Hualde, Simonet, and Nadeu (2011), and Simonet et al. (2012) measured the intensity minimum (i.e., when the acoustic energy is falling the fastest) and the intensity maximum (i.e., when the acoustic energy is rising the fastest). A similar measurement is that of the speed of consonantal release into the following vowel, which is measured as the difference between the minimum and maximum values of the consonant as a function of time (Ortega-Llebaria, 2004). Larger values correspond to more strengthened realizations since they indicate a steeper change in intensity between the consonant and the following vowel.

Acoustic duration is also an important cue to the realization of voiced stops because the duration and the constriction degree seem to be highly correlated in the case of Spanish voiced stops (Soler & Romero, 1999; see also Parrell, 2011; Parrell & Narayanan, 2018; see Section 4.2.2.1). Most studies have measured the duration of the consonant (Colantoni & Marinescu, 2010; Martínez Celdrán, 2013; Soler & Romero, 1999), although that of the consonant plus its flanking vowels has also been considered (Martínez Celdrán, 2013).

Simonet et al. (2012) used a measurement of spectral tilt, that is, "the degree to which intensity drops off as frequency increases" (Gordon & Ladefoged, 2001, p. 397; cf. Section 4.3.1.1). Specifically, Simonet et al. measured the difference in acoustic energy between a low frequency band (50–500 Hz) and a high frequency one (500–5000 Hz) for the consonant /d/. The greater the articulatory constriction, the less energy that is to be expected in the high frequency band. Consequently, higher values of difference in spectral tilt indicate more strengthened realizations.

Colantoni and Marinescu (2010) visually inspected spectrograms to manually categorize realizations into stops or approximants.[3] Martínez Celdrán (2013) and Martínez García (2017) also inspected spectrograms and took the presence of a release burst (which would reflect full articulatory closure) as a manifestation of a stop. Polo Cano and Elordieta (2016; see also Polo Cano, 2015) categorized the sounds studied as approximants when they did not present a release burst, their duration was inferior to 55 ms, and their intensity at the lowest point was higher than 60 dB, and as stops otherwise. These approaches are problematic when used in isolation (i.e., without using other measurements as well, such as of duration or intensity) because realizations are grouped into two categories, in spite of their continuous nature. Thus, although it can be argued that realizations that present a release burst can be unequivocally categorized as stops, the category of approximants may encompass realizations manifesting different degrees of acoustic strengthening/weakening.

4.2.2.3 Sources of variation

The acoustic characteristics of voiced stops have been investigated in several Spanish varieties (see Section 4.2.2.2). Studies on Yucatecan Spanish will be reviewed in Section 4.2.3. Overall, these studies on several Spanish varieties show that many linguistic and nonlinguistic factors influence the realization of voiced stops. In what follows, the two most important linguistic factors, prosodic position and lexical stress, are discussed first, followed by an overview of vowel quality of flanking vowels, speech rate, lexical frequency, as well as speaker-specific factors.

The articulatory studies reviewed in Section 4.2.2.1 show that strengthened realizations are found in the vicinity of a prosodic boundary, especially for IP-initial position; the acoustic studies have yielded similar results for IP-initial position (Eddington, 2011), as well as for other constituents. Thus, Polo Cano and Elordieta (2016; see also Polo Cano, 2015) found a higher number of stop realizations (versus approximant ones) at the beginning of Phonological Phrases for European Peninsular Spanish.

In phrase-medial position, some studies have provided evidence for differences between the realization of voiced stops in word-initial versus word-medial positions, whereas others have found no differences. Lavoie (2001), in a study of Mexican Spanish, found a significant effect of prosodic position on duration, with stops in word-initial position being more strengthened in terms of duration than in word-medial position. However, Lahoz Bengoechea (2015), who also investigated Mexican Spanish, did not find such an effect for

[3] Colantoni and Marinescu (2010) did not explain the criteria on which they based their categorization.

the same prosodic positions, neither in terms of acoustic duration nor of intensity. Cole et al. (1999) found no strengthening in terms of intensity for prosodic position in their study of /g/ in European Spanish. Eddington (2011) found more strengthening in terms of intensity for /b/ and /d/ in word-initial position compared to word-medial position; finally, Carrasco et al. (2012) found the same strengthening effect for the three stops in the Costa Rican variety, but not in the European Spanish one.

In read speech, strengthened realizations are more frequent in stressed syllables, whereas in spontaneous and scripted speech, the effect of lexical stress is either not present at all or the results are mixed. The following studies reviewed did not considered the possible role of accent (see Section 2.2.1.1). Some read speech studies have found more strengthened realizations, in terms of intensity as a function of lexical stress, in Caribbean Spanish (Ortega-Llebaria, 2004) and in Costa Rican and European Peninsular Spanish (Carrasco et al., 2012). For Mexican Spanish, both Lavoie (2001) and Lahoz Bengoechea (2015) found more strengthened realizations in terms of duration (as well as in terms of relative intensity; Lahoz Bengoechea, 2015). Additionally, Lavoie (2001) reported a significant interaction between prosodic position (word-initial vs. word-medial) and lexical stress for /b/ (but not for /d/ or /g/), with /b/ in word-initial position and in a stressed syllable having longer duration.

In spontaneous speech, Colantoni and Marinescu (2010) found more strengthened realizations in stressed syllables in terms of relative intensity, although not in terms of duration, for Argentinian Spanish. Eddington (2011), who examined several varieties, found more strengthened realizations for /b/ and /d/ (but not for /g/) in terms of relative intensity. Finally, Simonet et al. (2012) found no strengthening effect of lexical stress in terms of intensity and spectral tilt in their study of Spanish-dominant bilingual Spanish–Catalan speakers.

The effect of the vowel quality of flanking vowels has been studied in read and spontaneous speech as well, with contradictory results. In read speech, Cole et al. (1999) found for European Spanish that /g/ was more strengthened (in terms of relative intensity) when flanked by the vowels /i e/ and /a/. Ortega-Llebaria (2004) found, for Caribbean Spanish, more strengthened realizations for /b/ and /d/ flanked by /a/ in terms of the speed of consonantal release, but not for /i/, which patterned with /u/ in being the most weakened.

In spontaneous speech in Argentinian Spanish, Colantoni and Marinescu (2010) found no differences in duration related to the preceding vowel and no clear patterns for the following vowel. They also pointed out that realizations for /g/ (but not for /b/ and /d/) were more strengthened in terms of intensity when followed by /e/ or /a/. Simonet et al. (2012), in their study of Spanish-dominant bilingual Spanish-Catalan speakers, found that realizations of /d/ were more strengthened in terms of intensity and spectral tilt when preceded by /i/ or /e/.

The effect that speech rate may have on strengthened realizations was investigated by Soler and Romero (1999). They found that realizations were strengthened by means of longer duration and lower intensity values in the "very slow" speaking rate when compared to the "moderately fast" one (p. 483).

The effect of lexical frequency on strengthening is not clear. Kingston (2008) investigated whether there were differences in relative intensity as a function of word frequency by means of an experiment with the three voiced stops in two groups of words, one of high-frequency words and another of low-frequency words. Results showed that there were no differences between the two groups. Furthermore, contrary to the author's expectations, consonants in the high-frequency group had longer duration. Additionally, Colantoni and Marinescu (2010) and Eddington (2011), who worked with spontaneous speech, indicated that /d/ presented more weakened realizations than /b/ or /g/, and also higher rates of elision. This elision phenomenon is common in all varieties of European and American Spanish (e.g., Cano Aguilar, 1992; Moreno Fernández, 2009). Eddington (2011) reported that /d/ (but not /b/ or /g/) was more weakened in high-frequency words; this was still significant after excluding the tokens that appeared in the extremely frequent *de* 'of' preposition. Also, Eddington pointed out that /d/ tokens that appear in the past participle suffixes (*-ado, -ada, -ados, -adas*), which are rather frequent in Spanish, have a tendency to have more weakened realizations or even be deleted, whereas he did not find this effect for the also very common past imperfect suffix *-aba*.

Bilingualism, gender, as well as other speaker-specific factors were considered in a study of Guatemalan Spanish as it is spoken by Spanish–Kaqchikel Maya bilinguals (McKinnon, 2020; see also Section 3.1.1). Like Yucatec Maya, Kaqchikel Maya has no voiced stops. Results indicated that female and older speakers produced more strengthened realizations than male and younger speakers.[4]

[4] In McKinnon (2020), as well as in Martínez García (2017; see Section 4.2.3), linguistic dominance was calculated by means of the Bilingual Language Profile (BLP) score (Birdsong et al., 2012; see Section 3.1.2).

In sum, there are several linguistic and speaker-specific factors that play a role in the realization of voiced stops in Spanish. The following section discusses these linguistic factors, as well as speaker-specific ones, in relation to Yucatecan Spanish.

4.2.3 Voiced stops in Yucatecan Spanish

Yucatecan Spanish reportedly differs from other varieties of Spanish in that voiced stops may be produced as stops intervocalically and voiceless stops seem to be aspirated (Lope Blanch, 1987; Michnowicz & Carpenter, 2013; Suárez Molina, 1945/1996). Furthermore, most studies on voiced stops in Yucatecan Spanish are sociolinguistic in nature. Thus, the focus is on social factors that may help to explain the distribution of stop and approximant realizations, and the preferred source of data is sociolinguistic interviews. The stop/approximant categorization of tokens is usually based on auditory analysis; in fact, none of the studies reviewed here have taken into account duration or intensity measures of voiced stops. Studies have been conducted in the states of Yucatán (García Fajardo, 1984; Lope Blanch, 1987; Michnowicz, 2009, 2011; Rosado Robledo, 2011) and Quintana Roo (Lope Blanch, 1987; Martínez García, 2017; Pérez Aguilar, 2002).[5] These studies indicate that stop realizations are frequent in intervocalic contexts, especially for /b/ and /d/: 65 % and 70 % in García Fajardo (1984); 42 % and 32 % in Michnowicz (2009), and 20 % and 22.6 % in Martínez García (2017), respectively.

The linguistic factors that have been considered in relation to the stop/approximant categorization are mainly position in the word (word-initial vs. word-medial) and lexical stress (stressed vs. unstressed syllables). Michnowicz (2011) reported that the percentage of stop realizations was higher in word-initial than in word-medial position for /b/ (46 % vs. 39 %) and /d/ (38 % vs. 27 %), although results only reached statistical significance for /d/. Martínez García (2017) found a larger number of stop realizations for the three stops in word-initial and utterance-initial position.[6] Nevertheless, instances in utterance-initial position (usually produced as stops) were also included in the analyses for word-initial position, which leaves it unclear whether the same results for phrase-medial, word-initial voiced stops would have been obtained. On the other hand, Rosado Robledo (2011) found no differences in the ratio of stop realizations in word-initial position as compared to those in word-medial position.

[5] The Spanish spoken in both states is assumed to be the same variety (see Section 1.1).

[6] In this work, *word* refers to the orthographic word. In the other works reviewed here, it is not made clear what is meant by it (e.g., the orthographic word, the lexical word).

For lexical stress, Michnowicz (2011) reported that there were more stop realizations for consonants in stressed syllables than in unstressed ones, whereas Martínez García (2017) reported a similar result, but only for the dentialveolar stop, and Rosado Robledo (2011) did not find that effect. Finally, García Fajardo (1984) reported the existence of weakening and elision of /b/ and /d/, especially for the past imperfect suffix forms for /b/ (*-aba, -aban*) and the past participles suffixes for /d/ (*-ado, -ada*), as well as in very frequent words (*todo* 'everything', *todavía* 'still', *cada* 'every', *de* 'of').

Several speaker-specific factors have been studied adopting various sociolinguistic perspectives and methods. A crucial issue that has been the focus of much research is whether Yucatec Maya has influenced Yucatecan Spanish in a way that results in more stop realizations of voiced stops. Some authors have attributed the high ratio of stop realizations to Maya influence (Suárez Molina, 1945/1996), whereas others doubt that this is the case (Alvar, 1969; Lope Blanch, 1987). Nevertheless, it is unclear how this influence (if existent) takes place. For example, Suárez Molina (1945/1996) did not provide any explanation of how this influence could have taken place, especially considering that Yucatec Maya has no voiced stops (see Section 1.1). Michnowicz (2009), who found that bilingual speakers[7] produced more stop realizations that monolinguals, argued that the higher rate of stop realizations was due to language contact, but not to direct influence from Yucatec Maya (pp. 80–82). Michnowicz argued that a higher rate of stop realizations is due to second language learning of Spanish by speakers with a different first language (Yucatec Maya), the output of which is then passed on to the next generation. However, other works have yielded results to the contrary. For example, Rosado Robledo (2011) found no differences between bilinguals (defined as those who speak Maya) and monolinguals. Similarly, Martínez García (2017) found no effect of Yucatec Maya linguistic dominance.

Other speaker-specific factors that have been investigated are gender, age, and sociocultural group. For the effect of gender, Pérez Aguilar (2002) reported that female speakers produced a slightly higher number of stop realizations than male ones, whereas García Fajardo (1984), Rosado Robledo (2011), and Martínez García (2017) did not find such an effect. Regarding age, authors agree that the youngest generation produces fewer stop realizations. García Fajardo (1984) reported that the middle-aged group (25–54 years of age) was the group that produced the most stop realizations, whereas Rosado Robledo (2011) indicated that the eldest generation (over 56 years of age) was the group that

[7] Michnowicz (2009) considered bilinguals those participants who speak Maya regularly and those who can understand it.

produced more stop realizations. Pérez Aguilar (2002) and Michnowicz (2009) reported the higher number of stop realizations as coming from the two older generations (30–49, and over 50 years of age, respectively). For sociocultural groups, García Fajardo (1984), Pérez Aguilar (2002), and Rosado Robledo (2011) reported that the lowest sociocultural group was the one that produced more cases of stop realizations.

In conclusion, voiced stops in Yucatecan Spanish have been studied using different sociolinguistic perspectives and methods that discuss a binary stop-versus-approximant realization; however, they are lacking in phonetic detail. Some of the studies reviewed suggest that intervocalic voiced stops may be produced more frequently as stops in word-initial position and in stressed syllables than in word-medial position and in unstressed syllables, but other studies contradict these claims. The evidence for an influence of social factors such as language knowledge and gender is also mixed. In sum, further research is needed that may shed some light on the linguistic and speaker-specific variation factors that play a role in the realization of voiced stops in Yucatecan Spanish.

4.3 Nonmodal phonation and glottalization of word-initial vowels

In this section, phonation types and their main acoustic characteristics are presented, particularly those of creaky voice (Section 4.3.1), followed by the description of creaky vowels in Yucatec Maya (Section 4.3.2). An overview is then provided of glottalization of word-initial vowels in several languages (Section 4.3.3) and in Spanish and Yucatecan Spanish (Section 4.3.4).

4.3.1 Phonation types

4.3.1.1 The continuum of phonation types

Ladefoged (1971) and Gordon and Ladefoged (2001) provide a widely used classification of *phonation types*, or *modes of phonation*, which are characterized in terms of the aperture between the arytenoid cartilages. The phonation types are placed in a continuum that goes from the most open state of the glottis (voiceless phonation), then to breathy voice, modal voice, creaky voice, and finally to the most closed state of the glottis, glottal closure.

In the most open state of the glottis, *voiceless phonation*, the vocal folds are abducted. Many languages in the world use voiceless sounds, such as for example voiceless stops (Gordon & Ladefoged, 2001, p. 384). In *breathy voice*, the vocal folds are rather abducted. Although the vocal folds are close enough to vibrate, there is a gap between them, which allows the passage of an audible turbulent airflow (Laver, 1994, p. 198). Auditorily, breathy voice has been described as "voice mixed in with breath" (Catford, 1977, p. 99). In *modal voice*, which is thought to be used in all languages, the arytenoid cartilages are completely adducted in the most efficient way to produce voiced sounds (Edmondson & Esling, 2006, p. 160). In *creaky voice*, the vocal folds are tightly adducted, but they are open enough along part of their length so that voicing is possible (Gordon & Ladefoged, 2001, p. 386). Auditorily, the impression may be one of "a rapid series of taps, like a stick being run along a railing" (Catford, 1964, p. 32). Finally, the maintenance of glottal closure, that is, a *glottal stop*, is the other endpoint of the continuum. The vocal folds are held tight and there is no airflow (Laver, 1994, p. 188). Phonological glottal stops may be produced "as creaky phonation on neighboring sounds rather than with complete glottal closure" (Gordon & Ladefoged, 2001, p. 391).

Although this continuum is useful in that it provides a classification of phonation types, it is important to keep in mind that it is a simplification, as Gordon and Ladefoged point out (2001, p. 384). The categories in the continuum are relative to each other, and the difference between, for example, breathy and modal voice in a given language may be different from the difference that another language may establish between them. Furthermore, the aperture of the glottis is not the only factor at the voice source that determines the resulting phonation types; the transfer function, that is, the filtering effect of the vocal tract, also plays a role in the resulting speech signal (Fant, 1960; Gobl & Ní Chasaide, 2010, pp. 378–379). For example, the degree and type of vocal fold tension or the respiratory effort also play a role in the variation at the voice source (Gobl & Ní Chasaide, 2010, p. 379), while the configuration adopted by supraglottal structures of the larynx, such as the ventricular folds, can also influence the resulting phonation types (Edmondson & Esling, 2006; Esling & Edmondson, 2011).

Languages can make use of the phonation types in several ways. In some languages, nonmodal phonation may be used to signal prosodic boundaries (e.g., Dilley et al., 1996, for English, and Di Napoli, 2018, for Italian); this has been referred to in several studies as *glottalization* (see Section 4.3.3). In other languages, phonation types are used to make phonological contrasts. For example, some languages make a contrast between breathy and modal consonants (e.g., nasal consonants in Hindi; Gordon & Ladefoged, 2001, p. 385), whereas others make the contrast between modal and creaky vowels (e.g., Yucatec Maya, Yalálag Zapotec, and Santa María Ocotepec Mixe; Avelino, 2016), or between breathy, modal, and creaky vowels (e.g., Jalapa Mazatec; Blankenship, 2002).[8] Finally, phonation types may be used to characterize the speech of individual speakers, that is, as voice quality. *Voice quality* has been defined as "those characteristics which are present more or less all the time that a person is talking" (Abercrombie, 1967, p. 91). This would include the study of the phonation types breathy, modal, and creaky voice (understood as voice qualities), but also of other voice qualities in which supraglottal articulators are involved (e.g., harsh voice, falsetto, sweet voice). For example, a breathy voice quality in English may be associated with sexiness and intimacy (Esling & Edmondson, 2011, p. 139), whereas sweet voice is used in Japanese to convey femininity (Starr, 2015).

4.3.1.2 Acoustic cues to breathy, modal, and creaky phonation

There are several acoustic cues that characterize breathy, modal, and creaky phonation, and which have been related to changes in the state of the glottis. The acoustic measurements of these cues are usually made directly on the speech signal. Inverse filtering of the speech signal is also possible, but it requires very technical, high-quality acoustic recordings (Gobl & Ní Chasaide, 2010, pp. 380–384; Hanson & Chuang, 1999, pp. 1064–1065). Breathy and creaky voice share some characteristics that set them apart from modal voice, such as a decrease in overall acoustic energy and a lower F0. Furthermore, duration may help to distinguish between modal and nonmodal (breathy and creaky) phonation types, as it seems that, at least in some languages, nonmodal vowels are longer than modal ones (Gordon & Ladefoged, 2001; see also Blankenship, 2002, for Mazatec and Tagalog).

The most important parameter in identifying correlates of voice quality is *spectral tilt* or *spectral slope* (Esling & Edmondson, 2011, p. 132; Hillenbrand et al., 1994). Several measures of spectral tilt have been used to characterize modes of phonation. Thus, the amplitude of the first harmonic, H1, is measured relative to that of harmonics of higher frequencies, such as the second harmonic (H2) or the third harmonic (H3). The amplitude of H1 may also be measured relative to the first-formant spectral peak, that is, the highest harmonic in that formant (A1), the second-formant spectral peak (A2), or the third-formant spectral peak (A3).[9] These measures are usually corrected to remove the filtering function of the vocal tract, which may boost their amplitudes, in order to bring the values closer to those of the actual source spectrum (Hanson & Chuang, 1999, p. 1066). The

[8] See Gordon and Ladefoged (2001) and Ladefoged and Maddieson (1996) for an overview of languages that mark contrasts by means of phonation types.

[9] These are not the only measures possible. Others, such as H2–H4 or H4–H2 kHz, where H2 kHz is the harmonic closest to 2000 Hz, have also been used (see the discussion in Garellek, 2019).

corrected measures are indicated by means of an asterisk (e.g., H1*). Figure 4.1 presents the spectrograms and fast Fourier transform (FFT) spectra of examples of breathy, modal, and creaky voice in the realizations of a Spanish-dominant female speaker of Yucatecan Spanish. The measurements, which were obtained through *PraatSauce* (Kirby, 2018–2019), were made on portions of the vowel /a/ in different positions in the utterance that were conducive to these types of phonation (breathy: utterance-final position; modal: phrase-medial position; creaky: word-initial position). The location of H1 and H2 were added by hand to the images obtained through *PraatSauce*. In line with the description in Gordon and Ladefoged (2001, p. 387), breathy voice is characterized by the presence of turbulent noise in the spectrogram, which makes it difficult to see the "vertical striations" (p. 387) that reflect individual pitch pulses, whereas in creaky voice the distance between the pitch pulses is further apart and more irregular than in modal voice.[10]

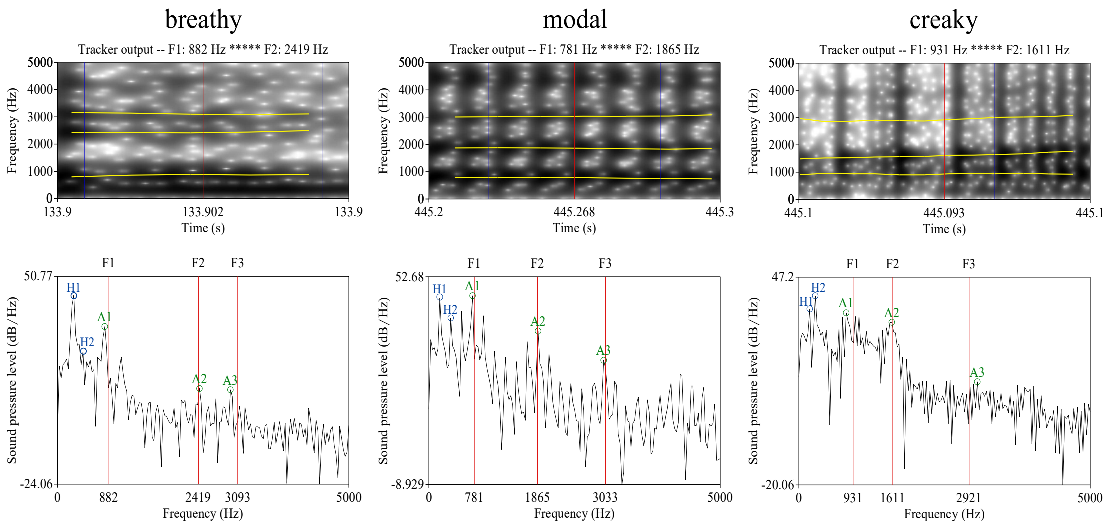

Figure 4.1. Spectrograms (top) and FFT spectra (bottom) of breathy, modal, and creaky voice realizations of portions of the vowel /a/ produced by speaker inf09 (female, Spanish dominant). Measurements were obtained over a 25 ms window whose limits are indicated by the blue bars in the spectrograms. Values for harmonics (H1, H2), formant spectral peaks (A1, A2, A3), and the location of formants (F1, F2, F3) are shown in the spectra.

These parameters of spectral tilt, which are believed to be acoustic correlates of articulatory parameters of the voice source, help to differentiate between phonation types. The acoustic measure H1–H2 is related to the *open quotient* (OQ), that is, the proportion of the glottal cycle during which the glottis is open (Gobl & Ní Chasaide, 2010, p. 391; Hanson & Chuang, 1999, p. 1066; Holmberg et al., 1995). In breathy voice, the OQ is greater than in

[10] Not all types of creaky voice have this kind of visualization in the spectrogram. For example, aperiodic voice does not. See Section 4.3.1.3.

modal voice, which is reflected in greater values for H1 than for H2, whereas for creaky voice, in which the OQ is relatively small, the value of H2 is higher than for H1; the values for modal voice are between those of breathy and creaky voice (Gobl & Ní Chasaide, 2010, pp. 393–394; Gordon & Ladefoged, 2001, pp. 397–399). Measures that compare H1 to the amplitude of the highest harmonic in a formant (e.g., A1, A3) seem to be related to the closing velocity of the vocal folds (Keating & Esposito, 2007). Stevens (1977; in Gordon & Ladefoged, 2001, p. 399) proposed that the less abrupt the closure, the lesser the difference in amplitude between H1 and higher harmonics would be. A less abrupt closure is a correlate of breathiness, whereas creaky voice is characterized by a more abrupt closure (Gordon & Ladefoged, 2001, p. 399).

Both breathy and creaky voice present more spectral noise compared to modal voice, but the origin of the noise differs. In breathy voice, the noise (which appears especially at higher frequencies) is due to the leakage of air through the glottis (Laver, 1980, pp. 120–122), whereas in creaky voice, it is due to the signal being quite irregular (Gobl & Ní Chasaide, 2010, p. 401; Gordon & Ladefoged, 2001, p. 397). This irregular vibration of the vocal folds during creaky voice can be perceived as noise (Keating et al., 2015). Two measures of periodicity/noise based on cepstra are commonly used: harmonics-to-noise ratio (HNR) and cepstral peak prominence (CPP). *Harmonics-to-noise ratio* is a measure of the amount of periodic energy relative to noise energy in the signal (de Krom, 1993; Esling & Edmondson, 2011, p. 142). *Cepstral peak prominence* measures cepstral peak amplitude relative to the amplitude of the cepstral signal (Hillenbrand et al., 1994, p. 772). To calculate CPP, a linear regression line that relates quefrency to cepstral magnitude is fitted first, thus normalizing the cepstral peak for overall amplitude. CPP is then calculated as the relative prominence of the cepstral peak compared to the value on the linear regression line that is immediately below the peak. A highly periodic signal has a clearer cepstral peak because of its clear harmonic structure, whereas a noisy signal does not (Hillenbrand et al., 1994, p. 772). Because of the greater presence of noise, both breathy and creaky voice are expected to show lower values of HNR and CPP than modal voice. Table 4.2 shows the expected direction of the values of the parameters of spectral tilt for breathy and creaky voice when compared to modal voice, as well as the values for F0.

Table 4.2. Expected direction of the values of several acoustic measures for phonation types. Breathy and creaky voice are compared to modal voice.

Acoustic measure	**breathy**	**modal**	**creaky**
F0	<	·	<
Spectral tilt			
H1–H2	>	·	<
H1–A1, H1–A3	>	·	<
Noise			
HNR	<	·	<
CPP	<	·	<

4.3.1.3 Types of creaky voice

In the definition by Gordon and Ladefoged (2001; see Section 4.3.1.1), there is voicing during creaky voice, whereas in Keating et al. (2015) this is not necessarily so. Keating et al. (2015) classify creaky voice in six categories (see Table 4.3). *Prototypical creaky voice* has low and irregular F0, with low values of H1–H2, which indicate a constricted glottis. *Vocal fry* has low but regular F0, with low values of H1–H2. The high damping of the pulses makes it possible to hear them separately. *Multiply pulsed voice* has alternating longer and shorter pulses, in pairs (double pulsing) or in higher multiples, with low values of H1–H2. *Aperiodic voice* is defined by a vibration of the vocal folds so irregular that there is no periodicity. *Nonconstricted creak* has low and irregular F0, but the glottis is not constricted, so values of H1–H2 are not low. In *tense/pressed voice*, F0 is not low and it is not irregular, with low values of H1–H2. Tense/pressed voice can function phonologically as creaky voice in languages where high tone and creaky phonation can co-occur.

Table 4.3. Characteristics of the types of creaky voice. The table is a simplified adaptation of the one in Keating et al. (2015, p. 3). Empty cells indicate that the acoustic correlate is either variable or unknown.

Property	**low F0**	**irregular F0**	**glottal constriction**
Main acoustic correlate	**low F0**	**high noise**	**low H1–H2**
Type			
prototypical	✓	✓	✓
vocal fry	✓		✓
multiply pulsed		✓	✓
aperiodic	NO	✓	✓
nonconstricted	✓	✓	NO
tense	NO		✓

The basis for this classification can be considered both phonological and perceptual. Thus, aperiodic creaky voice (with no voicing) and nonconstricted creaky voice (without lowering of F0) are still classified as creaky because they can be the phonetic instantiation of sounds that are creaky in phonological terms, and which may also be perceived as creaky. Garellek (2019) discusses this issue as follows:

> Given that creaky voice represents a cluster of vocal fold articulations and voice qualities, one might be inclined to ask whether phoneticians should retain the more general term 'creaky' at all. I believe we should, because it is useful to have a word for the abstract phonological category, which may be realized (in different phonological environments and/or by different speakers) using different articulations. (p. 81)

In fact, there are other (nonphonological) classifications of creaky voice that have also taken perception into account. For example, one category in the study of American English by Redi and Shattuck-Hufnagel (2001) is what they defined as *glottal squeak,* that is, "a sudden shift to relatively high sustained f_0, which was usually very low amplitude" (p. 414).

4.3.2 Creaky vowels in Yucatec Maya

Yucatec Maya makes a phonemic contrast between long-high vowels /íi ée óo úu/ and long-high creaky vowels /íḭ éḛ óo̰ úṵ/ (Avelino, 2011; Avelino et al., 2011; Frazier, 2009; see also Section 1.1). In Yucatec Maya, creaky vowels are also referred to as *glottalized, laryngealized*, or *rearticulated*. The term *rearticulated,* which is commonly used in grammatical studies (Avelino et al., 2011, p. 2), means that vowels are transcribed as [vʔv] because they have traditionally been considered to be a long modal vowel with a glottal stop in the middle (e.g., Briceño Chel & Can Tec, 2017; Pike, 1946).

When nonmodal phonation is used to mark phonological contrasts in vowels, it may be limited to only part of the vowel (Gobl & Ní Chasaide, 2010, p. 405; Gordon & Ladefoged, 2001, p. 393). Furthermore, Silverman (1997) discusses several languages that present contrastive phonation (breathy, modal, and creaky vowels) as well as tone (Otomanguean, Jalapa Mazatec, Comaltepec Chinantec, and Copala Trique). Silverman argues that "tone and non-modal phonation may be sequenced with respect to each other, so that tone may be realized with modal voice" (p. 236).[11] This is also the case in Yucatec Maya, where the high tone occurs in the first part of the vowel and is then followed by creaky voice (Frazier, 2009, 2013). Both Frazier (2009, 2013) and Avelino et al. (2011) describe three main phonetic realizations for creaky vowels in Yucatec Maya: [vv̰v], [vv̰], and [vʔv]. The first

[11] There are languages that do not show this pattern, such as Mpi (Silverman, 1997).

one is the most common in the studies by Frazier (2009, 2013), who analyzed words uttered in isolation by 24 speakers, and words in carrier sentences that were read by another 27 speakers, all of them from Yucatán. The second one, [vv̰], is the most frequent, according to Avelino et al. (2011), who analyzed a read speech corpus of verb forms produced by 8 speakers (3 female, 5 male) from Yucatán. The production with a glottal stop ([vʔv]) appears only sporadically (Avelino et al., 2011; Frazier, 2009, 2013). Frazier (2009) also reports that a small number of tokens were produced with creaky voice throughout, or with a dip in amplitude that was perceived as creaky voice.

Measures of spectral tilt show differences between modal and creaky vowels in Yucatec Maya. The values of H1–H2 seem to consistently distinguish between the two vowel types (Avelino et al., 2011). Thus, the first portion of creaky vowels was produced with modal voice (indicated by positive values of H1–H2), followed by creaky voice (indicated by negative values of H1–H2). This finding was replicated by Avelino (2016), who used corrected measures of the harmonics (H1*–H2*). Moreover, Avelino (2016) found that H1*–A3* was a better index of differences between modal and creaky vowels than H1*–H2*, with H1*–A1* not yielding conclusive results.

4.3.3 Glottalization of word-initial vowels

There is crosslinguistic evidence that suggests that a creaky realization of vowels in word-initial position may be used to mark prosodic boundaries. This phenomenon is usually referred to as *glottalization* (see the discussion in Garellek, 2013, p. 5, for other definitions). In this dissertation, *glottalization* refers to glottal stops and perceptually equivalent realizations (produced with creaky voice), which may signal prosodic boundaries. Moreover, Dilley et al. (1996) propose that "glottalization of word-initial vowels at prosodically significant locations may represent a strengthening of the articulatory gesture associated with the onset of the prosodic constituent or prominence" (p. 438). Thus, glottalization of word-initial vowels could be considered an acoustic cue to prosodic strengthening, which may help to cue prosodic boundaries.

Glottalization of word-initial vowels has been investigated by means of acoustic studies for several languages, including American English, Mexican Spanish, Standard German, and Polish. Overall, these studies suggest that glottalization may cue phrase-level boundaries, but it seems to be associated even more frequently with prominent positions. For American English, Pierrehumbert and Talkin (1992) suggested that glottalization was favored by the presence of an Intonational Phrase (IP) boundary, whereas Pierrehumbert

(1995) showed that glottalization was favored by pitch accents in word-initial position. In line with these studies, Dilley et al. (1996) analyzed a small corpus of radio news stories and concluded that prosodic position and the presence of a pitch accent explained the presence of glottalization to a great extent. Thus, word-initial vowels were more frequently glottalized after a pause or at IP boundaries and at intermediate phrase (ip) boundaries than in phrase-medial position, and also more frequently in pitch accented syllables than in unaccented ones. Whether the preceding segment was a vowel or a consonant did not have an effect on glottalization. Conversely, the study by Garellek (2014) showed no effect of prosodic position. Garellek (2014) conducted an electroglottographic (EGG) study for American English and Mexican Spanish in order to examine glottal constriction as a cue to prosodic strengthening in vowels. The participants in the Spanish study were all English–Spanish bilinguals who came from either Mexico City or from the Los Angeles area. Additionally, Garellek took H1*–H2* measures of the vowels. The results of the acoustic measure were in agreement with those of the EGG study. Token vowels were measured in Utterance-initial, IP-initial, ip-initial, and ip-medial positions. Half of them were in unstressed syllables, while the other half were in stressed ones; none bore a pitch accent. Overall, the results showed the same pattern for English and Spanish: the degree of glottal constriction was greater for vowels in stressed syllables, particularly in Utterance-initial and IP-initial positions, but prosodic position did not play a role.

In Standard German, glottalization appears to be more frequent (i) the lower the speech rate is, (ii) in low than in nonlow vowels, (iii) and for vowels in stressed syllables (Malisz et al., 2013). Unlike in some of the studies on American English, for which the presence of IP and ip boundaries increased the frequency of glottalization, glottalization in German seems to occur equally frequently in phrase-initial and phrase-medial positions. Additionally, word class does not seem to play a role (Malisz et al., 2013). Results for Polish mirror those of German, with two exceptions: glottalization in Polish is more frequent phrase initially than phrase medially, and also more frequent for tokens in lexical words in phrase-initial position than in phrase-medial position (Malisz et al., 2013).

In sum, these studies suggest that glottalization of word-initial vowels may be an acoustic cue to prosodic strengthening that may mark prosodic boundaries at the phrase level in American English and Polish, although probably not in Mexican Spanish and German. Moreover, for the four languages, glottalization seems to be associated more frequently with prominent positions, whether at the word or at the phrase level.

4.3.4 Glottalization of word-initial vowels in Spanish and Yucatecan Spanish

The glottalization of word-initial vowels in Spanish is not a pan-Hispanic phenomenon, being described for only a handful of varieties. It has been characterized as a word-related feature, and primarily as the insertion of a glottal stop before the vowel. However, it is unclear what *word* may refer to in these studies, although it is reasonable to assume that they refer to the orthographic word. Overall, prosodic constituency and prosodic strengthening are not taken into account, either at the Prosodic Word level or at higher levels (an exception to this is Garellek, 2014, reviewed in Section 4.3.3 above; see also Michnowicz & Kagan, 2016, discussed below).

The studies that discuss glottalization in Spanish usually provide a general description of glottal stop insertion before vowels. For several Spanish varieties, the insertion of a glottal stop has been treated as the product of a contact situation with languages that have a phonological glottal stop.[12] This is the case for Yucatecan Spanish (Lope Blanch, 1987; Michnowicz & Kagan, 2016), Guatemalan Spanish, which is in contact with Kaqchikel Maya (McKinnon, 2018), Paraguayan Spanish, in contact with Guaraní (de Granda, 1982, p. 158; Trawick & Michnowicz, 2019), and Philippine Spanish, which is in contact with several Philippine languages (Lipski, 1987, pp. 215–216; Lipski does not specify which Philippine languages). Nevertheless, some Spanish varieties for which glottalization is reported are not in contact with other languages: this is the case for Nicaraguan Spanish (Chappell, 2013) and Mexico City Spanish, for which Avelino (2018) points out that a glottal stop may be "inserted as onset to onsetless syllables" (p. 224). Only a handful of these studies provide information about the data on which they have based their conclusions (Chappell, 2013; Mckinnon, 2018; Michnowicz & Kagan, 2016; Trawick & Michnowicz, 2019). The vowels are usually labeled for the presence of a glottal stop based on auditory analysis, sometimes accompanied by the visual inspection of waveforms and spectrograms, which reveal that the perceived glottal stops can also be produced as creaky voice. These studies do not provide measures of acoustic parameters.

The studies on consonantal and vocalic resyllabification that were reviewed in Section 2.2.1.3 suggest that many factors have an effect on lack of resyllabification in Spanish. One of them is the glottalization of word-initial vowels. Glottalization as a cue to

[12] A glottal stop as a consonantal allophone has been reported in several Spanish varieties. In Nicaraguan Spanish, it is a word-final allophone of /s/ when it is followed by a vowel (Chappell, 2013). In the speech of Afro-Costa Ricans in the province of Limón, it is an allophone of /x/ (Zimmer, 2011, pp. 193–195). On the Pacific coast of Colombia, it is an allophone of /k/ (Correa, 2012, pp. 49–50).

lack of resyllabification is briefly mentioned in two studies on European Peninsular Spanish (Aguilar, 2003; Strycharczuk & Kohlberger, 2016), although it is unclear to what degree speakers of this variety use glottalization. Aguilar (2003) mentioned creaky voice as one strategy for maintaining the hiatus of vowel sequences across words (and thus there being no resyllabification; p. 2114), but provided no further information about how frequently used creaky voice was.[13] In their study of consonantal resyllabification, Strycharczuk and Kohlberger (2016) mentioned that there was a small number of tokens with lack of resyllabification, in which speakers introduced a pause, which was followed by glottalization of the following vowel (p. 6).

The literature on Yucatecan Spanish has treated glottalization as one of the most characteristic features of the variety. Furthermore, in these studies glottalization of both vowels and consonants is considered. Because of the lack of phonetic information, it is unclear what this glottalization may actually be. Suárez Molina (1945/1996, p. 64) indicates that vowels may be glottalized in word-medial position in emphatic pronunciation; Barrera Vásquez (1946/1977, p. 342) and Suárez Molina (1945/1996, pp. 63–64) argue that vowels may also be glottalized word finally. In the case of consonants, it is unclear if they are glottalized consonants or ejectives (see Lope Blanch, 1987; Martín Butragueño, 2014; Suárez Molina, 1945/1996).

Michnowicz and Kagan (2016) is the first and only study to date on the glottalization of word-initial vowels in Yucatecan Spanish in a similar sense to how glottalization is understood in the present dissertation. In their study, they analyzed acoustic recordings of sociolinguistic interviews conducted with 18 participants (9 female, 9 male) from Mérida city, Yucatán, and the surrounding areas. Some of the participants were monolingual, while others were Yucatec Maya–Yucatecan Spanish bilinguals. Michnowicz and Kagan (2016) grouped speakers into monolinguals and bilinguals by asking the participants about their own knowledge of Yucatec Maya and Yucatecan Spanish, the languages spoken by their relatives, and some other linguistic indicators (p. 223). Tokens were grouped into those that presented a full glottal stop, creaky voice, or "non-insertion" of a glottal stop (p. 224). The authors classified the first 100 instances per speaker of (phrase-medial) vowel-initial words by inspecting the corresponding waveforms and spectrograms. The results showed that 11 % of all cases were glottalized, with 4.5 % produced as glottal stops and 6.5 % as having creaky

[13] The strategies for maintaining a hiatus across words mentioned by Aguilar (2003) are: "separation by means of a tonal movement, insertion of a pause, lengthening of the word-final vowel, appearance of creaky voice between the two vowels forming the group" (p. 2114).

voice. Glottalization was more frequent (i) for vowels in stressed syllables, (ii) in syllables preceded by longer words, and (iii) for the vowel /o/. Whether the preceding sound was a vowel or a consonant did not play a role. Moreover, glottalization was more frequent among bilingual speakers (recall that, in Yucatec Maya, glottal stops can be the first or the last sound of a word; see Section 2.3). Differences between female and male speakers did not reach statistical significance, but older speakers produced more glottalized tokens.[14]

Michnowicz and Kagan (2016) argue that the insertion of a glottal stop blocks resyllabification: the glottal stop "then occupies the onset of the word-initial syllable" (p. 221). This claim for Yucatecan Spanish, which is not new (see, e.g., Lipski, 1996, p. 302; Lope Blanch, 1987, p. 35), must not be understood as resyllabification across Prosodic Words, but across orthographic words (cf. Martín Butragueño, 2014, p. 251; see Section 1.1). Because of the frequent insertion of glottal stops to block resyllabification, Yucatecan Spanish differs from the descriptions usually given for Spanish in general, a language in which glottalization is infrequent (e.g., Bissiri et al., 2011) and resyllabification is widespread (see Section 2.2.1.3). Michnowicz and Kagan (2016) also argue that the insertion of glottal stops may help to mark word boundaries, thus facilitating the recognition of words (p. 235).

In sum, although the glottalization of word-initial vowels has been attested for Yucatecan Spanish, there are several points in need of further research, such as (i) what its phonetic nature is, (ii) how it relates to prosodic structure, and (iii) whether knowledge of Yucatec Maya and gender play a role.

4.4 Summary

Chapter 4 has presented a literature review of prosodic strengthening of voiced stops and glottalization of vowels, necessary in order to address the research questions introduced in Section 1.1.

In Section 4.2, an overview of the theoretical framework that groups realizations into stops and approximants in Spanish was provided. This was followed by a description of articulatory and acoustic studies on several Spanish varieties, which showed that the phonetic variation of voiced stops is better understood as a continuum of realizations. A

[14] Incidentally, Colazo-Simón (2007), in a study of the ejectives [p' t' k'] and the voiceless plosives [p t k] of Yucatec Maya, suggests that glottalization across words that end and begin with a vowel is pervasive in Yucatecan Spanish. Colazo-Simón also makes the observation that glottalization seems to be favored in sequences of an unstressed vowel followed by a stressed vowel of different quality, such as for example between *cuatro* and *años* in *cuatro años* 'four years' (p. 151).

description of the sources of variation in the acoustic realization of voiced stops, both linguistic and speaker-specific, showed that there is some evidence for prosodic position and lexical stress having an effect on strengthening, especially in word-initial position and in stressed syllables, although there are differences between read and spontaneous speech. Other factors may have an effect as well, but they have been examined in few studies. The survey of studies on voiced stops in Yucatecan Spanish has shown that they have focused on social factors to explain variation, and also that they lack in phonetic detail. Furthermore, the effect of prosodic position and lexical stress on strengthening may be similar to that of the other Spanish varieties examined, while the effect of knowledge of Yucatec Maya is not clear.

In Section 4.3, breathy, modal, and creaky voice have been introduced as categories of a continuum of phonation types, and their acoustic characteristics and a classification of creaky voice types has been provided (Section 4.3.1). We have seen that nonmodal phonation is used in Yucatec Maya to make phonological contrasts between creaky and modal vowels (Section 4.3.2), followed by how the glottalization of word-initial vowels may be a manifestation of prosodic strengthening that signals prosodic boundaries in several languages (Section 4.3.3). Finally, glottalization of word-initial vowels in Spanish and Yucatecan Spanish has been discussed (Section 4.3.4).

The research questions introduced in Section 1.1 will be addressed by means of two corpus speech studies (Part II) and two read speech studies (Part III). Additionally, Chapter 10 in Part IV will examine the role of pitch accent in relation to lexical stress.

Part II

Corpus speech studies

Chapter 5

Voiced stops — corpus speech study

5.1 The present study

The present study aims to explore the prosodic strengthening of Yucatecan Spanish voiced stops in the Prosodic Word (PW) domain in a corpus of sociolinguistic interviews. The studies reviewed in Section 4.2 showed that the realizations of voiced stops in Spanish can be placed in a continuum that ranges from full stop realizations to vowel-like approximant realizations, that is, from strengthened to weakened ones. In Yucatecan Spanish, strengthened realizations in intervocalic context are rather frequent according to existing studies on the subject, which categorize realizations into stops and approximants based primarily on auditory analyses. However, an analysis of the acoustic parameters of strengthened realizations is still lacking. Based on previous research on Spanish voiced stops, two of the acoustic cues to prosodic strengthening are longer acoustic duration and, at the far end of the continuum, the presence of a release burst (see Section 4.2.2.2); these two acoustic parameters are examined in this chapter.[1] The presence of a release burst will be used to draw a distinction between two categories, stops (with release burst) and approximants. Although it is true that a binary stop/approximant categorization may not be entirely appropriate because it does not capture the continuum of voiced stop realizations, it does reflect which realizations are produced with a full oral closure (i.e., the ones that can be categorized as *stops*). Moreover, the category analysis can complement the duration analysis. This is so because the realizations of Spanish approximants may be so vowel-like (i.e., weakened) in terms of acoustic duration that they cannot reliably be measured (Martínez Celdrán, 2013). Although an analysis of duration for these tokens would not be possible, they could still be perceived and categorized as *approximants*.

The present study examines the realizations of the bilabial and dentialveolar voiced stops in syllabic onset position in CV and CVC syllables preceded by a vowel. The velar voiced stop is excluded on the basis of its low frequency in Spanish (Rojo, 1990) and also based on the conclusions of a previous study (Martínez García, 2017; see also Section 4.2),

[1] Measurements of intensity could not be made due to low recording amplitude.

in which the number of velar tokens was insufficient to fit statistical models.[2]

Because previous studies on Yucatecan Spanish suggest that strengthened realizations of voiced stops may appear more frequently word initially and in stressed syllables (see Section 4.2.3), I aim to investigate whether prosodic strengthening is more frequent in Prosodic Word (PW) initial position and in lexically stressed syllables. Three positions in the PW domain are considered: lexical-word (LW) medial (e.g., *empeza**b**a* 'it started'), LW-initial (e.g., (*de* (***d**escanso*)$_{LW}$)$_{PW}$ 'of relaxation'), and PW-initial (e.g., (***d**e descanso*)$_{PW}$). I expect that there will be more strengthening of tokens in PW-initial position than in LW-medial position because strengthening of tokens in PW-initial position may serve to mark the prosodic boundary corresponding to the PW. Moreover, based on the studies that suggest that strengthening is more frequent word initially, I expect that tokens at LW-initial position will be more strengthened than those in LW-medial position. In terms of lexical stress, I expect to find more strengthening in tokens that appear in stressed syllables than in unstressed ones. I also aim to examine whether there is an interaction between prosodic position (in the PW domain) and lexical stress. Tentatively, I hypothesize that there will be more strengthening of tokens in stressed than in unstressed syllables in PW-initial > LW-initial > LW-medial positions. Finally, because PWs may start with a FW (e.g., (***d**e descanso*)$_{PW}$) or a LW (e.g., (***d**escanso*)$_{PW}$), I also seek to determine whether there is a difference in strengthening as a function of word class (only for the dentialveolar voiced stop, and excluding stressed FWs; see Section 5.2.5.3). Tentatively, I expect the strengthening of tokens in (unstressed) FWs, which in this study always appear in PW-initial position, to be similar to that of LWs in PW-initial position due to being in PW-initial position.

Language dominance will be assessed by means of the Bilingual Language Profile (BLP; see Section 3.1.2). Studies of speaker-specific variation for voiced stops in Yucatecan Spanish have investigated the realizations of Spanish monolinguals and Maya–Spanish bilinguals, with some indicating that bilinguals produce more stop realizations than monolinguals, while others did not find such a difference (see Section 4.2.3). If bilingual speakers do indeed produce more strengthened realizations, such a result may be amenable to an explanation in terms of assimilation as it is conceived of in the Speech Learning Model

[2] Martínez García (2017) examined the recordings of 12 speakers, 8 of whom are also considered in the present study. Utterance-initial stops were included and only presence/absence of a release burst was examined. The approximate percentage of recordings used in Martínez García (2017) that are also used in the data set examined in this chapter is 100 % for speakers inf15, inf21, and inf14; 80 % for inf13; 70 % for inf22; and 50 % for inf18, inf20, and inf07.

(SLM; Flege, 1995, 1999, 2002; see Section 3.1.1). Yucatec Maya, unlike (Yucatecan) Spanish, lacks voiced stops, but nonetheless has a rather rich inventory of stops (see Section 1.1). Bilingual speakers whose first language (L1) is Yucatec Maya would assimilate the second-language (L2) voiced stops, that is, they would interpret the phonetic categories of Spanish in terms of Yucatec Maya. The bilinguals would then create a merged L1–L2 category (cf. McKinnon, 2020, discussed in Section 3.1.1, for bilingual Kaqchikel–Spanish speakers). Furthermore, whereas most studies of Yucatecan Spanish group speakers into bilinguals and monolinguals, by using the BLP I will obtain a continuous measure that will reflect speaker-specific variation in language dominance.

For gender, speakers will be grouped as female or male on the basis of their self-reporting on belonging to one of the two groups. Because most of the studies did not find an effect of gender, I seek to confirm that there is no evidence in support of this effect. Table 5.1 presents a summary of the expected direction of the results.[3]

Table 5.1. Expected results for strengthening of voiced stops in the corpus speech study.

Variables	Expected results
1. PROSODIC POSITION	PW-initial > LW-initial > LW-medial
2. LEXICAL STRESS	stressed > unstressed syllables
3. PROSODIC POSITION × LEXICAL STRESS	PW-initial stressed > unstressed > LW-initial stressed > unstressed > LW-medial stressed > unstressed
4. LANGUAGE DOMINANCE	Yucatec Maya > Yucatecan Spanish
5. GENDER	no effect
6. WORD CLASS (/d/)	LWs = FWs

5.2 Methods

5.2.1 Sociolinguistic interviews

The speech materials that were used in this study are part of a corpus of sociolinguistic interviews about culture and language. The interviews were conducted by a native speaker of Northwestern European Spanish (the author of the dissertation) and an L2 speaker of

[3] This table has its counterpart in Table 5.3 in Section 5.4, where the summary of results will be presented. Similar tables with *expected results* (e.g., Table 5.1) and with *results* (e.g., Table 5.3) will be included in the following chapters as well.

Spanish.[4] The presence of a non-Yucatecan Spanish interlocutor was necessary because of the characteristics of the sociolinguistic interview: questions such as 'What is your city like?', 'What kind of music is typical of this area?', and the like could only be interpreted as meaningful by the participants if the interviewer was not from the area and not an acquaintance. For the present study, at first only answers given to questions about climate and food were used as speech materials (e.g., 'What is the climate like in Felipe Carrillo Puerto?', 'What are typical dishes from this area?', 'What do you like to cook?'). However, answers to other topics were included as well for some speakers in order to obtain the number of tokens required by this study (see Section 5.2.2).

5.2.2 Speech materials

The recorded sociolinguistic interviews were first transcribed by a native speaker of Yucatecan Spanish and then further checked and corrected (mostly for orthography) where necessary. Text files were created for each sound file. The sound files (wav) and the text files (txt) were then automatically aligned using the BAS Pipeline online service (Kisler et al., 2017). The pipeline G2P (grapheme-to-phoneme) → MAUS → PHO2SYLL (phoneme-to-syllable) was used. As the G2P conversion of BAS Pipeline is based on standard European Spanish, the G2P mapping table was slightly adapted to better suit Yucatecan Spanish (e.g., by excluding the dental fricative consonant, which does not exist in Yucatecan Spanish). The G2P conversion was fed into MAUS, which estimates the most likely pronunciation of the phonemes created in the previous step. This was in turn fed into the PHO2SYLL conversion. The final files were Praat TextGrid files (Boersma & Weenink, 2019) with several tiers that contained the orthographic words, the phonemic annotation, and the syllabic annotation. Misalignments of whole utterances were corrected manually in Praat (Boersma & Weenink, 2019). The resulting tiers were used as guidance to find the voiced stops, which were then segmented manually.

The initial goal was to annotate 90 tokens per speaker of the bilabial and dentialveolar stops pooled together.[5] The tokens considered for the study appeared in the onset position of CV and CVC syllables that were preceded by a vowel. In this way, consonants in utterance-initial position were excluded. Prosodic domains above the PW were not

[4] These recordings were conducted together with Melanie Uth as part of the project *A05: Prominence marking and language contact in Spanish, CRC 1252 Prominence in Language*. The total number of participants recorded was 41.

[5] The total number of tokens for one speaker, inf33, was 86 because four observations had to be discarded after the labeling of tokens.

considered. Voiced stops in the final syllable of an utterance were excluded from the analysis in order to avoid final lengthening effects. Instances at the end of an utterance were those followed by a pause longer than 250 ms. Additionally, instances that appeared in stretches of speech with laryngealized voice quality or with background noise were also excluded. Table A1 in Appendix A presents information about the number of tokens per voiced stop and per speaker. The final number of observations considered in this study differed per speaker (range = 86–233, mean = 115) and per stop (bilabial = 867, dentialveolar = 1497).

The bilabial stop can be represented in Spanish orthography by both <b> and v> (cf. Section 4.2.2.1). In the present study, tokens whose orthography was either <b> or <v> were included. For the dentialveolar stop (also marginally for the bilabial stop), some tokens were excluded from the analysis, specifically instances of auditorily perceived elision (e.g., *pue* instead of *puede* 'he/she can', *to* instead of *todo* 'all', past participles ending in *-ado* > *ao*). The deletion of the dentialveolar sound may be explained not in terms of weakening, but of lexical frequency effects and sociolinguistic meaning, to the point that reduced forms of words with the dentialveolar may have become conventionalized (Hualde, Simonet, & Nadeu, 2011; see Sections 4.2.2.3 and 4.2.3 for further discussion). Tables A2 (for the bilabial stop) and A3 (for the dentialveolar stop) in Appendix A show the elided tokens.

5.2.3 Speakers

The interviews with 20 speakers (10 female, 10 male) were selected based on several criteria.[6] First, younger speakers, that is, those under 30 years of age, were excluded because they did not engage as much in the conversation as older speakers, thus providing fewer speech materials. This circumstance was evident at the time of the recordings. Second, there was an equal number of female and male participants. Third, a selection of speakers who presented a range of language dominance scores, from Maya-dominant to Spanish-dominant, was made (see below). Finally, the low quality of some recordings led to the exclusion of some participants' recordings.

All participants gave their informed consent, completed a personal information questionnaire, and filled in the BLP questionnaire (Birdsong et al., 2012). The questionnaire comprises questions about personal language history, use, proficiency, and attitudes, which

[6] Although the optimal choice would have been to include the maximum number of speakers in the study, a selection was necessary because of time constraints.

are then used to obtain a score (the BLP score) that quantifies language dominance (see Section 3.1.2 for discussion). The two ends of the continuum correspond to the highest dominance in the two languages. In the present dissertation, the negative end (–218) corresponds to the highest dominance in Yucatec Maya, whereas the positive end (+218) corresponds to the highest dominance in Yucatecan Spanish. A score of 0 indicates balanced bilingualism (Amengual & Simonet, 2019). Table 5.2 provides the BLP score of each participant, as well as other personal information. Scores closer to the highest Mayan dominance (e.g., –210, –150) would correspond to speakers who spoke little or no Spanish, which explains the seemingly unbalanced distribution of the BLP scores in favor of the highest dominance in Spanish. The resulting BLP scores, which can have up to three decimals, were rounded off to whole numbers before any statistical analyses were performed in this and the other studies in the dissertation.

Table 5.2. Participants' Bilingual Language Profile (BLP) scores, speaker ID, gender, age, state of origin, and highest level of formal education: corpus speech study on voiced stops.

BLP score	Speaker ID	Gender	Age	State of origin	Highest level of formal education
-61	inf34	male	68	Quintana Roo	primary education
-46	inf17	female	30	Quintana Roo	high school
-26	inf15	female	43	Quintana Roo	university
-25	inf22	male	38	Yucatán	high school
-24	inf19	female	34	Quintana Roo	university
-24	inf18	female	38	Quintana Roo	university
-20	inf20	male	47	Quintana Roo	high school
3	inf21	male	40	Quintana Roo	university
10	inf14	male	46	Quintana Roo	high school
15	inf16	male	57	Quintana Roo	high school
23	inf27	female	58	Yucatán	high school
96	inf08	male	72	Yucatán	secondary education
99	inf23	female	37	Yucatán	high school
111	inf33	male	30	Quintana Roo	university
160	inf29	male	66	Yucatán	secondary education
177	inf13	female	36	Quintana Roo	university
179	inf12	male	34	Yucatán	university
180	inf31	female	34	Yucatán	university
195	inf11	female	67	Yucatán	high school
202	inf07	female	54	Quintana Roo	university

The Spanish version of the Spanish–English BLP questionnaire, which is available on the website of the BLP,[7] was modified to suit the studies in the present dissertation. Specifically, the questions about English were modified to refer to Yucatec Maya, and the Spanish names of the levels of education were in some instances changed to the Mexican Spanish names.

The personal information questionnaire and the BLP questionnaire included questions about gender, age, place of origin, highest level of formal education, and other languages spoken. All participants, who at the time resided in Felipe Carrillo Puerto, were either from the states of Quintana Roo or Yucatán. At the time of the recordings, all speakers had lived in Felipe Carrillo Puerto for at least 6 years, and in Quintana Roo for at least 16 years. Only one participant (inf21) reported speaking a language other than Spanish or Maya (English). Speakers were paid for their participation.

5.2.4 Recordings

Acoustic recordings were carried out in a quiet room in the city of Felipe Carrillo Puerto using an AKG C 544 L head-mounted microphone connected to a Presonus Audiobox USB. The speech signal was digitized at a sampling rate of 44,100 Hz and a 16-bit resolution. The quality of the recordings was fair overall and was satisfactory enough for the purposes of the present study.

5.2.5 Acoustic analysis

This section presents the criteria for the annotation of duration (Section 5.2.5.1), the labeling of category (i.e., stop vs. approximant; Section 5.2.5.2), and the annotation of prosodic factors (Section 5.2.5.3).

5.2.5.1 Criteria for annotating duration

The duration of the oral constriction of the voiced stop realizations was manually annotated in Praat (Boersma & Weenink, 2019) following Turk et al. (2006). Thus, the onset and release of the oral constriction were manually located by means of visual inspection of the spectrogram and subsequent inspection of the waveform for a fine-grained segmentation. The onset is indicated by "a decrease in overall amplitude" and "cessation of all but the lowest formant and harmonic energy" (Turk et al., 2006, p. 6), whereas the release is associated with a release burst or, if the burst is not evident, the location of the release is set

[7] https://sites.la.utexas.edu/bilingual/using-the-blp/access-testing-materials/

close to the F2 onset (Turk et al., 2006, p. 8). The precise locations for the onset and the release were determined by means of the *move cursor to nearest zero crossing* option in Praat (Boersma & Weenink, 2019).

Turk et al. (2006) provide guidelines on how to segment voiced stops, but not on how to segment voiced approximants such as the ones present in Spanish and in this dataset. Nevertheless, they provide suggestions on how to proceed when uncertainty in segmentation arises: for cases where it is evident that the boundary between segments can be placed somewhere within a range of 5–10 ms, they suggest that common criteria be used to segment all doubtful cases.[8] For cases where the boundary cannot be located with certainty, or not at all, they recommend excluding those tokens from the analysis. In the present study, I used Turk et al.'s suggestions in order to annotate approximants. For a sizeable number of approximant tokens, the duration could not be annotated because the beginning and end boundaries could not be made out from the flanking vowels; also, whereas in some cases a drop in intensity corresponding to the consonant could be observed, in others it was absent from the spectrogram and waveforms, even though the consonants were clearly perceivable. Approximants whose duration could be measured were included in the analyses of duration that will be discussed in the following sections. Approximants whose duration could not be manually measured with some degree of certainty were excluded from the analyses of duration, but included in the analyses of *category*, that is, in the stop/approximant categorization of tokens based on the presence/absence of a release burst (see Section 5.2.5.2). Figure 5.1 provides an example of the annotation of duration for a voiced stop realization and several approximant ones.

5.2.5.2 Criteria for labeling category

The voiced stops were labeled as stops if they presented a release burst in the spectrogram, and as approximants otherwise (cf. Figure 5.1 below). The presence of a release burst indicates that the realizations are produced with a full oral closure. All tokens were labeled as stop or approximant, independently of whether their duration could be measured or not.

[8]"One way . . . is to annotate them (e.g. with *?*), and to segment them according to a chosen policy of either 'when in doubt, place the boundary earlier', or 'when in doubt, place the boundary later', to be applied throughout the dataset" (Turk et al., 2006, p. 16).

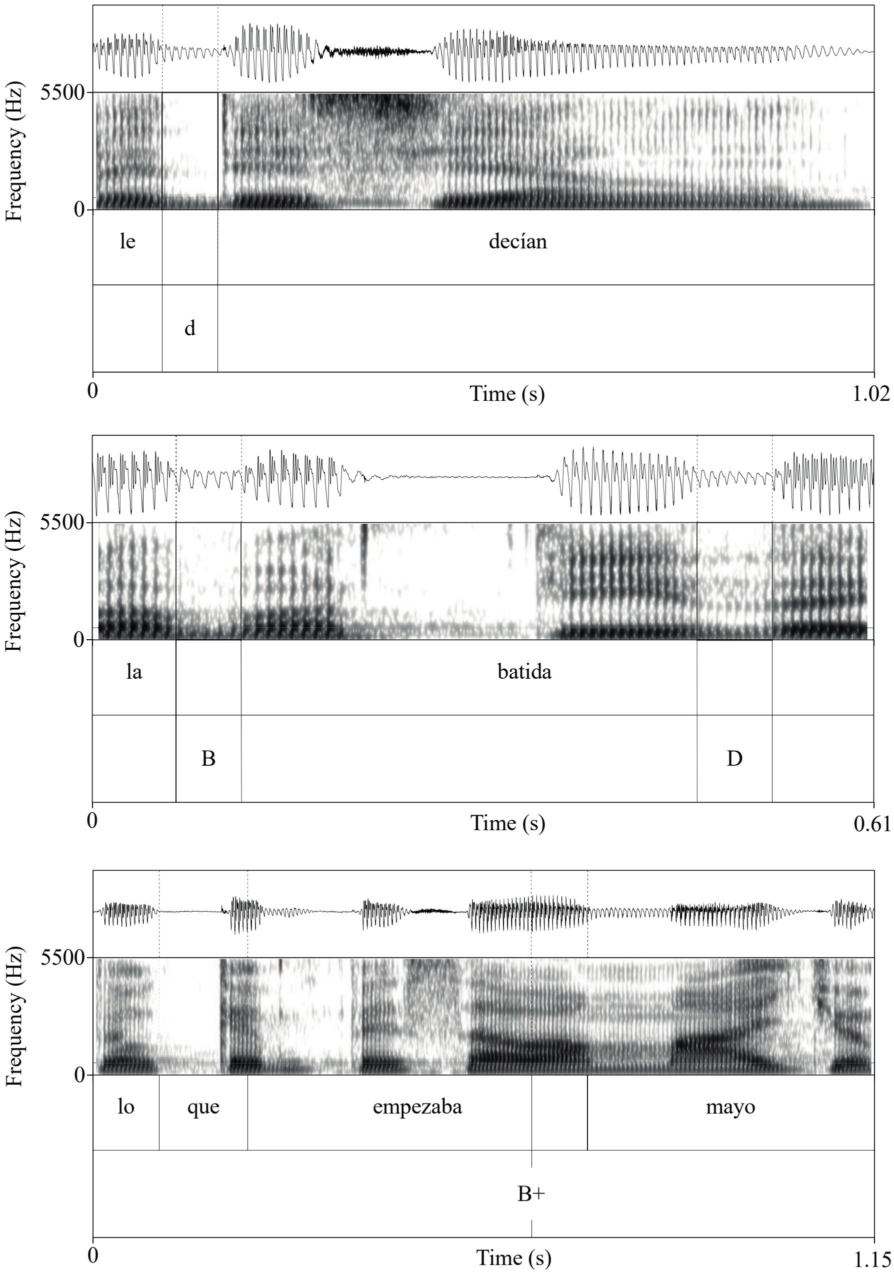

Figure 5.1. Example of the annotation of duration of voiced stops. Top: a stop realization, indicated by ‘d’, in the sequence *le decían* ‘(they) told him’ (speaker inf34, male). Center: an approximant realization of /b/ (indicated by ‘B’) and another of /d/ (‘D’), whose beginning and end points are annotated, in the sequence *la batida* ‘the battue’ (speaker inf34, male). Bottom: an approximant realization of /b/ (indicated by ‘B+’) whose duration was not measured, in the sequence *lo que empezaba mayo* ‘when May started’ (speaker inf31, female).

5.2.5.3 Annotation of prosodic factors

PROSODIC POSITION (i.e., position in the PW domain), LEXICAL STRESS, and WORD CLASS were annotated. For the factor PROSODIC POSITION, the voiced stop could appear in *LW-medial* position (e.g., *empeza**ba*** ‘started’), in *LW-initial* position (e.g., *le **d**ecían* ‘they told him’), or in *PW-initial* position (e.g., ***b**anda* ‘band’, ***d**e descanso* ‘of relaxation’).

For LEXICAL STRESS, the tokens could appear as the onset of a *stressed* syllable or of an *unstressed* one. For WORD CLASS, the target consonant could be in a *LW* (e.g., ***d**escanso*) or in a *FW* (e.g., ***d**e*). Only unstressed *FWs* were taken into consideration. Although the words in which the consonants appeared were not annotated, an examination of the data showed that the only *unstressed FWs* in which the consonants appeared were *de* 'of', *del* 'of + the.DET.M.SG', and *desde* 'from'. In sum, tokens that appeared in *FWs* all involved the dentialveolar stop and appeared exclusively in *PW-initial* position.[9]

5.2.6 Statistical analysis

5.2.6.1 Bayesian inference

The core of Bayesian inference is "reallocation of credibility [probability] across possibilities" (Kruschke, 2015, p. 22). One key advantage of Bayesian inference over a frequentist approach is that it is possible to incorporate prior knowledge or beliefs into a model. For example, in a study on the relationship between the number of people in Cologne's parks and the amount of sunlight per day (measured in minutes of sunlight), we could investigate how much the number of people increases as a function of the amount of sunlight per day, but also how certain we can be about it.[10] After stating the research questions, we would define a model that is relevant to our research questions with the data that we have collected. Thus, we would include in the model the dependent variable (the number of people in Cologne's parks) and the independent variable (the amount of sunlight per day), and also some random variation. We would then specify a *prior probability distribution* of the parameters of the model that reflects our previous knowledge. For example, we may have reasons to believe that the number of people may be in the hundreds, so we could fit a prior that takes into account this belief: this would constitute an *informative prior*. Conversely, we might choose to fit a prior that would simply disallow infinite numbers: this would be a *vague prior*. The result of fitting the Bayesian model would be the joint posterior distribution of the model parameters, which comes from the prior probability distribution and a likelihood function. If the data that we have collected are sufficient, the data (and not the priors) will largely determine the *posterior distribution*

[9] For the bilabial stop, there were only three instances of FWs, which were excluded due to being so few in number. Consequently, the subcorpus for the bilabial stop only contains LWs. For the dentialveolar stop, three instances of stressed FWs were excluded, as well as 57 tokens of unstressed FWs that were within the PW (e.g., *lo **d**e la convocatoria* 'about the announcement').

[10] This example is a rather simplified adaptation of the weight–height example in Kruschke (2015, pp. 25–30).

(Vasishth et al. 2018, p. 152). The posterior distribution will show the most credible values of the parameter and also the uncertainty in the estimated values. In other words, the posterior distribution shows what we should believe about the value of the parameters, given the data (Sorensen et al. 2016, p. 178). Thus, we would interpret the posterior distribution of our example model and conclude, for example, that the most probable number of people in Cologne's parks is 50 people for each increase in sunlight of a minute. Moreover, we could also establish the range of values that are most probable, for example those that correspond to the *95 % highest density interval* of the posterior distribution. In this fictional case, the result that we could obtain is that the increase in the number of people per minute may range from 35 to 65, with a most likely number of 50.

5.2.6.2 Bayesian models

A series of Bayesian mixed-effects models were fitted with the `brms` package (Bürkner, 2017) in the R programming environment (R Core Team, 2020), in the Stan computational framework (Carpenter et al., 2017), with the R package `RStan` (Stan Development Team, 2020). `brms` is a package that uses a syntax similar to the package `lme4` (Bates et al., 2015).

For the bilabial voiced stop, two models were fitted. A linear mixed model was fitted for the dependent variable DURATION (log-transformed). The predictors were PROSODIC POSITION (levels *LW-medial*, *LW-initial,* and *PW-initial*), LEXICAL STRESS (levels *unstressed* and *stressed*), BLP SCORE (SCALED), [11] and GENDER (levels *female* and *male*), with PROSODIC POSITION × LEXICAL STRESS as an interaction term. SPEAKER was included as a random intercept. The likelihood function used was Gaussian. Weakly informative priors were specified for the intercept (*Normal* (0, 10)), for the parameters representing the effects of the predictors and the interaction (*Normal* (0, 10)), for the standard deviation of the random effect (*Normal* (0, 10)), and for the standard deviation of the residual error (*Normal* (0, 10)). The range of the duration values (nontransformed) lay between 11 and 274 ms, which in log-scale corresponds to 2.4 and 5.6, respectively. Since it is highly unlikely that any single parameter will take on a value more extreme than the most extreme values of the data, the priors should capture the whole range of possible values of the parameters.

The second model for the bilabial stop was a logistic mixed model, which was fitted for the categorical dependent variable CATEGORY (treatment coded, levels *approximant* and *stop*). The predictors were PROSODIC POSITION (levels *LW-medial*, *LW-initial,* and

[11] The scaling was performed on the rounded BLP values. This procedure was followed for all studies in this dissertation.

PW-initial), LEXICAL STRESS (levels *unstressed* and *stressed*), BLP SCORE (SCALED), and GENDER (levels *female* and *male*), with PROSODIC POSITION × LEXICAL STRESS as an interaction term. SPEAKER was included as a random intercept. The likelihood function used was Bernoulli. Weakly informative priors were specified for the intercept (*Normal* (0, 3)), for the parameters representing the effects of the predictors and the interaction (*Normal* (0, 3)), and for the standard deviation of the random effect (*Normal* (0, 5)). *Normal (0, 3)* indicates that we expect a normal distribution of the model parameters with a mean of 0 and standard deviation of 3, which means that the parameters are believed to lie between -6 and +6 with a 95 % probability. Because the model uses a logistic link function, the priors thus defined should be able to account for at least 95 % of all probability. In fact, the range of possible values that the parameters could take on covers even rather extreme distributions of CATEGORY: thus, a value of +6 (/-6) on the log-odds scale is equivalent to a value of 99.75 % (/0.0025) on the probability scale.

For the dentialveolar voiced stop, two models parallel to those of the bilabial stop (i.e., one for DURATION, another for CATEGORY) were fitted. Two additional models were fitted to examine the effect of WORD CLASS. As indicated in Section 5.2.5.3, tokens in *FWs* appeared in *PW-initial* position. Consequently, a subset of the data was created for tokens in *PW-initial* position. A Bayesian linear mixed-effects model was fitted with DURATION (log-transformed) as the dependent variable and WORD CLASS (levels *FW* and *LW*) as the predictor, with SPEAKER as the random intercept. A parallel logistic mixed model was fitted with CATEGORY as the dependent variable (levels *approximant* and *stop*), and the same predictor (WORD CLASS) and random intercept (SPEAKER) as those in the model for DURATION. The priors for the DURATION model were the same as those for the previous DURATION models. The range of the duration values (nontransformed) lay between 10 and 194 ms, which in log-scale corresponds to 2.3 and 5.3, respectively. The priors for the CATEGORY model were the same as those for the previous CATEGORY models.

All models were fitted with four chains and 10,000 iterations, 2000 of which constituted the warm-up phase. Convergence was evaluated by means of the $\hat{R}$ value ($\hat{R}$ = 1.00 for all models) and the visual inspection of trace plots and the effective sample size (ESS) measures (Bulk ESS and Tail ESS). Good fit of the models was assessed via posterior predictive checks.

The package `emmeans` (Lenth, 2020) was used to obtain estimated marginal means and to conduct pairwise comparisons of the levels of the predictors of the interaction. This was done in order to assess main effects and partial effects for all of the possible combinations in the interactions.

Inference about the model parameters was carried out by summarizing and displaying posterior distributions.[12] Along with the parameter estimates, summarizing the posterior distribution also means displaying the *95 % credible interval* (CI) of each parameter. A 95 % CI is a property of the parameter that expresses our uncertainty about its value (Nicenboim et al., 2016, p. 7). The 95 % CI is "the range over which we can be 95 % certain that the true values of the parameter lie, given these particular data and the model" (Vasishth et al., 2018, p. 152). In the context of confirmatory hypotheses, testing the 95 % CI is usually interpreted as follows: if 0 is *not* included in it, there is evidence for an effect (Kruschke et al., 2012). Additionally, we can make an estimation of the probability of the parameter being smaller than or greater than 0. If 0 is included within the CI, it could still be the case that there is some weak evidence for an effect if most of the probability of the posterior distribution is either less or greater than 0 (e.g., Nicenboim et al., 2016).

5.3 Results

Overall information on the data set and model results are first provided for the bilabial stop (Section 5.3.1), followed by the overall information on the data sets and results for the models of the dentialveolar stop (Section 5.3.2).

5.3.1 Bilabial stop

36.22 % of the tokens that were included in the CATEGORY model had to be excluded from the DURATION model because their duration could not be measured reliably (cf. Section 5.2.5.1). Consequently, the number of tokens in the DURATION analysis ($N = 553$) is smaller than that in the CATEGORY analysis ($N = 867$). Table A4 in Appendix A provides the raw count of tokens per speaker that were included in the DURATION and CATEGORY models. Figure 5.2 provides the duration values of tokens according to PROSODIC POSITION and CATEGORY. In line with the literature on Spanish (e.g., Martínez Celdrán, 2013; see Section 4.2.2.2), the duration of *stop* realizations is longer than that of *approximant* ones.

[12] There are other approaches to carrying out inference in Bayesian statistics, such as computing a Bayes factor to compare a model with a certain fixed effect of interest with a model without that effect, or evaluating the predictive performance of a model against another by means of *k*-fold cross validation or leave-one-out (LOO) validation (Vasishth et al., 2018).

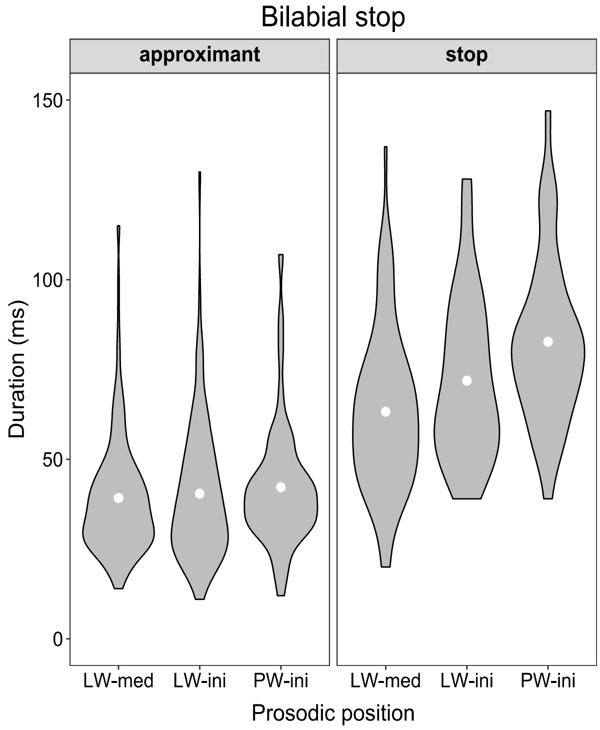

Figure 5.2. Violin plot (with mean points) of the duration of *approximant* and *stop* realizations of the bilabial stop ('LW-med' – *LW-medial* position; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position): corpus speech study.

5.3.1.1 Duration

In order to examine whether there was a main effect of PROSODIC POSITION, pairwise comparisons of the levels of the predictor were obtained. The results show no clear evidence that the levels *LW-medial* and *LW-initial* positions differ from each other since 0 is (just barely) included in the 95 % CI of the difference estimate (-0.09, CI [-0.18, 0.00]). However, since most of the probability mass is below 0, it might be reasonable to conclude that there is some weak evidence for a difference between these groups, with a shorter DURATION of tokens in the former level. On the other hand, there is evidence that both *LW-medial* and *LW-initial* tokens differ from those in *PW-initial* position in that they present shorter duration, since 0 is not included in the 95 % CI (contrast with *LW-medial*: -0.27, CI [-0.38, -0.16]; contrast with *LW-initial*: -0.19, CI [-0.31, -0.06]).

Similarly, the effect of LEXICAL STRESS was assessed by means of a pairwise comparison between the two levels of the predictor. Results show that, for the duration of tokens in *stressed* and *unstressed* syllables, the upper boundary of the CI is close to 0 (-0.09, CI [-0.18, 0]), although because most of the probability mass is below 0, we may conclude that there may be some weak evidence for a difference between the two groups, with tokens in *unstressed* syllables being shorter than in *stressed* ones.

Estimated marginal means and pairwise comparisons for the levels of the predictors of the interaction (PROSODIC POSITION × LEXICAL STRESS) were obtained (see Table A5 in Appendix A). Figure 5.3 is a plot of the conditional effects for the interaction. In it, the contrast between the predictors' levels can be observed.[13] There is evidence that the level *LW-medial* × *unstressed* differs from the other five levels in that it has shorter duration. Additionally, three levels (*LW-medial* × *stressed, LW-initial* × *unstressed,* and *LW-initial* × *stressed*) differ from *PW-initial* × *stressed* in that they have shorter duration. *PW-initial* × *stressed* and *PW-initial* × *unstressed* do not differ significantly. For pairwise comparisons in which 0 is just barely included in the 95 % CI of the difference estimate, the results show no clear evidence that the level *PW-initial* × *unstressed* differs from *LW-medial* × *stressed* and *LW-initial* × *stressed.* However, since most of the probability mass is above 0, it might be reasonable to conclude that there is weak evidence for a difference between these groups, with tokens in *PW-initial* × *unstressed* having a slightly greater duration than in the other two levels.

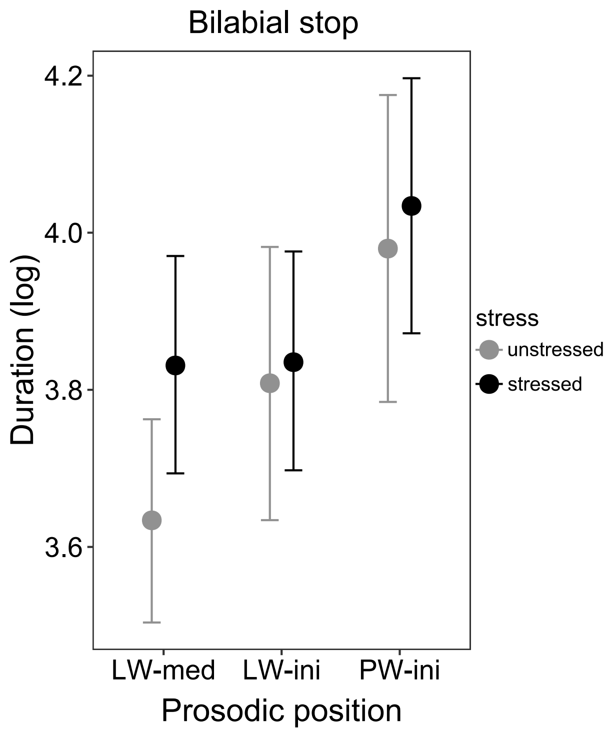

Figure 5.3. Interaction plot. Effect of PROSODIC POSITION conditional on LEXICAL STRESS ('LW-med' – *LW-medial* position; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position): DURATION model for the bilabial stop in the corpus speech study.

13 The plot shows predicted probabilities, whereas pairwise comparisons are on the logit scale.

Thus, as an interim summary for PROSODIC POSITION × LEXICAL STRESS, we can conclude that (i) the level *LW-medial* × *unstressed* differs greatly from the other levels in that the duration of bilabial tokens is considerably shorter; (ii) the tokens in the level *PW-initial* × *stressed* have considerably greater duration than those in *LW-medial* and *LW-initial* positions; and (iii) there is some weak evidence that *PW-initial* × *unstressed* may differ from the *stressed* groups of *LW-medial* and *LW-initial* positions.

The results for the BLP SCORE (SCALED) predictor indicate that 0 is included within the CI ($\hat{\beta}$ = -0.08, CI [-0.17, 0.00]). Based on the posterior samples, the estimated probability of the parameter being less than 0 is 0.97. This suggests that there is some weak evidence that the duration of the bilabial decreases the higher the BLP SCORE is, that is, the more Spanish dominant it becomes, although we cannot be certain about this. Figure 5.4 shows the duration of bilabial tokens for each speaker.

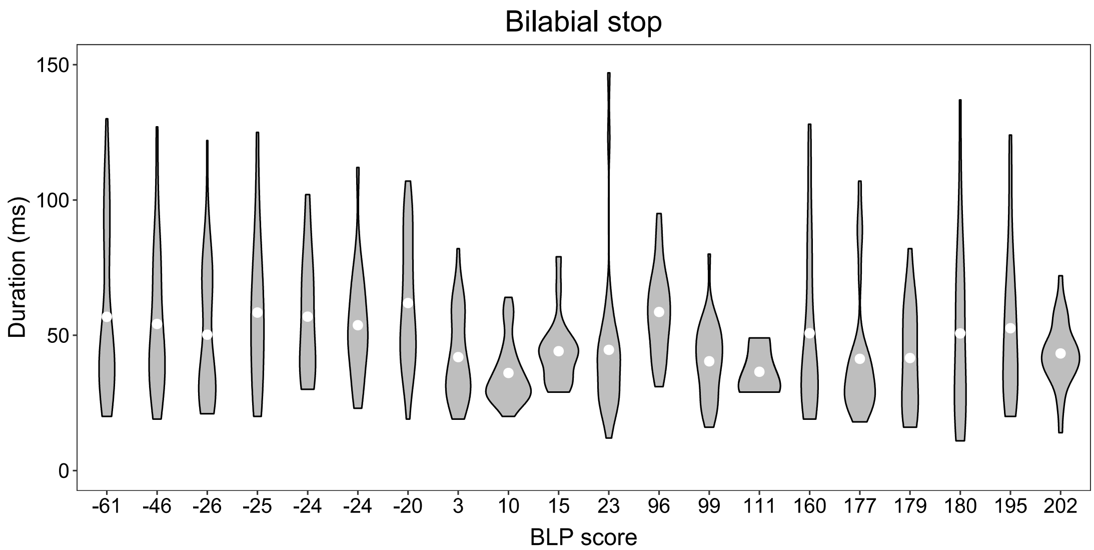

Figure 5.4. Violin plot (with mean points) of the duration of bilabial tokens per speaker in the read speech study. Each BLP score on the x-axis corresponds to one speaker.

The results for the GENDER predictor indicate that 0 is included within the CI ($\hat{\beta}_{genderM}$ = -0.00, CI [-0.18, 0.17]). Based on the posterior samples, the posterior probability of the parameter being lesser than zero is 0.5, that is, there is a 0.5 probability that *male* speakers produce the bilabial sound with shorter duration than *female* speakers, given the available data and the model. Consequently, we can draw the inference that the DURATION of the bilabial does not differ as a function of GENDER.

5.3.1.2 Category

15.47 % of the total percentage of tokens in *LW-medial* position were *stops*, while in *LW-initial* position it was 26.39 %, and in *PW-initial* position, 33.89 %. Table A4 in Appendix A shows the number of tokens per speaker for the bilabial voiced stop that were included in the CATEGORY model (*N* = 867, of which *stops* = 180, i.e., 20.76 %). Figure 5.5 provides the percentages of *stop* realizations per speaker. Speaker inf33 (BLP score = 111) only produced *approximant* realizations.

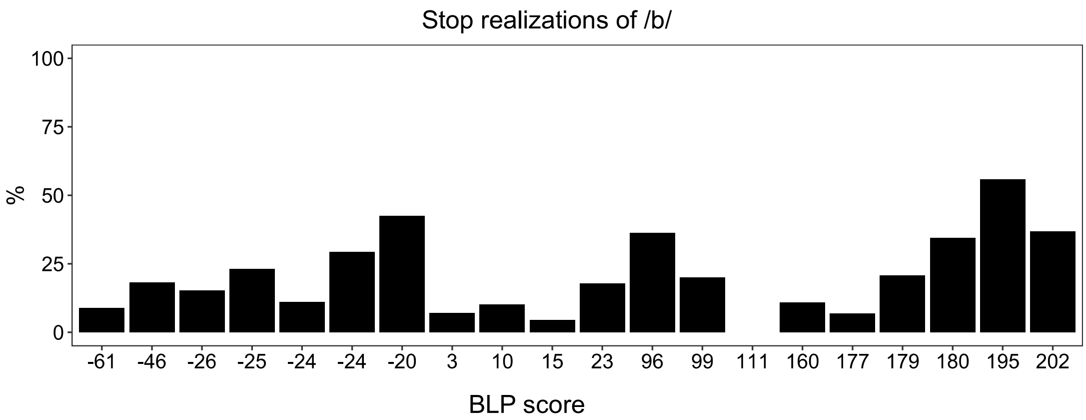

Figure 5.5. Percentage of *stop* realizations of the bilabial stop per speaker: CATEGORY model in the corpus speech study. Each BLP score on the x-axis corresponds to one speaker.

In order to examine whether there was a main effect of PROSODIC POSITION, pairwise comparisons of the levels of the predictor were obtained. Results show that there is evidence to consider that the level *LW-medial* differs from *LW-initial* and *PW-initial* in that there is a smaller probability of *stop* realizations in the former since 0 is not included in the CI (contrast with *LW-initial*: -0.67, CI [-1.12, -0.21]; contrast with *PW-initial*: -1.13, CI [-1.62, -0.60]). The evidence for the comparison of *LW-initial* and *PW-initial* indicates that these two groups do not differ considerably since 0 is included in the CI (-0.45, CI [-1.02, 0.13]).

Similarly, the effect of LEXICAL STRESS was also assessed by means of a pairwise comparison between the two levels of the predictor LEXICAL STRESS. The results show that the duration of tokens in *stressed* and *unstressed* syllables does not seem to differ substantially since 0 is included in the CI (-0.21, CI [-0.63, 0.2]).

Estimated marginal means and pairwise comparisons were obtained for the levels of the predictors of the interaction PROSODIC POSITION × LEXICAL STRESS (see Table A6 in Appendix A for the pairwise comparisons). Figure 5.6 is a plot of the conditional effects of the interaction between PROSODIC POSITION and LEXICAL STRESS. The level

LW-medial × *unstressed* differs from the other five levels in that it has fewer stop realizations. Additionally, there is evidence that indicates that *PW-initial* × *unstressed* differs from *LW-medial* × *stressed* in that the former has a higher probability of stop realizations.

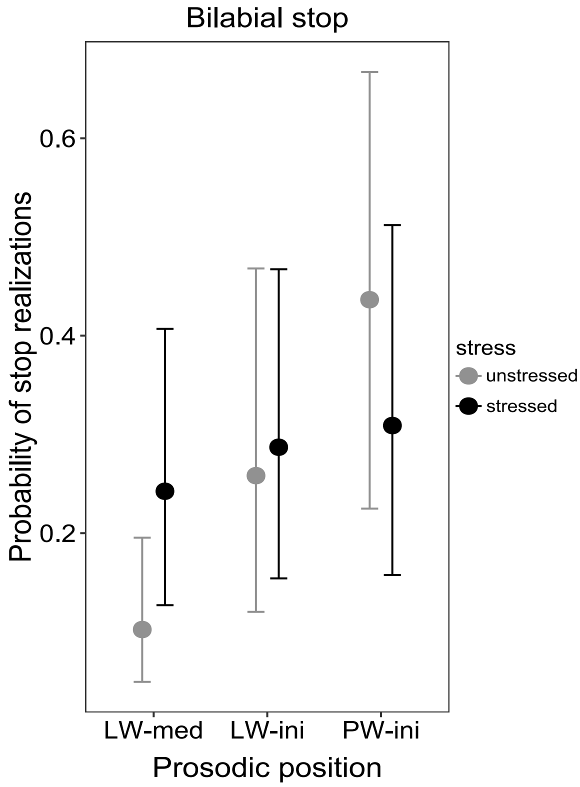

Figure 5.6. Interaction plot. Effect of PROSODIC POSITION conditional on LEXICAL STRESS ('LW-med' –*LW-medial* position; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position): CATEGORY model for the bilabial stop in the corpus speech study.

In sum, (i) the level *LW-medial* × *unstressed* differs considerably from the other levels in that the probability of *stop* realizations of the bilabial is smaller, and (ii) *LW-initial* × *stressed* differs from *PW-initial* × *unstressed*.

The results for the BLP SCORE (SCALED) predictor indicate that 0 is included within the CI ($\hat{\beta}$ = 0.16, CI [-0.35, 0.66). Based on the posterior samples, the estimated probability of the parameter being greater than 0 is 0.74. Consequently, we cannot with sufficient certainty draw the inference that the ratio of *stop* realizations of the bilabial increases the higher the BLP SCORE is, that is, the more Spanish dominant it becomes.

The results for the GENDER predictor indicate that 0 is included within the CI ($\hat{\beta}_{genderM}$ = -0.54, CI [-1.58, 0.46]). Based on the posterior samples, the posterior probability of the parameter being lesser than zero is 0.86, that is, there is a 0.86 probability that *male* speakers produce less *stop* realizations of the bilabial than *female* speakers, given the data available and the model. Because 0 is included within the CI, we cannot with sufficient certainty draw the inference that the ratio of stop realizations of the bilabial differs as a function of GENDER.

5.3.2 Dentialveolar stop

56.11 % of the tokens that are included in the CATEGORY model had to be excluded from the DURATION model because their duration could not be measured reliably. Consequently, the number of tokens for the DURATION analysis ($N = 657$) is smaller than that for the CATEGORY analysis ($N = 1497$). Table A7 in Appendix A provides the raw count of tokens per speaker for each analysis. Figure 5.7 provides the duration values of tokens according to PROSODIC POSITION and CATEGORY. Similarly to the results for the bilabial stop, the duration of *stop* realizations of the dentialveolar is longer than that of *approximants*.

5.3.2.1 Duration

In order to examine whether there was a main effect of PROSODIC POSITION, pairwise comparisons of the levels of the predictor were obtained. The results show evidence that the levels *LW-medial* and *LW-initial* positions differ, since 0 is not included in the 95 % CI of the difference estimate (-0.22, CI [-0.30, -0.13]), with duration being shorter in *LW-medial* than in *LW-initial* position (see Figure 5.7). Similarly, there is evidence that the level *LW-medial* differs from the level *PW-initial* in that, for the former, the duration of bilabial tokens is shorter (-0.11, CI [-0.21, 0.00]). For the comparison of *LW-initial* and *PW-initial* positions, the results show no clear evidence that the levels differ since 0 is just barely included in the 95 % CI of the difference estimate (0.11, CI [-0.01, 0.22]). However, since most of the probability mass is above 0, it might be reasonable to conclude that there is some weak evidence for a difference between these groups in that the duration is longer in *LW-initial* than in *PW-initial* position.

Similarly, the effect of LEXICAL STRESS was assessed by means of a pairwise comparison between the two levels of the predictor. Results show that, for the duration of tokens in *stressed* and *unstressed* syllables, there is no clear evidence that they differ much since 0 is included in the CI (-0.06, CI [-0.15, 0.02]), although because most of the probability mass is below 0, we may conclude that there is some weak evidence for a difference between the two groups, with tokens in *unstressed* syllables being shorter than in *stressed* ones.

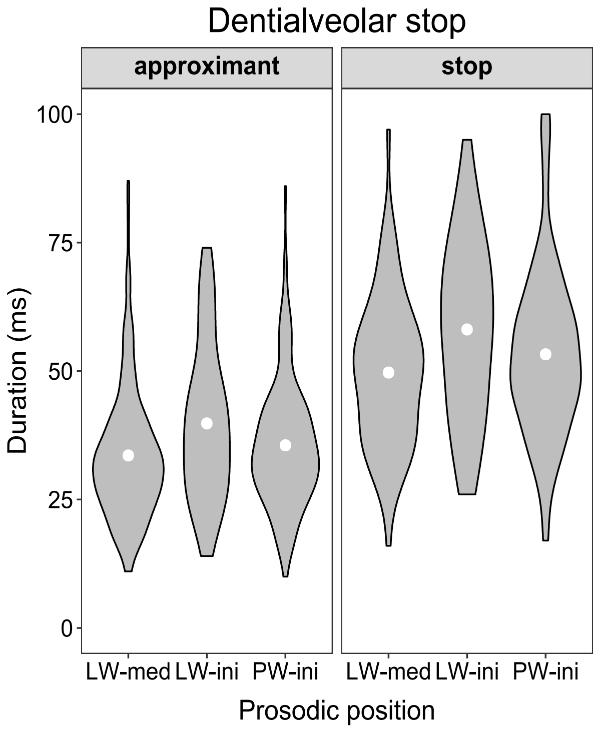

Figure 5.7. Violin plot (with mean points) of the duration of *approximant* and *stop* realizations of the dentialveolar stop ('LW-med' – *LW-medial* position; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position): corpus speech study.

Estimated marginal means and pairwise comparisons for the levels of the predictors of the interaction PROSODIC POSITION × LEXICAL STRESS were obtained (see Table A8 in Appendix A). Figure 5.8 is a plot of the conditional effects of PROSODIC POSITION × LEXICAL STRESS. The level *LW-medial* × *unstressed* differs from the other five levels in that it presents shorter duration. Additionally, there is evidence that tokens in *LW-initial* × *unstressed* are longer than in *LW-medial* × *stressed* and longer than in *PW-initial* × *unstressed*. Also, there is some evidence that tokens in *LW-medial* × *stressed* are shorter than those in *LW-initial* × *stressed* position, and that tokens at *PW-initial* × *unstressed* are shorter than those in *LW-initial* × *stressed*. The levels *PW-initial* × *stressed* and *PW-initial* × *unstressed* do not differ considerably since 0 is included in the CI.

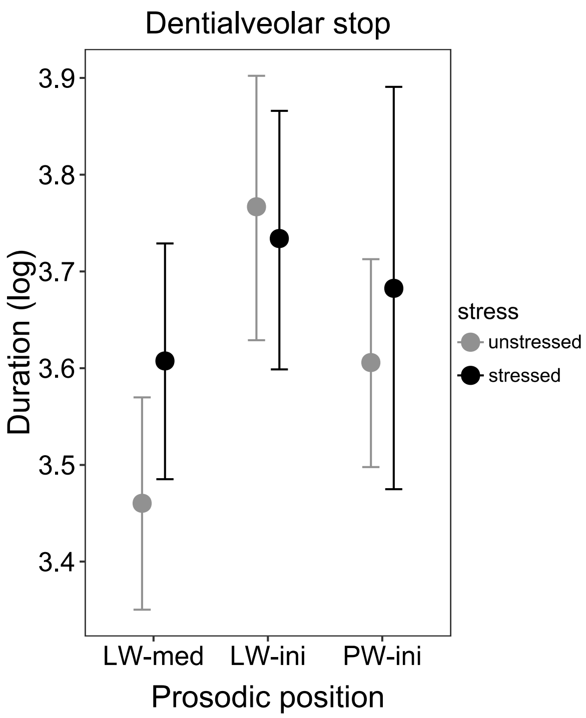

Figure 5.8. Interaction plot. Effect of PROSODIC POSITION conditional on LEXICAL STRESS ('LW-med' – *LW-medial* position; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position): DURATION model for the dentialveolar stop in the corpus speech study.

As an interim summary for PROSODIC POSITION × LEXICAL STRESS, we can conclude that the level *LW-medial* × *unstressed* differs considerably from the other levels in that the duration of dentialveolar tokens is shorter. The differences between levels in the interaction indicate that *LW-initial* differs from *LW-medial* and *PW-initial* positions, although the evidence for a difference with *PW-initial* positions is rather weak. Additionally, tokens in *LW-medial* position differ from those in *PW-initial* in having shorter duration.

The results for the BLP SCORE (SCALED) predictor indicate that 0 is not included within the CI ($\hat{\beta}$ = -0.10, CI [-0.17, -0.03]). Based on the posterior samples, the estimated probability of the parameter being less than 0 is 0.99. Thus, there is evidence that the duration of dentialveolar tokens decreases the higher the BLP SCORE is, that is, the greater the Yucatecan Spanish dominance is. Figure 5.9 shows the duration of dentialveolar tokens for each speaker (ordered by BLP SCORE).

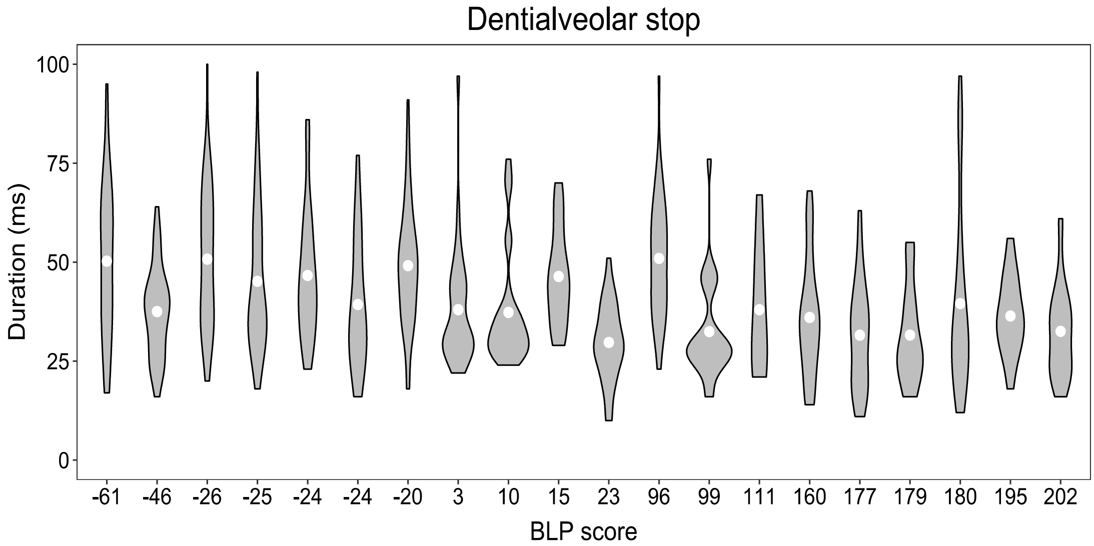

Figure 5.9. Violin plot (with mean points) of the duration of dentialveolar tokens per speaker in the read speech study. Each BLP score on the x-axis corresponds to one speaker.

The results for the GENDER predictor indicate that 0 is included within the CI ($\hat{\beta}_{genderM}$ = 0.09, CI [-0.05, 0.23]). Based on the posterior samples, the posterior probability of the parameter being greater than zero is 0.9, that is, there is a 0.9 probability that *male* speakers produce the dentialveolar with longer duration than *female* speakers, given the available data and the model. Because most of the probability mass is above 0, there may be some weak evidence for a difference between the two groups.

5.3.2.2 Category

Table A7 in Appendix A shows the number of tokens of the dentialveolar voiced stop per speaker that were included in the CATEGORY model (N = 1497, of which *stops* = 231, i.e., 15.43 %). In *LW-medial* position, 9 % of the total percentage of tokens were *stops*, while in *LW-initial* position it was 30.37 %, and in *PW-initial* position, 19.96 %. Figure 5.10 provides the percentages of *stop* realizations per speaker.

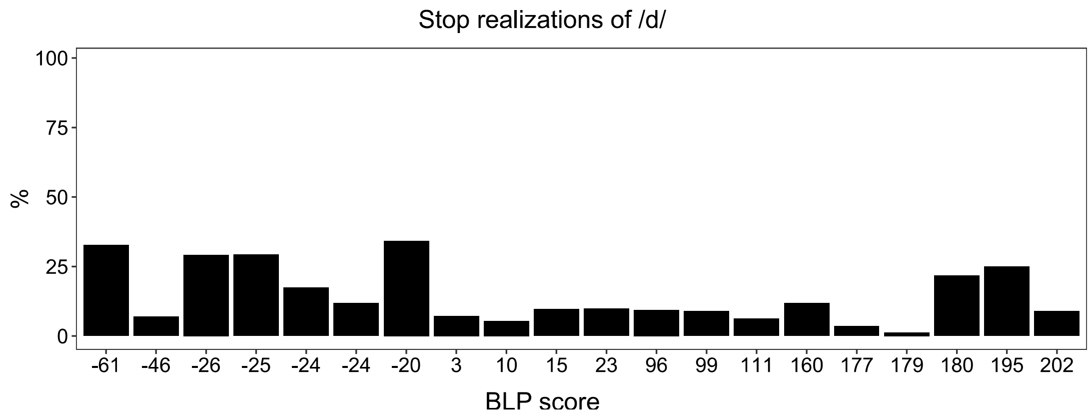

Figure 5.10. Percentage of *stop* realizations of the dentialveolar stop per speaker: CATEGORY model in the corpus speech study. Each BLP score on the x-axis corresponds to one speaker.

In order to examine whether there was a main effect of PROSODIC POSITION, pairwise comparisons of the levels of the predictor were obtained. The results show evidence that the *LW-medial* and *LW-initial* positions differ from each other since 0 is not included in the 95% CI of the difference estimate (-1.36, CI [-1.79, -0.92]), with fewer *stop* realizations in *LW-medial* than in *LW-initial* position. Similarly, there is evidence that the level *LW-medial* differs from the level *PW-initial* in that for the former, the probability of stop realizations is smaller (-1.00, CI [-1.51, -0.46]). For the comparison between *LW-initial* and *PW-initial* positions, the results show no evidence that the levels differ, since 0 is included in the 95% CI (0.36, CI [-0.19, 0.94]).

Similarly, the effect of LEXICAL STRESS was also assessed by means of a pairwise comparison between the two levels of the predictor. The results show that there is evidence that the number of *stop* realizations of tokens in *stressed* and *unstressed* syllables differs since 0 is not included in the CI (-0.85, CI [-1.27, -0.43]), with a lower probability for *unstressed* syllables.

Estimated marginal means and pairwise comparisons for the levels of the predictors of the interaction PROSODIC POSITION × LEXICAL STRESS were obtained (see Table A9 in Appendix A). Figure 5.11 is a plot of the conditional effects of the interaction. The level *LW-medial* × *unstressed* differs from the other five levels in that the probability of having *stop* realizations is lower. Additionally, the level *LW-medial* × *stressed* differs from *LW-initial* × *stressed* in that there is evidence that the probability of having stop realizations is lower in the former. Also, there is evidence for fewer *stop* realizations in *PW-initial* × *unstressed* than in *LW-initial* × *stressed.* There is one pairwise comparison in which 0 is just barely included in the 95% CI of the difference estimate, namely the one between *PW-initial* × *unstressed* and *LW-initial* × *stressed*. Since most of the probability mass is below 0, it might be reasonable to conclude that there is some weak evidence for a difference between these groups, with a smaller probability of tokens in *PW-initial* × *unstressed* having *stop* realizations.

As an interim summary for the interaction PROSODIC POSITION × LEXICAL STRESS, we can conclude that the level *LW-medial* × *unstressed* differs considerably from the other levels in that the probability of having *stop* realizations of dentialveolar tokens is smaller. Also, *LW-initial* × *stressed* differs considerably from *LW-medial* × *stressed*, as well as from *PW-initial* × *unstressed*.

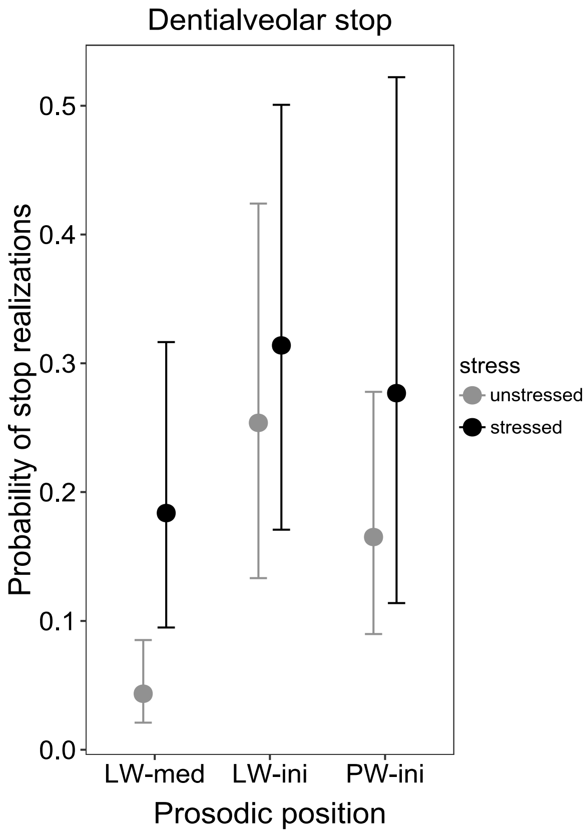

Figure 5.11. Interaction plot. Effect of PROSODIC POSITION conditional on LEXICAL STRESS ('LW-med' – *LW-medial* position; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position): CATEGORY model for the dentialveolar stop in the corpus speech study.

The results for the BLP SCORE (SCALED) predictor indicate that 0 is included within the CI ($\hat{\beta}$ = -0.30, CI [-0.77, 0.16). Based on the posterior samples, the estimated probability of the parameter being lesser than 0 is 0.9. Because most of the probability mass is below 0, there may be some weak evidence for the ratio of *stop* realizations decreasing the higher the BLP SCORE is, that is, the higher the Yucatecan Spanish dominance is, but we cannot be certain about it.

The results for the GENDER predictor indicate that 0 is included within the CI ($\hat{\beta}_{genderM}$ = -0.18, CI [-1.12, 0.75]). Based on the posterior samples, the posterior probability of the parameter being less than zero is 0.65, that is, there is a 0.65 probability that *male* speakers produce less instances of *stop* realizations for the dentialveolar than *female* speakers, given the available data and the model. Because of these results, we cannot with sufficient certainty draw the inference that the ratio of *stop* realizations of the dentialveolar differs as a function of GENDER.

5.3.2.3 Word class

In order to investigate whether there was a difference in terms of duration and stop/approximant realizations as a function of WORD CLASS, a subset of data with the dentialveolar tokens that appear in *PW-initial* position was used (see Section 5.2.5.3). Similarly to the data sets used for the analysis of duration and category for the dentialveolar stop, the number of tokens in the subset for the DURATION model (*N* = 256, of which 201 were *FWs*) was smaller than for the CATEGORY model (*N* = 506, of which 413 were *FWs*). Table A7 in Appendix A presents the number of *FWs* and *LWs* per speaker in the DURATION and CATEGORY models.

For the DURATION model, the results for the WORD CLASS predictor indicate that 0 is not included within the CI ($\hat{\beta}_{LW}$ = 0.17, CI [0.05, 0.29]), thus providing evidence for an effect of WORD CLASS. Based on the posterior samples, the posterior probability of the parameter being greater than zero is 1, that is, there is a probability of 1 that voiced stops in *LWs* are produced with greater duration than in *FWs*. The results for the CATEGORY model mirror those of the DURATION model ($\hat{\beta}_{LW}$ = 0.77, CI [0.18, 1.35]), with the posterior probability of the parameter being greater than zero being 0.99.

5.4 Discussion and summary

Table 5.3 is an overview of the results of the models presented in the present chapter. WORD CLASS will be considered after discussing the results for the interaction between PROSODIC POSITION and LEXICAL STRESS. The mathematical symbol for 'greater-than or equal to' is used here and in the next chapters to indicate that there may be some weak evidence of a difference between some levels of a variable, but that we cannot be certain about this.

Table 5.3. Results for strengthening of voiced stops in the corpus speech study. 'LW-med' – *LW-medial* position; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position; 'str' – *stressed*; 'uns' – *unstressed*.

		Results			
		Bilabial voiced stop		**Dentialveolar voiced stop**	
Variables	**Expected results**	**Duration**	**Category**	**Duration**	**Category**
1. PROSODIC POSITION	*PW-initial* > *LW-initial* > *LW-medial*	+/- *PW-initial* > (*LW-initial* ≥ *LW-medial*)	+/- (*PW-initial* & *LW-initial*) > *LW-medial*	× (*LW-initial* ≥ *PW-initial*) > *LW-medial*	+/- (*PW-initial* & *LW-initial*) > *LW-medial*
2. LEXICAL STRESS	*stressed* > *unstressed* syllables	? *stressed* ≥ *unstressed*	×	? *stressed* ≥ *unstressed*	✓
3. PROSODIC POSITION × LEXICAL STRESS	*PW-ini × str > uns > LW-ini × str > uns > LW-med × str > uns*	+/- *PW-ini × str* > all other levels (except *PW-ini × uns*)> *LW-med × uns*	× all other levels > *LW-med × uns*	× *LW-ini* > (*PW-ini* & *LW-med × str*)> *LW-med × uns*	× (*LW-ini* ≥ *LW-med × str; LW-ini* ≥ *PW-ini*) > *LW-med × uns*
4. LANGUAGE DOMINANCE	Yucatec Maya > Yucatecan Spanish	?	×	✓	?
5. GENDER	no effect	✓	? *females* ≥ *males*	? *males* ≥ *females*	✓
6. WORD CLASS (/d/)	LWs = FWs			× *LWs > FWs*	× *LWs > FWs*

The analyses of the data for both the bilabial and the dentialveolar voiced stops show that there is an effect of PROSODIC POSITION, although not in the direction expected, whereas the evidence for an effect of LEXICAL STRESS is not entirely clear overall.

The results for the overall effect of PROSODIC POSITION show that the expected tendency in terms of strengthening (i.e., *PW-initial* > *LW-initial* > *LW-medial* positions) is not fulfilled overall. It is clear from the results that *LW-medial* position is at the lower end of the tendency: tokens in *LW-medial* position have shorter duration and fewer stop realizations. This is clearly the case for the level *LW-medial* × *unstressed,* which differs from all other levels, and partially so for the level *LW-medial* × *stressed.* However, the order (in terms of strengthening) of the levels of *PW-initial* and *LW-initial* position is not straightforward.

The results for the overall effect of LEXICAL STRESS indicate that, when there is evidence for an effect, tokens in *stressed* syllables are likely to have greater duration (DURATION models) and a greater probability of stop realizations (CATEGORY models). However, the evidence for an effect of LEXICAL STRESS is mixed, with very weak or no support for an effect in the bilabial models and with some support in the dentialveolar models. This result is similar to that observed for several studies on voiced stops in Spanish that have also used spontaneous speech (see Section 4.2.2.3).

The results for the interaction of PROSODIC POSITION and LEXICAL STRESS indicate that there is no support for the expected results, since the partial evidence in support of such claim comes from only *LW-medial* × *unstressed* < all other levels in all models, and *PW-initial* × *stressed* > all other levels in the DURATION model for the bilabial stop. Importantly, for the other three models, tokens in *LW-initial* position (especially when *stressed*) have an overall similar or higher probability of presenting strengthened realizations than when in *PW-initial* position and *LW-medial* position. Moreover, there are no substantial differences in strengthening in *LW-initial* position due to STRESS. Apart from the model results, it must be kept in mind that position in higher prosodic domains was not controlled for, meaning that tokens in *PW-initial* position could also be at the initial position of higher prosodic domains, such as the intermediate phrase or the Intonational Phrase, for which more strengthened realizations could be expected due to a cumulative effect of domain-initial strengthening (see Section 2.4.2). However, the higher degree of strengthening found in the *LW-initial* position, but not in the *PW-initial* position, suggests that there was no additional strengthening due to higher prosodic boundaries.

These results point towards a difference between items in *word-medial* and *word-initial* positions, where word-initial refers to both *PW-initial* and *LW-initial* positions in the analyzed data. Thus, this study provides evidence that supports a difference in terms of strengthening for tokens in word-initial versus word-medial position. This result is in line with the literature on Yucatecan Spanish (see Section 4.2.3). Whereas most previous studies have compared tokens in word-medial and word-initial positions without clearly stating what they take to be a word, the present study allows us to draw a more complex picture of where strengthening occurs. Moreover, the evidence obtained by means of this study is not only in terms of approximant/stop realizations, which is what previous studies of Yucatecan Spanish investigated, but also of duration.

The results show that there is a clear difference between tokens in *stressed* and *unstressed* tokens word medially. Although the lexical items in which the tokens appeared were not annotated, it is likely that a sizeable number of *unstressed* tokens corresponded to voiced stops in highly frequent words, such as those that include past imperfect suffix forms (*-aba, -aban*) and past participle suffixes (*-ado, -ada*), as well as frequent words such as *todo* 'everything' or *todavía* 'still', for which weakening and elision has been previously reported for Yucatecan Spanish (cf. Section 5.2.2).

In terms of WORD CLASS, which was only investigated for dentialveolar tokens in *PW-initial* position, there is strong evidence to support that tokens in *LWs* have longer duration and more stop realizations than those in *FWs*. This result is not in line with the expected results, that is, a similar strengthening due to PROSODIC POSITION. Nevertheless, this result may be due to the high lexical frequency of *FWs*: tokens appeared in the words *de* 'of', *del* 'of-the.M.SG', and *desde* 'from', all of them extremely frequent in Spanish (cf. Section 5.2.5.3). Another possible explanation could be the fact that, for *LWs*, both tokens in *stressed* and *unstressed* syllables were considered, whereas for *FWs* all tokens appeared in *unstressed* tokens. In other words, including tokens in *stressed* syllables (which could lead to a higher degree of strengthening) might play a role in the results.

Results for LANGUAGE DOMINANCE (calculated by means of the BLP SCORE (SCALED) predictor) show that there could be a slight effect OF LANGUAGE DOMINANCE in the expected direction, that is, that Yucatec Maya dominance favors more strengthened realizations. However, the evidence obtained is weak, and it is even contradictory for the CATEGORY model for the bilabial stop. The overall result is not surprising in light of previous studies on Yucatecan Spanish, which present mixed views on whether language dominance has an effect or not (see Section 4.2.3). Importantly, the weak evidence in this study in support of an effect of LANGUAGE DOMINANCE when compared to previous studies may be due to the choice of the BLP score as a continuous measure versus the grouping into monolinguals and bilinguals used in previous studies of Yucatecan Spanish (cf. Section 3.1.2). If the analysis of LANGUAGE DOMINANCE in the present study had been made in a similar way, that is, grouping speakers into Maya–Spanish bilinguals and Spanish monolinguals (with a cutoff point at 0, for example), language dominance may have turned out to have had a greater effect. Additionally, because the evidence for more strengthened realizations in bilingual speakers is weak and partially contradictory, it is rather unlikely that bilingual speakers have created a merged L1–L2 category for voiced stops in the sense of the Speech Learning Model (SLM).

For GENDER, the overall evidence obtained from the models suggests that there are no differences as a function of GENDER in the realization of the two voiced stops, neither in terms of duration nor in terms of allophonic variation. Thus, this study appears to confirm previous findings in the literature about the lack of any gender effect on Yucatecan Spanish voiced stops (see Section 4.2.3).

In the present study, the two phonetic parameters that manifest prosodic strengthening, namely greater (acoustic) duration and the presence of a release burst (used to group realizations into stop and approximant ones), were taken to be complementary. Although duration captures fine-grained variation and is consequently best suited for the study of prosodic strengthening, a considerable number of tokens had to be excluded due to the impossibility of measuring their duration in a reliable way (see Section 5.2.5.1). On the other hand, the analysis of stop/approximant realizations, although a simplification, allowed us to include all data points.

In summary, the present study of the strengthening of voiced stops in a corpus of spontaneous speech has shown that strengthening is favored when tokens appear in word-initial position, which includes both PW-initial and LW-initial positions, and that the evidence for an effect of lexical stress is mixed, although it is clear that tokens in unstressed syllables in word-medial position pattern differently from the rest in that they present less strengthening. Function words are less strengthened than LWs, which could be due to the higher frequency of the former or to the inclusion of stressed tokens only for LWs. In terms of speaker-specific variation, it may be that the Yucatec Maya dominant speakers tend to produce more strengthened realizations, but the evidence for such tendency is weak and partially contradictory. Finally, the results for gender suggest that female and male speakers produce strengthened realizations to the same extent.

Chapter 6

Lack of resyllabification and glottalization of word-initial vowels

6.1 The present study

The present study investigates lack of resyllabification and glottalization of (orthographic) word-initial vowels, which may appear within and across Prosodic Words (PWs) in Yucatecan Spanish. More specifically, it examines how lack of resyllabification and glottalization of word-initial vowels are related, and also the role of several linguistic and speaker-specific factors in the distribution of lack of resyllabification and glottalization.

The studies reviewed in Section 4.3.4 show that glottalization of word-initial vowels is rather frequent in Yucatecan Spanish and that it also takes place in other varieties, such as European Peninsular Spanish, although its frequency and distribution are unclear. For Yucatecan Spanish, Michnowicz and Kagan (2016) reported that 11 % of all phrase-medial word-initial sequences amenable to resyllabification were glottalized, also showing that glottalization was subject to prosodic, segmental, and speaker-specific effects. Michnowicz and Kagan claimed that glottalization helps mark word boundaries by blocking resyllabification. However, because the focus of their study was on glottalization and not on (lack of) resyllabification, we have no information as to whether there were other instances of lack of resyllabification unrelated to glottalization. Since some works on Spanish have shown that lack of resyllabification can take place independent of glottalization (e.g., Aguilar, 2003; Strycharczuk & Kohlberber, 2016), I first examine lack of resyllabification without reference to glottalization, and then I examine how glottalization relates to lack of resyllabification in Yucatecan Spanish.

The studies on Spanish resyllabification are not necessarily clear on what is meant by *word* (e.g., Michnowicz & Kagan, 2016), although it is likely that they refer to the orthographic word. Since the PW is defined as the domain of primary stress, unstressed function words (FWs) are clitics that merge with the following lexical word (LW), whereas stressed FWs and LWs can constitute PWs by themselves (see Section 2.2.1.2). According to Elordieta (2014), unstressed FWs may either be part of a recursive PW or they may adjoin directly to a Phonological Phrase (PhP). Because of the exploratory nature of the present study, two positions in relation to the PW domain are examined: within the PW and across PWs. *Within the PW*, the word-initial vowel may belong to (i) a LW preceded by one or more unstressed FWs (e.g., ((*en*) *el* ***agua***)$_{PW}$ '(in) the water'), (ii) a stressed FW preceded by one or more unstressed FWs (e.g., ((*la*) *de* ***una***)PW '(that) of one.F'), or (iii) an unstressed FW preceded by another (e.g., (*en* ***el*** *mercado*)PW 'in the market'). *Across PWs*, the word-initial vowel may be that of (i) an unstressed FW that is cliticized to a LW (e.g., (***en*** *el mercado*)PW, (***el*** *agua*)PW), (ii) a LW preceded by another LW (e.g., (*beben*)PW (***agua***)PW 'drink.PRS.3PL water'), or (iii) a stressed FW that is not preceded by an unstressed FW (e.g., *(toman)*PW (***una***)PW 'take.PRS.3PL one.F'). Tentatively, the *outer* position in the PW (i.e., across PWs) will be considered to favor more lack of resyllabification and glottalization than the *inner* position in the PW (i.e., within the PW) for two reasons: (i) because the former constitutes a larger unit than the latter, and (ii) because previous studies on Spanish (vocalic) resyllabification have found that lack of resyllabification is more common across LWs than between a FW and a LW (Aguilar, 2005, for European Peninsular Spanish; Alba, 2006, for New Mexico Spanish; see Section 2.2.1.3).

In the present study, the effect of lexical stress will be assessed, as well as whether there is an interaction between prosodic position and lexical stress. Lack of resyllabification in Spanish has been linked to stressed syllables (e.g., Aguilar, 2003, 2005; Alba, 2006; Navarro Tomás, 1918/1996, pp. 148–149; see Section 2.2.1.3), while Michnowicz and Kagan (2016) found more instances of glottalization of word-initial vowels in stressed syllables than in unstressed ones in Yucatecan Spanish. This result is in line with studies that have shown that glottalization is more frequently associated with prominent positions (see Section 4.3.3). Thus, there is more glottalization in pitch accented syllables for American English (Dilley et al., 1996) and in stressed syllables for American English and Central Mexican Spanish (Garellek, 2014) and German and Polish (Malisz et al., 2013). In sum, it seems likely that both lack of resyllabification and glottalization in Yucatecan Spanish will be favored in word-initial stressed syllables rather than in unstressed ones. Tentatively, I expect the highest frequency of lack of resyllabification and glottalization when the target vowel is in a stressed syllable across PWs.

The effect of language dominance will be assessed by means of the Bilingual Language Profile (BLP; see Section 3.1.2). In the study by Michnowicz and Kagan (2016), bilingual Yucatec Maya–Yucatecan Spanish speakers produced more instances of glottalization (and thus of lack of resyllabification) than monolinguals.

That bilingual speakers should produce more glottalization of vowels could be related to the fact that, in Yucatec Maya, morphological roots and words may start and end with a glottal stop (Frazier, 2009; see Section 2.3). However, diphthongs, which are postlexical in Yucatec Maya, may arise from the frequent postlexical deletion of the glottal consonants between vowels (Gussenhoven & Teeuw, 2008; see Section 2.3), which runs counter to a Yucatec Maya influence on Yucatecan Spanish in terms of glottalization and thus lack of resyllabification, because it could be argued that a postlexical diphthong across word boundaries is resyllabification. Nevertheless, there are no phonetic studies to confirm or reject these claims. Another factor that could help explain the glottalization of word-initial vowels in Yucatecan Spanish is the fact that there are (long-high) creaky vowels in Yucatec Maya (/íḭ éḛ óo̰ úṵ/). Thus, the glottalization of vowels in Yucatecan Spanish could be related to the existence of creaky vowels in Yucatecan Spanish. However, short vowels (/i e a o u/), which could be similar to the Spanish vowels, also exist in Yucatec Maya (see Section 1.1). Tentatively, I expect that speakers with greater scores of Yucatec Maya dominance will produce more tokens that present lack of resyllabification and glottalization, in line with the results obtained by Michnowicz and Kagan (2016).

Finally, Michnowicz and Kagan (2016) did not find an effect of gender (see Section 4.3.4). Consequently, a similar result is expected in the present study. Table 6.1 presents a summary of the expected direction of the results.

Table 6.1. Expected results for lack of resyllabification and glottalization of vowels in word-initial position: corpus speech study.

Variables	Expected results
1. Prosodic position	across PWs > within the PW
2. Lexical stress	stressed > unstressed syllables
3. prosodic position × lexical stress	across PWs × stressed > all other
4. Language dominance	Yucatec Maya > Yucatecan Spanish
5. Gender	no effect

The remainder of the chapter is organized as follows. Section 6.2 describes the characteristics of the speech materials (Section 6.2.1), the participants (Section 6.2.2), the annotation criteria that were used (Section 6.2.3), and the statistical analysis (Section 6.2.4). Section 6.3 presents an overview of the data (Section 6.3.1), the results for lack of resyllabification (Section 6.3.2), and the results for glottalization (Section 6.3.4). Section 6.4 discusses the two studies and also provides the reasons for a post hoc analysis. Section 6.5 introduces a post hoc analysis of the data with random forests, which includes an explanation of random forests (Section 6.5.1), the methods used (Section 6.5.2), the results (Section 6.5.3), and a discussion (Section 6.5.4). Section 6.6 provides a summary of the whole chapter.

6.2 Methods

6.2.1 Speech materials

The speech materials that were used in this study are part of the same corpus of sociolinguistic interviews about culture and language that was described in Section 5.2.1 for voiced stops. They were recorded, selected, transcribed, and aligned in the same manner as indicated in Sections 5.2.2 and 5.2.4. Similarly to the voiced stop study, the resulting tiers from the alignment process were used as guidance in locating the tokens. For the present study, the first 150 observations per speaker of vowels in resyllabification contexts were annotated ($N = 2400$).

6.2.2 Speakers

The interviews of 16 speakers (8 female, 8 male) were selected following the same criteria as the ones presented in Section 5.2.3. All participants in the present study except for two of them (inf24 and inf25) also took part in the voiced stops study of Chapter 5. The participants, who at the time resided in Felipe Carrillo Puerto, were either from the states of Quintana Roo or Yucatán (see Table 6.2). At the time of the recordings, all speakers had lived in Felipe Carrillo Puerto for at least 11 years and in Quintana Roo for at least 11 years as well. Only one participant (inf21) reported speaking a language other than Spanish or Maya (English).

6.2.3 Annotation

6.2.3.1 Criteria for labeling resyllabification and glottalization

Vowels in (orthographic) word-initial position that were in a resyllabification context (i.e., not preceded by a pause) were labeled by the author of the dissertation (*rater 1*; see below) for presence or lack of resyllabification by means of an auditory analysis. Similarly to Alba (2006), who argued that it is not always clear whether a vocalic sequence may be considered a hiatus or not (see Section 2.2.1.3), whether a sequence was or was not resyllabified in the present study was not always clear. Consequently, a conservative approach was taken, so that doubtful cases were labeled as *resyllabified*. A second rater (*rater 2*) was asked to code some of the tokens for resyllabification or lack thereof. *Rater 2* was a phonetician and native speaker of Uruguayan Spanish who coded 20 % of all tokens in the study ($N = 480$, 30 tokens per speaker) after a training session with 100 tokens (number of tokens per speaker = 6–7). The labeling by *rater 1* was the one submitted to statistical analyses (see Section 6.3.2 for the results of the interrater agreement between the two raters).

Table 6.2. Participants' Bilingual Language Profile (BLP) scores, speaker ID, gender, age, state of origin, and highest level of formal education: lack of resyllabification and glottalization study.

BLP score	Speaker ID	Gender	Age	State of origin	Highest level of formal education
-46	inf17	female	30	Quintana Roo	high school
-26	inf15	female	43	Quintana Roo	university
-24	inf18	female	38	Quintana Roo	university
-20	inf20	male	47	Quintana Roo	high school
3	inf21	male	40	Quintana Roo	university
10	inf14	male	46	Quintana Roo	high school
15	inf16	male	57	Quintana Roo	high school
35	inf25	female	44	Quintana Roo	high school
96	inf08	male	72	Yucatán	secondary education
99	inf23	female	37	Yucatán	high school
142	inf24	male	47	Quintana Roo	high school
160	inf29	male	66	Yucatán	secondary education
177	inf13	female	36	Quintana Roo	university
179	inf12	male	34	Yucatán	university
195	inf11	female	67	Yucatán	high school
202	inf07	female	54	Quintana Roo	university

Tokens were also coded by *rater 1* for presence of glottalization based on a visual inspection of spectrograms in Praat (Boersma & Weenink, 2019). In the majority of cases, glottalization was determined by an error in F0 tracking by the Praat algorithm with the pitch floor set at 75 Hz, and also by the visual appearance of the glottal pulses, which appeared further apart than in segments without glottalization (see Figure 6.1; cf. Figure 4.1).[1]

[1] There were 11 cases in which the vowels were preceded by a brief period of silence. These vowels were perceived as being preceded by glottal stops.

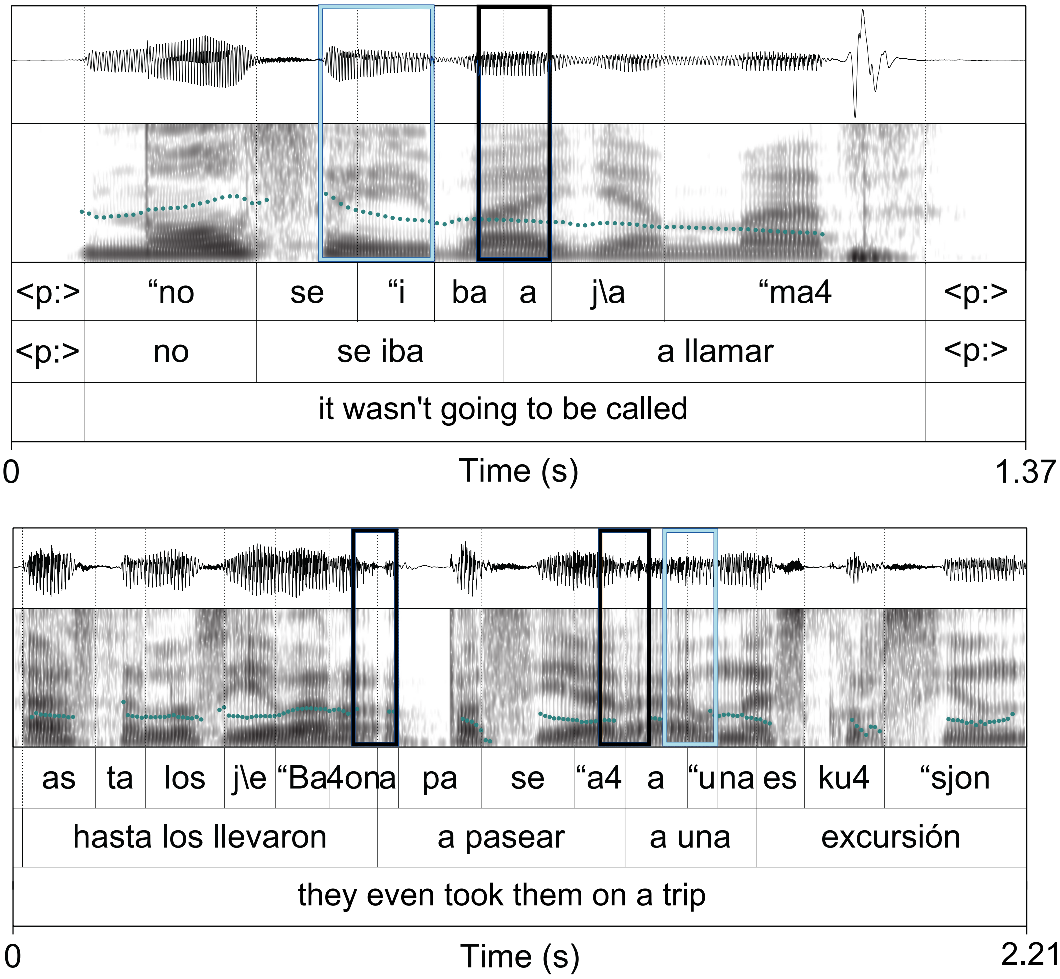

Figure 6.1. Examples of resyllabification (top) and lack of resyllabification (bottom) by speakers inf23 (female) and inf20 (male), respectively. All instances of lack of resyllabification present glottalization. The word-initial vowels appear within PWs (in light blue) or across PWs (in black).

6.2.3.2 Annotation of prosodic factors

Position in relation to the PW (within the PW or across PWs) and lexical stress (in a stressed or in an unstressed syllable) were annotated for word-initial vowels. As already noted in Section 6.1, for *within the PW*, tokens could belong to (i) a LW preceded by one or more unstressed FWs (e.g., ((*en*) *el* ***agua***)$_{\text{PW}}$ ‘(in) the water’), (ii) a stressed FW preceded by one or more unstressed FWs (e.g., ((*la*) *de* ***una***)$_{\text{PW}}$ ‘(that) of one.F’), or (iii) an unstressed FW preceded by another (e.g., (*en* ***el*** *mercado*)$_{\text{PW}}$ ‘in the market’). *Across PWs*, word-initial vowels could be those of (i) an unstressed FW in PW-initial position that was cliticized to a LW (e.g., (***en*** *el mercado*)$_{\text{PW}}$, (***el*** *agua*)$_{\text{PW}}$), (ii) a LW preceded by another LW (e.g., (*beben*)$_{\text{PW}}$ (***agua***)$_{\text{PW}}$ ‘(they drink water’), or (iii) a stressed FW that was not preceded by an unstressed FW (e.g., (*toman*)$_{\text{PW}}$ (***una***)$_{\text{PW}}$ ‘take-PRS.3PL one.F’).

Due to the unscripted nature of the recordings, it was impossible to predict whether there would be more instances of tokens within or across PWs, or in stressed or unstressed syllables. Tokens that appeared in stretches of speech with laryngealized voice quality or with background noise were excluded. Table B1 in Appendix B presents information about the number of tokens per speaker in relation to prosodic position, and whether tokens were resyllabified and glottalized or not.

6.2.4 Statistical analysis

Two Bayesian mixed-effects models were fitted with the `brms` package (Bürkner, 2017) in the R programming environment (R Core Team, 2020; cf. Section 5.2.6.2), one to investigate lack of resyllabification and another to investigate glottalization. The logistic mixed model fitted to examine lack of resyllabification included the dependent variable LACK OF RESYLLABIFICATION (treatment coded, reference level *resyllabified*; the other level was *not resyllabified*), the predictors PROSODIC POSITION (levels *within the PW* and *across PWs*), LEXICAL STRESS (levels *unstressed* and *stressed*), BLP SCORE (SCALED), and GENDER (levels *female* and *male*), with an interaction term for PROSODIC POSITION × LEXICAL STRESS, and a random intercept for SPEAKER. The logistic mixed model fitted to examine glottalization paralleled that for lack of resyllabification, with GLOTTALIZATION as the dependent variable (treatment coded, levels *no* and *yes*).

For both models, the likelihood function used was Bernoulli. Weakly informative priors were specified for the intercept (*Normal* (0, 3)), for the parameters representing the effects of the predictors and the interaction (*Normal* (0, 3)), and for the standard deviation of the random effect (*Normal* (0, 5)); see Section 5.2.6.2 for an explanation of these priors). The convergence of the models, as well as their fitting and the assessment of goodness of fit, were evaluated following the procedure described in Section 5.2.6.2. The package `emmeans` (Lenth, 2020) was used to obtain estimated marginal means and to conduct pairwise comparisons of the levels of the predictors of the interaction.

For the study on lack of resyllabification, the interrater agreement on the coding of 20 % of the data set was calculated by means of the function *kappaFleiss* from the package `KappaGUI` (Santos, 2018).

6.3 Results

6.3.1 Overview of the data

A total of 406 tokens out of 2400 tokens (i.e., 16.92 %) were not resyllabified. Figure 6.2 shows the percentage of realizations with lack of resyllabification *within the PW* and *across PWs* (versus the percentage of resyllabification *within the PW* and *across PWs*, respectively, not shown in the figure) for each speaker, represented on the x-axis by BLP scores. Participants inf23 and inf07, whose BLP scores are 99 and 202, respectively, did not produce any instances of lack of resyllabification *within the PW*.

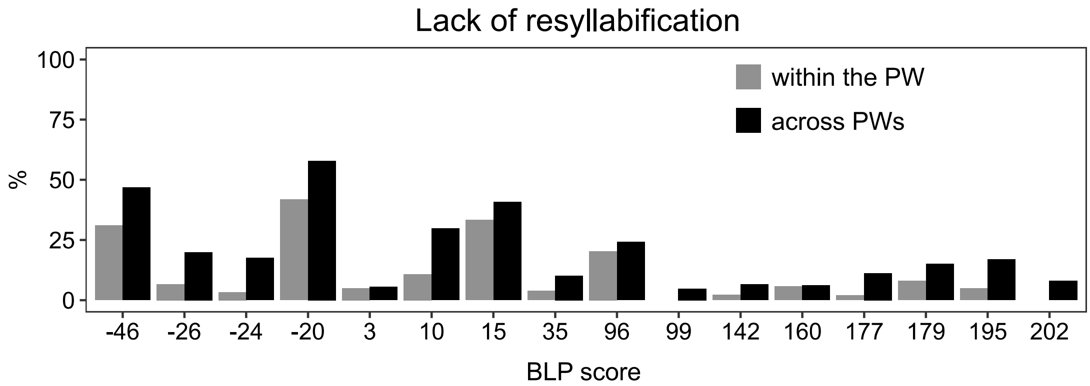

Figure 6.2. Realizations with lack of resyllabification (in %) *within the PW* and *across PWs* per speaker. Each BLP score on the x-axis corresponds to one speaker.

A total of 405 tokens (i.e., 16.88 % of the total number of tokens) presented cues to glottalization. Figure 6.3 shows the percentage of glottalized realizations *within the PW* and *across PWs* for each speaker, represented on the x-axis ordered by BLP scores.

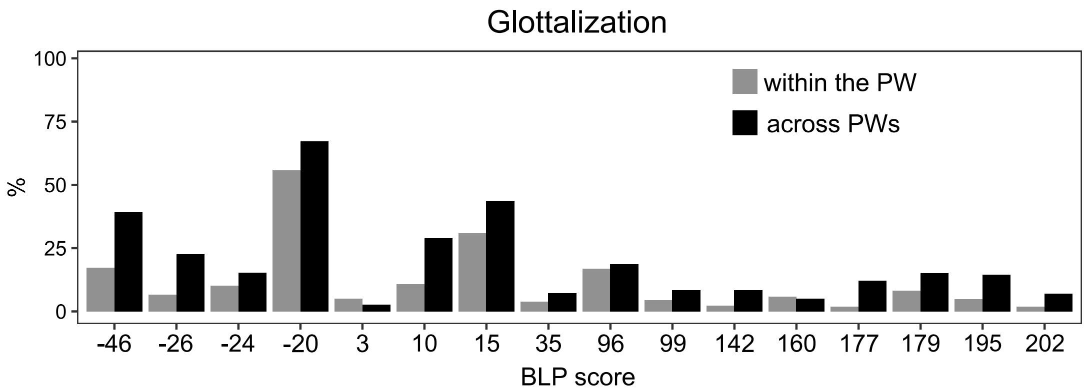

Figure 6.3. Glottalized realizations (in %) *within the PW* and *across PWs* per speaker. Each BLP score on the x-axis corresponds to one speaker.

The total number of tokens with lack of resyllabification (406) and the total with glottalization (405), and also the similarities between Figures 6.2 and 6.3, may be misleading in that it may seem as if all glottalized tokens were also tokens with lack of resyllabification: this was not the case. Nevertheless, a high number of tokens with lack of resyllabification were also glottalized (346 out of 406, i.e., 85.22 %), which indicates a sizeable overlap between the two. Additionally, of the total of 1994 resyllabified tokens, 59 presented glottalization. Conversely, of the total of 406 tokens with lack of resyllabification, 60 did not present glottalization. In other words, for 119 tokens (4.96 %) it is not the case that lack of resyllabification corresponds to glottalization. Table 6.3 presents additional information about the relationship between (lack of) resyllabification and glottalization.

Table 6.3. Resyllabified/not resyllabified and glottalized/not glottalized tokens according to prosodic position. The percentages shown refer to the total number of tokens in the data set ($N = 2400$).

resyllabified		**not resyllabified**	
within PW (*n* = 749)		**within PW (*n* = 89)**	
glottalized	not glottalized	glottalized	not glottalized
21 (0.88 %)	728 (30.33 %)	71 (2.96 %)	18 (0.75 %)
across PWs (*n* = 1245)		**across PWs (*n* = 317)**	
glottalized	not glottalized	glottalized	not glottalized
38 (1.58 %)	1207 (50.29 %)	275 (11.46 %)	42 (1.75 %)

6.3.2 Lack of resyllabification

In order to examine whether there was a main effect of PROSODIC POSITION, a pairwise comparison of the levels of the predictor was obtained. The results show evidence that the levels *within the PW* and *across PWs* differ from each other since 0 is not included in the 95 % CI of the difference estimate (-0.91, CI [-1.2, -0.63]), with a smaller probability of lack of resyllabification in the former. The same procedure was followed to examine whether there was a main effect of LEXICAL STRESS. The results show that there is evidence that the levels *unstressed* and *stressed* differ from each other (-0.68, CI [-0.96, -0.39]), with a smaller probability of lack of resyllabification in the former.

Estimated marginal means and pairwise comparisons for the levels of the predictors of the interaction (PROSODIC POSITION × LEXICAL STRESS) were obtained (see Table B2 in Appendix B). Figure 6.4 is a plot of the conditional effects of the interaction between PROSODIC POSITION and LEXICAL STRESS. The results of the pairwise comparisons show that the probability of lack of resyllabification for the level *within the PW* × *unstressed* is smaller than that of the other combinations of PROSODIC POSITION and LEXICAL STRESS, since 0 is not included in the CI. *Within the PW* × *stressed* has a smaller probability of lack of resyllabification than *across PWs* × *stressed*, whereas for the other combinations, 0 is included within the CI, which indicates that there is no clear evidence for greater or smaller probability of lack of resyllabification for those comparisons.

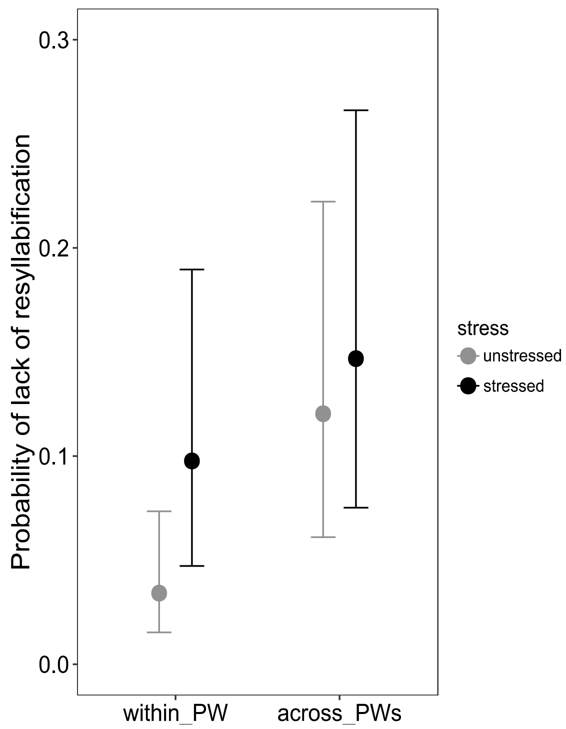

Figure 6.4. Interaction plot for lack of resyllabification in the corpus speech study. Effect of PROSODIC POSITION conditional on LEXICAL STRESS.

The results for the BLP SCORE (SCALED) predictor indicate that 0 is not included within the CI ($\hat{\beta}$ = -0.57, CI [-1.09, -0.06]). Based on the posterior samples, the estimated probability of the parameter being less than 0 is 0.98. This suggests that there is evidence that the probability of lack of resyllabification decreases the higher the BLP SCORE is, that is, with Yucatecan Spanish dominance (cf. Figure 6.2).

The results for the GENDER predictor indicate that 0 is included within the CI ($\hat{\beta}_{genderM}$ = 0.49, CI [-0.54, 1.51]). Based on the posterior samples, the probability of the parameter being greater than zero is 0.84, that is, there is a 0.84 probability that *males* resyllabify less than *female* speakers, given the available data and the model. Because most of the probability mass is above 0, there may be some weak evidence for a difference between the two levels, but we cannot be certain about this.

To sum up, the combination *within the PW* × *unstressed* differs considerably in that the probability of resyllabifying is higher than in the other combinations of the factors.

Additionally, for *within the PW* × *stressed*, the probability of resyllabifying is higher than for *across PWs* × *stressed*. There is evidence in support of more Yucatec Maya dominant speakers resyllabifying less often, but there is no clear support for differences between *female* and *male* speakers in terms of resyllabification.

So far, we have seen the results for the statistical model. As indicated in Section 6.2.3.1, part of the data were also coded for resyllabification or lack thereof by a second rater. Interrater agreement between *rater 1* and *rater 2* for 20 % of the data set was *moderate*: Fleiss' κ = 0.46 (see Landis & Koch, 1977, for the strength of the agreement associated with values of the statistic). *Rater 2* coded more tokens as *not resyllabified* than *rater 1* (*rater 1*: 91, *rater 2*: 171), which may indicate that *rater 1* adopted a more conservative approach to coding lack of resyllabification.

6.3.3 Glottalization

In order to examine whether there was a main effect of PROSODIC POSITION, we obtained a pairwise comparison of the levels of the predictor. The results show evidence that the levels *within the PW* and *across PW* differ from each other, since 0 is not included in the 95 % CI of the difference estimate (-0.78, CI [-1.06, -0.5]), with a smaller probability of glottalization in the former. The same procedure was followed to examine whether there was a main effect of LEXICAL STRESS. The results show that there is evidence that the levels *unstressed* and *stressed* differ from each other (-0.44, CI [-0.73, -0.17]), with a smaller probability of glottalization for the former.

Estimated marginal means and pairwise comparisons for the levels of the predictors of the interaction (PROSODIC POSITION × LEXICAL STRESS) were obtained (see Table B3 in Appendix B). The results for the pairwise comparisons show that the probability of glottalized realizations for the level *within the PW* × *unstressed* is smaller than that of the other combinations of PROSODIC POSITION and LEXICAL STRESS, since 0 is not included in the CI. For the remaining combinations, 0 is included in the CI, which means that there is no clear evidence for greater or smaller probability of glottalization for those comparisons. For the contrast between *across PWs* × *unstressed* – *within the PW* × *stressed* most of the probability mass is above 0, which indicates that it might be reasonable to conclude that there is weak evidence for a difference between the two, with tokens in *across PWs* × *unstressed* having a slightly greater probability of being glottalized, but we cannot be certain about this (cf. Figure 6.5).

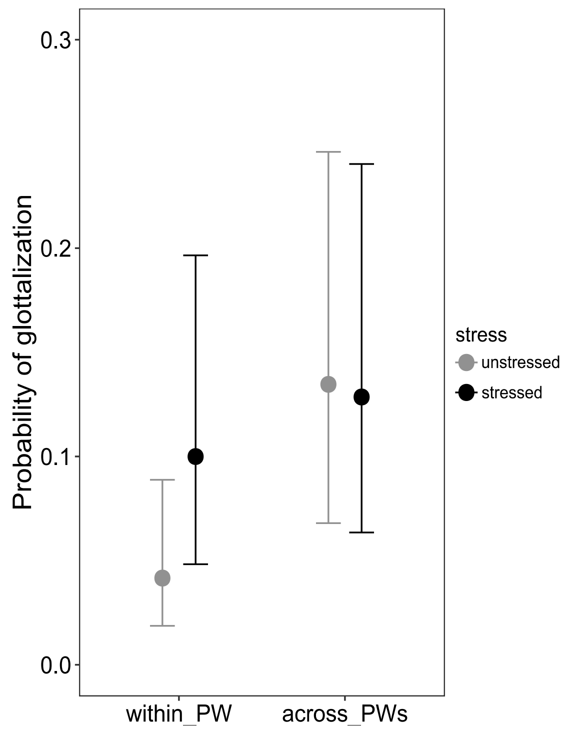

Figure 6.5. Interaction plot for glottalization in the corpus speech study. Effect of PROSODIC POSITION conditional on LEXICAL STRESS.

The results for the BLP SCORE predictor indicate that 0 is not included within the CI ($\hat{\beta}$ = -0.55, CI [-1.07, -0.02]). Based on the posterior samples, the estimated probability of the parameter being less than 0 is 0.98. This suggests that there is evidence that the probability of glottalization decreases the higher the BLP SCORE is, that is, with Yucatecan Spanish dominance (cf. Figure 6.3).

The results for the GENDER predictor indicate that 0 is included within the CI ($\hat{\beta}_{genderM}$ = 0.44, CI [-0.60, 1.45]). Based on the posterior samples, the posterior probability of the parameter being greater than zero is 0.81, that is, there is a 0.81 probability that *males* use glottalization more than *female* speakers. Thus, there may be some weak evidence for a difference between the two levels, but we cannot be certain about this.

To sum up, the level *within the PW* × *unstressed* differs from the other levels in that the probability of glottalization is lower for it. There is evidence in support of more Yucatec Maya dominant speakers using glottalization more often, and there is no clear support for differences in terms of glottalization between *female* and *male* speakers.

6.4 Discussion

The results of the Bayesian models for lack of resyllabification and glottalization, which mirror each other to a great degree, show that most of them agree with the expectations introduced in Section 6.1. Table 6.4 presents a comparison between the results expected and those obtained. The results for PROSODIC POSITION show that the probability of lack of resyllabification and glottalization is more frequent *across PWs* than *within the PW*, in line with the works on Spanish mentioned in Section 6.1, which found that lack of resyllabification is more frequent across LW + LW than across (unstressed) FW + LW sequences. Also in accordance with the works on Spanish and other languages mentioned in Section 6.1, the results for LEXICAL STRESS show that the probability of lack of resyllabification and glottalization is also more frequent for word-initial vowels in *stressed* syllables than in *unstressed* ones.

Table 6.4. Results for lack of resyllabification and glottalization of vowels in word-initial position in the corpus study.

Variables	Expected results	Results
1. Prosodic position	*across PWs* > *within the PW*	✓
2. Lexical stress	*stressed* > *unstressed* syllables	✓
3. Prosodic position × lexical stress	*across PWs* × *stressed* > all other	*within the PW* × *unstressed* < all other
4. Language dominance	Yucatec Maya > Yucatecan Spanish	✓
5. Gender	no effect	? *males* > *females*

These results need to be interpreted in light of the results of the interaction, which clearly show that the level *within the PW* × *unstressed*, which presents the least probability of lack of resyllabification and glottalization, differs from the other levels, which tend to group together. Tokens in the *within the PW* × *unstressed* level appear in (i) LWs or stressed FWs whose first syllable is unstressed and which are preceded by one or more FWs (e.g., (*en el **artículo***)$_{PW}$ 'in the article'), and (ii) unstressed FWs preceded by other unstressed FWs, with which they cliticize to LWs or stressed FWs in order to form a PW (e.g., *(en **el** mercado)*$_{PW}$ 'in the market', *(que **a** veces)*$_{PW}$ 'that some times'). The results suggest that unstressed FWs form a close connection with LWs/stressed FWs; importantly, this is only the case if the first syllable of the LW/stressed FW is unstressed; otherwise (e.g., (*en el **agua***)$_{PW}$ 'in the water'), the probability of lack of resyllabification and glottalization increases considerably. This result can be interpreted as LEXICAL STRESS having a strong effect *within the PW*. However, *across PWs*, there appears to be no differences for the levels *unstressed* and *stressed*. In sum, if we take the claim by Michnowicz and Kagan (2016) that glottalization helps mark word boundaries by blocking resyllabification and we apply it to the PW domain, the claim seems to receive support from the fact that glottalization and lack of resyllabification are more frequent at the left edge of the PW (i.e., *across PWs*) than *within the PW*. However, LEXICAL STRESS has an effect on lack of resyllabification and glottalization *within the PW*, whereas at the left edge of the PW it seems that PROSODIC POSITION alone is enough to favor lack of resyllabification and glottalization.

The results for LANGUAGE DOMINANCE (measured by means of the BLP SCORE) are in line with the expectations: the greater the Yucatec Maya dominance, the greater the probability of lack of resyllabification and glottalization. Conversely, GENDER appears to have no effect on the probability of lack of resyllabification and glottalization, although there may be some weak evidence in support of a greater probability of lack of resyllabification and glottalization for *males*, but we cannot be certain about this.

The results for LANGUAGE DOMINANCE deserve further commentary. It is unlikely that the Speech Learning Model (SLM; Flege, 1995, 1999, 2002; see Section 3.1.1) could be employed to make predictions about the glottalization of Yucatecan Spanish vowels. The SLM has three important limitations in the context of the present study: (i) it is based on "phonetic categories" (Flege, 1995, p. 239; cf. Section 3.1.1), that is, it does not consider prosodic phenomena, (ii) only vowels in word-initial position are examined in the present study (vs. independent from prosodic position), and (iii) participants are not grouped into monolinguals and bilinguals, as is common practice in studies that make use of the SLM (see Section 3.1.1). In terms of assimilation, a new merged L1–L2 category for Yucatecan Spanish vowels could be hypothesized. However, the literature on Yucatecan Spanish indicates that glottalization takes place in word-initial position, and even when there is glottalization, vowels are still more likely to be produced without it (see below). Moreover, the existence of short vowels in Yucatec Maya that are similar to the Spanish vowels may preclude the creation of a merged L1–L2 category between Yucatec Maya creaky vowels and Yucatecan Spanish vowels. In terms of dissimilation, it seems unlikely that the L2 speakers of Spanish have created a new /ʔ/ category. Had they done that, we would expect a higher rate of glottalization and no effect of PROSODIC POSITION. In sum, although the greater the Yucatec Maya dominance is, the greater the probability of lack of resyllabification and glottalization, the SLM model cannot account for these phenomena.

The results show that a sizeable number of tokens were not resyllabified (16.92 %) and present glottalization (16.88 %), which is in line with the study on Yucatecan Spanish by Michnowicz and Kagan (2016), who report 11 % instances of glottalization (and, in their view, of lack of resyllabification as well). Because both their study and the present one are based on corpora of sociolinguistic interviews, the ratio of lack of resyllabification and glottalization in Yucatecan Spanish may be around the percentages obtained, at least for the speech style found in sociolinguistic interviews.

The moderate agreement for lack of resyllabification between the two raters suggests that its coding based on auditory inspection may be a good approximation to the phenomenon, but not entirely reliable. Arguably, the fact that the data come from spontaneous speech is at the root of the disagreement between the raters due to the inherent variability of spontaneous speech. More importantly, it is unclear which were the acoustic cues used by the raters to determine the syllabic affiliation of word-final vowels and consonants. For example, speakers may use several strategies to avoid resyllabification (cf. Section 4.3.4). Glottalization seems an important cue to lack of resyllabification, a claim that is supported by the fact that most tokens with lack of resyllabification coded by rater 1 were also glottalized, but there was also a small number of tokens coded as lacking resyllabification that were not glottalized. Consequently, the results from the present study must be interpreted as preliminary.

The present study has explored some linguistic and some speaker-specific factors that may help explain the distribution of lack of resyllabification and glottalization in the PW domain. However, previous studies on Spanish indicate that other factors may also be involved. These factors include the quality of the vowel in word-final and word-initial positions (Alba, 2006; Aguilar, 2003, 2005; Michnowicz & Kagan, 2016), previous mention (Alba, 2006), lexical frequency (Alba, 2006), or age (Michnowicz & Kagan, 2016). In the next section, I explore the weight of some of these factors, as well as others, by means of a post hoc study using random forests.[2]

6.5 A post hoc analysis with random forests

6.5.1 Random forests

Random forests are an exploratory tool that can help find undiscovered patterns in the data. The main appeal of using random forests is that a large number of predictors can be included, even if they are collinear and nonorthogonal (Tagliamonte & Baayen, 2012). A random forest is a set of conditional inference trees that have been obtained through recursive partitioning. Thus, the trees provide "estimates of the likelihood of the value of the response variable . . . based on a series of binary questions about the values of the predictor variables" (Tagliamonte & Baayen, 2012, p. 159). Figure 6.6 illustrates this point by means

[2] The random forest analysis was inspired by the one for variation in glottalization in Italian by Di Napoli (2018), Section 6.

of a classification tree[3] that predicts lack of resyllabification on part of the data set examined in the previous sections.

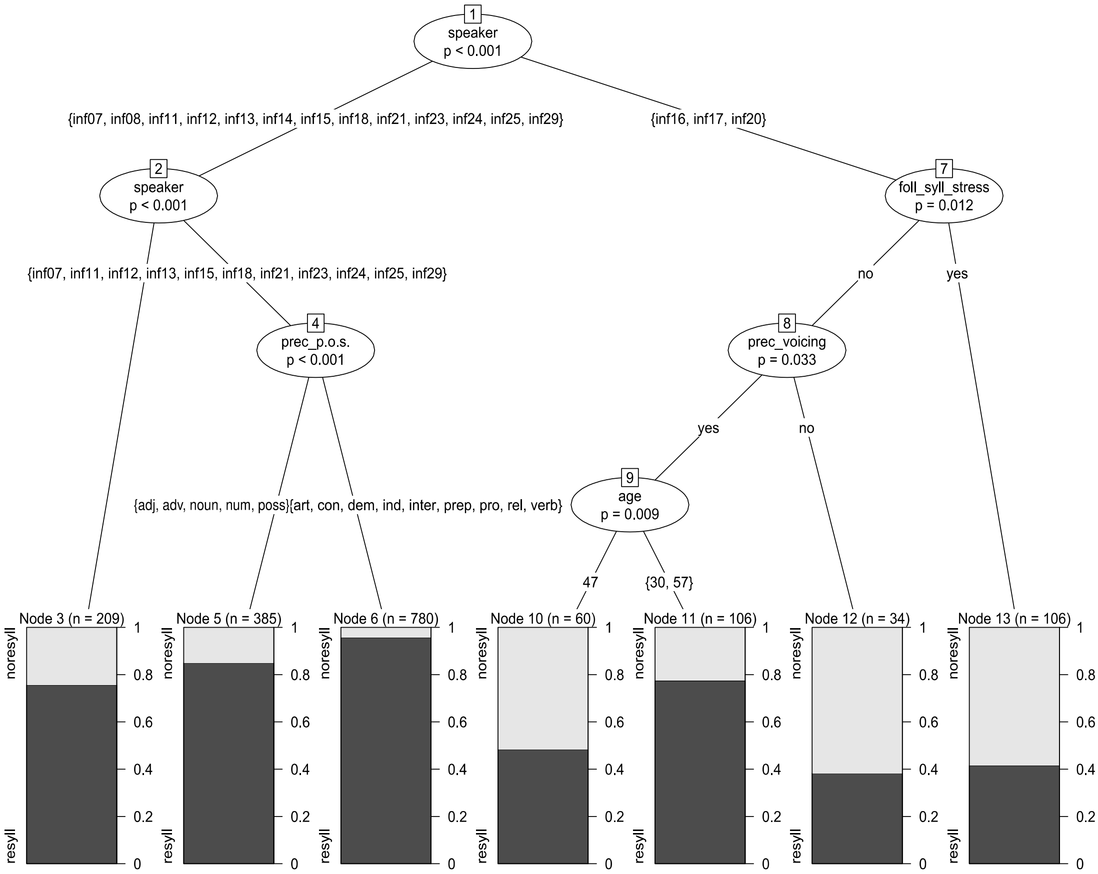

Figure 6.6. Classification tree predicting lack of resyllabification. See text for an explanation.

The tree was grown using the function *ctree* in the package `party` (Strobl, Hothorn, & Zeileis, 2009). In this classification tree, there are several variables that will be included in the training set of the lack of resyllabification random forest, which will be explained in Section 6.5.2. In this tree, from the total of 16 speakers, 13 speakers are grouped into one node (node 2), which includes 1374 tokens. This group is characterized by having more instances of resyllabification than the other group of 3 *speakers* (the relative frequencies of *noresyll*, i.e., lack of resyllabification, and *resyll*, i.e., resyllabification, are indicated in the bars). In this group (node 2), the tokens are further split into node 3 and node 4. Node 3 includes the *speakers* with the highest frequency of lack of resyllabification (within node 2), whereas node 4 includes the *speakers* with the lowest frequency of lack of resyllabification (within node 2). Node 4 is further split into nodes 5 and 6 on the basis of PRECEDING PART

[3] *Classification trees* correspond to categorical dependent variables, whereas *regression trees* correspond to continuous dependent variables (Crawley, 2013, p. 771).

OF SPEECH, with node 5 including five levels of the variable and node 6 including the other nine. The *p* values correspond to the partitioning at each node.

6.5.2 Methods

Two random forests were grown in order to explore other factors that may be involved in lack of resyllabification and glottalization. With the first one, the *lack of resyllabification forest*, I examine the importance of several variables in order to predict whether tokens would present lack of resyllabification or not ($N = 2400$, of which 406 present lack of resyllabification; see Section 6.3.1). The goal of the second random forest, the *glottalization forest*, is the same as that of the lack of resyllabification forest, but the data set used is different. For the glottalization forest, the data points are only those that were coded as *glottalized* in the study of glottalization presented in the previous sections ($N = 405$, of which not resyllabified = 346, resyllabified = 59; see Section 6.3.1). In this manner, I investigate predictors that may be relevant in determining whether a glottalized token is cueing lack of resyllabification.

The dependent variable (LACK OF RESYLLABIFICATION, with levels *lack of resyllabification* and *resyllabification*) and the predictors were the same for both forests. The 19 predictors were linguistic and speaker-specific predictors, some of which (PROSODIC POSITION, LEXICAL STRESS, BLP SCORE (SCALED), GENDER, and SPEAKER) had already been included in the Bayesian analyses discussed in the previous sections.

The linguistic variables included in the random forests were:

- PROSODIC POSITION (*within the PW, across PWs*)
- PRECEDING SYLLABLE STRESS (*no, yes*)[4]
- FOLLOWING SYLLABLE STRESS (*no, yes*)
- PRECEDING SYLLABLE TYPE (CV, CVC, CCV, etc.)
- FOLLOWING SYLLABLE TYPE (CV, CVC, CCV, etc.)
- ARTICULATION RATE (ratio of syllables per second excluding pauses)
- PRECEDING WORD STRESS (*no, yes*)
- FOLLOWING WORD STRESS (*no, yes*)

[4] In the previous sections, *lexical stress* was used to refer to lexical stress on the *word-initial* syllable. However, because in the random forests the predictor lexical stress on the *word-final* syllable is also included (referred to here as PRECEDING SYLLABLE STRESS), the former *lexical stress* variable is referred to here as FOLLOWING SYLLABLE STRESS. Additionally, all instances of *preceding* in the next variables refer to the sound, syllable, or word before the coded vowels, whereas instances of *following* refer to the sound, syllable, or word in which the coded vowels appeared.

- PRECEDING WORD CLASS (*FW*, *LW*)
- FOLLOWING WORD CLASS (*FW*, *LW*)
- PRECEDING PART OF SPEECH (adjective, adverb, article, conjunction, interjection, demonstrative, interrogative, noun, numeral, possessive, preposition, pronoun, relative, verb)[5]
- FOLLOWING PART OF SPEECH (see previous variable)
- PRECEDING VOICING (of the word-final sound: *voiced, unvoiced)*
- PRECEDING SOUND TYPE (*consonant, vowel*)
- FOLLOWING SOUND TYPE (/i e a o u/)

The speaker-specific variables included in the forests were:

- BLP SCORE (SCALED)
- GENDER (*female, male*)
- AGE
- SPEAKER (16 speakers)

It is evident that many of the variables are highly collinear (e.g., PRECEDING SYLLABLE STRESS and PRECEDING WORD STRESS). As indicated above, random forests allow us to examine such variables.

The random forests were fitted with the `ranger` package (Wright & Ziegler, 2017) in R (R Core Team, 2020). This package makes it possible to grow forests with a large number of trees. Following Baumann and Winter (2018), each forest was grown with a random subset of 70 % of the data (henceforth, *trained set*). Its predictions were then tested against the predictions of the remaining 30 % of the data (*untrained set*). The value for *mtry* (i.e., the number of predictors that are drawn to grow a tree) was set to 4, that is, the closest number to the square root of the number of predictors, following Strobl, Malley, and Tutz (2009). 50,000 trees were grown. Variable importance was calculated via permutation tests (Strobl, Malley, & Tutz, 2009).

6.5.3 Results

Results for lack of resyllabification and glottalization are shown graphically in Figures 6.7 and 6.8, respectively. The figures show the conditional permutation of variable importance of each predictor in relation to the others. Because the importance of explaining the data of each predictor is computed relative to that of the other predictors, it is best practice to

[5] Words were coded into these parts of speech following Quilis (1999) and Hualde (in press).

provide a description of their ranking instead of providing the resulting values (Strobl, Malley, & Tutz, 2009).

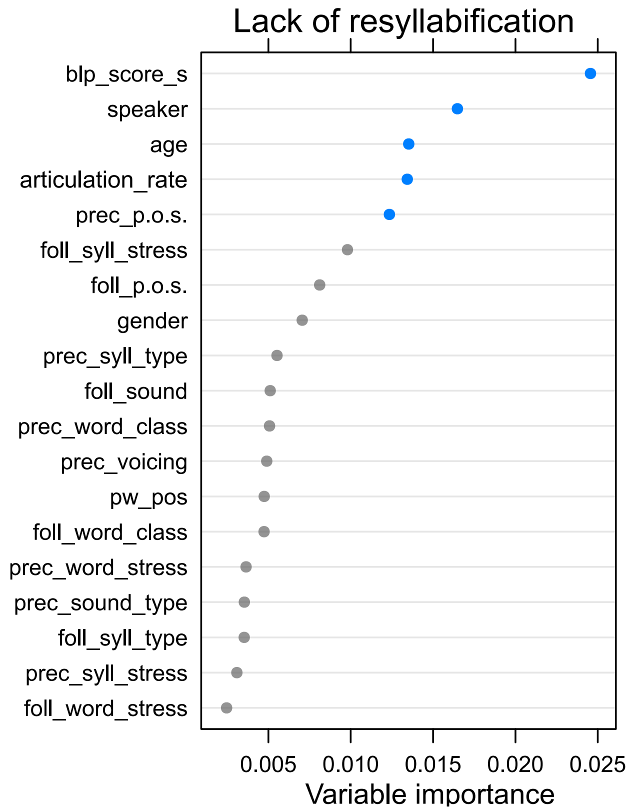

Figure 6.7. Variable importance of predictors in the lack of resyllabification random forest. The five most important predictors appear in blue.

For the lack of resyllabification forest (see Figure 6.7), five predictors turned out to be the most important: BLP SCORE (SCALED), SPEAKER, AGE, ARTICULATION RATE, and PRECEDING PART OF SPEECH, followed by FOLLOWING SYLLABLE STRESS (i.e., *lexical stress* in the Bayesian models) and the other predictors. These results are not surprising because the five predictors present high variability, meaning that there are quite different BLP SCORES, different AGES (cf. Table 6.2), and different patterns of lack of resyllabification among SPEAKERS. The same can be said about the ARTICULATION RATE, which was in the range of 4.7–6.42 syllables per second, and about the PRECEDING PART OF SPEECH.

The correlation matrix that results from the comparison of the trained set ($n = 1680$) with the untrained set ($n = 720$) shows that the trained set can predict 94.62 % of the resyllabified tokens (563 out of 595), but only 28.8 % of the tokens that present lack of resyllabification (35 out of 125). Additionally, the correlation between the two sets ($r = 0.30$) indicates that the variables included in the trained set can predict to some extent the presence or lack of resyllabification in the untrained set. This suggests that there may be other factors, not included in the forest, which are involved in the distribution of lack of resyllabification.

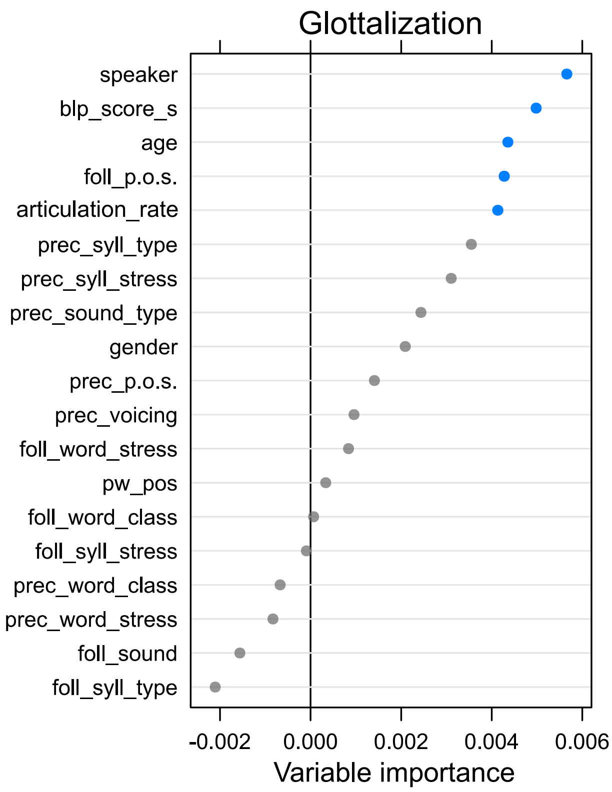

Figure 6.8. Variable importance of predictors of the glottalization random forest. Predictors with values higher than 0 are significant. The five most important predictors appear in blue.

For the glottalization forest, four out of the five most important predictors were also among the five most important in the lack of resyllabification forest (SPEAKER, BLP SCORE (SCALED), AGE, and ARTICULATION RATE). The remaining predictor, FOLLOWING PART OF SPEECH, had some importance in the lack of resyllabification forest. The predictors with values lower than 0 are inconsequential (see Figure 6.8). The correlation matrix that results from the comparison of the trained set ($n = 283$) with the untrained set ($n = 122$) shows that the trained set can predict 97.17 % of the tokens that present lack of resyllabification (102 out of 106), but only 18.75 % of the resyllabified ones (3 out of 16). The correlation between the two sets ($r = 0.06$) indicates that the variables included in the trained set can only slightly predict the presence or absence of resyllabification in the untrained set. This result is not surprising given the small number of tokens and that most tokens presented lack of resyllabification.

6.5.4 Discussion

The two random forests have provided some clues about the importance of several predictors in the data. Whereas the role of BLP SCORE and SPEAKER has already been considered in the present chapter by means of Bayesian modeling, AGE and ARTICULATION RATE appear as predictors that may help to describe the distribution of lack of resyllabification and glottalization in Yucatecan Spanish. It is important to keep in mind that random forests serve to explore the data, not to explain it. While random forests may help discover patterns in the data, the variable importance that certain predictors may receive may be due to the variability that they present.

The low correlation values between trained and untrained sets point to the existence of other factors that may be involved in lack of resyllabification and glottalization. Considering the relatively large number of predictors in the forests grown in this study (even if collinear predictors are excluded), lack of resyllabification and glottalization appear to be highly complex phenomena.

6.6 Summary

The studies in this chapter have shown that lack of resyllabification of vowels in word-initial position is a rather frequent phenomenon in Yucatecan Spanish. One of the most important cues to lack of resyllabification is glottalization, to the point that models for lack of resyllabification and glottalization pattern together. The results of Bayesian analyses have indicated that both are phenomena related to the PW domain and that they may serve to mark the boundaries of PWs, that they are favored by lexical stress, and that they are favored by a higher Yucatec Maya dominance. Gender does not seem to be a relevant factor. While the study of all the cues to lack of resyllabification is out of the scope of this dissertation, the present study has provided some evidence for glottalization of word-initial vowels as a prosodic strengthening phenomenon to mark prosodic boundaries. Moreover, a random forest approach has shown that lack of resyllabification and glottalization of word-initial vowels are two complex phenomena in which many factors may be involved.

Part III

Read speech studies

Chapter 7

Read speech task

7.1 Overview

This chapter details the read speech task that was used to obtain the data sets that will be presented and analyzed in Chapter 8 (voiced stops) and Chapter 9 (glottalization of word-initial vowels), as well as the characteristics of the speakers, the recordings, and the phrase-boundary annotation carried out prior to further annotation and analysis of the data sets. The information on the task is provided as a distinct chapter and not as part of a *Methods* section in Chapters 8 and 9 for two reasons: (i) because the task, the participants, the recordings, and the annotation of phrase boundaries are the same for both studies, and (ii) due to the rather large amount of information concerning the design and presentation of the task that has to be given. The task was designed taking into consideration several of the variables that were included in Chapters 5 and 6 (prosodic position, lexical stress, and word class), as well as new variables (position of the carrier word in the sentence and repetition).

The current chapter is organized as follows. The design of the task is presented in Section 7.2, which includes an overview (Section 7.2.1) and the design characteristics related to voiced stops (Section 7.2.2) and vowels in word-initial position (Section 7.2.3). The presentation of the task to the participants is introduced in Section 7.3. The main characteristics of the participants who completed the task are presented in Section 7.4, while the characteristics of the acoustic recordings are provided in Section 7.5. Finally, the annotation of phrase boundaries is presented in Section 7.6.

7.2 Design

7.2.1 Overall design

The read speech task included a familiarization phase and three sets of declarative sentences (with 18 sentences each), all of which ended in the word *porque* 'because' (see Tables C1–C4 in Appendix C for the whole list of sentences). The sentences were designed so that they would include at least one instance of a target voiced stop or vowel, although an effort was made to include as many target sounds as possible (stops and/or vowels) in the same sentence in order to reduce the number of sentences and thus the time needed to complete the task. The number of tokens per sentence ranged between two (e.g., *Re**b**eca taló **á**rboles en la selva* 'Rebeca cut down trees in the rain forest') and four (***B**ernardo ensayó la **ó**pera en **e**l pa**b**ellón* 'Bernardo rehearsed the opera in the pavilion'). Within the same orthographic word, up to two tokens could be included (***bod**ega* 'cellar' and ***deb**ate* 'debate').

The target voiced stops and vowels were organized according to position of the carrier word in the sentence, prosodic position, lexical stress, and word class (word class was not included for the bilabial stop because there are no commonly used unstressed FWs with this sound in Spanish). These variables, along with repetition, will be considered in the studies both of voiced stops and of the glottalization of word-initial vowels. Prosodic position, lexical stress, and word class will be presented separately for voiced stops and vowels in Sections 7.2.2 and 7.2.3, respectively.

Position of the carrier word in the sentence refers to whether the word to which the target stop belonged appeared at the beginning of the sentence (i.e., as the subject, which always corresponded to a proper name[1]), in medial position (i.e., in the object or adverbial after the verb), or in final position (i.e., in the sentence-final adverbial or modifier). Moreover, because all sentences ended with the word *porque* 'because', no final-lengthening effects were expected on tokens that appeared in words in sentence-final position. All sentences (and thus, tokens) had a first and a second *repetition*. Carrier words in sentence-initial position, that is, proper names (e.g., *la **B**onita* 'the Cute One', *el **É**tico* 'the Ethical One') appeared three times per set, meaning that they were in fact mentioned six times each.

Vowels and lexical frequency were taken into account when designing the task, although not as variables. *Vowels* refers to the vowels flanking the voiced stops or to the vowels that constituted the target sounds in the study of vowels in word-initial position. Only three vowels (/e/, /a/, /o/) were used, for two reasons: (i) to control for possible vowel quality effects (see Sections 4.2.2.3 for voiced stops and 4.3.4 for glottalization of word-initial vowels), and (ii) because they are the most common vowels in Spanish (Rojo, 1990), which meant that a large number of different carrier words could be used while still controlling for vowel quality effects to a certain degree.[2] With respect to *lexical frequency*, an effort was made to include words that would be familiar to the participants, although the lexical frequency of lexical words (LWs) was not controlled for. Also, it goes without saying

[1] Proper names and also nicknames were used as subjects. For ease of expression, both proper names and nicknames are referred to as *proper names*.

[2] There is one exception: the second bilabial stop in *bi**b**erón* 'feeding bottle', which is one of the target words in the study of voiced stops, is preceded by /i/.

that function words (FWs) are more frequent than LWs. Several possibilities of accounting for lexical frequency were considered during the design of the task, such as using existing corpora or online search engines, but they were problematic. The existing corpora on Spanish either focus on European Peninsular Spanish (Alameda & Cuetos, 1995; Sebastián Gallés et al., 2000) or, when including Mexican Spanish, they either have a very limited number of entries (Real Academia Española, 1997–2019) or it is not possible to search for specific Mexican regions, in our case, the Yucatán Peninsula or the state of Quintana Roo (Davies, 2016). Using lexical frequency data of Mexican Spanish instead of Yucatecan Spanish has the disadvantage of potentially providing partial results for some Spanish words; this could be the case if there are Yucatec Maya counterparts to Spanish words that are more frequently used by all or some of the Yucatecan Spanish speakers, regardless of their knowledge of Yucatec Maya. Consequently, the decision was taken to have the list of sentences checked by the fieldwork collaborator, who made sure that all words were known in Yucatecan Spanish.

7.2.2 Voiced stops

The read speech task was designed in order to include tokens of the bilabial (see Table 7.1) and the dentialveolar voiced stops (see Table 7.2) in syllabic onset position in CV or CVC syllables and preceded by a vowel (see Section 7.2.1). In the study of voiced stops in Chapter 5, the velar voiced stop was excluded; consequently, it was excluded from the read speech task in order to facilitate the comparison of results with those of Chapter 5. The bilabial voiced stop in Spanish is represented orthographically by both <b> and <v> (cf. Section 4.2.2.1). Historically, /v/ existed in Spanish, although it disappeared completely in the 16th century. The pronunciation of <v> as a labiodental fricative in contemporary Spanish is considered a hypercorrection error due to past prescriptive works that recommended its pronunciation as such (Real Academia Española & Asociación de Academias de la Lengua Española, 2010, pp. 91–92). Furthermore, some read speech studies have noted that <v> may be produced as a labiodental fricative (Dmitrieva et al., 2015; Lavoie, 2001), although it is not clear which factors may have motivated that pronunciation. Consequently, only instances of <b> were included in the task in order to avoid this possibility.

Table 7.1. Target bilabial voiced stops (in boldface).

Position in the sentence	Prosodic position	Lexical stress	Carrier word
initial	**PW-initial**	yes	***B**erto*
		no	***B**ernardo*
	LW-initial	yes	*la **B**ola*
		no	*la **B**onita*
	LW-medial	yes	*Re**b**eca*
		no	*Dé**b**ora*
medial	**PW-initial**	yes	***b**ombas*
		yes	***b**elga*
		no	***b**ebidas*
		no	***b**oletos*
	LW-initial	yes	*la **b**áscula*
		yes	*la **b**oca*
		no	*la **b**ecerra*
		no	*la **b**otana*
	LW-medial	yes	*a**b**ejas*
		yes	*ta**b**aco*
		no	*el bi**b**erón*
		no	*el a**b**anico*
final	**PW-initial**	yes	***b**árbaro*
		yes	***b**ásica*
		no	***b**ajito*
		no	***b**arato*
	LW-initial	yes	*de **b**odas*
		yes	*de **b**óvedas*
		no	*a la **b**odega*
		no	*a la **b**asura*
	LW-medial	yes	*en la ca**b**eza*
		yes	*de de**b**ate*
		no	*en la ca**b**ecera*
		no	*en el pa**b**ellón*

Table 7.2. Target dentialveolar voiced stops (in boldface).

Position in the sentence	Prosodic position	Lexical stress	Word class	Carrier word
initial	**PW-initial**	yes	LW	***D**enis*
		no	LW	***D**elfín*
	LW-initial	yes	LW	*la **D**óberman*
		no	LW	*la **D**octora*
	LW-medial	yes	LW	*A**d**ela*
		no	LW	*A**d**elino*
medial	**PW-initial**	yes	LW	***d**ólares*
		yes	LW	***d**ados*
		no	LW	***d**olores*
		no	LW	***d**esfiles*
		no	FW	***d**e casa*
		no	FW	***d**e la iglesia*
		no	FW	***d**e la plaza*
		no	FW	***d**el avión*
	LW-initial	yes	LW	*la **d**osis*
		yes	LW	*la **d**alia*
		no	LW	*la **d**efensa*
		no	LW	*la **d**enuncia*
	LW-medial	yes	LW	*ma**d**era*
		yes	LW	*po**d**ólogo*
		no	LW	*gra**d**erío*
		no	LW	*me**d**allones*
final	**PW-initial**	yes	LW	***d**ébil*
		yes	LW	***d**ócil*
		no	LW	***d**oméstico*
		no	LW	***d**eprisa*
		no	FW	***d**e élite*
		no	FW	***d**e la luna*
		no	FW	***d**e orégano*
		no	FW	***d**e bóvedas*
	LW-initial	yes	LW	*de **d**anza*
		yes	LW	*de **d**edos*
		no	LW	*de **d**onantes*
		no	LW	*de **d**ebate*
	LW-medial	yes	LW	*a la bo**d**ega*
		yes	LW	*de a**d**orno*
		no	LW	*para pe**d**agogos*
		no	LW	*ca**d**avéricos*

Prosodic position refers to the position of the target word in the Prosodic Word (PW) domain. The three possible positions were the same as those used in Chapter 5, that is, LW medial, LW initial, and PW initial. The tokens belonging to the carrier words in initial position in the sentence and in PW-initial position at the same time (i.e., ***B**erto*, ***B**ernardo,* ***D**enis*, and ***D**elfín*) were ultimately excluded from the analyses in order to focus on the study of prosodic strengthening in the PW domain only, with all analyzed tokens appearing in phrase-medial position.[3]

Lexical stress refers to whether the token appeared in a stressed or an unstressed syllable. The unstressed syllables of LWs that contained target segments appeared before a stressed syllable (e.g., ***b**oletos* [boˈletos] 'tickets').[4] In terms of *word class*, all bilabial tokens were LWs, whereas dentialveolar tokens included LWs in all prosodic positions, as well as the FWs *de* 'of' and *del* 'of-the.M.SG' in PW-initial position. The FWs that included the voiced stop could precede a LW (e.g., ***d**e casa* 'from home') or another FW (e.g., ***d**e la iglesia* 'from the church'). In total, the task included 30 tokens for the bilabial stop[5] and 38 for the dentialveolar stop, or 28 tokens and 36 tokens, respectively, after excluding the tokens belonging to carrier words in initial position in the sentence and in PW-initial position at the same time.

7.2.3 Vowels in word-initial position

The vowels in word-initial position considered for the glottalization study were /e/, /a/, and /o/ (cf. Section 7.2.1). These vowels could be (i) preceded by a word that ended in a consonant (/l/, /n/, or /s/), (ii) preceded by a word that ended in a vowel, which could be /e/, /a/, or /o/ in unstressed syllables and /o/ in stressed syllables, or (iii) in absolute sentence-initial position (***É**dison,* ***E**lena,* ***e**l Guapo* 'the Handsome One', and ***e**l Güero* 'the Blonde One').

As already indicated in Section 7.2.1, tokens appeared in carrier words placed in sentence-initial, sentence-medial, and sentence-final positions (see Tables 7.3, 7.4, and 7.5, respectively). The tokens in *absolute* sentence-initial position were ultimately excluded from the analyses in order to focus on the study of prosodic strengthening in the PW domain

[3] In line with phonological descriptions of the realization of Spanish voiced stops (see Section 4.2.1) and phonetic studies that have shown that voiced stops are articulated as full stops in IP-initial position (see Section 4.2.2.1), the tokens belonging to to the carrier words ***B**erto*, ***B**ernardo,* ***D**enis*, and ***D**elfín* were produced primarily as stops (bilabial tokens: 244 stops, 4 approximant realizations; dentialveolar tokens: 229 stops, 3 approximant realizations).

[4] There is one exception: *Dé**b**ora*.

[5] The word ***b**árbaro* was produced as ***b**árbara* by some speakers. Instances of both were pooled together.

only, with all target vowels appearing in phrase-medial position.[6]

Table 7.3. Target vowels (in boldface) in carrier words in sentence-initial position.

Prosodic position	Word class (stressed)	Carrier word
PW-initial	LW (yes)	**É**dison
	LW (no)	**E**lena
	FW	**e**l Guapo
	FW	**e**l Güero
LW-initial	LW (yes)	el **É**tico
	LW (no)	el **E**nano

Table 7.4. Target vowels (in boldface) in carrier words in sentence-medial position.

Prosodic position	Word class (stressed)	Carrier word
PW-initial	LW (yes)	**á**rabe
	LW (yes)	**á**rboles
	LW (no)	**a**bejas
	LW (no)	**a**ctores
	FW	**e**n el colegio
	FW	**a** su madre
	FW	**a** la asociación
LW-initial	LW (yes)	al **á**rbitro
	LW (yes)	el **ó**rgano
	LW (yes)	la **ó**rbita
	LW (yes)	la **ó**pera
	LW (no)	el **o**ficio
	LW (no)	el **a**banico
	LW (no)	la **o**veja
	LW (no)	la **o**reja
across FWs	FW	en **e**l colegio
	FW	con **e**l sombrero
	FW	en **e**l parque
	FW	para **e**l concurso
	FW	para **e**l concierto
	FW	desde **e**l domingo

[6] In fact, the perceptual assessment of whether vowels in absolute sentence-initial position were glottalized or not proved difficult for both coders (cf. Section 7.6 for the annotation of tokens in phrase-initial position).

Table 7.5. Target vowels (in boldface) in carrier words in sentence-final position.

Prosodic position	Word class (stressed)	Carrier word
PW-initial	LW (yes)	**á**cido
	LW (yes)	**á**spera
	LW (no)	**a**bajo
	LW (no)	**a**bierto
	FW	**e**n el cajón
	FW	**a** los turistas
	FW	**a**l banquero
LW-initial	LW (yes)	en **a**gua
	LW (yes)	con **o**dio
	LW (yes)	de **é**lite
	LW (yes)	de **é**bano
	LW (no)	en **o**toño
	LW (no)	con **a**ceite
	LW (no)	de **o**régano
	LW (no)	de **a**zúcar
across FWs	FW	en **e**l pabellón
	FW	con **e**l vestido
	FW	en **e**l cajón
	FW	para **e**l bebé
	FW	para **e**l cabello
	FW	desde **e**l amanecer

The target vowels could appear in three prosodic positions in relation to the PW: *PW-initial position, LW-initial position,* and *across FWs.* These positions are not equivalent to those in the previous study of vowels in Chapter 6, nor are the FWs included (in Chapter 6, stressed FWs were also considered). Tokens in *PW-initial position* appeared either at the beginning of a LW (e.g. ***a**bejas* 'bees') or of the FWs ***a*** 'to', ***e**l* 'the-M.SG,' or ***e**n* 'in'. These FWs were followed by either a LW (e.g., *Guapo* in ***e**l Guapo*) or by another FW plus a LW (e.g., *los turistas* in ***a** los turistas* 'to the tourists'). Tokens in *LW-initial position* appeared at the beginning of a LW preceded by one FW (e.g., *en **a**gua* 'in water'). Finally, tokens that appeared *across FWs* corresponded to the vowel of the FW *el*, which was placed between another FW and a LW (e.g., *en **e**l colegio* 'in the school', *con **e**l sombrero* 'with the hat'). Tokens across FWs appeared only in carrier words in sentence-medial and sentence-final positions.

Lexical stress refers to whether the token appeared in a stressed or an unstressed syllable. Since all FWs in the task were unstressed, this variable only applies to LWs. The task included the same number of tokens in stressed and in unstressed syllables of LWs. The unstressed syllables that contained target segments appeared before a stressed syllable, not after (e.g., ***o**reja* [oˈɾexa] 'ear').[7]

In total, there were 48 tokens in the task, 28 of which corresponded to vowels in LWs and 20 in FWs, or a total of 44, and 26 corresponding to vowels in LWs and 18 corresponding to vowels in FWs, after excluding the tokens belonging to carrier words in initial position in the sentence and in PW-initial position at the same time.

7.3 Presentation

The task was presented to the participants by the fieldwork collaborator, a Yucatec Maya–Yucatecan Spanish speaker who also recruited the participants. Before that, the author of the dissertation had prepared the technical equipment and then left the room. During the familiarization phase, the participants were instructed to talk naturally. They were also instructed to first read each sentence to themselves to make sure that its meaning was clear to them, and also to ask questions whenever they felt it necessary. Then the task proper would begin. The sets were presented to the participant by means of paper cards. Each of the three sets of sentences was assigned a different color. The collaborator would ask the participant to choose which color to start (and continue) with, thus providing a randomization of the sets. The participant would take a break after each set.

Participants had to read the same sentence twice, the first time without adding any extra words (e.g., *Rebeca taló árboles en la selva porque…* 'Rebeca cut down trees in the rain forest because…'), and the second time finishing it in whichever way they preferred (*Rebeca taló árboles en la selva porque necesitaba leña* 'Rebeca cut down trees in the rain forest because she needed wood'). Tokens in the *completion part* (i.e., after *porque*) were not included in the data sets analyzed in Chapters 8 and 9.

The sentences of a given set were presented to the participants in five stacks of cards. The first stack was made up of pictures of people with their names written at the bottom of the cards, such as, for example, *Rebeca* (see Table C5 in Appendix C). There were six proper names per set, which meant that each name was used in three sentences. For each set, a different group of proper names was used. Cards in the second stack had a verb written on

[7] There is one exception: the token in ***a**banico* 'fan', which appears two syllables before the stressed syllable ([abaˈniko]).

them (e.g., *taló*); the verb was always in the past preterite form, 3rd person singular. Cards in the third stack had an object and/or an adverbial written on them (e.g., *árboles*), whereas those in the fourth one had an adverbial or a modifier (e.g., *en la selva*). Cards in the fifth stack had the word *porque* written on them. Either the participant or the collaborator would turn the cards upside down; whether it was the collaborator or the participant doing so was decided by the participant in advance. The collaborator had been instructed beforehand that either way was possible as long as the participant had sufficient time to read the sentence and avoided making noises by manipulating the cards during the recording.

In the original design of the task, there was a memorization part before the reading of the sentences in order to bring some spontaneity to the task and to prevent participants from getting bored, thus avoiding a repetitive reading style. Therefore, participants were expected to memorize the proper names of the six people that would appear in each set of cards, and then the task proper would start. However, although the pilot study had shown that this was indeed a valid approach,[8] the cognitive load of the task turned out to be too high, causing the first participant to hesitate and to ask the collaborator on multiple occasions whether the name said out loud was the right one. As a result, I decided that participants would be given the people' images with the proper names written on them. In the end, the completion part proved challenging enough that some degree of spontaneity was achieved and a repetitive reading style avoided overall.

7.4 Speakers

A total of 26 speakers were recorded (12 female, 14 male). Eighteen of them had already participated in the sociolinguistic interviews that were the basis for the studies presented in Chapters 5 and 6. All participants gave their informed consent and completed a questionnaire about personal information. Only new participants filled in the BLP questionnaire (Birdsong et al., 2012); they were grouped into female or male on the basis of their self-reporting.

Participants took 40–70 minutes to complete the task. The recordings of 5 speakers were discarded for one of the following reasons: (i) the task was presented differently (first speaker; see Section 7.3), (ii) there were technical issues in the recording process (3 speakers), or (iii) the speaker used a dental prosthesis (1 speaker). Consequently, only the acoustic recordings of 21 speakers (11 female, 10 male) were analyzed further (see Table 7.6).

[8] The pilot study was conducted with linguists familiar with experimental tasks.

Table 7.6. Participants' Bilingual Language Profile (BLP) scores, speaker ID, gender, age, state of origin, and highest level of formal education of participants: read speech task.

BLP score	Speaker ID	Gender	Age	State of origin	Highest level of formal education
-46	inf17	F	33	Quintana Roo	high school
-25	inf22	M	40	Yucatán	high school
-24	inf19	F	36	Quintana Roo	university
-24	inf18	F	40	Quintana Roo	university
3	inf21	M	42	Quintana Roo	university
15	inf16	M	59	Quintana Roo	university
23	inf27	F	61	Yucatán	high school
25	inf45	M	47	Quintana Roo	university
66	inf04	F	33	Quintana Roo	university
71	inf43	F	51	Yucatán	university
147	inf44	M	33	Quintana Roo	university
148	inf49	M	52	Yucatán	secondary school
148	inf47	F	44	Yucatán	university
157	inf30	M	48	Quintana Roo	university
166	inf46	M	41	Tabasco	university
171	inf48	M	38	Yucatán	university
177	inf28	F	65	Guerrero	university
177	inf13	F	38	Quintana Roo	university
179	inf12	M	35	Yucatán	university
199	inf09	F	32	Campeche	high school
202	inf07	F	56	Quintana Roo	university

Fourteen speakers had already participated in the recordings of sociolinguistic interviews (speakers inf04 to inf30).[9] All participants, who at the time resided in Felipe Carrillo Puerto, were from one of the three states of the Yucatán peninsula, with two exceptions. The participant from Tabasco (inf46) had lived in the state of Quintana Roo for 25 years (the last 15 years in Felipe Carrillo Puerto), whereas the participant from Guerrero (inf28) had been a Felipe Carrillo Puerto resident for 48 years. Due to their long residence in Quintana Roo, the recordings of these two speakers were not excluded. At the time of the recordings, all speakers had lived in Quintana Roo for at least 15 years, and in Felipe Carrillo Puerto for

[9] It bears noting that, even though speakers inf04 to inf30 took part in the sociolinguistic interviews, not all of them were included in the studies in Chapters 5 and 6 (e.g., inf28).

at least 12 years.[10] Six participants (inf04, inf19, inf21, inf30, inf43, and inf45) reported speaking a language other than Spanish or Maya (English).

7.5 Recordings

Recordings were carried out in a quiet room in Felipe Carrillo Puerto, Quintana Roo, using an AKG C 544 L head-mounted microphone connected to a Presonus Audiobox USB. The speech signal was digitized at a sampling rate of 44,100 Hz and a 32-bit resolution.

7.6 Phrase-boundary annotation

The sentences in the read speech task were designed so that they could be produced as a single Intonational Phrase (IP), at least up to the word *porque*. However, there are some studies on Spanish that show that speakers may phrase IPs into intermediate phrases (ips). This phrasing is extremely variable, although phrasing the subject as a single ip is rather common (e.g., D'Imperio et al., 2005; Elordieta et al., 2005; Feldhausen et al., 2010). Consequently, the possibility that participants would phrase the sentences differently, either as several ips or IPs, was considered. The difference between an ip and an IP in Spanish is of a perceptual nature, with a stronger boundary perceived between IPs than between ips (see Section 2.2).

Before the speech materials were further annotated and analyzed, the phrase-initial boundaries introduced by the participants were annotated. Thus, *phrase-initial position* refers to the initial position of both the intermediate phrase (ip) and the Intonational Phrase (IP). A token in PW-initial position that appears in phrase-initial position (e.g., ((*Rebeca*)$_{\text{PW}}$ (*taló*)$_{\text{PW}}$)$_{\textbf{Phrase}}$ **((*árboles*)**$_{\text{PW}}$ (*en la selva*)$_{\text{PW}}$)$_{\textbf{Phrase}}$…) rather than in phrase-medial position (e.g., ((*Rebeca*)$_{\text{PW}}$ (*taló*)$_{\textbf{PW}}$ **(*árboles*)**$_{\textbf{PW}}$ (*en la selva*)$_{\text{PW}}$)$_{\textbf{Phrase}}$…) may be more strengthened due to the cumulative effect of domain-initial strengthening (see Section 2.4.2). Therefore, in order to study strengthening in the PW domain only, only tokens that appeared in phrase-medial position were considered.

All phrase boundaries were annotated by the author of the study and by a native speaker of German proficient in Spanish, based on auditory analysis. The agreement between coders was very high, with only 49 instances of disagreement, of which seven corresponded to voiced stops and 42 to vowels (out of the possible maximum number of 3360 instances of voiced stops and 2016 instances of vowels in phrase-medial position, according to the design of the task). These 49 sentences, plus 11 additional ones without

[10] Speakers inf12, inf13, inf49 did not report how long they had been living in Felipe Carrillo Puerto.

disagreements, were sent to a third coder, the fieldwork collaborator, who listened to them and indicated how the speaker was dividing them up into chunks. The boundaries set by the three coders were pooled together and tokens annotated as appearing in phrase-initial position were excluded. Tables C6, C7, and C8 in Appendix C present all instances where participants introduced a phrase boundary before a bilabial stop token, a dentialveolar token, or a vowel token, respectively. Almost all of the voiced stop tokens were produced as stops (105 out of 107), in agreement with phonological descriptions that posit a stop allophone after a pause (see Section 4.2.1) and phonetic studies that have shown that the phonological voiced stops are articulated with full closure in IP position (e.g., Lavoie, 2001; Parrell, 2011; see Section 4.2.2.1). The perceptual assessment of glottalization for tokens in phrase-initial position proved difficult for both coders, who were unsure about whether a considerable number of them presented glottalization or not. Consequently, a total of 127 tokens in phrase-initial position were excluded.

Chapter 8
Voiced stops — read speech study

8.1 The present study

The present study aims to further investigate the prosodic strengthening of Yucatecan Spanish bilabial and dentialveolar voiced stops in the Prosodic Word (PW) domain, in this case on the basis of read speech. The speech materials examined were obtained by means of the read speech task described in Chapter 7.

The main acoustic cues to strengthening that are examined are acoustic duration, (change in) acoustic intensity, and presence of a release burst. Thus, longer acoustic duration, a greater change in intensity from the beginning of the voiced stop to its lowest point, and presence of a release burst would correspond to strengthened realizations of the voiced stops. In Chapter 5 (henceforth referred to as *the previous study* for ease of expression), acoustic duration was measured by means of manual segmentation of the speech signal, a procedure that has also been used in previous studies on Spanish voiced stops (Colantoni & Marinescu, 2010; Martínez Celdrán, 2013; Soler & Romero, 1999; see Section 4.2.2.2). Manual segmentation, if done properly and consistently, is a valid approach to annotating speech sounds, although it has certain disadvantages, such as discrepancies that may arise from the annotations of two different coders, or the fact that some tokens cannot be annotated reliably. In the present study, the automatic method developed by Ennever et al. (2017) will be used to measure duration and change in intensity. Additionally, like in the previous study, the presence of a release burst will be used to draw a distinction between stop and approximant realizations.

The linguistic factors that will be investigated are the same as in the previous study, that is, prosodic position (in the PW domain), lexical stress, the interaction between those two factors, and word class. The speaker-specific factors investigated, language dominance and gender, are also the same as in the previous study. By examining the same factors but in a more controlled way I hope to assess whether the results from the corpus study can be replicated, thus providing more evidence in support of the strengthening effects already discussed. Additionally, repetition (first versus second mention) will also be included.[1]

The results of the previous study were interpreted as showing that there was a distinction in terms of strengthening between tokens in word-initial position (more strengthened) and in word-medial position (less strengthened). Importantly, *word-initial* included PW-initial (e.g., (***d**e* (*descanso*)$_{LW}$)$_{PW}$ 'of relaxation') and lexical-word (LW) initial (e.g., (*de* (***d**escanso*)$_{LW}$)$_{PW}$). The results for the effect of lexical stress were mixed, although when present, there was more strengthening for tokens in stressed syllables than in unstressed ones. Because the present study investigates read speech, I expect an effect of lexical stress, similarly to other works on Spanish voiced stops that have used read speech (see Section 4.2.2.3). Tentatively, in terms of the interaction between the two variables, I expect the following strengthening pattern: *word-initial* × *stressed* > *word-initial* × *unstressed* > *word-medial* × *stressed* > *word-medial* × *unstressed* (*word-initial* includes tokens in PW-initial and LW-initial positions). The most robust result in the previous study was that *word-medial* × *unstressed* tokens presented the least degree of strengthening. Consequently, I hypothesize that this will also be the case in the present study. Additionally, in line with the results in the previous study, I expect that tokens in function words (FWs) will be less strengthened than those in lexical words (LWs) in the same prosodic position (PW-initial position only).

The previous study showed mixed results for the effect of language dominance, suggesting little to no effect. Tentatively, for ease of argumentation, I expect that the probability of strengthened realizations will increase with Yucatec Maya dominance. For gender, the result expected is that this variable will not have an effect on the strengthening of voiced stops, as was the case in the previous study.

The reduction of repeated mentions has been studied for several languages (e.g., Clopper & Turnbull, 2018, for English; Wiener et al., 2012, for Mandarin; Kaland & Himmelmann, 2019, for Papuan Malay). Overall, these studies indicate that second mentions of words tend to be acoustically reduced when compared to their first mentions. The cues to this reduction are F0, duration, and intensity, among others. A study on Yucatecan Spanish (Martínez García & Kaland, 2019) showed that second mentions of LWs

[1] POSITION OF THE CARRIER WORD IN THE SENTENCE, which could be *initial, medial*, or *final* (Section 7.2.2), was also examined to confirm that it did not have an effect on the DURATION, CHANGE IN INTENSITY, and presence of a release burst (CATEGORY) of the tokens analyzed. Because this was indeed the case, this variable is not discussed further. Table D1 in Appendix D presents the results of the statistical analyses that were conducted in order to examine this variable.

were reduced in terms of duration, whereas second mentions of FWs were *longer* than first mentions, which was interpreted as the result of having included both stressed and unstressed FWs in the data.[2] In the present study, I seek to investigate whether tokens in second mentions are reduced, which in the context of this study means less strengthened. If second mentions of *words* are reduced, it seems reasonable to assume that this reduction will also be present in their segments (in this case, the voiced stops). In the task employed in the present study, participants had to read the same sentences twice, with the first reading of a sentence followed by a second one (see Section 7.3). In other words, the second reading was highly redundant, which may have favored the reduction of the whole word and its sounds. Thus, I expect voiced stops to show less strengthening in their second mention.

In the previous study, tokens of the dentialveolar stop in unstressed FWs, which appeared in PW-initial position only, were less strengthened than those in LWs in the same prosodic position. Frequency effects were suggested as a possible way of explaining the differences in strengthening between FWs and LWs, with tokens in higher frequency words (i.e., FWs) being less strengthened than those in less frequent words (i.e., LWs). The present study seeks to determine whether the same result is obtained in read speech. Table 8.1 shows the expected results for the variables included in this study.

Table 8.1. Expected results for the strengthening of voiced stops in the read speech study.

Variables	Expected results
1. Prosodic position	word-initial (= PW-initial & LW-initial) > LW-medial
2. Lexical stress	stressed > unstressed syllables
3. Prosodic position × lexical stress	word-initial × stressed > word-initial × unstressed > word-medial × stressed > word-medial × unstressed
4. Language dominance	Yucatec Maya > Yucatecan Spanish
5. Gender	no effect
6. Repetition	first > second mention
7. Word class (/d/)	LWs > FWs

Speech rate has been found to influence the realization of voiced stops in Spanish in that more strengthened realizations are more frequent with slower speech rate (Soler & Romero, 1999; see Section 4.2.2.3). Although speech rate is not included as a variable, we may expect differences in strengthening between the results of the previous study, which was conducted on spontaneous speech and in which participants spoke at their normal speech rate, and the

[2] The data in Martínez García & Kaland (2019) were taken from the recordings of 10 male speakers of Yucatecan Spanish, which are part of the corpus of sociolinguistic studies described in Section 5.2.1.

present study, which is based on read speech. Moreover, speaking style differs between the two studies, with read speech probably favoring a more careful pronunciation. Thus, I expect that the frequency of strengthened realizations will be higher in this study. I will assess this possibility by comparing the results of the two studies, although no statistical analyses will be performed. Finally, it bears noting that speech rate and repetition may not be independent from each other. It is possible that participants produced the second reading of the sentences at a higher speech rate than the first, meaning that reduction of second mentions may be due to an increase in speech rate.

The remainder of the chapter is organized as follows. Section 8.2 introduces the speech materials (Section 8.2.1), the automatic annotation procedure that was used to measure duration and change in intensity (Section 8.2.2), and the statistical analyses that were performed (Section 8.2.3). Section 8.3 presents the results of the statistical analysis for duration, change in intensity, and category (i.e., stop/approximant realizations) for the bilabial stop (Section 8.3.1) and for the dentialveolar stop (Section 8.3.2). Section 8.4 discusses the results, while Section 8.5 provides a summary of the chapter.

8.2 Methods

8.2.1 Speech materials

Based on the design of the read speech task and the fact that recordings by 21 speakers constitute the speech materials (see Chapter 7), the maximum number of phrase-medial data points possible was 1512 for the bilabial stop (4 proper names × 6 repetitions × 21 speakers; 24 tokens × 2 repetitions × 21 speakers) and 1848 for the dentialveolar stop (4 proper names × 6 repetitions × 21 speakers; 32 tokens × 2 repetitions × 21 speakers).[3] As indicated in Section 7.6, all instances that appeared after phrase boundaries introduced by the participants (i.e., tokens in phrase-initial position) were excluded. There were also 20 tokens in phrase-medial position that were produced as fricatives (bilabial: 3; dentialveolar: 17), which were also excluded. Several data points were excluded due to being in a syllable that bore postlexical secondary stress (Section 2.2.1.1), due to errors, background noise, or hesitations.

[3] Recall that proper names (e.g., *la **B**onita*) appeared three times per set of sentences (see Section 7.2.1), meaning that they were in fact mentioned six times each. However, given the fact that repetitions of sentences were produced in pairs, with the second mention in a sentence being produced right after the first one, tokens in proper names were coded as *first* or *second* mention.

There were four data sets created, two for the bilabial stop and two for the dentialveolar stop; one data set for each stop corresponded to the analysis of duration and change in intensity and the other corresponded to the analysis of category. The number of observations analyzed was slightly smaller in the duration and change in intensity analyses than in the category analyses because the errors (79 tokens) and failures to demarcate stops (159) yielded by the automated annotation procedure used were excluded from the duration and change in intensity analyses, but kept for the category ones (see Section 8.2.2 below). Tables D2 and D3 in Appendix D provide the raw count of tokens that were obtained for duration/change in intensity and category for the bilabial stop and for the dentialveolar stop, respectively. In total, the data sets for the bilabial stop comprised 1335 tokens (duration/change in intensity) and 1417 tokens (category), while the data sets for the dentialveolar stop comprised 1611 tokens (duration/change in intensity) and 1726 tokens (category).

In order to investigate whether there were differences in terms of duration, change in intensity, and stop/approximant realizations as a function of word class, a subset of data was created with the dentialveolar tokens that appeared in *PW-initial* position. This was done because FWs only appeared in PW-initial position and only for the dentialveolar stop. The number of tokens for the duration and change in intensity models ($N = 564$, FWs = 283) was smaller than for the category model ($N = 614$, FWs = 300). While all tokens in FWs appeared in *unstressed* syllables, roughly half the tokens in LWs were *stressed* (duration/change in intensity models: $N = 281$, of which *stressed* = 134; category model: $N = 200$, of which *stressed* = 145; see Table D3 in Appendix D).

8.2.2 Automatic annotation of duration and intensity

The annotation of the acoustic duration and the (relative) intensity of voiced stops was made using an R script (R Core Team, 2020) to measure stops in intervocalic position, as introduced in Ennever et al. (2017). This method has been used for languages such as Basque (Hualde et al., 2019), Campidanese Sardinian (Katz & Pitzanti, 2019), and Gurindji (Ennever et al., 2017). Its advantages are several: (i) it is automatic, (ii) it makes reproducibility of other studies' results possible, and, most importantly for the present study, (iii) stop-like sounds and approximant-like sounds are measured following the same procedure, meaning that they can be compared, thus being suitable for the study of Spanish voiced stops. Moreover, Ennever et al.'s method explicitly links the acoustic measures obtained to articulatory landmarks. Ennever et al. (2017) demonstrated that change in intensity was positively correlated with the duration of the stops in Gurindji.[4] This result is in agreement with the claim by Parrell and Narayanan (2018) that the spatial magnitude of the gesture (of Spanish voiced stops) is conditioned by its duration (see Section 4.2.2 and references therein). Thus, for the Yucatecan Spanish stops, we would expect that more strengthened realizations would have longer duration and greater values of change in intensity.

Duration is measured from the start of the closing gesture to the end of the opening gesture of the stop, while change in intensity (within the stop) is measured from the start of the closing gesture to the lowest point in intensity in the stop. Change in intensity is calculated from a continuous function, *relative intensity over time* (*i(t)*), which is taken to be "a broad proxy for articulator height" (Ennever et al., 2017, p. 17). Duration is obtained by setting thresholds on a first derivative of *i(t)*. In the current study, the data were obtained from the analysis of the frequency band 400–1200 Hz with a smoothing parameter of 0.7 and the default 0.6 threshold ratio (see Ennever et al., 2017, for details).

The method described proved useful in the annotation of the Yucatecan Spanish data set. An additional advantage of this method is that it yields Praat TextGrids (Boersma & Weenink, 2019) where the boundaries of the voiced stop are annotated, meaning that the researcher can visually assess the annotation. All TextGrids demarcated by the script were inspected, which led to the exclusion from the analysis of 79 tokens that had been incorrectly annotated and of 159 that the script failed to demarcate. While these tokens were not considered in the analysis of duration and change in intensity, they were considered in the category analysis.

8.2.3 Statistical analysis

Three sets of models were fitted for the variables DURATION, CHANGE IN INTENSITY, and CATEGORY.

For each voiced stop, a separate Bayesian linear mixed-effects model was fitted for the dependent variable DURATION (log-transformed) with the `brms` package (Bürkner, 2017) in the R programming environment (R Core Team, 2020). The predictors were PROSODIC POSITION (levels *LW-medial*, *LW-initial,* and *PW-initial*), LEXICAL STRESS (levels *unstressed* and *stressed*), BLP SCORE (SCALED), GENDER (levels *female* and *male*), and REPETITION (levels *1* and *2*), with PROSODIC POSITION × LEXICAL STRESS as an interaction term. SPEAKER and

[4] See also Ennever et al. (2017) for an in-depth discussion about the relationship between articulatory and acoustic measures.

WORD were included as random intercepts. The likelihood function used was Gaussian. Weakly informative priors were specified for the intercept (*Normal* (0, 10)), for the parameters representing the effects of the predictors and the interaction (*Normal* (0, 10)), for the standard deviation of the random effect (*Normal* (0, 10)), and for the standard deviation of the residual error (*Normal* (0, 10)). The range of the duration values (non-transformed) lay between 23 and 184 ms for the bilabial stop, which in log-scale corresponds to 3.13 and 5.22, respectively, and between 30 and 191 ms for the dentialveolar stop, which in log-scale corresponds to 3.4 and 5.25, respectively. Since it is highly unlikely that any single parameter would take on a value more extreme than the most extreme values of the data, these priors should capture the whole range of possible values of the parameters. An additional model was fitted for DURATION (log-transformed) in order to examine the effect of WORD CLASS (levels *FW* and *LW*) for a subset of the dentialveolar stop data. The data set used corresponded to tokens in *PW-initial* position exclusively, the only position in which FWs appeared. SPEAKER and WORD were included as random intercepts. The priors used were the same as in the models above, except for the prior specified on the intercept, which was made slightly more informative (*Normal* (0, 8)) due to convergence issues.

For each voiced stop, a separate Bayesian linear mixed-effects model was fitted for the dependent variable CHANGE IN INTENSITY. The predictors and random intercepts were the same as the ones in the DURATION models. The data for CHANGE IN INTENSITY were normally distributed; consequently, no transformation of the variable was needed. The likelihood function used was Gaussian. Weakly informative priors were specified for the intercept (*Normal* (0, 30)), for the parameters representing the effects of the predictors and the interaction (*Normal* (0, 15)), for the standard deviation for the random effect (*Normal* (0, 15)), and for the standard deviation of the residual error (*Normal* (0, 15)). The range of the CHANGE IN INTENSITY values lay between 0.03 and 31.19 dB for the bilabial stop and between 0.41 and 29.87 dB for the dentialveolar stop, which means that the prior specification can be considered generous and should allow for the whole range of values that the parameters could possibly take on. An additional model was fitted with CHANGE IN INTENSITY as the dependent variable and WORD CLASS (levels *FW* and *LW*) as the predictor for the subset of dentialveolar tokens in *PW-initial* position, with SPEAKER and WORD as random intercepts and the same priors as those in the other CHANGE IN INTENSITY models.

For each voiced stop, a separate Bayesian logistic mixed-effects model was fitted for the dependent variable CATEGORY (treatment coded, levels *approximant* and *stop*). The predictors and random intercepts were the same as the ones in the DURATION and CHANGE

IN INTENSITY models. The likelihood function used was Bernoulli. Weakly informative priors were specified for the intercept (*Normal* (0, 3)), for the parameters representing the effects of the predictors and the interaction (*Normal* (0, 3)), and for the standard deviation of the random effect (*Normal* (0, 5); see Section 5.2.6.2 for an explanation of these priors). An additional model was fitted with CATEGORY as the dependent variable and WORD CLASS (levels *FW* and *LW*) for the subset of dentialveolar tokens in *PW-initial* position, with SPEAKER and WORD as random intercepts. The priors used were the same as those in the other CATEGORY models.

The convergence of the models, as well as their fitting and the assessment of goodness of fit, were evaluated following the procedure described in Section 5.2.6.2. The maximum tree depth was increased from the default one (10) to a higher one (20) in some models due to warnings (DURATION models and the CHANGE IN INTENSITY model for the bilabial stop). The package `emmeans` (Lenth, 2020) was used to obtain estimated marginal means and to conduct pairwise comparisons of the levels of the predictors of the interaction.

8.3 Results

8.3.1 Bilabial stop

The number of tokens included in the data set of the DURATION/CHANGE IN INTENSITY analyses (N = 1335) is smaller than in that of the CATEGORY analysis (N = 1417), which includes 546 *stops* and 871 *approximants* (see Section 8.2.1). Figure 8.1 illustrates the relationship between duration, change in intensity, and category for the data set used in the DURATION/CHANGE IN INTENSITY analyses. Clearly, the figure shows a positive correlation between duration and change in intensity (r = 0.75), which is in line with the argumentation in Ennever et al. (2017) and others. Moreover, it also shows that tokens that were categorized as *stops* have longer duration and a greater change in intensity than those categorized as *approximants*. These results are consistent with the acoustic characteristics of strengthening of voiced stops in Spanish (see Sections 4.2.2 and 8.2.2).

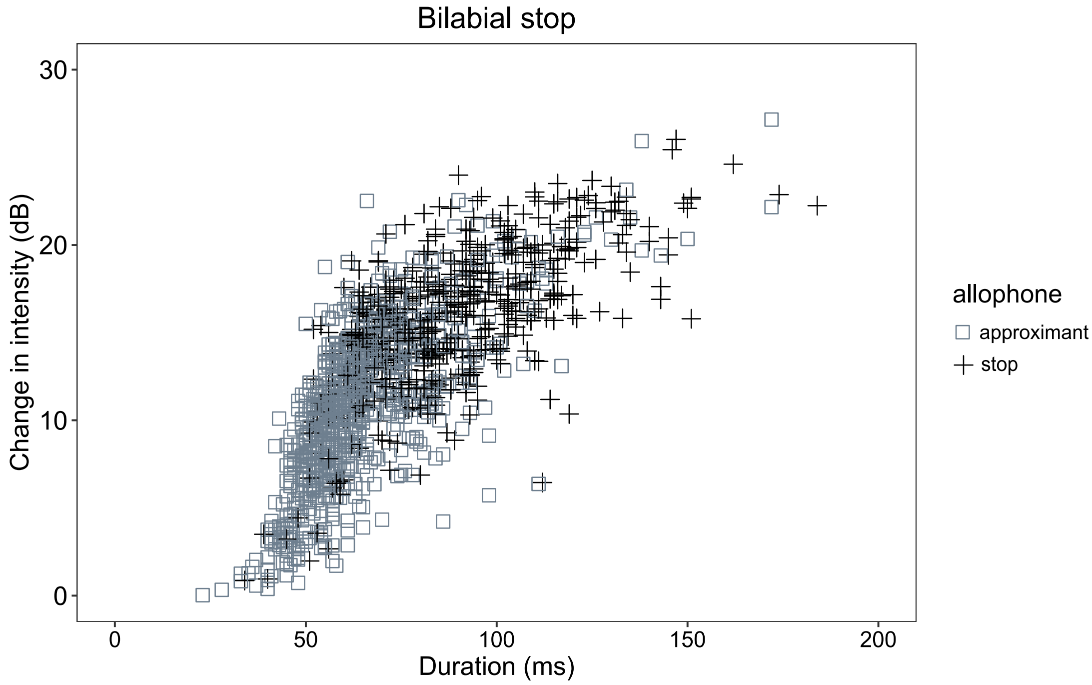

Figure 8.1. Duration (ms) and change in intensity (dB) of *approximant* and *stop* realizations of the bilabial stop within the DURATION/CHANGE IN INTENSITY data set.

The following sections present the results of the DURATION, CHANGE IN INTENSITY, and CATEGORY models for the bilabial stop.

8.3.1.1 Duration

In order to examine whether there was a main effect of PROSODIC POSITION, pairwise comparisons of the levels of the predictor were obtained. The results show evidence that the levels *LW-medial* and *LW-initial* differ from each other (-0.16, CI [-0.26, -0.07]), with a shorter duration of tokens in the former level, as do the levels *LW-medial* and *PW-initial* (-0.25, CI [-0.36, -0.15]), with a shorter duration of tokens in the former level. Finally, the results show no clear evidence that the levels *LW-initial* and *PW-initial* differ from each other (-0.09, CI [-0.20, 0.02]). Nevertheless, because most of the probability mass is below 0, there may be weak evidence in support of tokens in *LW-initial* position being shorter than in *PW-initial* position, although we cannot be certain about this. The effect of LEXICAL STRESS was assessed in the same manner. The duration of tokens in *unstressed* and *stressed* syllables differs (-0.2, CI [-0.29, -0.12]), with tokens in *unstressed* syllables being shorter than in *stressed* ones.

Estimated marginal means and pairwise comparisons for the levels of the predictors of the interaction (PROSODIC POSITION × LEXICAL STRESS) were obtained (see Table D4 in Appendix D). According to the model, tokens corresponding to the level *LW-medial* ×

unstressed differ considerably from those in the other levels (with the exception of *LW-medial* × *stressed*) in that the tokens are shorter. Tokens in the level *LW-medial* × *stressed* pattern with *LW-initial* × *unstressed* and *PW-initial* × *unstressed*. Finally, tokens in *LW-initial* × *stressed* pattern with those in *PW-initial* × *stressed*. Tokens in *PW-initial* × *stressed* have the longest duration (see Figure 8.2).

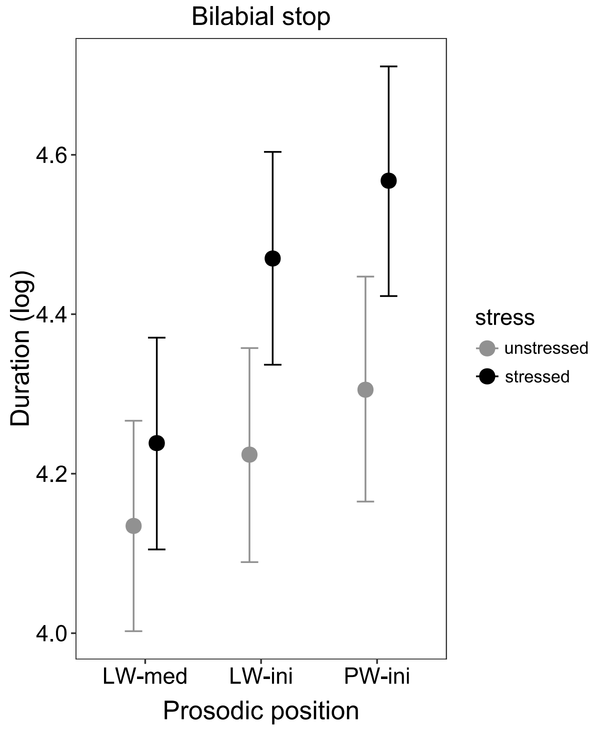

Figure 8.2. Interaction plot. Effect of PROSODIC POSITION conditional on LEXICAL STRESS ('LW-med' – *LW-medial* position; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position): DURATION model for the bilabial stop in the read speech study.

The results for the BLP SCORE (SCALED) and GENDER predictors indicate that 0 is included within the CI (BLP SCORE (SCALED): $\hat{\beta} = -0.03$, CI [-0.10, 0.04]; GENDER: $\hat{\beta}_{genderM} = -0.07$, CI [-0.21, 0.07]). Based on the posterior samples, the estimated probability of the parameter being less than 0 is 0.81 for BLP SCORE (SCALED) and 0.87 for GENDER. This suggests that there may be some weak evidence that the duration of the bilabial stop decreases the higher the BLP SCORE is (i.e., in the direction of Yucatecan Spanish dominance), and that *male* speakers produce shorter realizations than *female* speakers, although we cannot be certain about this. Figure 8.3 is a plot of the duration of bilabial tokens by speaker.

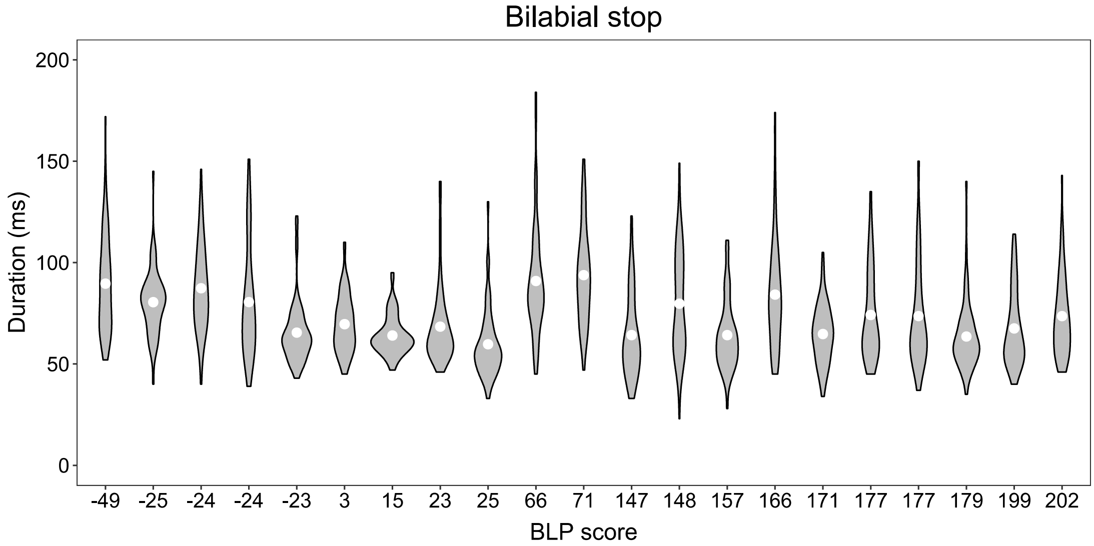

Figure 8.3. Violin plot (with mean points) of the duration of bilabial tokens per speaker in the read speech study. Each BLP score on the x-axis corresponds to one speaker.

The results for the REPETITION predictor show evidence that *first* and *second* mentions differ ($\hat{\beta}_{rep2}$ = -0.03, CI [-0.05, -0.01]). The posterior probability of the parameter being less than zero is 1, that is, there is a probability of 1 that *second* mentions have a shorter duration than *first* mentions. Consequently, we can draw the inference that the DURATION of bilabial tokens differs as a function of REPETITION.

8.3.1.2 Change in intensity

Pairwise comparisons of the levels of the predictor PROSODIC POSITION show evidence that the levels *LW-medial* and *LW-initial* differ from each other (-2.37, CI [-4.09, -0.61]), with a smaller change in intensity of tokens in the former level; the same is true of the levels *LW-medial* and *PW-initial* (-4.11, CI [-5.99, -2.23]), whereas there is no evidence that the levels *LW-initial* and *PW-initial* differ from each other (-1.74, CI [-3.59, 0.18]). In this last comparison, because most of the probability mass is below 0, there is some weak evidence in support of tokens in *LW-initial* having a smaller change in intensity than those in *PW-initial*. Results for LEXICAL STRESS show that change in intensity of tokens in *unstressed* and *stressed* syllables differs substantially (-3.08, CI [-4.57, -1.61]), with tokens in *unstressed* syllables having a smaller change in intensity.

Estimated marginal means and pairwise comparisons for the levels of the predictors of the interaction PROSODIC POSITION × LEXICAL STRESS were obtained (see Table D5 in Appendix D). The comparisons show that the level *LW-medial* × *unstressed* differs greatly from the other levels (except *LW-medial* × *stressed* and *LW-initial* × *unstressed*) in that the change in intensity of tokens is considerably smaller. The level *LW-medial* × *stressed* patterns with *LW-initial* × *unstressed* and *PW-initial* × *unstressed*. Finally, *LW-initial* × *stressed* patterns with *PW-initial* × *stressed* (see Figure 8.4).

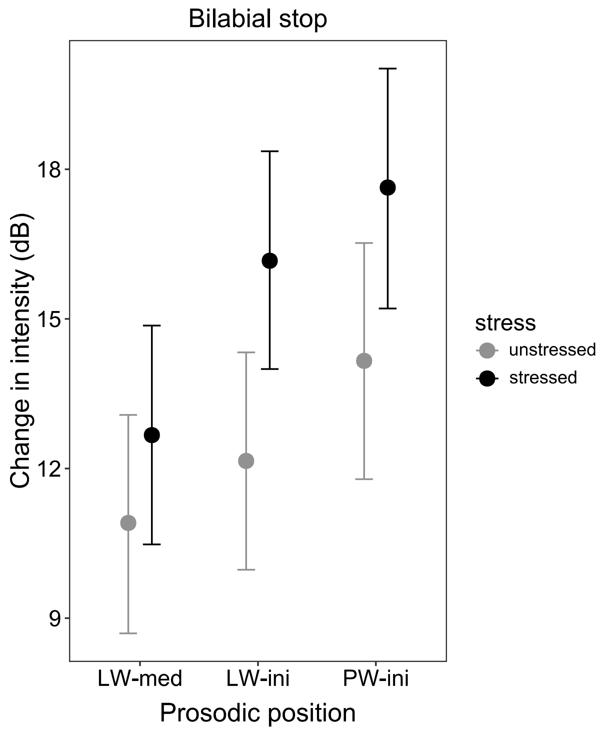

Figure 8.4. Interaction plot. Effect of PROSODIC POSITION conditional on LEXICAL STRESS ('LW-med' – *LW-medial* position; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position): CHANGE IN INTENSITY model for the bilabial stop in the read speech study.

Similarly to the results for DURATION, the results for the BLP SCORE (SCALED) and GENDER predictors indicate that 0 is included within the CI (BLP SCORE (SCALED): $\hat{\beta} = -0.84$, CI [-1.88, 0.17]; GENDER: $\hat{\beta}_{genderM} = -1.33$, CI [-3.34, 0.77]). The posterior probability of the parameter being less than 0 is 0.95 for BLP SCORE (SCALED) and 0.9 for GENDER. This suggests that there may be some weak evidence that the change in intensity decreases the higher the BLP SCORE is, and that tokens produced by *male* speakers have a smaller CHANGE IN INTENSITY than those of *female* speakers, although we cannot be certain about this. Figure 8.5 plots the change in intensity by speaker.

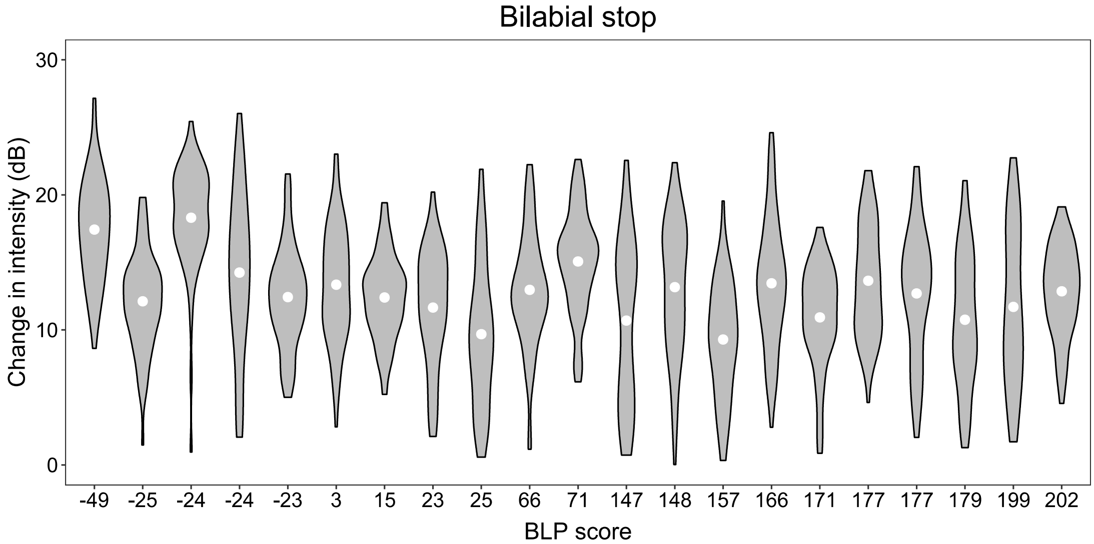

Figure 8.5. Violin plot (with mean points) of the change in intensity of bilabial tokens per speaker in the read speech study. Each BLP score on the x-axis corresponds to one speaker.

The results for the REPETITION predictor indicate that *first* and *second* mentions differ ($\hat{\beta}_{rep2}$ = -0.76, CI [-1.17, -0.36]). The posterior probability of the parameter being less than zero is 1, that is, there is a probability of 1 that *second* mentions have a smaller CHANGE IN INTENSITY than *first* mentions. Consequently, we can draw the inference that CHANGE IN INTENSITY differs as a function of REPETITION.

8.3.1.3 Category

The total percentage of *stop* realizations in the data set was 38.53 %. In *LW-medial* position, 23.6 % of the total percentage of tokens in this position were *stops*, while in *LW-initial* position it was 48.19 %, and in *PW-initial* position, 49.62 %. These results contrast with those of the previous study, which show a lower number of realizations categorized as *stops* (total: 20.76 %; in *LW-medial* position: 15.47 %; in *LW-initial* position: 26.39 %; in *PW-initial* position: 33.89 %; see Section 5.3.1.2). Figure 8.6 shows the percentage of stop realizations per speaker.

Pairwise comparisons for the levels of PROSODIC POSITION show evidence that the levels *LW-medial* and *LW-initial* differ from each other (-1.63, CI [-2.42, -0.84]), with a smaller probability of *stop* realizations in the former level, as do the levels *LW-medial* and *PW-initial* (-2.08, CI [-2.93, -1.23]). The results show no clear evidence that the levels *LW-initial* and *PW-initial* positions differ from each other (-0.44, CI [-1.26, 0.40]). Similarly, the results for the effect of LEXICAL STRESS show a smaller probability of *stop* realizations in *unstressed* syllables (-2.04, CI [-2.75, -1.37]).

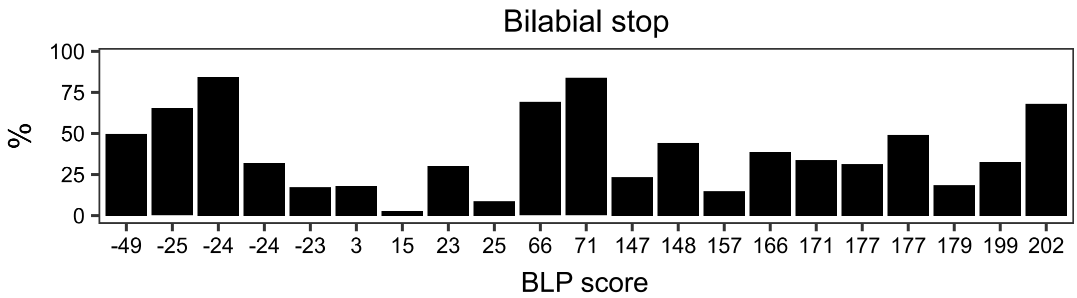

Figure 8.6. Percentage of *stop* realizations of the bilabial stop per speaker in the read speech task. Each Bilingual Language Profile (BLP) score represents a speaker.

Estimated marginal means and pairwise comparisons for the levels of the predictors of the interaction (see Table D6 in Appendix D) show that the level *LW-medial* × *unstressed* differs greatly from all of the other levels in that the probability of *stop* realizations is lower. The *stressed* levels of *LW-initial* and *PW-initial*, which pattern together, differ from the corresponding *unstressed* levels (which also pattern together) in that they have a higher probability of *stop* realizations. The level *LW-medial* × *stressed* patterns with *LW-initial* × *unstressed* and *PW-initial* × *unstressed* (see Figure 8.7).

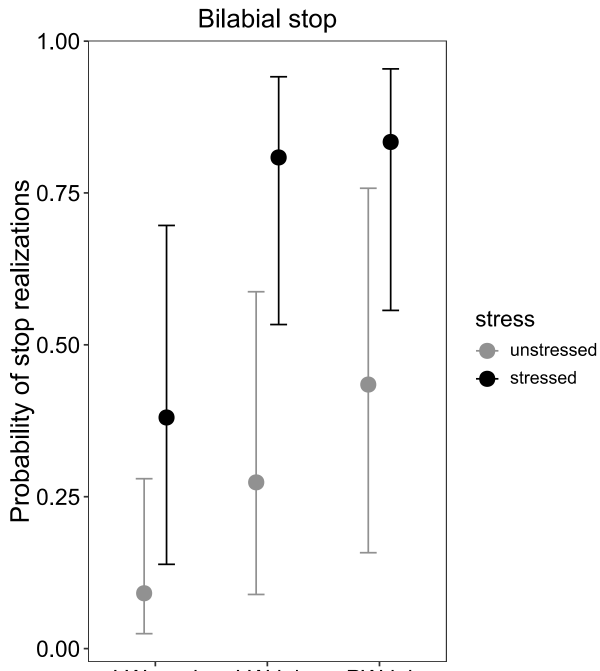

Figure 8.7. Interaction plot. Effect of PROSODIC POSITION conditional on LEXICAL STRESS ('LW-med' – *LW-medial* position; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position): CATEGORY model for the bilabial stop in the read speech study.

The model results for the BLP SCORE (SCALED) predictor ($\hat{\beta}$ = 0.09, CI [-0.75, 0.93]) and the GENDER predictor ($\hat{\beta}_{genderM}$ = -0.82, CI [-2.44, 0.84]) show that 0 is included within the CI.

The posterior probability of the parameter being greater than zero is 0.59 for BLP SCORE (SCALED) and less than 0 for GENDER is 0.85. This last result indicates that there may be some weak evidence to draw the inference that the probability of *stop* realizations is less frequent for *male* than for *female* speakers, but we cannot be certain about this.

The results for the REPETITION predictor indicate that *first* and *second* mentions differ ($\hat{\beta}_{rep2}$ = -0.30, CI [-0.59, -0.01]). The posterior probability of the parameter being less than zero is 0.98, that is, there is a 0.98 probability that *second* mentions are produced less frequently as *stops* than *first* mentions. Consequently, we could draw the inference that the probability of *stop* realizations differs as a function of REPETITION.

8.3.1.4 Summary

Overall, the results show that the probability of strengthening is greater in *word-initial* position (vs. *word-medial* position) and in *stressed* syllables (vs. *unstressed* syllables). Moreover, the *stressed* levels in *word-initial* position (i.e., *PW-initial* and *LW-initial*) pattern together, as do the *unstressed* levels in *word-initial* position. Also in accordance with the expected results, there is a greater degree of strengthening for the tokens in *stressed* syllables. In *word-medial* position, there is a tendency for *unstressed* tokens to present less strengthening than in the other levels, although this level may pattern with *word-medial* × *stressed* and/or *word-initial* × *unstressed*.

The evidence in support of more strengthening as a function of Yucatec Maya dominance is unclear, with some weak evidence only in the DURATION and CHANGE IN INTENSITY models; the results for the effect of GENDER are also mixed, with weak evidence in support of a greater degree of strengthening by *female* speakers. Finally, there is a greater degree of strengthening for *first* than *second* mentions in the three models, which is in line with the expected results.

8.3.2 Dentialveolar stop

Similarly to the bilabial stop (see Section 8.3.1), the number of tokens included in the data set of the DURATION/CHANGE IN INTENSITY analyses (N = 1611) is smaller than in that of the CATEGORY analysis (N = 1726), which includes 913 *stops* and 813 *approximants*. The relationship between the three variables (DURATION, CHANGE IN INTENSITY, and CATEGORY) in the data set used in the DURATION/CHANGE IN INTENSITY analyses (Figure 8.8) also mirrors that of the bilabial stop (see Figure 8.1), showing a positive correlation between DURATION and CHANGE IN INTENSITY (r = 0.74).

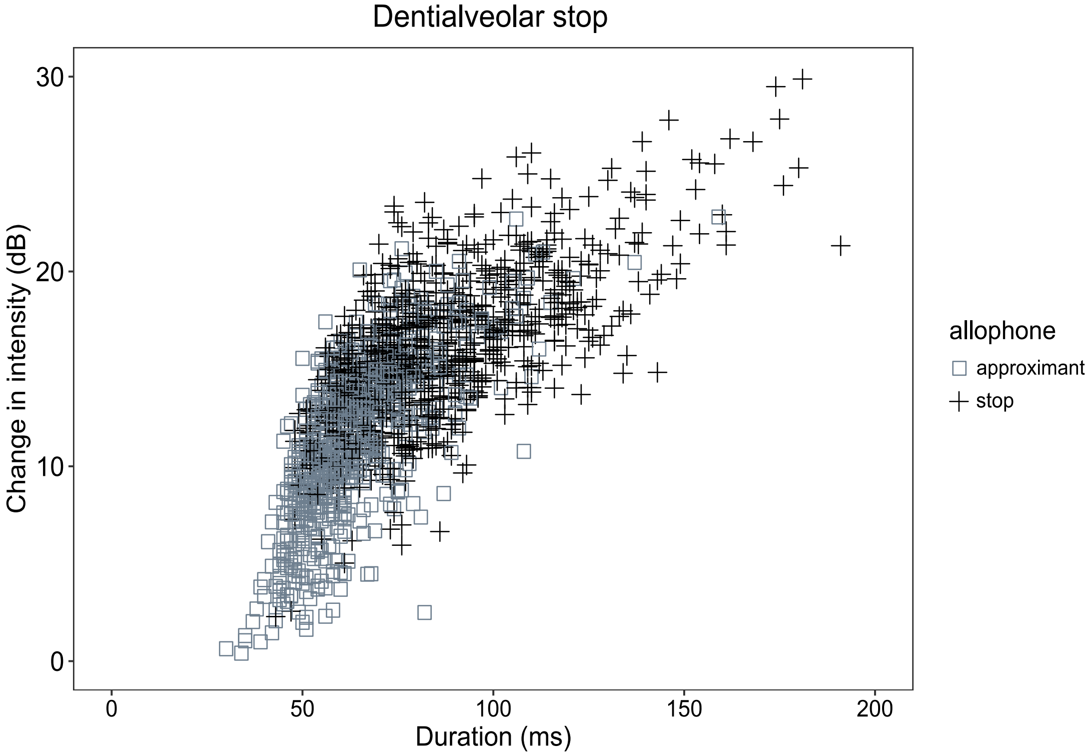

Figure 8.8. Duration and change in intensity of *approximant* and *stop* realizations of the dentialveolar stop within the DURATION/CHANGE IN INTENSITY data set.

The following sections present the results of the DURATION, CHANGE IN INTENSITY, and CATEGORY models corresponding to the dentialveolar stop, including the models that examine WORD CLASS.

8.3.2.1 Duration

In order to examine whether there was a main effect of PROSODIC POSITION, pairwise comparisons of the levels of the predictor were obtained. The results show evidence that the levels *LW-medial* and *LW-initial* do not differ from each other (-0.07, CI [-0.14, 0.01]); neither do the levels *LW-initial* and *PW-initial* (-0.07, CI [-0.15, 0.00]). In both instances, most of the probability mass is below 0, suggesting there is some weak evidence in support of tokens in the former level of the comparison having shorter duration, although we cannot be certain about this. There is evidence that the levels *LW-medial* and *PW-initial* differ from each other (-0.14, CI [-0.22, -0.07]), with shorter duration in the former level. The results for the effect of LEXICAL STRESS show that the duration of tokens in the *unstressed* and *stressed* levels differs (-0.24, CI [-0.3, -0.18]), with shorter duration in the *unstressed* level. Estimated marginal means and pairwise comparisons for the levels of the predictors of the interaction (see Table D7 in Appendix D) show that the three *unstressed* levels pattern together, as do two *stressed* levels, *LW-medial* × *stressed* and *LW-initial* × *stressed*. All other levels differ from each other (see Figure 8.9).

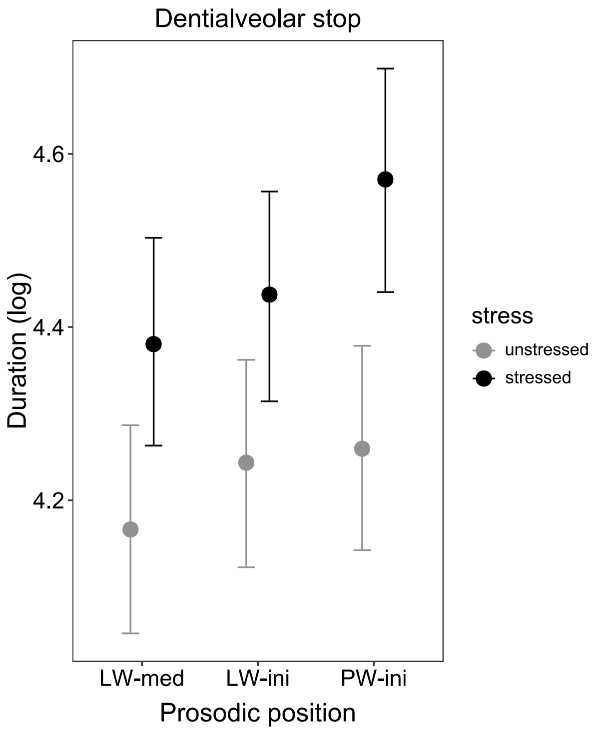

Figure 8.9. Interaction plot. Effect of PROSODIC POSITION conditional on LEXICAL STRESS ('LW-med' – *LW-medial* position; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position): DURATION model for the dentialveolar stop in the read speech study.

The results for the BLP SCORE (SCALED) predictor indicate that 0 is included within the CI ($\hat{\beta}$ = -0.04, CI [-0.12, 0.03]). The estimated probability of the parameter being less than 0 is 0.9. This suggests that there may be some weak evidence that the duration of tokens decreases the higher the BLP SCORE is, that is, in the direction of Yucatecan Spanish dominance, although we cannot be certain about this. Figure 8.10 shows the duration of the dentialveolar tokens for each speaker.

The results for the GENDER predictor indicate that 0 is included within the CI ($\hat{\beta}_{genderM}$ = -0.04, CI [-0.18, 0.10]). The posterior probability of the parameter being smaller than zero is 0.7, that is, there is a 0.7 probability that *male* speakers produce shorter tokens of the stop than *female* speakers, although we cannot be certain about this.

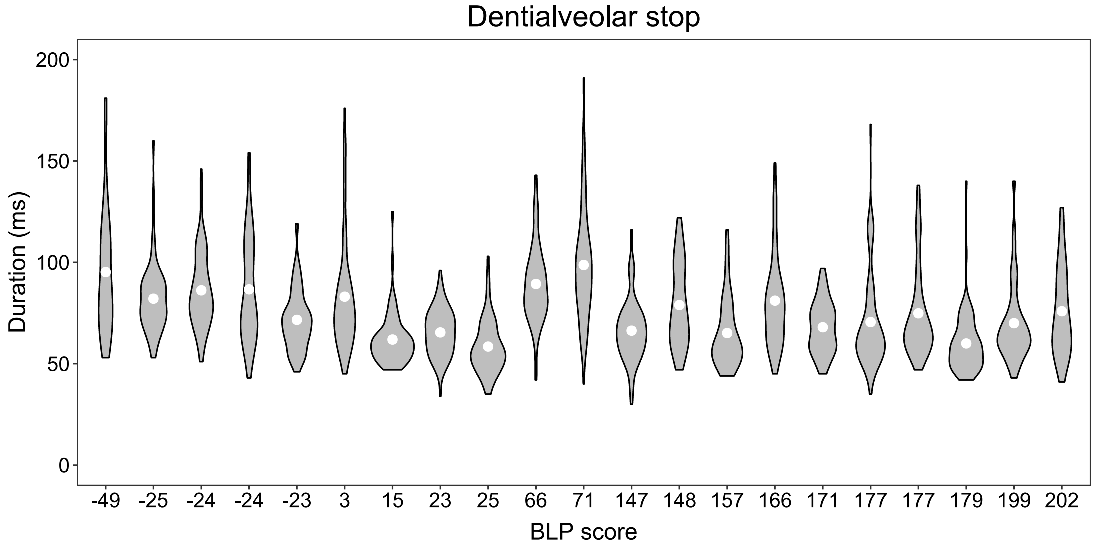

Figure 8.10. Violin plot (with mean points) of the duration of dentialveolar tokens per speaker in the read speech study. Each BLP score on the x-axis corresponds to one speaker.

The results for the REPETITION predictor indicate that 0 is not included in the CI ($\hat{\beta}_{rep2}$ = -0.04, CI [-0.06, -0.02]). Furthermore, the posterior probability of the parameter being less than 0 is 1, that is, there is a probability of 1 that *second* mentions have a shorter duration than *first* mentions. Thus, the duration of the dentialveolar tokens seems to differ as a function of REPETITION.

The results for the WORD CLASS predictor in the WORD CLASS model indicate that 0 is included within the CI ($\hat{\beta}_{LW}$ = 0.26, CI [-0.02, 0.55]). The posterior probability of the parameter being greater than zero is 0.97, which shows that there is a 0.97 probability that the tokens in *LWs* are produced with greater DURATION than in *FWs*, but we cannot be certain about this.

8.3.2.2 Change in intensity

The results of the pairwise comparisons of the levels of PROSODIC POSITION show evidence that the levels *LW-medial* and *LW-initial* do not differ from each other (-0.72, CI [-1.87, 0.37]). Conversely, the levels *LW-medial* and *PW-initial* do differ from each other (-2.22, CI [-3.36, -1.07]), as do the levels *LW-initial* and *PW-initial* (-1.50, CI [-2.63, -0.29]), with a smaller change in intensity in the former levels of each comparison. Similarly, the results for LEXICAL STRESS show that the change in intensity of tokens in *unstressed* and *stressed* syllables differs substantially (-3.04, CI [-3.97, -2.07], with tokens in *unstressed* syllables manifesting a smaller change in intensity.

Estimated marginal means and pairwise comparisons for the levels of the predictors of the interaction (see Table D8 in Appendix D) indicate that most levels differ from the others, but the only level that clearly does so is *PW-initial* × *stressed,* with a greater CHANGE IN INTENSITY. The levels *LW-medial* × *unstressed* and *LW-initial* × *unstressed* pattern together, as do the levels *LW-medial* × *stressed* and *LW-initial* × *stressed*; *LW-initial* × *stressed* also patterns with *PW-initial* × *unstressed* (Figure 8.11).

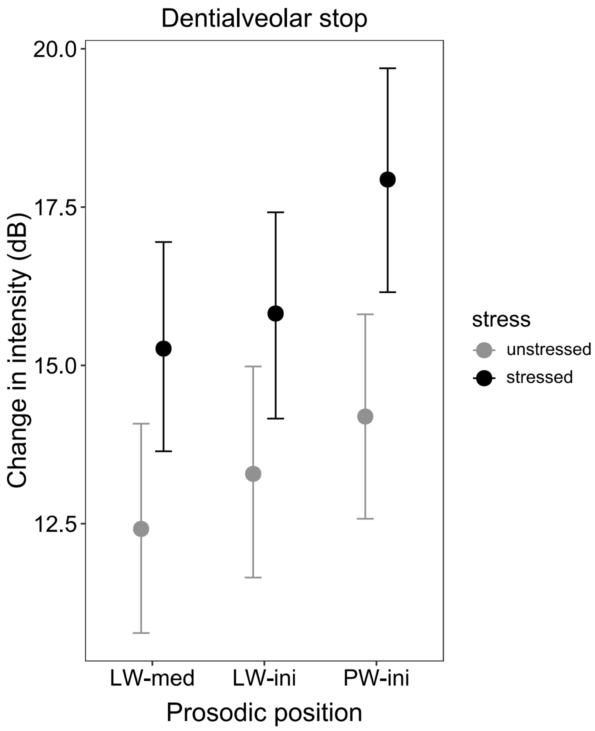

Figure 8.11. Interaction plot. Effect of PROSODIC POSITION conditional on LEXICAL STRESS ('LW-med' – *LW-medial* position; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position): CHANGE IN INTENSITY model for the dentialveolar stop in the read speech study.

The results for the BLP SCORE (SCALED) predictor indicate that 0 is not included within the CI ($\hat{\beta}$ = -0.96, CI [-1.87, -0.04]). Based on the posterior samples, the estimated probability of the parameter being less than 0 is 0.98. Thus, there is evidence that suggests that the change in intensity decreases the higher the BLP SCORE is, that is, in the direction of Yucatecan Spanish dominance. Figure 8.12 shows the DURATION of the dentialveolar tokens for each speaker.

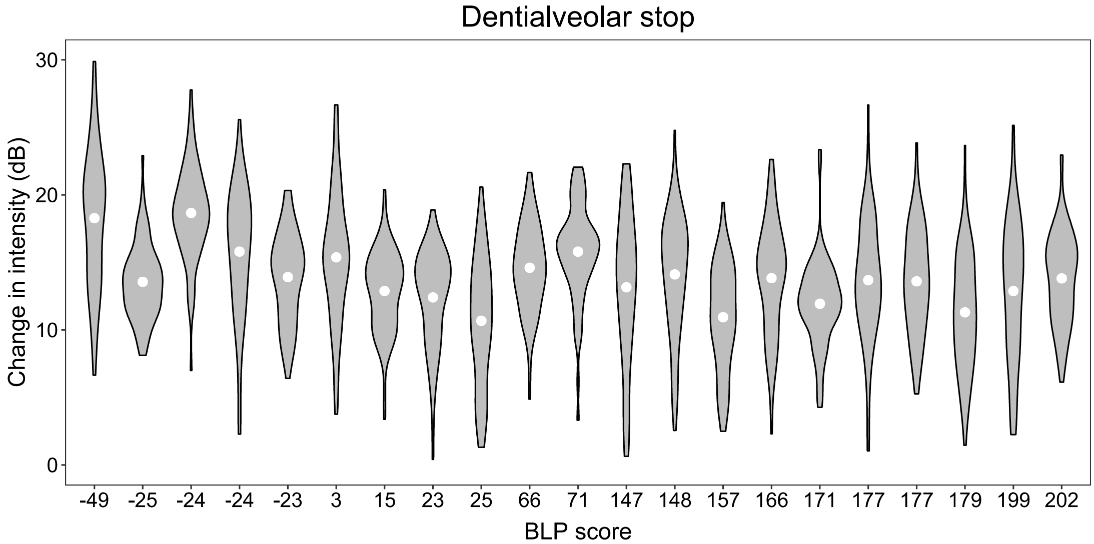

Figure 8.12. Violin plot (with mean points) of the change in intensity of dentialveolar tokens per speaker in the read speech study. Each BLP score on the x-axis corresponds to one speaker.

The results for the GENDER predictor indicate that 0 is included within the CI ($\hat{\beta}_{genderM}$ = -0.84, CI [-2.67, 1.03]). The posterior probability of the parameter being less than zero is 0.82, that is, there is a 0.82 probability that *male* speakers produce the stop with a smaller change in intensity than *female* speakers, but we cannot be certain about this.

The results for the REPETITION predictor indicate that 0 is not included in the CI ($\hat{\beta}_{rep2}$ = -0.73, CI [-1.08, -0.38]). The posterior probability of the parameter being less than zero is 1, that is, there is a probability of 1 that *second* mentions have a smaller change in intensity than *first* mentions, thus leading to the inference that the change in intensity differs as a function of REPETITION.

The results for the WORD CLASS predictor in the WORD CLASS model indicate that 0 is included within the CI ($\hat{\beta}_{LW}$ = 3.30, CI [-0.21, 6.71]). The posterior probability of the parameter being greater than zero is 0.97, that is, there is a 0.97 probability that voiced stops in *LWs* are produced with a greater change in intensity than in *FWs*, but we cannot be certain about this.

8.3.2.3 Category

The total percentage of *stop* realizations in the data set was 52.9 %. In *LW-medial* position, 47.84 % of the total percentage of tokens were *stops*, while in *LW-initial* position it was 55.76 %, and in *PW-initial* position, 54.89 %. These results contrast with those of the previous study, which showed lower numbers of realizations categorized as *stops* (total: 15.43 %; in *LW-medial* position: 9 %; in *LW-initial* position: 30.37 %; in *PW-initial* position: 19.96 %; see Section 5.3.2.2). Figure 8.13 shows the percentage of stop realizations per speaker.

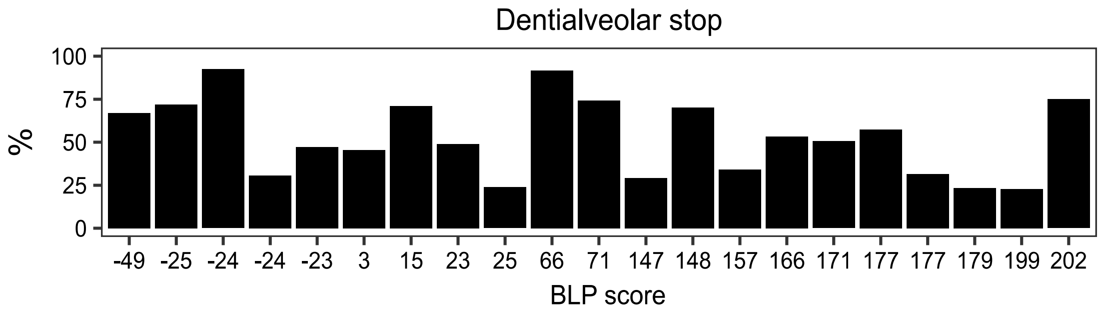

Figure 8.13. *Stop* realizations of the dentialveolar stop (in %) per speaker in the read speech study. Each BLP score on the x-axis corresponds to one speaker.

Pairwise comparisons of the levels of the predictor PROSODIC POSITION show evidence that the levels *LW-medial* and *LW-initial* do not differ much from each other (-0.47, CI [-1.33, 0.43]). Conversely, there is evidence that the levels *LW-medial* and *PW-initial* differ from each other (-1.75, CI [-2.69, -0.79]), as do the levels *LW-initial* and *PW-initial* (-1.29, CI [-2.27, -0.35]), with a smaller probability of *stop* realizations in the former levels of the comparisons. Similarly, the results for the effect of LEXICAL STRESS show a smaller probability of *stop* realizations in *unstressed* syllables than in *stressed* ones (-2.19, CI [-2.95, -1.41]).

Estimated marginal means and pairwise comparisons for the levels of the predictors of the interaction (PROSODIC POSITION × LEXICAL STRESS) were obtained (see Table D9 in Appendix D). The probability of *stop* realizations is highest for the level *PW-initial* × *stressed*. The three *unstressed* levels, which pattern together, present the lowest probability of *stop* realizations. *LW-medial* × *stressed* and *LW-initial* × *stressed*, which are similar to each other, pattern together with *PW-initial* × *unstressed*. *PW-initial* × *unstressed* also patterns with *LW-medial* × *stressed* and *LW-initial* × *stressed* (see Figure 8.14).

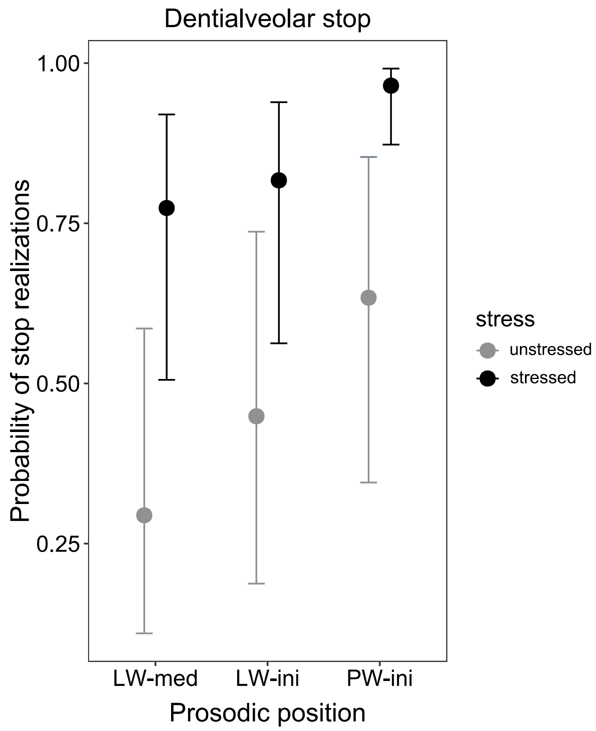

Figure 8.14. Interaction plot. Effect of PROSODIC POSITION conditional on LEXICAL STRESS ('LW-med' – *LW-medial* position; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position): CATEGORY model for the dentialveolar stop in the read speech study.

The results for the BLP SCORE (SCALED) and GENDER predictors indicate that 0 is included within the CI ($\hat{\beta}$ = -0.36, CI [-1.02, 0.31]; $\hat{\beta}_{genderM}$ = -0.68, CI [-2.01, 0.65]). Both for BLP SCORE (SCALED) and GENDER, the posterior probability of the parameter being less than 0 is 0.86. Consequently, there may be some weak evidence to infer that the probability of *stop* realizations decreases the higher the BLP SCORE is, and also for *male* speakers producing less instances of *stop* realizations than *female* speakers, although we cannot be certain about this.

The results for the REPETITION predictor indicate that 0 is not included in the CI ($\hat{\beta}_{rep2}$ =-0.39, CI [-0.63, -0.15]). The posterior probability of the parameter being less than 0 is 1, that is, there is a probability of 1 that *second* mentions are produced less frequently as *stops* than *first* mentions. Consequently, we can draw the inference that the probability of *stop* realizations differs as a function of REPETITION.

The results for the WORD CLASS predictor in the WORD CLASS model indicate that 0 is included within the CI ($\hat{\beta}_{LW}$ = 2.19, CI [-0.13, 4.35]). Based on the posterior samples, the posterior probability of the parameter being greater than zero is 0.97, that is, there is a 0.97 probability that *LWs* are produced with more instances of *stop* realizations than *FWs*, which suggests that there may be more *stop* realizations in the case of *LWs*, but that we cannot be certain about this.

8.3.2.4 Summary

The probability of strengthening is greater in *PW-initial* versus *LW-initial* and *LW-medial* (which can be grouped together), as well as in *stressed* syllables (versus *unstressed* ones). Moreover, tokens in the level *PW-initial* × *stressed* present the highest probability of strengthening, followed by both tokens in *stressed* syllables in *LW-medial* and *LW-initial* positions. In other words, the two levels included in the *word-initial* (vs. *word-medial*) category do not pattern together, which is contrary to expectations. Furthermore, there is a tendency for *unstressed* tokens in all prosodic positions (but especially in *LW-medial* and *LW-initial*) to pattern together.

The evidence in support of more strengthening as a function of Yucatec Maya dominance is unclear, with some stronger evidence for it in the CHANGE IN INTENSITY model. The results for the effect of GENDER are not clear-cut, with rather weak evidence for *females* producing more strengthened realizations. Consequently, given these results, it can be argued that GENDER seems to have no effect on strengthening in the case of the dentialveolar voiced stop. Finally, there is a greater degree of strengthening for *first* than *second* mentions in the three models, in line with the expected results. Differences in terms of strengthening between tokens in *FWs* and *LWs* are not strong, although *LWs* seem to present a greater degree of strengthening, which is in line with the expected results.

8.4 Discussion

The present study provides further support for some of the findings of the study based on spontaneous speech (Chapter 5). Table 8.2 is an overview of the results of the models. WORD CLASS will be considered after discussing the results for PROSODIC POSITION and LEXICAL STRESS.

PROSODIC POSITION and LEXICAL STRESS have a partially different effect on strengthening for each voiced stop, although there are two common characteristics: (i) LEXICAL STRESS favors strengthened realizations, and (ii) tokens in *LW-medial position* × *unstressed* are the least strengthened of all the possible combinations of the two variables. Compared to the expected results (which reflect the results from the previous study), the effect of LEXICAL STRESS on strengthening is confirmed, whereas the result for PROSODIC POSITION differs. Thus, for the bilabial stop, tokens in *word-initial* position (i.e., *PW-initial* and *LW-initial*) are more strengthened than those in *word-medial* position, in line with the results of the previous study. For the dentialveolar stop, it is tokens in *PW-initial* position that are more strengthened, while the levels *LW-initial* and *LW-medial* tend to pattern

Table 8.2. Results for strengthening of voiced stops in the read speech study. 'LW-med' – *LW-medial* position; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position; 'str' – *stressed*; 'uns' – *unstressed.*

Variables	Expected strengthening	Duration	Change in intensity	Category	Stop
1. PROSODIC POSITION	*word-ini (= PW-ini & LW-ini) > LW-med*	+/- (*PW-ini ≥ LW-ini) > LW-med*	+/- (*PW-ini ≥ LW-ini) > LW-med*	✓	/b/
		+/- *PW-ini* ≥ (*LW-ini* ≥ *LW-med*)	× *PW-ini* > (*LW-ini & LW-med*)	× *PW-ini* > (*LW-ini & LW-med*)	/d/
2. LEXICAL STRESS	*stressed > unstressed* syllables	✓	✓	✓	/b/
		✓	✓	✓	/d/
3. PROSODIC POSITION × LEXICAL STRESS	(1) *word-ini × str* > (2) *word-ini × uns* > (3) *word-med × str* > (4) *word-med × uns*	+/- 1 > 2 & 3 ≥ 4	+/- 1 > 2 & 3 ≥ 4	+/- 1 > 2 & 3 ≥ 4	/b/
		× 1 (*PW-ini*) > (1 (*LW-ini*) & 3 > 2 + 4	× 1 (*PW-ini*) > (1 (*LW-ini*) & 3 ≥ 2 + 4	× 1 (*PW-ini*) > (1 (*LW-ini*) & 3 ≥ 2 + 4	/d/
4. LANGUAGE DOMINANCE	Yucatec Maya > Yucatecan Spanish	?	?	×	/b/
		?	✓	?	/d/
5. GENDER	No effect	? *females ≥ males*	? *females ≥ males*	? *females ≥ males*	/b/
		? *females ≥ males*	? *females ≥ males*	? *females ≥ males*	/d/
6. REPETITION	*1st > 2nd* mention	✓	✓	✓	/b/
		✓	✓	✓	/d/
7. WORD CLASS (/d/)	*LWs > FWs*	? *LWs ≥ FWs*	? *LWs ≥ FWs*	? *LWs ≥ FWs*	/d/

together in terms of strengthening. Moreover, while in the previous study the effect of LEXICAL STRESS was somewhat mixed (although it showed a tendency towards tokens in *stressed* syllables being more strengthened), in the present study the effect of LEXICAL STRESS is clear for both stops, especially for the dentialveolar stop. In sum, the effect of PROSODIC POSITION is secondary to that of LEXICAL STRESS in the read speech data.

It could be argued that the different strengthening patterns of the two voiced stops could be partially due to the presence of FWs in the dentialveolar data set. However, the results of the WORD CLASS models do not support such a claim.[5] The evidence for tokens belonging to LWs being more strengthened than tokens belonging to FWs is not that clear. This result is rather surprising because it contradicts the results of the previous study, in

[5] Recall that the subsets used to study the effect of WORD CLASS comprised only dentialveolar tokens in *PW-initial* position. In this position, *stressed* syllables corresponded to LWs, whereas *unstressed* syllables corresponded to either LWs or FWs.

which there was strong evidence for more strengthened realizations of the tokens in LWs. The result might be explained in terms of the different speech materials that were used in the two studies. Thus, a read speech task may favor a rather hyperarticulated pronunciation of both LWs and FWs, meaning that more strengthened realizations of tokens in FWs could be produced, whereas in spontaneous speech such an effect may be absent (other differences due to the different speech materials will be further discussed below). In sum, although it is possible that the inclusion of FWs may have led to less strengthening of tokens in *unstressed* syllables in *PW-initial* position in the dentialveolar set than in those of the bilabial set, this does not seem to be the case.

The results for the effect of language dominance, which was measured by means of the BLP SCORE (SCALED) variable, show that, although Yucatec Maya dominance may be related to more strengthened realizations, the evidence for this claim is not that strong. Moreover, the fact that 7 participants in the present study did not take part in the previous one (see Section 7.4) suggests that the results are rather robust, especially considering that the data sets are rather different. In sum, it is likely that the strengthening of voiced stops described is a characteristic of Yucatecan Spanish, independent of bilingualism.

The results for the effect of GENDER show that there is little evidence for an effect of GENDER on strengthening, evidence that points towards *female* speakers producing more strengthened realizations. In the previous study, there was also some weak evidence for an effect of GENDER in two models, but in one of them *females* produced more strengthened realizations while in the other it was *male* speakers who seemed to do so. In sum, the contradictory and weak evidence suggests that GENDER does not play a role in the strengthening of voiced stops.

In line with the expectations, *second* mentions were less strengthened than *first* mentions. This result was obtained for all models, which suggests that it is a robust result. Thus, the claim that reduction, which was equated with less strengthening (see Section 8.1), may also be observable in segments, seems to be confirmed. In light of the results and previous literature on voiced stops in Spanish (Soler & Romero, 1999; see Section 8.1), the possibility that *second* mentions may be less strengthened due to the second reading of the sentences being produced at a higher speech rate seems likely. However, speech rate was not measured, meaning that this claim would need to be tested further. In addition, the results in the present study are in line with the evidence that suggests that, in Yucatecan Spanish, *second* mentions of words are reduced in terms of DURATION (see Section 8.1).

Overall, there is more evidence for strengthening than in the previous study, which may be due to differences in speech rate and/or differences in the type of data set (see Section 4.2.2.3).[6] The evidence for a greater strengthening effect in the tokens analyzed comes from the higher percentage of *stop* realizations in the CATEGORY models (see Sections 8.3.1.3 and 8.3.2.3).[7] In the previous study, participants spoke at their normal speech rate, whereas in this study they were reading. Participants were instructed to read as if they were talking in a natural way, but reading and speaking in a conversation clearly differ. It is likely that the task itself favored a hyperarticulated pronunciation, at least for some speakers, which would help to explain the higher number of stop realizations in the present study than in the previous one.

The method used to measure duration and change in intensity (see Section 8.2.2) proved to be useful in the study of the Yucatecan Spanish data. Nevertheless, many tokens could not be measured using either this method or the manual annotation used in the previous study. This is to be expected because of the great degree of weakening that Yucatecan Spanish voiced stops as well as Spanish voiced stops in general can manifest. In terms of results, given the positive correlation between duration and change in intensity, it is unsurprising that the models for both dependent variables yielded similar results for each stop. Moreover, the category results are also in line with those for duration and change in intensity. Overall, Figures 8.1 and 8.10 showed that the tokens categorized as *stops* and *approximants* corresponded to more strengthened and more weakened realizations, respectively. Consequently, the category analysis can be considered complementary to those of duration and change in intensity.

8.5 Summary

The present study has investigated the strengthening of Yucatecan Spanish voiced stops by means of a read speech task. Strengthening was studied by taking automatic measures of acoustic duration and change in intensity, as well as by the presence of a release burst. Several linguistic parameters (prosodic position, lexical stress, word class, and repetition) and speaker-specific parameters (language dominance and gender) were considered. Results

[6] The role of accentuation will be discussed in Chapter 10.

[7] The comparison of the DURATION results of the two sets of studies is problematic because DURATION was segmented differently. Thus, it is unsurprising that the range of values differs considerably between the corpus study (bilabial: 11–274 ms, dentialveolar: 10–94 ms) and the read speech study (bilabial: 23–184 ms, dentialveolar: 30–191 ms). Recall that CHANGE IN INTENSITY was only examined in the read speech study.

show that there is evidence for a strong effect of lexical stress and also, to a lesser extent, an effect of prosodic position. The strengthening of tokens in FWs and LWs words does not differ greatly, whereas tokens in second mentions are less strengthened than in first mentions. The effect of language dominance and gender is unclear, suggesting that they may not have an effect on strengthening.

Chapter 9

Glottalization of word-initial vowels — read speech study

9.1 The present study

The goals of the present study are twofold: (i) to examine the effect that several linguistic and speaker-specific factors have on the glottalization of word-initial vowels in Yucatecan Spanish, and (ii) to investigate the phonetic nature of this glottalization.[1] As already noted in Section 4.3.3, in this dissertation, *glottalization*, which is a manifestation of prosodic strengthening, refers to glottal stops and perceptually equivalent realizations, usually produced with creaky voice. Glottalization is an acoustic cue to prosodic strengthening, which may help cue prosodic boundaries. By means of these goals, I hope to obtain evidence for the research questions related to vowels, specifically regarding where prosodic strengthening occurs, whether language dominance and gender are sources of speaker-specific variation that explain the extent and distribution of strengthening, and what the phonetic nature of the strengthening is (see Section 1.1). The first goal of the study is addressed in Section 9.2, while the second goal is addressed qualitatively in Section 9.3 and acoustically in Section 9.4. The data for the studies in the present chapter come from the read speech task described in Chapter 7.

In the second part of this chapter, Section 9.2, the effect of several variables on glottalization is examined. Some of these variables have already been considered in the previous study (i.e., prosodic position, lexical stress, language dominance, and gender), while others are new (repetition and word class).[2] By investigating the same factors but in a more controlled way, I hope to assess whether the previous results can be replicated, thus providing more evidence in support of the effects already found. The previous study showed that glottalization is a rather frequent phenomenon in spontaneous speech in Yucatecan Spanish. Moreover, it showed that glottalization was indeed an important cue to lack of resyllabification. Prosodic position in the Prosodic Word (PW) and lexical stress had an effect on glottalization in that glottalization was more frequent for vowels across PWs, whether in unstressed or stressed syllables (e.g., (**e**n el mercado)$_{PW}$ 'in the market',

[1] For ease of expression, *glottalization of word-initial vowels* will be shortened to *glottalization*. Additionally, the study reported in Chapter 6 will be referred to as *the previous study.*

[2] Word class was examined in Chapter 6, but only in the random forest analyses.

(**a**gua)$_{\text{PW}}$ ‘water’), and for vowels in stressed syllables within the PW (e.g., (*en el **agua***)$_{\text{PW}}$ ‘in the water’). Conversely, glottalization was less frequent when the word-initial vowel appeared in an unstressed syllable within the PW (e.g., (*en **el** mercado*)$_{\text{PW}}$). In sum, the previous study provided some evidence for glottalization as a prosodic strengthening phenomenon to mark boundaries between PWs.

In the present study, prosodic position is categorized differently in order to better account for differences that may appear within the PW and that were not considered in the previous study; these differences may be related to word class. Specifically, prosodic position is categorized as follows: (i) in PW-initial position ((***en*** *el agua*)$_{\text{PW}}$, (***agua***)$_{\text{PW}}$), which corresponds to *across PWs* in the previous study, (ii) in lexical-word (LW) initial position (e.g., (*en el* (***agua***)$_{\text{LW}}$)$_{\text{PW}}$), and (iii) across function words (FWs) ((*en **el** agua*)$_{\text{PW}}$). The last two labels correspond to *within the PW* in the previous study. In PW-initial and LW-initial position, vowels can appear in stressed and in unstressed syllables, whereas they can only appear in unstressed syllables *across FWs*, since all FWs included in the read speech task were unstressed (see Section 7.2.3). Tentatively, I expect the most instances of glottalization in PW-initial position (PW-initial > LW-initial > across FWs) and in stressed syllables (versus unstressed ones). In terms of the interaction between prosodic position and lexical stress, the expected results cannot take the results of the previous study as their basis because of the different categories used for prosodic position. Tentatively, I expect the following pattern, from more to less glottalization: PW-initial × stressed/unstressed > LW-initial × stressed > LW-initial × unstressed > across FWs.

The results corresponding to the two speaker-specific variables already investigated in the previous study, language dominance and gender, showed that Yucatec Maya dominance favored glottalization, while gender (most likely) had no effect on it. Consequently, I expect that the results in the present study will be similar.

The two new variables examined are repetition (first versus second mention) and word class. In Chapter 8, there was a clear effect of repetition on the strengthening of voiced stops, with second mentions being less strengthened. The present study will test whether there is a similar effect for glottalization, with second mentions presenting fewer instances of glottalization. Finally, word class is considered by examining glottalization of LWs and FWs in PW-initial position. Tentatively, and for ease of argumentation, I expect no differences in terms of glottalization between LWs (which may be stressed or unstressed in the read speech task) and FWs (unstressed). Table 9.1 shows the expected results for the variables included in the present study.

Table 9.1. Expected results for glottalization in the read speech study.

Variables	Expected results
1. PROSODIC POSITION	PW-initial > LW-initial > across FWs
2. LEXICAL STRESS	stressed > unstressed syllables
3. PROSODIC POSITION × LEXICAL STRESS	PW-initial × stressed/unstressed > LW-initial × stressed > LW-initial × unstressed > across FWs
4. LANGUAGE DOMINANCE	Yucatec Maya > Yucatecan Spanish
5. GENDER	no effect
6. REPETITION	first > second mention
7. WORD CLASS	no effect

In the third and fourth parts of the present study, the acoustic manifestation of glottalization in Yucatecan Spanish is examined by means of two complementary analyses following Di Napoli (2018), one qualitative (Section 9.3) and another quantitative (Section 9.4). The qualitative analysis consists of a classification of the types of glottalization found in the data. This classification follows that of creaky voice proposed by Keating et al. (2015), with the addition of a category for glottal stops (see Section 4.3.1.3). The quantitative analysis of glottalization takes into consideration two acoustic measures of spectral tilt (H1*–H2* and H1*–A3*) and one of periodicity/noise in the signal (cepstral peak prominence, or CPP), which were introduced in Section 4.3.1.2. These parameters have been investigated both in terms of their articulatory correlates and of how they relate to phonation types in several languages (see Di Napoli, 2018, pp. 77–93, for discussion), also for Yucatec Maya (see Section 4.3.2). More importantly, they have also been studied as acoustic correlates of glottalization (e.g., Di Napoli, 2018, for Italian; Garellek, 2012a, for American English). Furthermore, when examining the role of creaky phonation in Yucatec Maya, that is, a language in which there is phonological contrast between modal and creaky vowels, it was indicated that nonmodal phonation can be limited to only part of the vowel. In fact, this can also be the case when glottalization is a marker of phrase boundaries. For example, in Italian, glottalization in phrase-final position seems to be restricted to the second half of the vowel (Di Napoli, 2018). Consequently, the goals of the quantitative analysis of glottalization are (i) to investigate how the acoustic measures serve to best differentiate between glottalized and nonglottalized (henceforth, *modal*) vowels, and (ii) to examine where in the vowel there is acoustic evidence for glottalization.

The remainder of the chapter is organized as follows. Section 9.2 examines the effect of several linguistic and speaker-specific variables on the presence of glottalization. Section 9.3 presents an analysis of the types of glottalization that were found in the data set. This is followed by an analysis of several acoustic measures of glottalization (H1*–H2*, H1*–A3*, and CPP) in Section 9.4. Finally, Section 9.5 summarizes the findings of the previous chapters.

9.2 Presence of glottalization

9.2.1 Methods

9.2.1.1 Speech materials

Based on the design of the read speech task and the fact that recordings by 21 speakers were taken into account (see Section 7.4), the maximum number of phrase-medial data points possible was 2016 (2 proper names × 6 repetitions × 21 speakers; 42 tokens × 2 repetitions × 21 speakers).[3] As indicated in Section 7.6, all instances that appeared after boundaries introduced by the participants (i.e., tokens in phrase-initial position) were excluded. Several data points were also excluded due to being in a syllable that bore postlexical secondary stress (see Section 2.2.1.1), due to errors, background noise, or hesitations. In total, the data set analyzed comprised 1723 tokens.

9.2.1.2 Annotation of glottalization

Two coders, one the author of this dissertation and the other a native speaker of German and proficient in Spanish, assessed whether or not glottalization was perceived for each token. Following Dilley et al. (1996) and Redi and Shattuck-Huffnagel (2001), only tokens for which there was auditory and acoustic evidence were considered for further analysis. In terms of auditory evidence, the interrater agreement for perceived glottalization was *almost perfect* (Fleiss' $\kappa = 0.84$; see Landis & Koch, 1977, for the strength of the agreement associated with values of the statistic). Consequently, the annotations by the author of the dissertation were followed. The acoustic evidence came from further inspection of the waveforms and spectrograms, which showed that tokens perceived as presenting glottalization did indeed present acoustic cues to glottalization.

[3] Recall that proper names (e.g., *el Ético*) appeared three times per set of sentences (see Section 7.2.1), meaning that they were in fact mentioned six times each. However, given the fact that repetitions of sentences were produced in pairs, with the second mention in a sentence being produced right after the first one, tokens in proper names were coded as *first* or *second* mention.

The vowels were segmented from their onset to their offset, examining changes in amplitude (sudden increase/decrease) in the waveform and formant trajectories in the spectrogram. The precise locations for the onset and the offset were determined by means of the *move cursor to nearest zero crossing* option in Praat (Boersma & Weenink, 2019).

9.2.1.3 Statistical analysis

In order to examine the presence of glottalization, two logistic mixed models were fitted with the `brms` package (Bürkner, 2017) in the R programming environment (R Core Team, 2020). The first model was fitted for the categorical dependent variable GLOTTALIZATION (with the levels *modal* and *glottalized*). The predictors were PROSODIC POSITION (levels *LW-initial* and *PW-initial*),[4] LEXICAL STRESS (levels *unstressed* and *stressed*), BLP SCORE (SCALED), GENDER (levels *female* and *male*), and REPETITION (levels *first* and *second*), with an interaction term for PROSODIC POSITION × LEXICAL STRESS. SPEAKER and WORD were included as random intercepts.

The second model examined GLOTTALIZATION for tokens in *PW-initial position* in order to investigate the effect of the predictor WORD CLASS (levels *FW* and *LW*). SPEAKER and WORD were included as random intercepts.[5]

For both models, the likelihood function used was Bernoulli. Weakly informative priors were specified for the intercept (*Normal* (0, 3)), for the parameters representing the effects of the predictors and the interaction (*Normal* (0, 3)), and for the standard deviation of the random effect (*Normal* (0, 5)) (see Section 5.2.6.2 for an explanation of the same prior specification). The convergence of the models, as well as their fitting and the assessment of goodness of fit, were evaluated following the procedure described in Section 5.2.6.2. The package `emmeans` (Lenth, 2020) was used to obtain estimated marginal means and to conduct pairwise comparisons of the levels of the predictors of the interaction.

9.2.2 Results

Of the total of 1723 tokens, 314 (18.22 %) presented cues to glottalization. Table E1 in Appendix E provides the raw count of tokens per speaker. The number of glottalized tokens

[4] There were only 4 instances of glottalization *across FWs* (see Section 9.2.2). Consequently, these tokens, as well as the level *across FWs*, were excluded from the statistical analyses presented in Section 9.2.

[5] POSITION OF THE CARRIER WORD IN THE SENTENCE, which could be *initial, medial*, or *final* (see Section 7.2.3), was also examined in a model parallel to that of WORD CLASS to confirm that it did not have an effect on the rate of glottalization. Results show that this was indeed the case ($\hat{\beta}_{intercept}$ = -1.69, CI [-4.28, 0.90]; $\hat{\beta}_{medial}$ = -0.68, CI [-3.25, 1.94]; $\hat{\beta}_{final}$ = 0.06, CI [-2.51, 2.67]). Consequently, this variable is not discussed further.

across FWs was 4 out of 470, that is, 0.85 % of the tokens in that position. Consequently, tokens *across FWs* were excluded from the statistical analyses in Section 9.2, as already indicated, which led to a data set with 1253 tokens. In *LW-initial* position, 203 out of 826 tokens were labeled as glottalized (i.e., 24.58 %), while in *PW-initial* position 107 tokens out of 427 were labeled as such (i.e., 25.06 %). These percentages clearly show that glottalization is almost absent from the level *across FWs*, while it appears in a sizeable number of tokens in the other two levels of PROSODIC POSITION.

Figure 9.1 shows the percentage of glottalized vowels per PROSODIC POSITION and SPEAKER, ordered according to their BLP SCORES. Participants inf27, inf04, and inf49, whose BLP scores are 23, 66, and 148, respectively, did not produce glottalized tokens for one of the two levels of PROSODIC POSITION.[6]

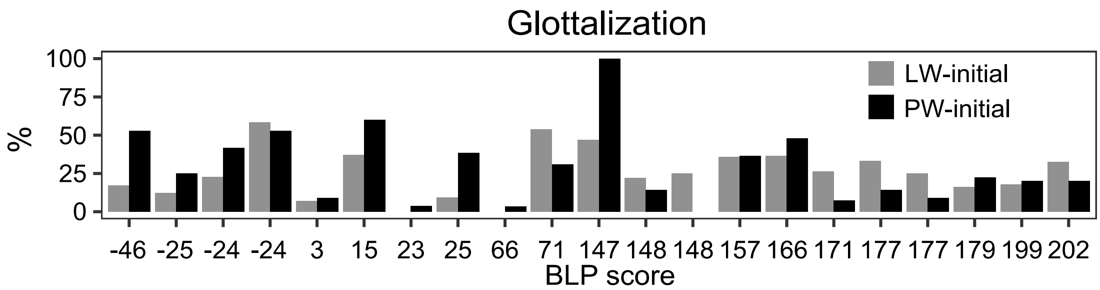

Figure 9.1. Percentage of realizations that present glottalization in *LW-initial* and *PW-initial* position per speaker in the read speech task. Each BLP score on the x-axis corresponds to one speaker.

The results corresponding to PROSODIC POSITION show that there is evidence that the levels *LW-initial* and *PW-initial* differ from each other ($\hat{\beta}_{\text{PW-initial}} = 1.61$, CI [0.24, 3.01]), with a higher probability of *glottalized* tokens in the latter. The posterior probability of the parameter being greater than 0 is 0.99, that is, there is a 0.99 probability of glottalization being more likely to be produced in *PW-initial* than in *LW-initial* position. For LEXICAL STRESS, the probability of *glottalized* tokens is higher in *stressed* syllables than in *unstressed* ones ($\hat{\beta}_{\text{stressed}} = 3.50$, CI [2.20, 4.85]). The posterior probability of the parameter being greater than 0 is 1, that is, there is a probability of 1 of glottalization being more likely to be produced in *stressed* than in *unstressed* syllables. Consequently, we can draw the inference

[6] The large proportion of *glottalized* tokens in *PW-initial* position produced by speaker inf44 (BLP score = 147) also deserves comment. The result may be due to the fact that approximately two thirds of his recordings had to be discarded because of strong background noise. The inclusion of the missing tokens (maybe produced more frequently without glottalization?) might have yielded a different result (i.e., a lower percentage of *glottalized* tokens), more consistent with the overall results.

that the probability of *glottalization* is greater in *PW-initial* position and in *stressed* syllables.

Pairwise comparisons for the levels of the predictors of the interaction (see Table E2 in Appendix E) show that all levels differ from each other, with the exception of the pair *LW-initial* × *stressed* – *PW-initial* × *stressed*, which shows that these two levels pattern together. However, since most of the probability mass for this comparison is below 0 (-1.28, CI [-2.92, 0.22]), it might be reasonable to conclude that there is some weak evidence for a difference between the levels, with a smaller probability of glottalized tokens in *LW-initial* × *stressed*, although we cannot be certain about it (see Figure 9.2).

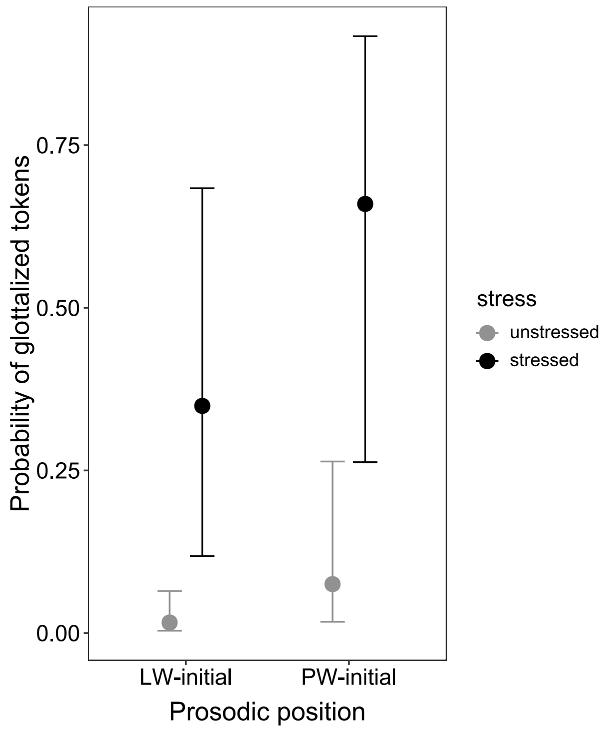

Figure 9.2. Interaction plot. Effect of PROSODIC POSITION conditional on LEXICAL STRESS: GLOTTALIZATION model in the read speech study.

The results for the BLP SCORE (SCALED), GENDER, and REPETITION predictors indicate that 0 is included within the CI ($\hat{\beta} = 0.08$, CI [-0.70, 0.87]; $\hat{\beta}_{genderM} = 0.19$, CI [-1.38, 1.77]; $\hat{\beta}_{rep2} = -0.22$, CI [-0.58, 0.14]). The estimated probability of the parameter being less than 0 is 0.42 for BLP SCORE (SCALED), 0.4 for GENDER, and 0.88 for REPETITION, which suggests that there is no evidence for an effect of BLP SCORE (SCALED) or GENDER and some weak evidence in support of *second* mentions being produced with *glottalization* less often than *first* mentions, but that we cannot be certain about this.

The second model examined glottalization for tokens in *PW-initial position* in order to investigate the effect of the predictor WORD CLASS. Of the total of 427 tokens in this position, 237 tokens appeared in *LW*s (of which 87 were *glottalized*, i.e., 36.71 % of all *LW*s) and 190 in *FW*s (20 *glottalized*, i.e., 10.53 %). *LW*s could be *stressed* (113, of which 66 were *glottalized*, i.e., 49.62 %) or *unstressed* (124, of which 21 were *glottalized*, i.e., 16.94 %), while *FW*s could only be *unstressed*.

The results for WORD CLASS indicate that 0 is included within the CI ($\hat{\beta}_{LW} = 1.59$, CI [-1.31, 4.28]). The posterior probability of the parameter being greater than zero is 0.88, that is, there is a 0.88 probability that *LW*s are produced with more instances of tokens with *glottalization* than in the case of *FW*s, thus suggesting that there may be some weak evidence for an effect of WORD CLASS, but that we cannot be certain about this.

9.2.3 Discussion

In terms of the distribution of glottalization, some results confirm the findings of Chapter 6, whereas others contradict them. Table 9.2 provides an overview of the models' results.

Table 9.2. Results for the presence of glottalization in the read speech study.

Variables	Expected results	Results
1. Prosodic position	*PW-initial* > *LW-initial* > *across FWs*	✓
2. Lexical stress	*stressed* > *unstressed* syllables	✓
3. Prosodic position × lexical stress	*PW-initial* × *stressed/unstressed* > *LW-initial* × *stressed* > *LW-initial* × *unstressed* > *across FWs*	× (*PW-initial* ≥ *LW-initial* × *stressed*) > *PW-initial* × *unstressed* > *LW-initial* × *unstressed* > *across FWs*
4. Language dominance	Yucatec Maya > Yucatecan Spanish	×
5. Gender	no effect	✓
6. Repetition	*first* > *second* mention	? *first* ≥ *second* mention
7. Word class	no effect	? LWs ≥ FWs

Overall, the results indicate that LEXICAL STRESS is the main factor that determines presence of glottalization, with tokens in *stressed* syllables having a greater probability of presenting glottalization than tokens in the *unstressed* ones. There is also some evidence of an overall effect of PROSODIC POSITION, with a higher probability of glottalization in *PW-initial* position than in *LW-initial* position (and in *LW-initial* position than *across FWs*). These results are in line with the findings of previous studies on other languages, which have found that glottalization is more frequently associated with prominent positions than with prosodic boundaries (see Section 4.3.3).

The interaction between PROSODIC POSITION and LEXICAL STRESS indicates that the higher probability of glottalization in *PW-initial* position is mostly due to the greater probability in the case of *unstressed* tokens in *PW-initial* position than of the *unstressed* ones in *LW-initial* position; a similar effect of PROSODIC POSITION is observed for the *stressed* tokens, although to a lesser extent. Contrary to the expectations based on the previous study, the two *stressed* levels tend to pattern together, which indicates the predominant effect of LEXICAL STRESS. In sum, the results indicate that glottalization in Yucatecan Spanish is a manifestation of prosodic strengthening that serves to mark the left edge of the PW and, more importantly, to signal lexical stress.[7]

The present study found no support for an effect of LANGUAGE DOMINANCE. This result contradicts the findings of Chapter 6, in which Yucatec Maya dominance favored more glottalization. In sum, it is unclear if this variable has an effect on glottalization. In line with the previous study, GENDER had no effect on the presence of glottalization, thus lending support to the statement that glottalization is independent of GENDER. Finally, there may some weak evidence in support of *first* mentions presenting more instances of glottalization than *second* mentions, but further studies are needed to confirm or reject this possibility.

For WORD CLASS, the results of the comparison between *FWs* and *LWs* in *PW-initial* position show some weak evidence that *LWs* may present more glottalization than *FWs*. A possible explanation may be that glottalization was more frequent in *stressed* than in *unstressed LWs*, thus lending further support for the effect of LEXICAL STRESS on glottalization. This appears to be the case. Although LEXICAL STRESS could not be fitted into the model for WORD CLASS because *FWs* could only be *unstressed*, the overall data presented in Section 9.2.2 point to more glottalization in *stressed LWs* than in *unstressed LWs* and *FWs*.

Finally, a comparison of the ratio of glottalization in the previous study (16.62 %) and the present (18.22 %) suggests that glottalization in Yucatecan Spanish may be independent of the speech materials used, although further research would be needed to support this claim.

[7] See Chapter 10 for a further interpretation of these results in terms of accent.

9.2.4 Summary

In this section, the effect of several variables on the presence of glottalization was examined. Lexical stress and prosodic position have an effect on glottalization, with more instances primarily in stressed syllables and in PW-initial position. As such, glottalization signals lexical stress and it also serves to mark the boundaries of PWs. The results indicate that language dominance, gender, and word class have no effect on glottalization, while there is some weak evidence in support of an effect of repetition, with more glottalization in first than second mentions.

9.3 Types of glottalization

9.3.1 Methods

In the present study, the 314 tokens that were labeled as *glottalized* in the study of Section 9.2 were further analyzed in order to provide a qualitative analysis of the types of glottalization found in the data. Since several studies have shown that acoustic cues to glottalization can be present across word boundaries (e.g., Dilley et al., 1996, and Garellek, 2012a, for American English; Kohler, 2001, for German), it was also investigated whether the sound preceding the vowels studied presented cues to glottalization.

In the present study, *types of glottalization* encompass different manifestations of creaky voice and glottal stops. The classification of creaky voice types in Keating et al. (2015) is followed (see also Section 4.3.1.3), with the addition of the category *glottal stop*. Following Garellek (2012a), a glottal stop differs from creaky voice by the presence (in the former) of at least a two-pulse period of silence after the burst.[8] This characterization is due to the fact that, in some instances of vowel–vowel sequences that present irregular F0, it is not always clear whether the correct classification should be glottal stop or creaky voice.

9.3.2 Results

Table 9.3 shows the types of glottalization that were found in the data, with the exception of the glottal stop. The main acoustic characteristics that can be observed in the waveform and that were used to classify the tokens are also provided (after Keating et al., 2015).[9]

[8] See also Figure 9.3e.

[9] The original table in Keating et al. (2015), as well as Table 4.3 in the present dissertation (which is a simplification of Keating et al.'s table as well), include low H1–H2 as a correlate of glottal constriction for all the types of creaky voice that are found in Yucatecan Spanish. Because the classification presented here was based on the visual analysis of waveforms and spectrograms, the measure is not included in the table.

Table 9.3. Main characteristics of the types of glottalization that were found in the Yucatecan Spanish data (with the exception of the glottal stop). The table is based on the one in Keating et al. (2015, p. 3). Empty cells indicate that the acoustic correlate is either variable or unknown.

Property **Type**	**low F0**	**irregular F0**
prototypical creaky voice	✓	✓
vocal fry	✓	
multiply pulsed voice		✓
aperiodic voice	NO	✓

Figure 9.3 shows an example of each type of glottalization. In practice, *prototypical creaky voice* and *vocal fry* were not always easy to distinguish. According to Keating et al. (2015), F0 is regular in vocal fry, but it can also be irregular. The defining characteristic of vocal fry is its great degree of damping of the pulses, which results in them being perceived as individual pulses. Consequently, this perceptual cue was used to distinguish between doubtful cases of prototypical creaky voice versus vocal fry. *Multiply pulsed voice*, which presents alternating longer and shorter pulses, presented double pulsing, that is, two simultaneous periodicities (triple-pulsed creak, which is also possible, according to Keating et al., 2015, was not found in the data). Finally, in *aperiodic voice*, there is no periodicity, due to the irregular vibration of the vocal folds (Keating et al., 2015).

Figure 9.4 presents the frequency of each type of glottalization. Clearly, the most common manifestation in the data is prototypical creaky voice, followed by vocal fry. The frequency of the other three categories, multiply pulsed, aperiodic, and glottal stop, is much lower. Twenty-five tokens (7.96 %) presented characteristics of two different categories (*prototypical creaky voice + vocal fry, prototypical creaky voice + multiply pulsed,* or *aperiodic + prototypical*). In Figure 9.4, these tokens appear in both categories. Table E3 in Appendix E presents the raw count number of each type.

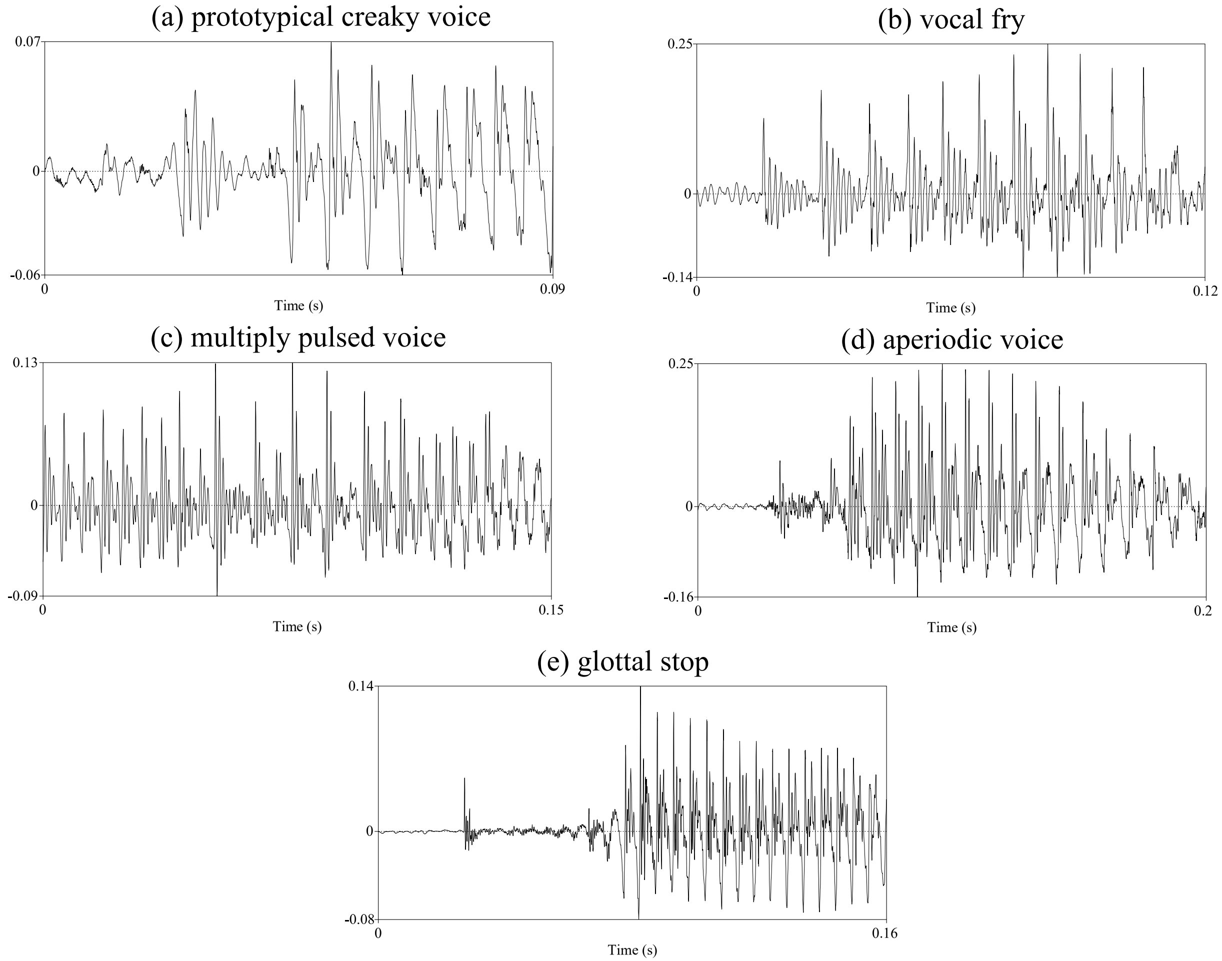

Figure 9.3. Examples of types of glottalization (in the first portion of the vocalic segment displayed): (a) prototypical creaky voice (***a*** *los turistas*; speaker inf18, female), (b) vocal fry (***á****cido*; speaker inf48, male), (c) multiply pulsed voice (***á****cido*; speaker inf27, female), (d) aperiodic voice (***á****cido*; speaker inf16, male), (e) glottal stop (***á****rbitro*; speaker inf19, female).

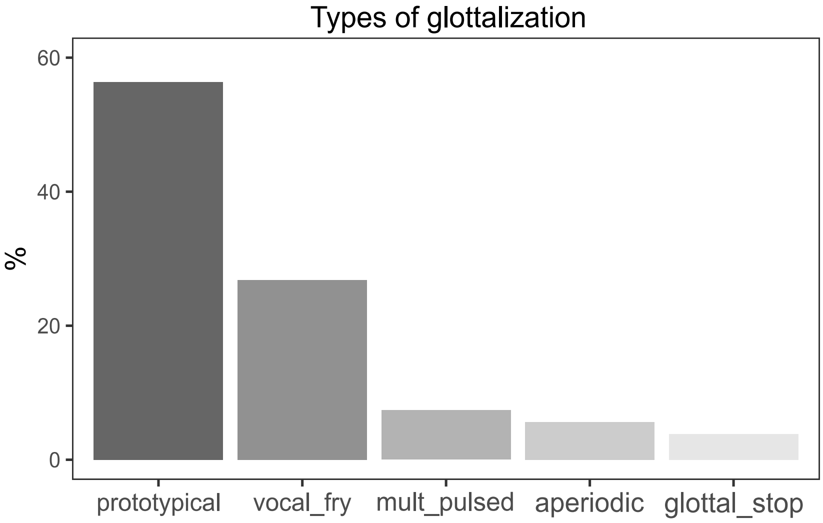

Figure 9.4. Frequency (in %) of each type of glottalization in the read speech study. ‘Prototypical’ – *prototypical creaky voice*; ‘mult_pulsed’ – *multiply pulsed voice*; ‘aperiodic’ – *aperiodic voice.*

As indicated above, the acoustic cues to glottalization can be present across word boundaries; that is, the word-final sound that precedes the vowel analyzed may also show evidence of glottalization. This was indeed the case for 232 tokens out of the total of 314 (i.e., 73.89 %). The word-final sounds that presented cues to glottalization corresponded to the word-final vowels, /l/, and /n/, that is, there were no cues to glottalization for word-final /s/, which is expected because it is rare for fricative sounds to co-occur with nonmodal phonation (Ladefoged & Maddieson, 1996, p. 178).

9.3.3 Discussion and summary

The analysis in this section has shown that the glottalization of word-initial vowels in Yucatecan Spanish can be manifested through different types of glottalization. This result is in line with the findings of Redi and Shattuck-Hufnagel (2001), who showed that glottalization can be produced in multiple ways, as well as the results of studies on other languages (e.g., Di Napoli, 2018; Garellek, 2013). Prototypical creaky voice is the most common one, followed by vocal fry, multiply pulsed voice, aperiodic voice, and glottal stop. Furthermore, glottalization of the sound at the end of the preceding word is also frequent, in line with literature on other languages.

9.4 Acoustic measures of glottalization

9.4.1 Methods

The goal of this study is to investigate how several acoustic measures can distinguish between glottalized vowels and modal vowels in Yucatecan Spanish, as well as examining where in the vowel glottalization takes place. Three parameters that have been used to distinguish between glottalized vowels and modal vowels are examined (see Section 4.3.1.2). Two of these are measures of spectral tilt (H1*–H2* and H1*–A3*), while the third one is one of periodicity/noise (cepstral peak periodicity, or CPP). The amplitude of the first harmonic relative to the second one, H1*–H2*, is a correlate of open quotient, whereas the amplitude of the first harmonic relative to the third formant spectral peak, H1*–A3*, is considered a correlate of the closing velocity of the vocal folds.[10] Glottalized vowels should have lower values of these two measures than modal vowels. Furthermore, the types of glottalization that were found in the Yucatecan Spanish data, which were presented in the previous section, are described as having low values of H1–H2 in Keating et al. (2015). The measure of periodicity/noise, CPP, covers the whole range of frequencies (Shue et al., 2011).[11] Noise in glottalized (creaky) realizations is due to the irregular vibration of the vocal folds (e.g., Gobl & Ní Chasaide, 2010, p. 401). Because glottalized vowels present weaker peaks in the ceptrsum than modal vowels due to the noise that originates from the irregular vibration of the vocal folds in the former, the values of CPP should be lower for glottalized than for modal vowels.

PraatSauce (Kirby, 2018–2019) was used to obtain the above-mentioned acoustic measures. *PraatSauce* is a set of Praat scripts, similar to *VoiceSauce* (Shue, 2010) in terms of the measures that it can produce. For each vowel, 5 equidistant points that covered the first two thirds of each vowel were measured, that is, measurements were taken at the 0 %, 17 %, 33 %, 50 %, and 67 % of the vowel. There were two reasons for the decision to measure the first two thirds of the vowel. First, the visual inspection of the waveforms and spectrograms in the study in Section 9.3 showed that glottalization was predominantly limited to the beginning of the vowel. Second, by excluding the last third of the vowel, the influence that the beginning of the gesture for the next syllable may have on the vowel is excluded. In

[10] Recall from Section 4.3.1.2 that the asterisks indicate that the measures have been corrected to remove the filtering function of the vocal tract.

[11] Another measure of periodicity/noise, *harmonics-to-noise ratio* (HNR), was also introduced in Section 4.3.1.2. I chose to examine only CPP because it seemed redundant to use two measures for the same phenomenon. Moreover, HNR is calculated over frequency ranges set by the researcher (de Krom, 1993).

conclusion, by measuring two thirds of the vowel, we should still be able to obtain evidence of whether there is a change in the acoustic parameters over the course of the vowel.

Each of the 5 points was measured with a window of 25 ms. The F0 range was set at 50–300 Hz for female speakers and at 50–500 Hz for male speakers. Formant tracking was performed in the 0–5,500 Hz range. The estimated values of the formants were 550 Hz for F1, 1550 Hz for F2, and 3000 Hz for F3.[12] The spectral measures were calculated through *Praat* internal functions. Measures of spectral tilt were corrected according to the method in Iseli et al. (2007). Because of the time normalization of the measures, only vowels whose duration was within one standard deviation from the mean duration of the glottalized vowels were analyzed, following Garellek (2012b). Instances where F0 and F3 had not been properly tracked were discarded since the validity of the measures of spectral tilt depends on their correct tracking. The total number of observations included in the analyses amounted to 1049 measures corresponding to modal vowels, and 554 measures corresponding to glottalized vowels.

9.4.1.1 Statistical analysis

A linear mixed model for each of the three variables (H1*–H2*, H1*–A3*, and CPP) was fitted with the `brms` package (Bürkner, 2017) in the R programming environment (R Core Team, 2020). Because the effect of several linguistic and nonlinguistic factors has already been investigated for the whole data set (Section 9.2), the models for the three acoustic variables include only VOWEL TYPE (levels *modal* and *glottalized*), MEASURE, which refers to the points in the vowels where the acoustic measures were taken (levels *0%, 17%, 33%, 50%*, and *67%*), and an interaction term for the two predictors. SPEAKER and WORD were included as random intercepts. The likelihood function, the priors, the evaluation of convergence of the models, their fitting and assessment of goodness of fit, as well as the calculation of the pairwise comparisons between the levels of the predictors, were obtained following the procedure indicated in Section 8.2.3 for the CHANGE IN INTENSITY models of the voiced stops. The maximum tree depth was kept at its default value of 10. The package `emmeans` (Lenth, 2020) was used to obtain estimated marginal means and to conduct pairwise comparisons of the levels of the predictors of the interaction.

[12] In a first analysis of the data, F3 was kept at its default value in *PraatSauce* (2500 Hz). The default value is in line with the values of F3 of Spanish vowels (see Albalá et al., 2008). However, an inspection of the data showed that F3 had not been properly tracked in many instances. Further inspection of the spectrograms showed that the values of F3 for the majority of speakers were around 3000 Hz. Consequently, the estimated value of F3 was changed to 3000 Hz, which significantly improved the tracking.

9.4.2 Results

9.4.2.1 H1–H2**

The results of the model provide evidence that the values of H1*–H2* do not differ greatly as a function of VOWEL TYPE ($\hat{\beta}_{glottalized}$ = -0.59, CI [-2.38, 1.21]). The posterior probability of the parameter being less than 0 is 0.74, that is, there is a 0.74 probability of lower values of H1*–H2* for *glottalized* vowels than for *modal* ones, but we cannot be certain about this.

Pairwise comparisons between the levels of VOWEL TYPE × MEASURE indicate that the MEASURE values are rather similar *within* each VOWEL TYPE (e.g., *modal* × *0 %* compared to *modal* × *17 %*), since 0 is included in the 95 % CI of the difference estimates of the comparisons (see Table E4 in Appendix E). Furthermore, the pairwise comparisons of the levels of MEASURE and VOWEL TYPE (e.g., *modal* × *0 %* compared to *glottalized* × *0 %*) suggest that the values of H1*–H2* are very similar for each contrast of levels. This is also reflected in Figure 9.5, which plots the conditional effects of the interaction between VOWEL TYPE and MEASURE for H1*–H2*.

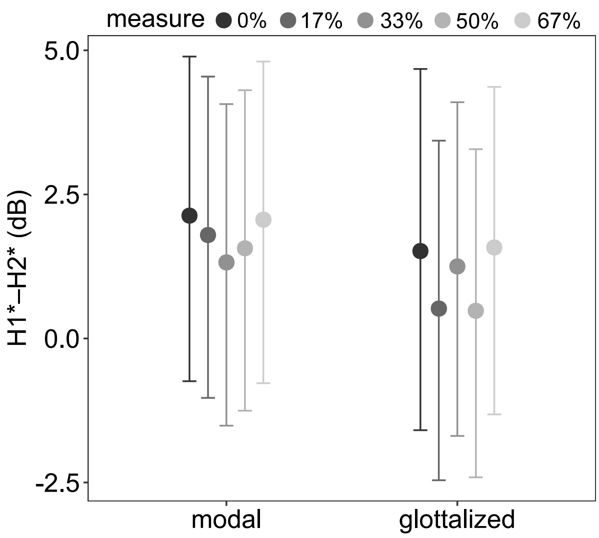

Figure 9.5. Interaction plot. Effect of VOWEL TYPE conditional on MEASURE for H1*–H2*.

9.4.2.2 H1–A3**

The results of the model provide strong evidence that the values of H1*–A3* differ as a function of VOWEL TYPE ($\hat{\beta}_{glottalized}$ = -3.42, CI [-5.91, -0.92]), with lower H1*–A3* values in the case of *glottalized* vowels. The posterior probability of the parameter being less than 0 is 1, that is, there is a probability of 1 of lower values of H1*–A3* for *glottalized* vowels than for *modal* ones.

The pattern within each VOWEL TYPE is rather alike. The levels of MEASURE within *modal* vowels present similar values among themselves according to pairwise comparisons of the levels of the interaction, as do the levels within *glottalized* vowels, although there is some weak evidence for *glottalized* × *17 %* having lower values than *glottalized* × *33 %*, since most of the probability mass is below 0 (see Table E5 in Appendix E). Importantly, there is evidence for a difference between *modal* and *glottalized* vowels for *0 %* and *17 %*, with lower values of H1*–A3* for the *glottalized* levels. These differences are depicted in Figure 9.6.

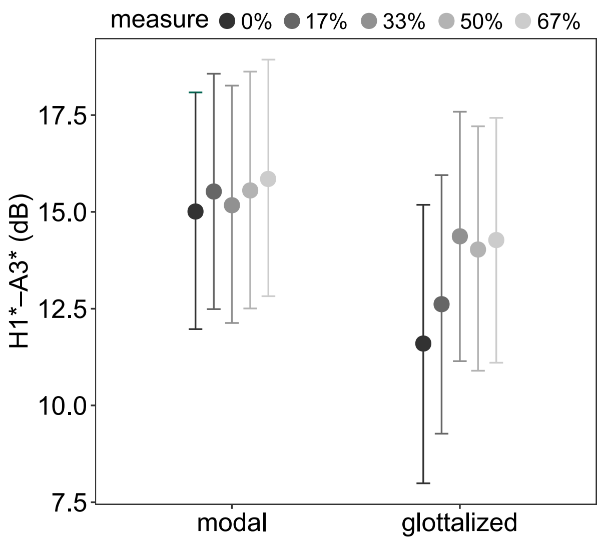

Figure 9.6. Interaction plot. Effect of VOWEL TYPE conditional on MEASURE for H1*–A3*.

9.4.2.3 Cepstral peak prominence

The results of the model provide strong evidence in support of the values of CPP differing as a function of VOWEL TYPE ($\hat{\beta}_{glottalized}$ = -4.63, CI [-5.81, -3.45]), with *glottalized* vowels having lower CPP values. The posterior probability of the parameter shows that there is a probability of 1 of lower values of CPP for *glottalized* vowels than for *modal* ones.

The results of the pairwise comparisons of the *modal* × MEASURE indicate that the CPP values increase from *0 %* to *17 %*. Pairwise comparisons of the *glottalized* × MEASURE indicate that the CPP values increase not only from *0 %* to *17 %*, but also from *17 %* to *33 %* and from *33 %* to *50 %*. The comparison of the levels of MEASURE between *modal* and *glottalized* vowels shows that the first four levels differ across VOWEL TYPES (i.e., *0 %, 17 %, 33 %,* and *50 %*). These differences are depicted in Figure 9.7.

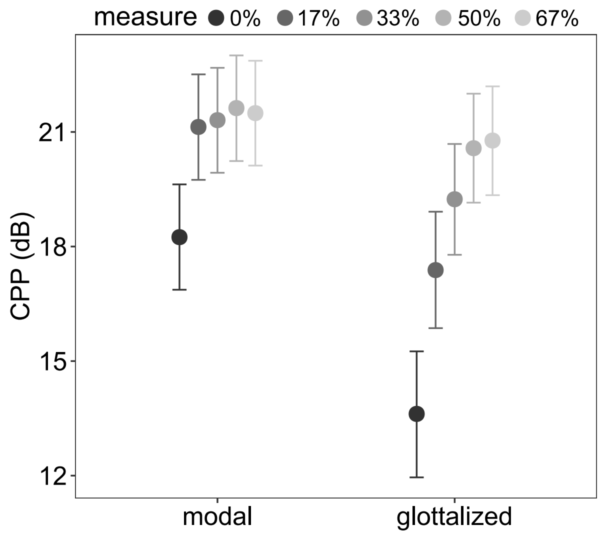

Figure 9.7. Interaction plot. Effect of VOWEL TYPE conditional on MEASURE for CPP.

9.4.3 Discussion and summary

Overall, the results show that modal and glottalized vowels clearly differ in terms of two acoustic parameters, H1*–A3* and CPP, but not in terms of H1*–H2*. Moreover, the evidence shows that glottalization takes place at the beginning of the vowel, more specifically in the first third of the vowel in the case of H1*–A3* and in the first half of the vowel in the case of CPP. Thus, these acoustic measures show that the glottalization of word-initial vowels in Yucatecan Spanish is restricted to part of the vowel, which is in agreement with the findings on glottalization as a marker of phrase boundaries in other languages, such as English (Garellek, 2012a) or Italian (Di Napoli, 2018).

The results of H1*–H2* are surprising for several reasons. First, it is well-attested crosslinguistically that H1*–H2* serves to distinguish between modal and nonmodal (both breathy and creaky) phonation (see Di Napoli, 2018, p. 82, for a thorough list of languages). Second, Avelino (2016) found that this measure of spectral tilt served to distinguish between modal and creaky vowels in Yucatec Maya (although H1*–A3* was a better index of differences between modal and creaky vowels; see Section 4.3.2). Finally, the results of the qualitative analysis in Section 9.3 showed that the most common glottalization type was prototypical creaky voice, followed by several others. With the exception of glottal stops, which constituted a small group of observations, all of the other types were expected to manifest lower values of H1*–H2* than modal vowels, according to the classification by Keating et al. (2015). In other words, glottalized vowels in the current data should present lower values of H1*–H2* than modal vowels, at least at their beginning. However, no strong evidence for such difference was found. Further research is needed to clarify whether there are variables that could explain this result or whether Yucatecan Spanish is an exception to a well-attested crosslinguistic finding.

9.5 Summary

The three studies described in this chapter provide evidence of glottalization in word-initial vowels in Yucatecan Spanish. While the first study examines further the factors that may have an effect in its distribution, following the results obtained in Chapter 6, the second and third studies provide qualitative and quantitative acoustic analyses of glottalization, respectively, for the first time in the study of Yucatecan Spanish.

The results indicate that glottalization in Yucatecan Spanish is a manifestation of prosodic strengthening that is used primarily to signal lexical stress and secondarily to mark the boundaries of words (PWs and LWs in the present study). Glottalization appears to be independent of language dominance, gender, word class, and probably also of repetition.

Acoustically, the most common type of glottalization is prototypical creaky voice, followed by vocal fry, multiply pulsed voice, aperiodic voice, and glottal stop. Glottalized vowels present lower values of H1*–A3* and CPP than modal values in their first half, but not of H1*–H2*.

Part IV

Further discussion and conclusion

Chapter 10

Lexical stress and accent in Yucatecan Spanish

10.1 The present study

The previous chapters in Parts II and III examined prosodic strengthening for voiced stops and vowels by means of a corpus of sociolinguistic interviews and a read speech task. A comparison of the results between the two speech styles shows a different effect of lexical stress on prosodic strengthening.

In corpus speech, in word-medial position, voiced stops in stressed syllables had longer duration and a higher number of stop realizations than in unstressed syllables, whereas in lexical-word-initial (LW-initial) and in Prosodic Word-initial (PW-initial) positions, lexical stress did not appear to have an effect (i.e., values for duration and stop realizations were similar for tokens in stressed and unstressed syllables). Tokens in word-initial position (meaning both LW-initial and PW-initial positions) were more strengthened than those in word-medial position, indicating that prosodic position had an effect on strengthening. For vowels, there were more instances of glottalization across PWs than within the PW, and there were more instances of glottalization when the vowels were in stressed syllables, although the main factor was still prosodic position.

In read speech, there was a clear effect of lexical stress: voiced stops in stressed syllables had longer duration, a greater change in intensity, and a higher number of stop realizations than voiced stops in unstressed syllables, while glottalization of vowels was more frequent for tokens in stressed syllables, both in LW-initial and PW-initial positions. In sum, whereas in corpus speech the effect of lexical stress appeared to be secondary to that of prosodic position, in read speech it was lexical stress, not prosodic position, that was the most important variable of the two.

The divergence between the results for corpus speech and for read speech begs the question of whether there may be another variable involved, namely pitch accent. In Sections 2.2.1.1 and 2.4.3, we saw that there is evidence in several languages (e.g., Dutch) that sounds in accented syllables may have longer duration than unaccented ones (both stressed and unstressed). However, it was also pointed out that this may not be the case in Spanish. In Section 2.2.1.1, the relationship between accent and stress in Spanish was discussed. In a nutshell, the few studies that have tried to disentangle the effects of accent and stress on duration and intensity in Spanish show that accent has no effect on duration, while it is unclear if there is an effect of accent on intensity (Ortega-Llebaria & Prieto, 2007; Torreira et al., 2014). It is worth noting that the data analyzed in those studies corresponded to read speech in an experimental setting, and that only European Peninsular varieties of Spanish were examined.

At the same time, several studies on Spanish (see Section 2.2.1.1) indicate that there are differences between spontaneous speech and read speech in terms of deaccenting, being considerably more frequent in spontaneous speech, with up to 30 % of instances being deaccented in one study (Face, 2003), while deaccentuation was only marginally present in read speech. Deaccentuation in Spanish is manifested by a lack of rise in F0 accompanying the stressed syllable (Face, 2003) and/or "when any type of F0 movement is absent from its stressed syllable" (Rao, 2007, p. 203). For the declarative sentences that Ortega-Llebaria and Prieto (2007) examined, "F0 is flat across the utterance and shows no pitch accent" (p. 158). These studies only considered pitch; thus, we have no information about whether deaccented syllables may be less strengthened as well, for example, in terms of duration. Moreover, with the exception of the preliminary study by Rao (2005), with one speaker of Mexico City Spanish, all the other studies referenced in Section 2.2.1.1 were conducted on European Peninsular Spanish.

If accent had an effect on prosodic strengthening in Yucatecan Spanish (which in our data would be manifested through longer duration of voiced stops or more presence of glottalization for vowels in accented syllables, for example), this would suggest that Yucatecan Spanish can be grouped with those languages for which there is a difference in terms of strengthening between tokens in accented and unaccented syllables, and not with Spanish, at least as far as the current studies show. Based on the literature, we would expect (i) that deaccented syllables[1] were more common in spontaneous speech than in read speech, and (ii) that accent would lead to an increase in prosodic strengthening, in line with the results for other languages, although not Spanish. Thus, deaccented syllables should be less strengthened than accented ones.

The present study aims to examine the role of accent in spontaneous (corpus) and read speech in the studies included in the present dissertation. Because accent was not a variable controlled for in the read speech studies (or in the corpus speech ones, for obvious reasons), the present study is preliminary. By taking a closer look at a selection of the data, this

[1] Recall that, in Spanish, stressed syllables are generally considered to be accented (see Section 2.2.1.1). Consequently, the term *deaccented* is preferred here over *unaccented*.

examination of accent may help to clarify why the effect of lexical stress on prosodic strengthening differed between the two speech styles. The questions that will guide this study are: (i) how frequent is pitch deaccenting in each speech style? (ii) which are the most common positions of pitch accents (e.g., prenuclear, nuclear) in each speech style? and (iii) is there any evidence that suggests that accent has an effect on prosodic strengthening?

10.2 Materials

In order to facilitate the comparison between corpus speech and read speech, only recordings of speakers who had participated in all four studies were considered. The number of participants was set to 4 (2 female, 2 male), with two speakers being Spanish-dominant ones (inf12 and inf13) and the other two being Maya–Spanish bilingual speakers (inf16 and inf17; see Table 10.1).

Table 10.1. Participants in the study of accent.

BLP score	Speaker ID	Gender	State of origin
79	inf12	male	Yucatán
177	inf13	female	Quintana Roo
-25	inf16	male	Quintana Roo
-46	inf17	female	Quintana Roo

Lexically stressed syllables from the data corresponding to the studies on voiced stops and those on vowels were examined. Tokens for /b/ and /d/ were pooled together. Importantly, word-medial syllables were not examined for two reasons: (i) to facilitate the comparison between the data taken from the analysis of voiced stops and vowels (recall that, for vowels, there were no word-medial tokens), and (ii) because the differences observed in terms of the effect of lexical stress between corpus and read speech appeared in word-initial position.

The stressed syllables were coded as *deaccented* or *accented*. Following Rao (2007; see Section 10.1), stressed syllables were coded as deaccented if there was no observable F0 movement on the stressed syllable. Some word-initial syllables were glottalized to some degree, which made it difficult to assess whether there was a change in F0 that could be related to a pitch accent (especially if the following sound was unvoiced, e.g., *el **É**tico* 'the ethical one').

Within the accented category, three positions of accents were considered: prenuclear, intermediate, and nuclear. *Prenuclear* refers to pitch accents that appeared phrase medially; *intermediate* refers to pitch accents that appeared at the end of the intermediate phrase (ip), whereas *nuclear* refers to pitch accents at the end of the Intonational Phrase (IP). Nuclear pitch accents on the relevant segments were only found in corpus speech, not in read speech. This was the case because the read speech task was designed so that none of the sounds studied would appear in IP-final position (see Section 7.3). In corpus speech, the presence of a rather long pause after the phrase in some instances raised doubts about classifying a certain pitch as being in ip-final or in IP-final position. The decision was taken perceptually: if the sentence was perceived as unfinished, that is, if there was a perception of continuation, it was classified as *intermediate* (cf. Section 2.2).

In Yucatecan Spanish, the low pitch accent is rather frequent in *nuclear* position (i.e., L*, as well as H+L*) in broad focus statements (e.g., Martín Butragueño, 2017; Uth, 2018). To date, no study appears to have investigated the types of pitch accents that may appear in *prenuclear* position in Yucatecan Spanish. Due to the frequency of (H+)L* in nuclear configurations in Yucatecan Spanish and the fact that L* can also appear in prenuclear positions in other varieties (see de-la-Mota et al., 2010, for Mexico City Spanish), it is likely that this pitch accent appeared in prenuclear position, meaning that, for a number of instances, it was not evident whether a syllable was deaccented or whether there was a low pitch accent. Consequently, the results for deaccented syllables must be considered with care. Instances with postlexical secondary stress had already been excluded from the read speech data (see Sections 8.2.1 and 9.2.1.1). Finally, the stressed syllables in the corpus speech appeared in sentences with many pragmatic meanings, whereas in the read speech all sentences were declarative.

10.3 Results

The results show that deaccented syllables were clearly more frequent in corpus speech than in read speech (14.93 % versus 0.63 %, respectively, in the data from voiced stops; 26.07 % versus 1.96 %, respectively, in the data from vowels; see Table 10.2). This result, which is in line with previous studies on Spanish, is to be expected considering that reading may have favored a hyperarticulated pronunciation, which was perceptible in the read speech task. Conversely, the speaking style in a sociolinguistic interview is more natural and allows room for more variation. Deaccented tokens were common for verbal forms of the verbs *ser* 'to be' (e.g., *es* 'is') and *haber* 'to be, to have' (e.g., *hay* 'is'), as well as with determiners (e.g., *esas* 'those-F'), similarly to the results in the study by Face (2003) discussed in Section 2.2.1.1.

Table 10.2. Raw count number (and percentage) of deaccented and accented syllables in the corpus and read speech studies. In these syllables, a voiced stop or a vowel appeared in word-initial position. The position of the accents (prenuclear, intermediate, and nuclear) is also provided.

	Corpus speech		**Read speech**	
	Stops	**Vowels**	**Stops**	**Vowels**
Deaccented	10 (14.93 %)	67 (26.07 %)	1 (0.63 %)	2 (1.96 %)
Accented	67 (85.07 %)	190 (73.93 %)	158 (99.37 %)	102 (98.04 %)
of which: prenuclear p.	44	151	36	22
intermediate p.	16	25	122	80
nuclear p.	7	14	—	—

Figure 10.1 provides an example of deaccentuation in corpus speech in a parenthetical sentence. Also, in some instances, such as in the example displayed in the following figure (Figure 10.2), the F0 contour of long stretches of speech appeared to be flat, even if new information was presented, without noticeable tonal excursions (except for the intermediate or the nuclear accent). Consequently, stressed syllables in such cases were categorized as deaccented. Figure 10.3 shows an example of accentuation in the read speech corpus produced by the same speaker as in the example in Figure 10.1. A rising F0 movement is clearly visible for all stressed syllables.

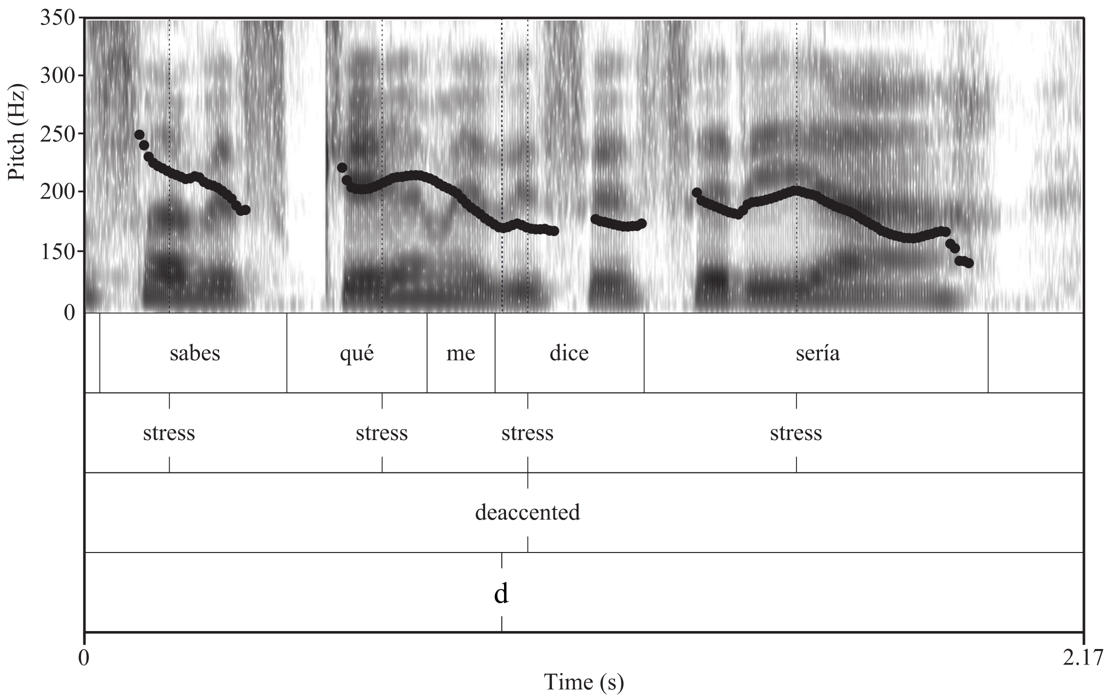

Figure 10.1. Example of a deaccented syllable (*di*) in corpus speech (voiced stops study, speaker inf17). The sentence is: «*¿Sabes qué?*» *—Me dice.* «*Sería…*» "'Do you know what?" He tells me. "It would be…"'.

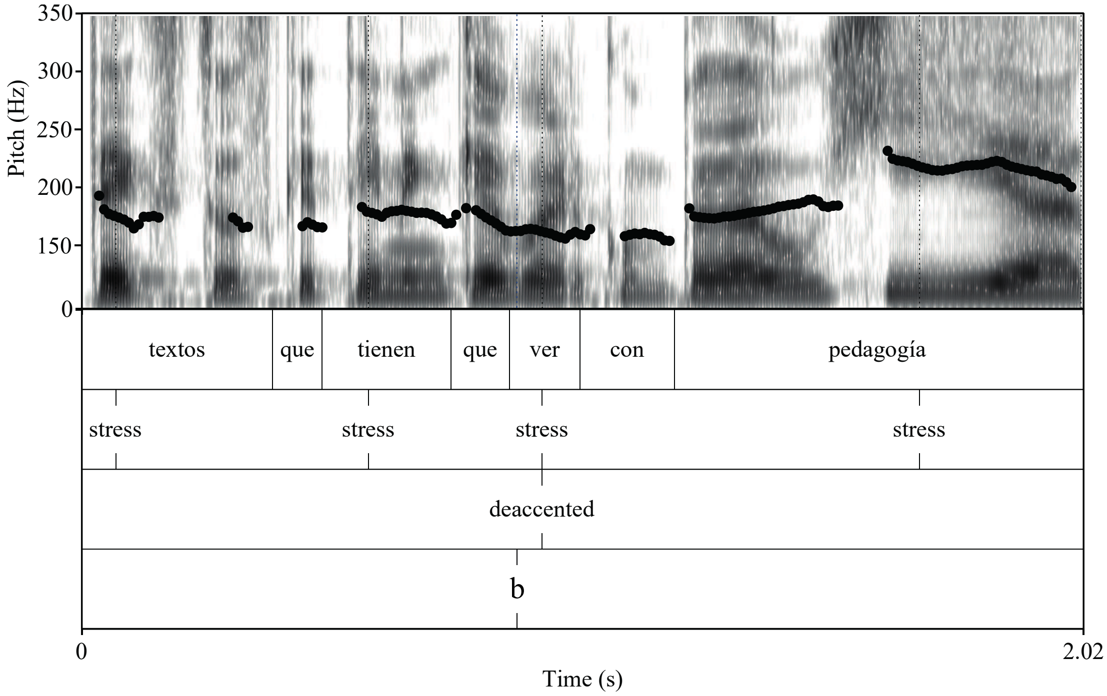

Figure 10.2. Example of flat F0 contour up to the word *pedagogía,* for which there is an accent in intermediate position (voiced stops study, speaker inf13). The sentence is: *Textos que tienen que ver con pedagogía* ‘Texts that have to do with teaching’.

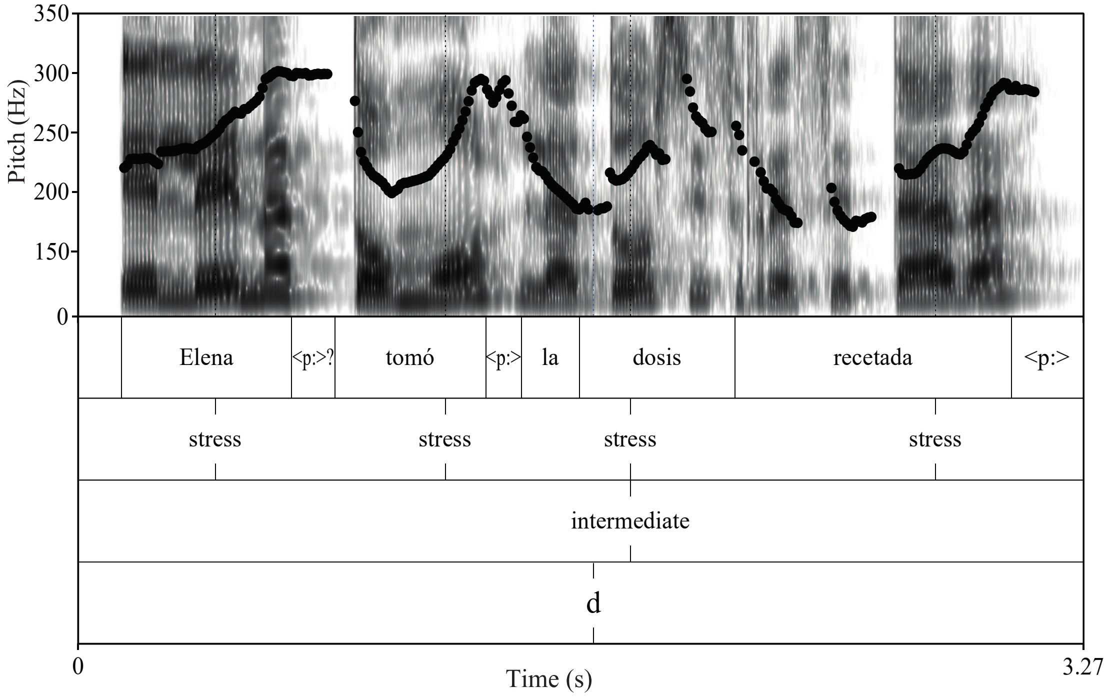

Figure 10.3. Example of accented syllables in read speech (voiced stops study, speaker inf17). The sentence is: *Elena tomó la dosis recetada* ‘Elena took the prescribed dose’.

The position of the accents present in corpus and read speech differed. In corpus speech, the most common position containing the target segments was the prenuclear, with some instances of accents in intermediate and nuclear positions (see Table 10.2). The position of the accents was not controlled for in the corpus speech for obvious reasons, nor are these data intended to suggest that a similar tendency may be observed in the data of the other participants whose speech is not examined here. The crucial point is that there was more variation in terms of the position of the accents in corpus speech than in read speech. In read speech, accents were found primarily in intermediate position. In other words, there was a tendency to phrase the read speech utterances into smaller sized constituents, that is, into ips, which were frequently followed by a pause (see Figure 10.3).

So far, we have seen some evidence that suggests that, in corpus speech, deaccentuation is more common, and also that there is more variation in the position of the accents that can be found. In order to investigate whether deaccentuation and this variation may play a role in terms of prosodic strengthening in corpus speech, I examine whether in accented syllables (i) voiced stops were longer, (ii) voiced stops had more stop realizations (versus approximant ones), and (iii) there were more instances of glottalization of vowels. A similar study was not conducted for read speech because of the low number of deaccented syllables (see Table 10.2).

Figure 10.4 shows the results for the duration of voiced stops. In these data, deaccented tokens were overall shorter than accented ones, which suggests that accent does have an effect on prosodic strengthening (cf. also the weakened realizations of the stops in Figures 10.1 and 10.2 to the strengthened realization that appears in Figure 10.3). In terms of the position of the accents, there seems to be a tendency towards longer voiced stops in nuclear position, followed by accents in intermediate position, and then by prenuclear accents. Nevertheless, given the small number of data points and the rather small differences, this tendency should be taken as a starting point for further study.

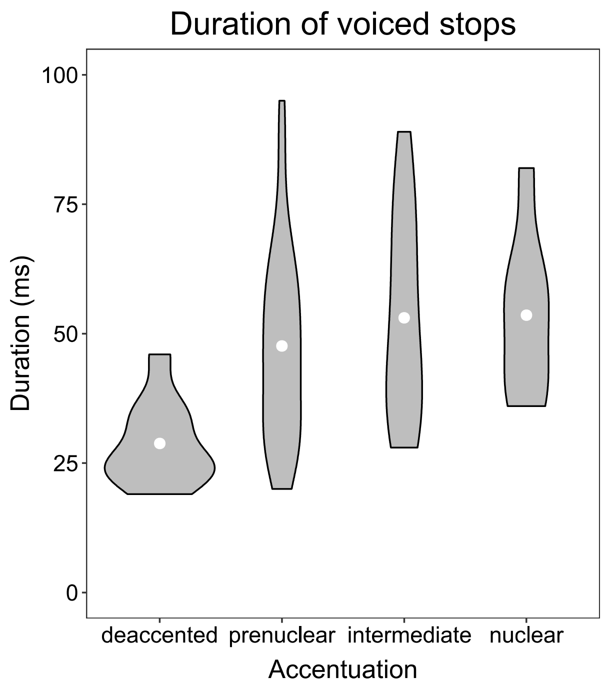

Figure 10.4. Violin plot (with mean points) of the duration of voiced stops in corpus speech according to accentuation: *deaccented* and accented (accents in *prenuclear*, *intermediate*, and *nuclear* position).

Similarly to the results for the duration of voiced stops, there were more stop realizations, as well as more instances with the presence of glottalization, in accented syllables than in deaccented syllables (Table 10.3). Therefore, these results strongly suggest that accent has an effect on prosodic strengthening in Yucatecan Spanish.

Table 10.3. Raw count number of *strengthened/not strengthened* realizations in corpus speech (stop/approximant for voiced stops, glottalized/not glottalized for vowels). The percentage of *strengthened* realizations within each cell appears in parentheses.

	Stops study	**Vowels study**
Deaccented	0/10 (0 %)	7/60 (10.45 %)
Accented	16/51 (23.88 %)	80/110 (42.11 %)

In summary, deaccenting is frequent in corpus speech and anecdotal in read speech. Due to the variation inherent in corpus speech and also to the rather controlled design and hyperarticulated production of the read speech task, pitch accents appear in several positions in the IP in corpus speech, whereas they tend to be concentrated at the end of the ip in read speech. Crucially, there is some evidence that suggests that accent may have an effect on prosodic strengthening for both voiced stops and vowels in Yucatecan Spanish.

10.4 Discussion

This preliminary study has provided some evidence that suggests that postlexical prominence by means of pitch accents favors strengthened realizations of voiced stops and vowels in Yucatecan Spanish. Moreover, the effects of accent and stress seem to be cumulative on the acoustic correlate examined (duration), so that tokens in accented syllables are more strengthened than in deaccented stressed syllables, which in turn are presumably more strengthened than those in unstressed syllables. In the light of these preliminary results, the effect of lexical stress that was presented in previous chapters needs to be reassessed.

Because it seems that virtually all of the tokens analyzed in the read speech data appeared in accented syllables, the prosodic strengthening that may be due to accent cannot be separated from that corresponding to lexical stress. The comparison of this data to the corpus speech data, where there is frequent deaccentuation, suggests that the effect found for lexical stress in read speech may indeed be due to accent. Moreover, in the corpus speech studies, it is likely that the effect of accent was obscured by analyzing tokens in deaccented and accented syllables together (and perhaps also by pooling together pitch accents that appeared in different phrasal positions). Tokens in deaccented syllables would present less strengthening than tokens in accented ones, meaning that the overall effect on strengthening of the variable lexical stress would be diminished. Consequently, although there is a tendency in the data towards more strengthening in stressed syllables (in word-initial position) in corpus speech, it does not provide the clear picture that appears in read speech. Whether this is indeed the case would need to be corroborated by further analyses of the data in which accent and lexical stress are clearly separated.

The present study was conducted with tokens in word-initial position. In word-medial position, there was evidence for a greater degree of prosodic strengthening in stressed than in unstressed syllables in corpus and read speech. Similarly to the tokens analyzed in word-initial position, tokens in stressed syllables in word-medial position could presumably be either accented or deaccented, and yet, there was an effect of the variable lexical stress. On the other hand, prosodic position (word-initial vs. word-medial) had an effect on prosodic strengthening, as has already been pointed out in the corresponding chapters and in Section 10.1, an effect that appears to be independent of that of (post)lexical prominence.

In sum, there is a high probability that the results associated with lexical stress in the current dissertation may be at least partially due to accentuation. Further studies may shed more light on the effect of accent on prosodic strengthening in Yucatecan Spanish.

Chapter 11

Conclusion

11.1 Overview

The main goal of this dissertation was to examine whether the phonetic realization of phonologically voiced stops and word-initial vowels in Yucatecan Spanish could be explained in terms of prosodic strengthening in relation to the Prosodic Word (PW) domain and to prosodic prominence. Thus, acoustic parameters corresponding to strengthening were investigated at the left edge of the PW and under lexical stress.

In particular, previous accounts of Yucatecan Spanish have argued that strong realizations of voiced stops and vowels are found at the beginning of the (orthographic) word and/or under lexical stress. This involves stop realizations of phonologically voiced stops (Michnowicz, 2011) and glottalized vowels (Martín Butragueño, 2014). This dissertation investigated the realization of voiced stops and vowels with the aim of ascertaining the extent to which these realizations could be situated within a crosslinguistic framework in which the position in the prosodic hierarchy plays a role in the realization of consonants and vowels (Dilley et al., 1996; Keating et al., 2004). Moreover, it provides insights into prosodic strengthening in Spanish, which has been investigated in only a few studies to date (Lavoie, 2001; Lahoz Bengoechea, 2015).

The relatively large number of participants, ranging from 16 to 21, made it possible to investigate the effect on prosodic strengthening of two speaker-specific factors, bilingualism (in terms of language dominance) and gender, both of which are rarely addressed in studies on prosodic strengthening because of the small numbers of participants, although previous studies have acknowledged inter- and intraspeaker variation (e.g., Keating et al., 2004; see Section 2.4.2).

This dissertation provides the first in-depth acoustic analysis of voiced stops and glottalization in Yucatecan Spanish. Previous analyses of both phenomena were based on auditory impressions with little acoustic detail (e.g., Lope Blanch, 1987; Michnowicz, 2009; Michnowicz & Kagan, 2016; Pérez Aguilar, 2002). Thus, a major contribution of this dissertation to the study of Yucatecan Spanish is to provide an acoustic description of voiced stops and glottalization of vowels in specific positions in the prosodic hierarchy.

This dissertation addresses the following questions:

1 Is there phonetic evidence for prosodic strengthening in Yucatecan Spanish?
   1.a For voiced stops?
   1.b For vowels?
2 If so, where does prosodic strengthening occur?
   2.a Initially in the Prosodic Word?
   2.b In stressed syllables?
3 What are the sources of speaker-specific variation in the extent and distribution of prosodic strengthening?
   3.a Language dominance?
   3.b Gender?

The following section presents a summary of results and a general discussion. The first three subsections discuss the research questions: whether there is phonetic evidence for prosodic strengthening (Section 11.2.1), where prosodic strengthening occurs (Section 11.2.2), as well as the sources of speaker-specific variation in prosodic strengthening (Section 11.2.3). The fourth subsection presents a proposal of how prosodic strengthening may help signal prosodic structure in the PW domain (Section 11.2.4). This section is then followed by the main conclusions of this dissertation (Section 11.3).

11.2 Summary of results and general discussion

A series of studies on voiced stops (Chapters 5 and 8) and glottalization of word-initial vowels (Chapters 6 and 9) addressed the above-mentioned questions in two speech styles, spontaneous speech (Chapters 5 and 6) and read speech (Chapters 8 and 9).

The studies on the bilabial and dentialveolar voiced stops examined several acoustic parameters that could indicate strengthening. The duration and presence of a release burst (which was used to draw a categorical distinction between stop and approximant sounds) was examined in both studies. Furthermore, the study conducted with read speech data took into account changes in intensity in the consonant. The effects that several factors could have on these parameters were examined by means of a series of Bayesian models. The factors were position in the PW domain, lexical stress, their interaction, word class (only for the dentialveolar stop), and repetition (only for the data in the read speech task), as well as language dominance (by means of the Bilingual Language Profile score) and gender.

The two studies on glottalization of word-initial vowels examined complementary aspects of the phenomenon. The study based on spontaneous speech provided an analysis of lack of resyllabification and glottalization because it has been argued that glottalization in Yucatecan Spanish blocks resyllabification (Michnowicz & Kagan, 2016). Thus, the presence/absence of resyllabification and the presence/absence of glottalization were annotated and then submitted to Bayesian analyses that included several factors: position in the PW domain, lexical stress, their interaction, language dominance, and gender. Furthermore, two random forest analyses were performed on the same data sets with many linguistic and nonlinguistic factors in order to explore new patterns in the data. Meanwhile, the study based on read speech provided a three-way analysis of glottalization, with a focus on its acoustic manifestation. First, it examined the same variables that were included in the Bayesian models in the corpus study, with the addition of repetition and word class. Second, it presented a classification of the types of glottalization that were found in the data. Third, three acoustic measures of glottalization were analyzed, namely two of spectral tilt (H1*–H2* and H1*–A3*) and one of periodicity/noise (cepstral peak prominence, or CPP).

11.2.1 Phonetic evidence for prosodic strengthening

Phonetic evidence was provided for the strengthening of voiced stops and vowels in word-initial position in Yucatecan Spanish. Voiced stops in Yucatecan Spanish presented a wide range of variation in their realizations, from vowel-like (weakened) realizations to full stops (strengthened). Strengthening was manifested by (i) longer duration, (ii) greater change in intensity (in this dissertation, measured from the start of the closing gesture corresponding to the consonant to its lowest point in intensity), and (iii) presence of a release burst, which indicates the greatest degree of strengthening (see Section 4.2.2.2). Parrell and Narayanan (2018) presented an articulatory undershoot account for phonologically voiced stops in Spanish (see Section 4.2.2.1), in line with previous works on prosodic strengthening: the spatial magnitude of the articulatory gesture is conditioned by its duration (Cho & Keating, 2001; Keating et al., 2004; Onaka, 2003; see Section 2.4.2). A longer duration of Spanish voiced stops is related to the consonantal gesture approximating its articulatory target, which may result in full closure of the articulators. In acoustic terms, this means that the strengthened realizations should have the three characteristics mentioned above. Moreover, it also means that duration and change in intensity should be positively correlated. Although this was not directly tested in this dissertation, the directionality of the results for duration, change in intensity, and even presence/absence of release burst in both spontaneous speech and read speech support this claim.[1] In sum, the Yucatecan Spanish data has provided acoustic evidence for this articulatory undershoot account, which agrees with the results of other articulatory and acoustic studies of Spanish (see Section 4.2.2).

Phonetic evidence for the prosodic strengthening of vowels in word-initial position came from the fact that a sizeable ratio of vowels in this position presented glottalization in their initial portion. This glottalization was both perceptually recognizable and, importantly, confirmed through a visual analysis of the data (by means of waveform and spectrographic information), an acoustic measure of spectral tilt that relates to the closing velocity of the vocal folds (H1*–A3*), and a measure of periodicity/noise (CPP), although there was no strong evidence for differences between glottalized vowels and modal ones in terms of H1*–H2*, which is the measure of spectral tilt related to the opening/closing of the glottis. One of the functions of prosodic strengthening at the syntagmatic level is to mark the disjuncture between two constituents (Cho, 2016; Cole et al., 2007; Georgeton & Fougeron, 2014; see Section 2.4.2). Glottalization in Yucatecan Spanish appears to have this function, thus supporting a prosodic strengthening interpretation of the phenomenon.

11.2.2 Occurrence of prosodic strengthening

In Yucatecan Spanish, prosodic strengthening occurred in word-initial position (meaning PW-initial and lexical-word, or LW-initial; see below) and under lexical stress. This is in line with some accounts of Yucatecan Spanish (e.g., Michnowicz, 2011) and also of other varieties of Spanish, which report more strengthened realizations in word-initial position and under lexical stress than in other contexts (see Section 4.2.2.3). The results in this dissertation also show that strengthened realizations in Yucatecan Spanish are much more frequent than what has been reported for other varieties of Spanish, especially with regards to glottalization, which seems to appear only occasionally in other varieties (see Section 4.3.4).

In Section 10.1, the interplay between prosodic position and lexical stress was assessed for voiced stops and vowels in relation to the two speech styles used in the studies, spontaneous and read speech. The conclusion reached was that, while prosodic position had a greater effect than lexical stress on prosodic strengthening in the spontaneous studies, the opposite was the case in read speech. Consequently, the possible role of accent as a variable

[1] Figures 5.2 (bilabial stop) and 5.7 (dentialveolar stop) illustrate how tokens that had been classified as *stop* in the spontaneous speech data have longer duration than those classified as *approximants.* Figures 8.1 (bilabial stop) and 8.8 (dentialveolar stop) illustrate the correlation between duration and change in intensity for tokens that had been classified as *approximant* or as *stop* in the read speech task.

involved in the divergence between the two speech styles was examined. The results of this examination strongly suggested that accent was at least partially responsible for the evidence found in support of lexical stress leading to prosodic strengthening, since the effects of accent and lexical stress were not separated in the data. For this reason, the descriptions of the effect of lexical stress that are presented below must be considered with caution.

In the case of voiced stops, strengthening occurred equally in PW-initial position (e.g., (***de*** (*descanso*)$_{LW}$)$_{PW}$ 'of relaxation') and in LW-initial position (e.g., (*de*(***d****escanso*)$_{LW}$)$_{PW}$) overall. In comparison, voiced stops in *word-medial* position were weakened. In general, strengthening also occurred under lexical stress, but there were some mixed results. In spontaneous speech, the effect of lexical stress was mixed, although it was clear that tokens in unstressed syllables in word-medial position were the least strengthened. In read speech, there was a strong effect of lexical stress, with tokens in stressed syllables being more strengthened, especially in PW-initial position. Finally, the results for the effect of word class were mixed. While in both spontaneous and read speech there was evidence for more strengthening for LWs than for FWs, in the read speech study the influence in terms of strengthening between tokens in FWs and in LWs was not that large.

The strengthening of word-initial vowels by means of glottalization occurred more frequently across PWs than within PWs, and also overall more frequently in stressed than in unstressed syllables, for both speech styles. Moreover, in the read speech study, the pattern of strengthening observed for the levels of the interaction between prosodic position and lexical stress (PW-initial and LW-initial × stressed > PW-initial × unstressed > LW-initial × unstressed > across FWs) indicates that, of the two variables, lexical stress had the greater effect on prosodic strengthening. The effect of word class, which was only examined in the read speech study, showed that there was a similar effect for both LWs and FWs in PW-initial position, although glottalization was slightly more frequent for LWs.

The effect of repetition, which was examined only in read speech for both voiced stops and vowels, showed that first mentions of voiced stops were more strengthened than second mentions, while a similar tendency was observed in terms of glottalization of the vowels, although the evidence in support of this interpretation was weak.

In sum, there is evidence of an effect of prosodic position and lexical stress on strengthening in Yucatecan Spanish. Thus, strengthened realizations of voiced stops and glottalization of vowels are manifestations of domain-initial strengthening that serve to mark the left edge(s)[2] of the PW and, more importantly, they are manifestations of the effect of lexical stress and probably also of accent.

11.2.3 Sources of speaker-specific variation

The results of the studies provided only weak evidence for an effect of language dominance on prosodic strengthening. The overall results for voiced stops (both in spontaneous and read speech) suggested that there was some weak evidence of an effect of language dominance, usually in the direction of Yucatec Maya dominance favoring more strengthened realizations in terms of duration, change in intensity, and presence of a release burst. The two studies on glottalization presented contradictory results: whereas in the study of spontaneous speech there was an effect of language dominance (once again meaning that Yucatec Maya favored glottalization), this effect was not found in the read speech task. All in all, the weak effect found for language dominance may be due to two methodological aspects of the studies of this dissertation. First, the number of participants may not have been large enough to provide a clear answer on the effect (or lack thereof) of language dominance. Thus, although the number of participants (16–21) is quite large, bilingualism is arguably a complex phenomenon, which is clearly manifested in this dissertation by the variety of BLP scores of the participants. Second, BLP scores were used as a continuous measure instead of being used to group participants into monolinguals and bilinguals, which is the approach taken in previous accounts of Yucatecan Spanish (e.g., Michnowicz, 2009; Michnowicz & Kagan, 2016; Rosado Robledo, 2011). While a "continuous" approach is clearly better suited to reflecting such a complex phenomenon, it may be the reason why no robust effect of language dominance was found. In sum, further studies that include a larger number of participants may provide a clearer picture of the effect of Yucatec Maya–Yucatecan Spanish bilingualism on prosodic strengthening.[3]

The results for gender provided very weak to no support for an effect, neither for voiced stops nor for vowels. For voiced stops, whenever the statistical models yielded an effect of gender, it was usually (but not always) in support of female speakers producing more strengthened realizations, while for the studies on glottalization, it was in support of male speakers. Overall, these results provide evidence in line with previous accounts of Yucatecan Spanish, which have claimed that there is no effect of gender on strengthened

[2] See Section 11.2.4.

[3] Moreover, in Chapters 5 and 6, it was shown that the Speech Learning Model (Flege, 1995, 1999, 2002; see Section 3.1.1) cannot account for either the strengthening of voiced stops or of vowels in Yucatecan Spanish.

realizations of voiced stops (García Fajardo, 1984; Martínez García, 2017; Rosado Robledo, 2011) or word-initial vowels (Michnowicz & Kagan, 2016), as well as evidence against studies that did claim there was such an effect (e.g., Pérez Aguilar, 2002, for voiced stops).

11.2.4 Strengthening in the Prosodic Word domain

This section discusses what the evidence of the effect of prosodic position obtained in this dissertation can contribute to the characterization of the PW domain in Yucatecan Spanish and in Spanish in general.

In the studies on voiced stops and the read speech study on vowels, a distinction was made between tokens in *PW-initial* position and in *LW-initial* position. This distinction did not reflect any theoretical claims about the inner structure of the PW in (Yucatecan) Spanish, because it is unclear what that structure is. While it is generally accepted that the Spanish PW can only have one primary stress (e.g., Elordieta, 2014; Hualde, 2009), the prosodization of unstressed FWs is unclear (see Section 2.2.1.2). Figure 11.1, which is an adaptation of Figure 2.2, illustrates this point.[4] PW-initial position corresponds to the beginning of the sequence *el árbol* 'the tree', whereas LW-initial position corresponds to the beginning of *árbol* (i.e., (el (árbol)$_{LW}$)$_{PW}$).

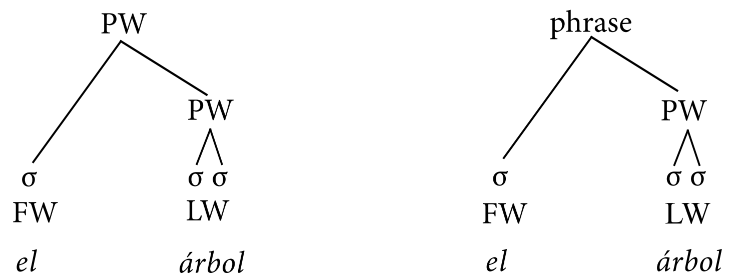

Figure 11.1. Two possible prosodic analyses of *el árbol* 'the tree'. 'FW' – *function word*; 'LW' – *lexical word.*

Overall, the results for the effect of prosodic position indicated that tokens in PW-initial position (regardless of whether the PW started with a FW or with a LW) and in LW-initial position could be grouped together. This was especially evident in the case of voiced stops. In word-medial position, tokens presented less strengthening. It follows, then, that strengthened realizations of voiced stops and glottalization of vowels are manifestations of domain-initial strengthening that serve to mark prosodic boundaries in the PW domain, as

[4] The two figures differ in that information about clitics is only provided in Figure 2.2, and in that *Phonological Phrase* has been substituted by *phrase*. Because of the unclear status of the Phonological Phrase as a constituent different from the intermediate phrase in Spanish (see Section 2.2) and because the focus of the dissertation is on the PW, I have opted to use *phrase.*

already indicated in Section 11.2.2. However, the results for prosodic strengthening can be interpreted as evidence for either of the possible structures of the PW represented in Figure 11.1. Thus, strengthening in PW-initial and LW-initial positions could mark the left edges of the *maximal* and *minimal PW*, respectively, in a recursive (or *extended*) PW (see Ito & Mester, 2009, for these denominations). In the present data, FWs that appeared in PW-initial position would be affixal clitics (see Section 2.2.1.2). Another possibility would be that strengthening helped mark the beginning of a phrase *and* a PW, which would be in agreement with the cumulative effect of domain-initial strengthening (Section 2.4.2). In this case, FWs would be free clitics. The second option can be rejected on the grounds that there is not enough evidence for a greater degree of strengthening in phrase-initial position (i.e., in PW-initial position) than in PW-initial position (i.e., LW-initial) in the present data. Consequently, a recursive PW account may fit the distribution of strengthening better (e.g., (**e**l (**á**rbol)$_{\text{PWmin}}$)$_{\text{PWmax}}$). Further investigation that specifically addresses phrasal-level prosody in Yucatecan Spanish is needed to determine the validity of this interpretation of prosodic strengthening.

An additional finding of the present dissertation relates to lack of resyllabification in Yucatecan Spanish. This dissertation provided evidence not only that lack of resyllabification is cued primarily by glottalization, but also that there are some restrictions to such lack of resyllabification. In particular, there were almost no cases of glottalization of vowels across (unstressed) FWs (e.g., (*en **el*** (*mercado*)$_{\text{PWmin}}$)$_{\text{PWmax}}$ '*in the market*'), whereas there were rather frequent at the left edges of the maximal PW (e.g., (**e**n (agua)$_{\text{PWmin}}$)$_{\text{PWmax}}$ 'in water') and minimal PW (e.g., (en (**a**gua)$_{\text{PWmin}}$)$_{\text{PWmax}}$). This suggests that the domain of resyllabification in Yucatecan Spanish cannot be the Intonational Phrase or the Utterance, as claimed by Nespor and Vogel (1986/2007) for Spanish (see Section 2.2.1.3), because of the frequent lack of resyllabification (cued by glottalization). Furthermore, lack of resyllabification appears to be related to a lower constituent than the IP, namely the PW. Lack of resyllabification in Yucatecan Spanish, which usually implies glottalization, is thus a PW-related phenomenon that can optionally be used to mark the left edges of the maximal and minimal PW. Thus, the domain of resyllabification in Yucatecan Spanish is the PW.

11.3 Conclusion

This dissertation has provided evidence for prosodic strengthening of both voiced stops and word-initial vowels in Yucatecan Spanish. The acoustic manifestations of prosodic strengthening of voiced stops are (i) longer duration, (ii) greater change in intensity, and, in extreme cases of strengthening, (iii) presence of a release burst. Strengthening of word-initial vowels is manifested through glottalization, which is present in the first portion of the vowel. Prosodic strengthening occurs in PW-initial position and especially under lexical stress, although accentuation may also play a role. Thus, prosodic strengthening is used to indicate (post)lexical prominence and boundaries at the PW level. In terms of speaker-specific variation, Yucatec Maya language dominance does not appear to favor more strengthened realizations either of voiced stops or word-initial vowels, while gender has no effect on the distribution of strengthened realizations. Finally, a proposal is made for the strengthening of voiced stops and glottalization of vowels being used to mark the left edges of a recursive PW in Yucatecan Spanish.

Appendices

Appendix A: Voiced Stops in the Corpus Study (Chapter 5)

Table A1. Raw data count of voiced stop tokens (bilabial and dentialveolar) per participant (*n = 20*).

BLP score	Speaker ID	Bilabial (*N* = 867)	Dentialv. (*N* = 1497)	Total (*N* = 2364)
-61	inf34	57	61	118
-46	inf17	44	100	144
-26	inf15	59	89	148
-25	inf22	26	82	108
-24	inf19	34	59	93
-24	inf18	27	69	96
-20	inf20	87	146	233
3	inf21	43	83	126
10	inf14	49	92	141
15	inf16	44	62	106
23	inf27	45	61	106
96	inf08	44	53	97
99	inf23	45	56	101
111	inf33	22	64	86
160	inf29	55	76	131
177	inf13	58	84	142
179	inf12	24	75	99
180	inf31	32	69	101
195	inf11	34	60	94
202	inf07	38	56	94

Table A2. Bilabial tokens excluded from the data set due to them being auditorily perceived as elided (in boldface). The words in which the tokens appeared are presented.

Speaker ID	Words	Glossing
inf29	*tra**b**aja**b**a,* *trabajá**b**amos*	work-PST.IPFV.3SG, work-PST.IPFV.1PL
inf31	*hu**b**o, sa**b**es*	be-PST.3SG, know-PRS.2SG

Table A3. Dentialveolar tokens excluded from the data set due to them being auditorily perceived as elided (in boldface). The words in which the tokens appeared are presented. There were no tokens excluded for speakers inf15 and inf18.

Speaker ID	Words	Glossing
inf07	*media**d**os*	mid-
inf08	*to**d**o*	all-M.SG
inf11	*encerra**d**os, to**d**a*	locked-up-M.PL, all-F.SG
inf12	*to**d**o* (× 2), *to**d**os* (× 2)	all-M.SG, all-M.PL
inf13	*to**d**os*	all-M.PL
inf14	*to**d**o*	all-M.SG
inf16	*merca**d**o*	market
inf17	*pregunta**d**o, to**d**o*	asked, all-M.SG
inf19	*pue**d**e* (× 2)	can-PRS.3SG
inf20	***d**e, **d**el, **d**onde,* *organiza**d**os, to**d**os*	of, of-the.M.SG, where, organized-M.PL, all-M.PL
inf21	*to**d**avía* (× 4)	still
inf22	***d**e, pue**d**es* (× 2)	of, can-PRS.2SG
inf23	*oportuni**d**ad, to**d**avía* (× 2)	opportunity, still
inf27	*pue**d**es, to**d**as*	can-PRS.2SG, all-F.SG
inf29	***d**e, **d**emás, (ha) esta**d**o, to**d**o* (× 2), *to**d**a, to**d**as* (× 4), *sali**d**a* (× 2), *to**d**avía, tra**d**icionales*	of, the rest, be-PTCP, all-M.SG, all-F.SG, all-F.PL, exit (noun), still, traditional.PL
inf31	*acompaña**d**o,* *acostumbra**d**os, to**d**o*	accompanied-M.SG, accustomed-to-M.PL, all-M.SG
inf33	***d**e, guarda**d**o, to**d**avía* (× 4), *to**d**o*	of, kept-M.SG, still, all-M.SG
inf34	*pue**d**es*	can-PRS.2SG

Table A4. Bilabial tokens included in the DURATION ($N = 553$) and CATEGORY ($N = 867$) models, per participant ($N = 20$). Tokens are grouped according to the PROSODIC POSITION in which they appeared: 'LW-med' – *LW-medial*; 'LW-ini' – *LW-initial*; 'PW-ini' – *PW-initial*.

		Duration			**Category**		
BLP score	**Speaker ID**	**LW-med**	**LW-ini**	**PW-ini**	**LW-med**	**LW-ini**	**PW-ini**
-61	inf34	25	11	2	41	13	3
-46	inf17	13	7	3	30	9	5
-26	inf15	23	8	2	46	11	2
-25	inf22	8	6	4	13	7	6
-24	inf19	16	3	3	22	6	6
-24	inf18	6	6	4	15	7	5
-20	inf20	34	16	7	57	17	13
3	inf21	16	7	5	28	9	6
10	inf14	15	14	5	21	23	5
15	inf16	9	2	7	28	3	13
23	inf27	13	12	3	19	20	6
96	inf08	22	7	12	25	7	12
99	inf23	19	7	5	30	10	5
111	inf33	2	1	1	18	2	2
160	inf29	14	13	4	33	17	5
177	inf13	15	8	8	36	13	9
179	inf12	5	10	4	6	13	5
180	inf31	11	7	3	17	11	4
195	inf11	16	9	3	21	9	4
202	inf07	21	7	4	24	9	5

Table A5. Pairwise comparisons for PROSODIC POSITION × LEXICAL STRESS: DURATION model for the bilabial stop. Estimates and 95 % Credible Intervals (CI) are presented. 'LW-med' – *LW-medial* position; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position; 'uns' – *unstressed*; 'str' – *stressed*. Pairwise comparisons for which 0 is not included in the CI are marked in bold.

Contrast	Estimate	Lower CI	Upper CI
LW-med, uns – LW-ini, uns	**-0.17**	**-0.32**	**-0.03**
LW-med, uns – PW-ini, uns	**-0.35**	**-0.52**	**-0.17**
LW-med, uns – LW-med, str	**-0.20**	**-0.30**	**-0.09**
LW-med, uns – LW-ini, str	**-0.20**	**-0.31**	**-0.10**
LW-med, uns – PW-ini, str	**-0.40**	**-0.53**	**-0.27**
LW-ini, uns – PW-ini, uns	-0.17	-0.38	0.04
LW-ini, uns - LW-med, str	-0.02	-0.17	0.13
LW-ini, uns – LW-ini, str	-0.03	-0.18	0.13
LW-ini, uns – PW-ini, str	**-0.23**	**-0.40**	**-0.06**
PW-ini, uns – LW-med, str	0.15	-0.03	0.33
PW-ini, uns – LW-ini, str	0.14	-0.03	0.33
PW-ini, uns – PW-ini, str	-0.05	-0.25	0.14
LW-med, str – LW-ini, str	0.00	-0.12	0.12
LW-med, str – PW-ini, str	**-0.20**	**-0.34**	**-0.06**
LW-ini, str – PW-ini, str	**-0.20**	**-0.34**	**-0.06**

Table A6. Pairwise comparisons for PROSODIC POSITION × LEXICAL STRESS: CATEGORY model for the bilabial stop. Estimates and 95 % Credible Intervals (CI) are presented. 'LW-med' – *LW-medial* position; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position; 'uns' – *unstressed*; 'str' – *stressed*. Pairwise comparisons for which 0 is not included in the CI are marked in bold.

Contrast	Estimate	Lower CI	Upper CI
LW-med, uns – LW-ini, uns	**-1.11**	**-1.82**	**-0.40**
LW-med, uns – PW-ini, uns	**-1.92**	**-2.69**	**-1.13**
LW-med, uns – LW-med, str	**-1.03**	**-1.56**	**-0.53**
LW-med, uns – LW-ini, str	**-1.26**	**-1.79**	**-0.72**
LW-med, uns – PW-ini, str	**-1.37**	**-2.01**	**-0.74**
LW-ini, uns – PW-ini, uns	-0.80	-1.74	0.12
LW-ini, uns – LW-med, str	0.09	-0.65	0.82
LW-ini, uns – LW-ini, str	-0.14	-0.87	0.61
LW-ini, uns – PW-ini, str	-0.25	-1.06	0.56
PW-ini, uns – LW-med, str	**0.88**	**0.09**	**1.68**
PW-ini, uns – LW-ini, str	0.66	-0.16	1.44
PW-ini, uns – PW-ini, str	0.55	-0.31	1.43
LW-med, str – LW-ini, str	-0.23	-0.77	0.32
LW-med, str – PW-ini, str	-0.34	-0.98	0.33
LW-ini, str – PW-ini, str	-0.11	-0.74	0.57

Table A7. Dentialveolar tokens included in the DURATION ($N = 657$) and CATEGORY ($N = 867$) models, per participant ($N = 1497$). Tokens are grouped according to the PROSODIC POSITION in which they appeared: 'LW-med' – *LW-medial*; 'LW-ini' – *LW-initial*; 'PW-ini' – *PW-initial*. Numbers in parentheses refer to tokens that appeared in *function words*.

		Duration			Category		
BLP score	**Speaker ID**	**LW-med**	**LW-ini**	**PW-ini**	**LW-med**	**LW-ini**	**PW-ini**
-61	inf34	12	6	17 (16)	23	11	27 (25)
-46	inf17	12	9	14 (11)	46	17	37 (27)
-26	inf15	21	8	14 (9)	43	16	30 (23)
-25	inf22	27	7	12 (11)	51	8	23 (22)
-24	inf19	14	4	5 (4)	41	5	13 (11)
-24	inf18	10	11	16 (12)	21	17	31 (26)
-20	inf20	27	13	42 (33)	74	14	58 (48)
3	inf21	11	3	17 (16)	40	8	35 (33)
10	inf14	16	2	10 (5)	61	10	21 (13)
15	inf16	9	12	7 (0)	31	14	17 (3)
23	inf27	10	6	13 (10)	29	11	21 (16)
96	inf08	14	7	16 (14)	23	9	21 (18)
99	inf23	12	4	8 (8)	27	6	23 (22)
111	inf33	4	1	2 (2)	43	6	15 (14)
160	inf29	7	6	16 (15)	36	10	30 (26)
177	inf13	19	2	7 (6)	52	2	30 (28)
179	inf12	13	4	7 (4)	49	9	17 (13)
180	inf31	16	5	9 (6)	42	7	20 (15)
195	inf11	10	8	14 (11)	33	8	19 (16)
202	inf07	16	3	10 (8)	35	3	18 (14)

Table A8. Pairwise comparisons for PROSODIC POSITION × LEXICAL STRESS: DURATION model for the dentialveolar stop. Estimates and 95 % Credible Intervals (CI) are presented. 'LW-med' – *LW-medial* position; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position; 'uns' – *unstressed*; 'str' – *stressed*. Pairwise comparisons for which 0 is not included in the CI are marked in bold.

Contrast	Estimate	Lower CI	Upper CI
LW-med, uns – LW-ini, uns	**-0.31**	**-0.42**	**-0.19**
LW-med, uns – PW-ini, uns	**-0.15**	**-0.22**	**-0.07**
LW-med, uns – LW-med, str	**-0.15**	**-0.24**	**-0.05**
LW-med, uns – LW-ini, str	**-0.27**	**-0.38**	**-0.16**
LW-med, uns – PW-ini, str	**-0.22**	**-0.41**	**-0.03**
LW-ini, uns – PW-ini, uns	**0.16**	**0.05**	**0.27**
LW-ini, uns – LW-med, str	**0.16**	**0.03**	**0.28**
LW-ini, uns – LW-ini, str	0.03	-0.10	0.17
LW-ini, uns – PW-ini, str	0.09	-0.12	0.30
PW-ini, uns – LW-med, str	0.00	-0.10	0.09
PW-ini, uns – LW-ini, str	**-0.13**	**-0.23**	**-0.02**
PW-ini, uns – PW-ini, str	-0.08	-0.26	0.12
LW-med, str – LW-ini, str	**-0.13**	**-0.25**	**-0.01**
LW-med, str – PW-ini, str	-0.07	-0.27	0.12
LW-ini, str – PW-ini, str	0.05	-0.15	0.26

Table A9. Pairwise comparisons for PROSODIC POSITION × LEXICAL STRESS: CATEGORY model for the dentialveolar stop. Estimates and 95 % Credible Intervals (CI) are presented. ‘LW-med’ – *LW-medial* position; ‘LW-ini’ – *LW-initial* position; ‘PW-ini’ – *PW-initial* position; ‘uns’ – *unstressed*; ‘str’ – *stressed*. Pairwise comparisons for which 0 is not included in the CI are marked in bold.

Contrast	Estimate	Lower CI	Upper CI
LW-med, uns – LW-ini, uns	**-2.01**	**-2.63**	**-1.43**
LW-med, uns – PW-ini, uns	**-1.47**	**-1.91**	**-1.05**
LW-med, uns – LW-med, str	**-1.60**	**-2.11**	**-1.08**
LW-med, uns – LW-ini, str	**-2.31**	**-2.93**	**-1.73**
LW-med, uns – PW-ini, str	**-2.13**	**-3.05**	**-1.14**
LW-ini, uns – PW-ini, uns	0.55	0.00	1.08
LW-ini, uns – LW-med, str	0.42	-0.19	1.03
LW-ini, uns – LW-ini, str	-0.30	-0.97	0.35
LW-ini, uns – PW-ini, str	-0.12	-1.11	0.91
PW-ini, uns – LW-med, str	-0.13	-0.58	0.31
PW-ini, uns – LW-ini, str	**-0.84**	**-1.36**	**-0.30**
PW-ini, uns – PW-ini, str	-0.66	-1.57	0.26
LW-med, str – LW-ini, str	**-0.71**	**-1.31**	**-0.10**
LW-med, str – PW-ini, str	-0.53	-1.50	0.41
LW-ini, str – PW-ini, str	0.18	-0.83	1.16

Appendix B: Lack of Resyllabification and Glottalization in the Corpus Speech Study (Chapter 6)

Table B1. Number of tokens with potential for resyllabification (*n = 150* per participant). Numbers in parentheses refer to *glottalized* tokens.

		Resyllabification		Lack of resyllabification	
BLP score	**Speaker ID**	**Within PW**	**Across PWs**	**Within PW**	**Across PWs**
-46	inf17	40 (0)	49 (4)	18 (10)	43 (32)
-26	inf15	70 (1)	60 (2)	5 (4)	15 (15)
-24	inf18	57 (5)	75 (2)	2 (1)	16 (12)
-20	inf20	25 (7)	45 (11)	18 (17)	62 (61)
3	inf21	38 (0)	104 (0)	2 (2)	6 (3)
10	inf14	41 (1)	73 (0)	5 (4)	31 (30)
15	inf16	28 (0)	64 (6)	14 (13)	44 (41)
35	inf25	50 (0)	88 (0)	2 (2)	10 (7)
96	inf08	47 (1)	69 (1)	12 (9)	22 (16)
99	inf23	67 (3)	79 (3)	0 (0)	4 (4)
142	inf24	43 (1)	99 (3)	1 (0)	7 (6)
160	inf29	49 (0)	92 (0)	3 (3)	6 (5)
177	inf13	51 (0)	87 (3)	1 (1)	11 (9)
179	inf12	34 (1)	96 (1)	3 (2)	17 (16)
195	inf11	58 (0)	74 (1)	3 (3)	15 (12)
202	inf07	51 (1)	91 (1)	0 (0)	8 (6)

Table B2. Pairwise comparisons for PROSODIC POSITION × LEXICAL STRESS: LACK OF RESYLLABIFICATION model, with estimates and 95 % Credible Intervals (CI). 'uns' – *unstressed*; 'str' – *stressed*. Pairwise comparisons for which 0 is not included in the CI are marked in bold.

Contrast	Estimate	Lower CI	Upper CI
within PW, uns – across PWs, uns	**-1.35**	**-1.78**	**-0.93**
within PW, uns – within PW, str	**-1.12**	**-1.61**	**-0.64**
within PW, uns – across PWs, str	**-1.58**	**-2.05**	**-1.14**
across PWs, uns – within PWs, str	0.23	-0.11	0.57
across PWs, uns – across PWs, str	-0.23	-0.52	0.05
within PW, str – across PWs, str	**-0.46**	**-0.83**	**-0.09**

Table B3. Pairwise comparisons for PROSODIC POSITION × LEXICAL STRESS: GLOTTALIZATION model, with estimates and 95 % Credible Intervals (CI). 'uns' – *unstressed*; 'str' – *stressed*. Pairwise comparisons for which 0 is not included in the CI are marked in bold.

Contrast	Estimate	Lower CI	Upper CI
within PW, uns – across PWs, uns	**-1.27**	**-1.68**	**-0.87**
within PW, uns – within PW, str	**-0.94**	**-1.41**	**-0.46**
within PW, uns – across PWs, str	**-1.22**	**-1.69**	**-0.80**
across PWs, uns – within PW, str	0.33	-0.01	0.69
across PWs, uns – across PWs, str	0.05	-0.24	0.35
within PWs, str – across PWs, str	-0.28	-0.68	0.10

Appendix C: Design of the Read Speech Task (Chapter 7)

Table C1. Sentences for the read speech task, familiarization phase, with English translation and an example. 'Medial' and 'final' refer to the position of the carrier words in the sentence.

Proper names

Laura	Rosario	la Gafas 'the Glasses'	Alberto	el Afro 'the Afro'	Paulino

Sentences

No.	Verb	Medial	Final	Translation
1	preparó	la comida	a las dos	prepared lunch at 2 o'clock
2	esculpió	una escultura	de madera	sculpted a wooden sculpture
3	disfrutó	del verano	intensamente	enjoyed summer immensely
4	viajó	por todo el mundo	alegremente	traveled happily around the world
5	pensó	cuidadosamente	sobre su vida	reflected carefully on her/his life
6	comió	el escabeche	velozmente	ate the stew quickly

Example

- First reading: *Laura preparó la comida a las dos porque*… 'Laura prepared lunch at 2 o'clock because…'
- Second reading: *Laura preparó la comida a las dos porque había desayunado muy tarde* 'Laura prepared lunch at 2 o'clock because she had had breakfast very late.'

Table C2. Sentences for the read speech task, set 1, with English translations and an example. The target tokens (bilabial and dentialveolar stops, and word-initial vowels) are marked in bold. 'Medial' and 'final' refer to the position of the carrier words in the sentence.

Proper names

A**d**elino	**B**erto	la **B**onita 'the Cute One'	Dé**b**ora	**E**lena	el **E**nano 'the Tiny One'

Sentences

No.	Verb	Medial	Final	Translation
1	jugó	en **e**l parque	con unos niños	played in the park with some kids
2	ensayó	para **e**l concierto	de guitarra	rehearsed for the guitar concert
3	trabajó	desde **e**l domingo	en la ca**b**ecera	had worked since Sunday at the municipality
4	tiró	la **b**otana	a la **b**asura	threw the snack into the trash
5	vendió	**b**oletos	**a** los turistas	sold tickets to the tourists
6	sintió	**d**olores	en **o**toño	felt pain in autumn
7	evitó	**b**ebidas	de **a**zúcar	avoided sugary drinks
8	estudió	jardinería	**b**ásica	studied basic gardening
9	le habló	al **á**rbitro	**d**ócil	spoke gently to the referee
10	salió	**d**e la plaza	con unos amigos	left the square with some friends
11	tocó	el **ó**rgano	**d**ébil	played the organ softly
12	tomó	la **d**osis	recetada	took the prescribed dose
13	se llenó	la **b**oca	**d**e **o**régano	filled her/his mouth with oregano
14	le cambió	**d**ólares	**a**l banquero	exchanged dollars with the banker
15	llevó	la **b**ecerra	a la **b**o**d**ega	brought the calf to the cellar
16	se subió	al gra**d**erío	**a**bierto	got onto the open stands
17	taló	**á**rboles	en la selva	cut down trees in the rain forest
18	usó	el **a**banico	de a**d**orno	used the fan as decoration

Table C3. Sentences for the read speech task, set 2, with English translations and an example. The target tokens (bilabial and dentialveolar stops, and word-initial vowels) are marked in bold. 'Medial' and 'final' refer to the position of the carrier words in the sentence.

Proper names

Adela	**B**ernardo	**D**elfín	la **D**óberman 'the Doberman'	el **É**tico 'the Ethical One'	**e**l Guapo 'the Handsome One'

Sentences

No.	Verb	Medial	Final	Translation
19	llamó	**a** su madre	**á**spera	called her/his mother harsh
20	se tocó	la **o**reja	**d**eprisa	quickly touched her/his ear
21	compró	un producto	para **e**l cabello	bought a hair product
22	ensayó	la **ó**pera	en **e**l pa**b**ellón	practiced the opera in the pavilion
23	estudió	**á**rabe	para pe**d**agogos	studied Arabic for teachers
24	salió	**d**el avión	rápidamente	got off the plane quickly
25	aprendió	**b**elga	en México	learned Belgian in Mexico
26	usó	la **b**áscula	con **e**l vestido	used the scale with the dress
27	alimentó	**ab**ejas	con **a**ceite	fed bees with oil
28	guardó	**d**ados	**en el** cajón	kept dice in the drawer
29	siguió	el **o**ficio	de su padre	followed her/his father's career
30	preparó	la **d**efensa	con el abogado	prepared the defense with the lawyer
31	escribió	la **d**enuncia	para el vecino	wrote the report for the neighbor
32	preparó	**d**esfiles	para Carnaval	prepared parades for Carnival
33	calentó	el bi**b**erón	para **e**l bebé	warmed up the feeding bottle for the baby
34	buscó	**a**ctores	ca**d**avéricos	looked for a ghostly looking actor
35	escribió	**a** la asociación	de **d**onantes	wrote to the donor organization
36	se dedicó	al diseño	**de b**óvedas	dedicated his/her life to designing vaults

Table C4. Sentences for the read speech task, set 3, with English translations and an example. The target tokens (bilabial and dentialveolar stops, and word-initial vowels) are marked in bold. 'Medial' and 'final' refer to the position of the carrier words in the sentence.

Proper names

la **B**ola	**D**enis	la **D**octora 'the Doctor'	**É**dison	**e**l Güero 'the Blonde One'	Re**b**eca

Sentences

No.	Verb	Medial	Final	Translation
37	echó	**b**ombas	con **o**dio	threw bombs with hate
38	cortó	ma**d**era	desde **e**l amanecer	had chopped wood since dawn
39	evitó	un problema	**d**oméstico	avoided a domestic problem
40	salió	**d**e casa	**b**árbaro/a	left the house savagely
41	fue	al po**d**ólogo	**b**arato	went to the cheap podiatrist
42	coleccionó	me**d**allones	de **é**bano	collected ebony medallions
43	salió	**d**e la iglesia	tranquilamente	left the church calmly
44	discutió	**e**n **e**l colegio	**b**ajito	discussed quietly at school
45	se apuntó	al programa	de **d**e**b**ate	enrolled in the debate program
46	buscó	una escuela	**d**e **é**lite	looked for an elite school
47	usó	una crema	de **d**edos	used a finger cream
48	comió	un mango	**á**cido	ate a sour mango
49	viajó	con **e**l sombrero	en la ca**b**eza	traveled with a hat on her/his head
50	dibujó	la **ó**rbita	**d**e la luna	drew the moon's orbit
51	practicó	para **e**l concurso	de **d**anza	practiced for the dancing competition
52	cortó	la **d**alia	**a**bajo	cut the dahlia at the bottom
53	cocinó	la **o**veja	de **b**odas	cooked the lamb for the wedding
54	metió	el ta**b**aco	en **a**gua	put the tobacco into water

Table C5. Images used for the characters in the read speech task. 'fam' = *familiarization phase.* Retrieved on August 6, 2019.

Name	Set	Retrieved from	Type of CC image
Laura	fam.	https://www.pexels.com/photo/actress-beautiful-blonde-blur-392748/	edit/use allowed, no credit necessary
Rosario	fam.	https://www.goodfreephotos.com/albums/people/female-face-woman-portrait.jpg	CC 0
La Gafas	fam.	https://www.flickr.com/photos/12978243@N07/16352326983/	CC BY-NC-SA 2.0
Alberto	fam.	https://ccsearch.creativecommons.org/photos/fd96e50b-2f15-40c6-93bc-822d45ec136	Public Domain Mark 1.0 (no copyright)
El Afro	fam.	https://www.flickr.com/photos/ccmanla/2682913717/	CC BY-NC 2.0
Paulino	fam.	https://www.flickr.com/photos/enthuan/8499107332/	CC BY-NC 2.0
La Bonita	1	https://pixabay.com/photos/woman-portrait-beautiful-ruda-girl-3247382/	edit/use allowed, no credit necessary
Elena	1	https://flic.kr/p/gkFs9L	CC BY-NC-SA 2.0
Débora	1	https://flic.kr/p/X92y2S	Public Domain Mark 1.0 (no copyright)
Adelino	1	https://www.flickr.com/photos/cgmely/47345905441/	CC BY-NC 2.0
El Enano	1	https://www.flickr.com/photos/geezaweezer/6546575885/	CC BY-NC 2.0
Berto	1	https://www.flickr.com/photos/ccmanla/2684123695/	CC BY-NC 2.0
Adela	2	https://pixabay.com/es/photos/ni%C3%B1a-inteligentes-mujeres-3240646/	edit/use allowed, no credit necessary
La Dóberman	2	https://flic.kr/p/29f2Qjo	CC BY-NC 2.0
El Ético	2	https://www.flickr.com/photos/krapow/2383813495/	CC BY-NC-SA 2.0
El Guapo	2	https://www.flickr.com/photos/giancarloaguilar/45497457282/	CC BY-NC-SA 2.0
Bernardo	2	https://www.flickr.com/photos/eole/15006753517/	CC BY-NC-SA 2.0
Delfín	2	https://www.flickr.com/photos/25228175@N08/2864976311/	CC BY-NC 2.0
La Doctora	3	https://www.flickr.com/photos/12533894@N06/14387367072	CC BY 2.0
La Bola	3	https://www.pexels.com/photo/portrait-photo-of-woman-718978/	edit/use allowed, no credit necessary
Rebeca	3	https://flic.kr/p/WZRmR8	Public Domain Mark 1.0 (no copyright)
Édison	3	https://www.flickr.com/photos/megan2300/442159511	CC BY 2.0
El Güero	3	https://www.flickr.com/photos/ccmanla/2682733115/	CC BY-NC 2.0
Denis	3	https://www.flickr.com/photos/gldar/23224579524/	CC BY-NC 2.0

Table C6. Excluded bilabial tokens (in boldface) in *phrase-initial* position. All instances ($n = 60$) were produced as *stop* realizations, except for one token in the word *barato* (speaker inf49), produced as an *approximant* realization.

Speaker ID	**Words (set 1)**	**Words (set 2)**	**Words (set 3)**
inf07	**b**oletos, **b**ásica	**b**elga	**b**árbaro (× 2), **b**arato
inf09	—	**b**elga	**b**árbaro, **b**arato
inf12	—	—	**b**árbaro
inf16	**b**otana	**b**elga	**b**árbaro, **b**arato
inf17	**b**oletos (× 2), be**b**idas	—	**b**árbaro
inf18	**b**oletos, be**b**idas	—	**b**ombas, **b**arato (× 2)
inf19	be**b**idas, **b**ásica	—	**b**ombas, **b**árbara (× 2), **b**arato
inf22	—	—	**b**árbaro (× 2)
inf27	—	—	**b**árbaro (× 2), **b**arato
inf28	—	—	**b**árbaro (× 2), **b**arato, **b**ajito (× 2)
inf30	—	—	**b**árbara
inf43	be**b**idas	—	**b**ombas, **b**árbaro, **b**arato
inf44	**b**oletos (× 2), be**b**idas (× 2)	**b**elga (× 2)	**b**ombas (× 2), **b**árbaro (× 2)
inf48	—	**b**elga	—
inf49	**b**oletos	—	**b**árbaro, **b**arato (× 2), **b**odas

Table C7. Excluded dentialveolar tokens (in boldface) in *phrase-initial* position. All instances (n = 47) were produced as *stop* realizations, except for one token in the word *doméstico* (speaker inf49), produced as an *approximant* realization.

Speaker ID	**Words (set 1)**	**Words (set 2)**	**Words (set 3)**
inf07	**d**ócil (× 2), **d**ébil, **d**e (in *de orégano*), **d**ólares	**d**eprisa, **d**ados	—
inf09	**d**ócil	**d**el	**d**e (in *de élite*)
inf13	**d**ócil	—	—
inf16	**d**ébil	**d**eprisa, **d**esfiles	—
inf17	**d**olores, **d**ólares	—	—
inf18	—	**d**ados (× 2)	**d**e (in *de casa*) (× 2)
inf19	**d**ócil (× 2)	**d**eprisa	**d**e (in *de la luna*)
inf22	**d**ébil, **d**e (in *de orégano*)	—	—
inf27	**d**ócil	—	—
inf28	**d**ócil, **d**ébil	**d**esfiles, **d**e (in *de bóvedas*)	**d**oméstico
inf30	—	**d**efensa	—
inf44	**d**olores, **d**ócil (× 2), **d**ólares (× 2)	**d**ados (× 2), **d**esfiles (× 2)	**D**octora
inf45	**d**e (in *de orégano*)	—	—
inf47	—	**d**el	—
inf48	—	**d**el	—
inf49	—	—	**d**oméstico

Table C8. Excluded vowel tokens (in boldface) in *phrase-initial* position (n = 127).

Speaker ID	Words (set 1)	Words (set 2)	Words (set 3)
inf07	**E**nano, **á**rboles (× 2), **a**l banquero	**á**rabe, **a** la asociación	**a**bajo
inf09	—	**a** la asociación, **á**spera	**é**lite, **a**bajo
inf12	**a**bierto	**á**spera (× 2), **a**ctores, **a** la asociación (× 2)	**a**bajo
inf13	—	**a** la asociación	**a**bajo
inf16	**a** los turistas, **o**régano, **á**rboles, **a**bierto	**e**n el cajón, **a**ctores	**a**bajo, **á**cido, **o**dio
inf17	—	**á**spera (× 2), árabe (× 2), **e**n el cajón, **a**ctores (× 2), **a** la asociación (× 2)	**a**bajo (× 2), **e**l amanecer
inf18	**á**rboles, **a**bierto (× 2)	**e**n el cajón, **a** la asociación	**e**n el colegio, **é**bano, **a**bajo (× 2)
inf19	**á**rboles, en **o**toño, **a**bierto (× 2)	**á**spera (× 2), **á**rabe (× 2), **a** la asociación, en **e**l pabellón	**e**n el colegio, **a**bajo
inf21	**a** los turistas	**á**spera (× 2), **a** la asociación	**á**cido
inf22	**o**régano, **a**l banquero	**á**spera, **a**ctores, **a** la asociación (× 2)	**e**n el colegio, **a**bajo
inf27	**a**l banquero	—	—
inf28	**a**l banquero, **a**bierto (× 2)	**á**spera, **á**rabe (× 2), **e**n el cajón, **a** la asociación	**a**bajo (× 2), **á**cido
inf30	—	**á**rabe, **a**bejas	en **e**l colegio, **a**bajo
inf43	**a** los turistas	**a** la asociación	—
inf44	—	**a**bejas, **a**ctores (× 2), **a** la asociación (× 2), **e**n el cajón (× 2)	**e**n el colegio (× 2)
inf45	**á**rboles (× 2)	**a** la asociación	—
inf46	**á**rboles	—	**e**n el colegio, **a**bajo
inf47	**á**rboles (× 2)	**a** su madre (× 2), **á**spera, **á**rabe (× 2)	—
inf49	**á**rbitro, **o**régano, **á**rboles	**á**rabe, **a**bejas	**a**bajo (× 2), **á**cido (× 2), **ó**rbita (× 2), **o**veja

Appendix D: Voiced Stops in the Read Speech Study (Chapter 8)

Table D1. Results of the Bayesian models for POSITION OF THE CARRIER WORD IN THE SENTENCE. The models included SPEAKER and WORD as random intercepts, with DURATION, CHANGE IN INTENSITY, and CATEGORY as the dependent variables. The priors were the same as those for the analysis of each dependent variable and voiced stop considered in the main text (Section 8.2.3). 0 is included within the Credible Intervals (CI), which indicates that the dependent variables are not affected by the POSITION OF THE CARRIER WORD IN THE SENTENCE.

Models	Bilabial stop	Dentialveolar stop
DURATION	$\hat{\beta}_{intercept}$ = 4.28, CI [4.08, 4.48] $\hat{\beta}_{medial}$ = 0.01, CI [-0.21, 0.23] $\hat{\beta}_{final}$ = -0.04, CI [-0.26, 0.18]	$\hat{\beta}_{intercept}$ = 4.33, CI [4.15, 4.52] $\hat{\beta}_{medial}$ = -0.10, CI [-0.29, 0.09] $\hat{\beta}_{final}$ = 0.02, CI [-0.17, 0.21]
INTENSITY	$\hat{\beta}_{intercept}$ = 13.01, CI [9.78, 16.19] $\hat{\beta}_{medial}$ = 0.31, CI [-3.15; 3.85] $\hat{\beta}_{final}$ = -0.76, CI [-4.23, 2.73)	$\hat{\beta}_{intercept}$ = 14.68, CI [12.21, 17.17] $\hat{\beta}_{medial}$ = -1.69, CI [-4.23, 0.83] $\hat{\beta}_{final}$ = 0.05, CI [-2.45, 2.58]
CATEGORY	$\hat{\beta}_{intercept}$ = -0.73, CI [-2.48, 1.04] $\hat{\beta}_{medial}$ = 0.12, CI [-1.71, 1.92] $\hat{\beta}_{final}$ = -0.11, CI [-1.93, 1.69]	$\hat{\beta}_{intercept}$ = 0.72, CI [-0.90, 2.31] $\hat{\beta}_{medial}$ = -0.72, CI [-2.35, 0.91] $\hat{\beta}_{final}$ = -0.21, CI [-1.83, 1.42]

Table D2. Number of bilabial tokens included in the DURATION/CHANGE IN INTENSITY ($N = 1335$) and CATEGORY ($N = 1417$) models per speaker ($N = 21$). Tokens are grouped according to the PROSODIC POSITION in which they appeared: 'LW-med' – *LW-medial*; 'LW-ini' – *LW-initial*; 'PW-ini' – *PW-initial*.

		Duration/Intensity			Category		
BLP score	**Speaker ID**	**LW-med**	**LW-ini**	**PW-ini**	**LW-med**	**LW-ini**	**PW-ini**
-46	inf17	27	28	10	27	28	11
-25	inf22	27	28	13	27	28	14
-24	inf19	24	24	8	27	26	9
-24	inf18	27	27	10	27	27	10
3	inf21	24	28	16	26	28	16
15	inf16	26	27	11	26	27	13
23	inf27	28	28	12	28	28	12
25	inf45	22	26	14	27	27	15
66	inf04	22	25	11	28	27	14
71	inf43	27	26	7	28	27	10
147	inf44	26	28	6	28	28	6
148	inf47	26	28	12	28	28	16
148	inf49	23	24	11	26	27	11
157	inf30	20	25	13	25	28	14
166	inf46	25	27	13	28	28	16
171	inf48	24	26	15	28	28	15
177	inf13	26	26	9	24	26	15
177	inf28	25	26	15	28	28	11
179	inf12	28	30	12	28	27	15
199	inf09	27	26	9	30	30	13
202	inf07	27	28	16	28	28	10

Table D3. Number of dentialveolar tokens included in the DURATION/CHANGE IN INTENSITY ($N = 1611$) and CATEGORY ($N = 1726$) models per speaker ($N = 21$). Tokens are grouped according to the PROSODIC POSITION in which they appeared: 'LW-med' – *LW-medial*; 'LW-ini' – *LW-initial*; 'PW-ini' – *PW-initial*. Numbers in parentheses refer to tokens that appeared in *function* words.

		Duration/Intensity			Category		
BLP score	**Speaker ID**	**LW-med**	**LW-ini**	**PW-ini**	**LW-med**	**LW-ini**	**PW-ini**
-46	inf17	27	25	29 (16)	27	25	30 (16)
-25	inf22	26	27	29 (14)	26	27	29 (14)
-24	inf19	21	21	22 (9)	23	25	27 (12)
-24	inf18	25	27	27 (15)	25	27	27 (15)
3	inf21	28	24	30 (15)	29	26	32 (16)
15	inf16	23	23	26 (14)	26	24	27 (15)
23	inf27	27	25	25 (11)	28	27	31 (16)
25	inf45	27	25	29 (14)	28	27	31 (15)
66	inf04	21	20	22 (9)	27	26	31 (15)
71	inf43	25	26	28 (15)	25	27	32 (16)
147	inf44	27	23	22 (15)	27	24	23 (16)
148	inf47	28	28	25 (9)	28	28	28 (12)
148	inf49	23	15	29 (14)	25	23	31 (16)
157	inf30	24	18	30 (16)	27	26	32 (16)
166	inf46	28	27	31 (15)	28	28	31 (16)
171	inf48	26	28	30 (15)	26	28	31 (15)
177	inf13	24	24	28 (14)	24	28	30 (15)
177	inf28	24	26	21 (12)	26	27	26 (15)
179	inf12	27	26	28 (13)	28	27	30 (14)
199	inf09	26	28	28 (3)	26	28	29 (14)
202	inf07	27	27	25 (15)	27	28	25 (15)

Table D4. Pairwise comparisons for PROSODIC POSITION × LEXICAL STRESS: DURATION model for the bilabial stop, with estimates and 95 % Credible Intervals (CI). 'LW-med' – *LW-medial* position; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position; 'uns' – *unstressed*; 'str' – *stressed*. Pairwise comparisons for which 0 is not included in the CI are marked in bold.

Contrast	Estimate	Lower CI	Upper CI
LW-med, uns – LW-ini, uns	-0.09	-0.23	0.05
LW-med, uns – PW-ini, uns	**-0.17**	**-0.31**	**-0.02**
LW-med, uns – LW-med, str	-0.10	-0.24	0.04
LW-med, uns – LW-ini, str	**-0.34**	**-0.48**	**-0.20**
LW-med, uns – PW-ini, str	**-0.43**	**-0.59**	**-0.29**
LW-ini, uns – PW-ini, uns	-0.08	-0.23	0.07
LW-ini, uns – LW-med, str	-0.01	-0.16	0.12
LW-ini, uns – LW-ini, str	**-0.24**	**-0.39**	**-0.11**
LW-ini, uns – PW-ini, str	**-0.34**	**-0.49**	**-0.20**
PW-ini, uns – LW-med, str	0.07	-0.07	0.22
PW-ini, uns – LW-ini, str	**-0.17**	**-0.31**	**-0.02**
PW-ini, uns – PW-ini, str	**-0.26**	**-0.42**	**-0.10**
LW-med, str – LW-ini, str	**-0.23**	**-0.37**	**-0.09**
LW-med, str – PW-ini, str	**-0.33**	**-0.48**	**-0.18**
LW-ini, str – PW-ini, str	-0.10	-0.26	0.05

Table D5. Pairwise comparisons for PROSODIC POSITION × LEXICAL STRESS: CHANGE IN INTENSITY model for the bilabial stop, with estimates and 95 % Credible Intervals (CI). 'LW-med' – *LW-medial* position; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position; 'uns' – *unstressed*; 'str' – *stressed*. Pairwise comparisons for which 0 is not included in the CI are marked in bold.

Contrast	Estimate	Lower CI	Upper CI
LW-med, uns – LW-ini, uns	-1.25	-3.71	1.17
LW-med, uns – PW-ini, uns	**-3.25**	**-5.94**	**-0.68**
LW-med, uns – LW-med, str	-1.77	-4.17	0.78
LW-med, uns – LW-ini, str	**-5.26**	**-7.72**	**-2.85**
LW-med, uns – PW-ini, str	**-6.72**	**-9.40**	**-4.03**
LW-ini, uns – PW-ini, uns	-2.01	-4.57	0.64
LW-ini, uns – LW-med, str	-0.52	-2.91	1.97
LW-ini, uns – LW-ini, str	**-4.01**	**-6.42**	**-1.57**
LW-ini, uns – PW-ini, str	**-5.49**	**-8.05**	**-2.81**
PW-ini, uns – LW-med, str	1.50	-1.27	3.99
PW-ini, uns – LW-ini, str	-2.01	-4.56	0.67
PW-ini, uns – PW-ini, str	**-3.47**	**-6.27**	**-0.72**
LW-med, str – LW-ini, str	**-3.49**	**-5.87**	**-0.96**
LW-med, str – PW-ini, str	-4.95	-7.65	-2.31
LW-ini, str – PW-ini, str	-1.46	-4.11	1.22

Table D6. Pairwise comparisons for PROSODIC POSITION × LEXICAL STRESS: CATEGORY model for the bilabial stop, with estimates and 95 % Credible Intervals (CI). 'LW-med' – *LW-medial* position; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position; 'uns' – *unstressed*; 'str' – *stressed*. Pairwise comparisons for which 0 is not included in the CI are marked in bold.

Contrast	Estimate	Lower CI	Upper CI
LW-med, uns – LW-ini, uns	**-1.33**	**-2.39**	**-0.22**
LW-med, uns – PW-ini, uns	**-2.05**	**-3.20**	**-0.85**
LW-med, uns – LW-med, str	**-1.81**	**-2.88**	**-0.73**
LW-med, uns – LW-ini, str	**-3.75**	**-4.92**	**-2.64**
LW-med, uns – PW-ini, str	**-3.92**	**-5.13**	**-2.67**
LW-ini, uns – PW-ini, uns	-0.72	-1.90	0.43
LW-ini, uns – LW-med, str	-0.48	-1.55	0.61
LW-ini, uns – LW-ini, str	**-2.42**	**-3.55**	**-1.38**
LW-ini, uns – PW-ini, str	**-2.59**	**-3.77**	**-1.40**
PW-ini, uns – LW-med, str	0.23	-0.94	1.40
PW-ini, uns – LW-ini, str	**-1.70**	**-2.90**	**-0.55**
PW-ini, uns – PW-ini, str	**-1.87**	**-3.13**	**-0.68**
LW-med, str – LW-ini, str	**-1.93**	**-3.01**	**-0.90**
LW-med, str – PW-ini, str	**-2.10**	**-3.26**	**-0.96**
LW-ini, str – PW-ini, str	-0.17	-1.33	1.00

Table D7. Pairwise comparisons for PROSODIC POSITION × LEXICAL STRESS: DURATION model for the dentialveolar stop, with estimates and 95 % Credible Intervals (CI). 'LW-med' – *LW-medial* position; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position; 'uns' – *unstressed*; 'str' – *stressed*. Pairwise comparisons for which 0 is not included in the CI are marked in bold.

Contrast	Estimate	Lower CI	Upper CI
LW-med, uns – LW-ini, uns	-0.08	-0.18	0.03
LW-med, uns – PW-ini, uns	-0.09	-0.19	0.01
LW-med, uns – LW-med, str	**-0.21**	**-0.32**	**-0.11**
LW-med, uns – LW-ini, str	**-0.27**	**-0.38**	**-0.16**
LW-med, uns – PW-ini, str	**-0.40**	**-0.52**	**-0.29**
LW-ini, uns – PW-ini, uns	-0.02	-0.12	0.08
LW-ini, uns – LW-med, str	**-0.14**	**-0.24**	**-0.03**
LW-ini, uns – LW-ini, str	**-0.20**	**-0.30**	**-0.08**
LW-ini, uns – PW-ini, str	**-0.33**	**-0.44**	**-0.21**
PW-ini, uns – LW-med, str	**-0.12**	**-0.22**	**-0.02**
PW-ini, uns – LW-ini, str	**-0.18**	**-0.28**	**-0.07**
PW-ini, uns – PW-ini, str	**-0.31**	**-0.42**	**-0.20**
LW-med, str – LW-ini, str	-0.06	-0.16	0.05
LW-med, str – PW-ini, str	**-0.19**	**-0.30**	**-0.08**
LW-ini, str – PW-ini, str	**-0.13**	**-0.25**	**-0.02**

Table D8. Pairwise comparisons for PROSODIC POSITION × LEXICAL STRESS: CHANGE IN INTENSITY model for the dentialveolar stop, with estimates and 95 % Credible Intervals (CI). 'LW-med' – *LW-medial* position; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position; 'uns' – *unstressed*; 'str' – *stressed*. Pairwise comparisons for which 0 is not included in the CI are marked in bold.

Contrast	Estimate	Lower CI	Upper CI
LW-med, uns – LW-ini, uns	-0.88	-2.52	0.70
LW-med, uns – PW-ini, uns	**-1.77**	**-3.30**	**-0.22**
LW-med, uns – LW-med, str	**-2.85**	**-4.41**	**-1.20**
LW-med, uns – LW-ini, str	**-3.41**	**-4.99**	**-1.76**
LW-med, uns – PW-ini, str	**-5.51**	**-7.22**	**-3.78**
LW-ini, uns – PW-ini, uns	-0.90	-2.47	0.65
LW-ini, uns – LW-med, str	**-1.97**	**-3.55**	**-0.36**
LW-ini, uns – LW-ini, str	**-2.54**	**-4.11**	**-0.90**
LW-ini, uns – PW-ini, str	**-4.64**	**-6.40**	**-2.95**
PW-ini, uns – LW-med, str	-1.08	-2.63	0.46
PW-ini, uns – LW-ini, str	**-1.64**	**-3.25**	**-0.10**
PW-ini, uns – PW-ini, str	**-3.74**	**-5.36**	**-1.99**
LW-med, str – LW-ini, str	-0.55	-2.07	1.05
LW-med, str – PW-ini, str	**-2.67**	**-4.40**	**-0.96**
LW-ini, str – PW-ini, str	**-2.10**	**-3.89**	**-0.47**

Table D9. Pairwise comparisons for PROSODIC POSITION × LEXICAL STRESS: CATEGORY model for the dentialveolar stop, with estimates and 95 % Credible Intervals (CI). 'LW-med' – *LW-medial* position; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position; 'uns' – *unstressed*; 'str' – *stressed*. Pairwise comparisons for which 0 is not included in the CI are marked in bold.

Contrast	Estimate	Lower CI	Upper CI
LW-med, uns – LW-ini, uns	-0.67	-1.92	0.52
LW-med, uns – PW-ini, uns	**-1.43**	**-2.58**	**-0.20**
LW-med, uns – LW-med, str	**-2.10**	**-3.32**	**-0.93**
LW-med, uns – LW-ini, str	**-2.37**	**-3.64**	**-1.12**
LW-med, uns – PW-ini, str	**-4.18**	**-5.65**	**-2.76**
LW-ini, uns – PW-ini, uns	-0.76	-2.02	0.45
LW-ini, uns – LW-med, str	**-1.44**	**-2.67**	**-0.15**
LW-ini, uns – LW-ini, str	**-1.71**	**-2.91**	**-0.43**
LW-ini, uns – PW-ini, str	**-3.52**	**-4.92**	**-2.06**
PW-ini, uns – LW-med, str	-0.67	-1.87	0.55
PW-ini, uns – LW-ini, str	-0.95	-2.17	0.26
PW-ini, uns – PW-ini, str	**-2.76**	**-4.14**	**-1.38**
LW-med, str – LW-ini, str	-0.28	-1.41	0.96
LW-med, str – PW-ini, str	**-2.09**	**-3.46**	**-0.68**
LW-ini, str – PW-ini, str	**-1.81**	**-3.27**	**-0.38**

Appendix E: Glottalization of Word-Initial Vowels in the Read Speech Study (Chapter 9)

Table E1. Number of tokens in the presence of glottalization study (*N* = 1723), per speaker (*N* = 21). Tokens are grouped according to the PROSODIC POSITION in which they appeared: *across FWs*; 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position.

		Modal			**Glottalized**		
BLP score	**Speaker ID**	**Across FWs**	**LW-ini**	**PW-ini**	**Across FWs**	**LW-ini**	**PW-ini**
-46	inf17	23	34	8	0	7	9
-25	inf22	24	36	15	0	5	5
-24	inf19	19	17	8	1	24	9
-24	inf18	20	24	7	1	7	5
3	inf21	24	40	20	0	3	2
15	inf16	19	22	8	0	13	12
23	inf27	23	43	26	0	0	1
25	inf45	23	39	16	0	4	10
66	inf04	22	44	27	0	0	1
71	inf43	22	18	18	0	21	8
147	inf44	20	9	0	0	8	4
148	inf47	24	32	18	0	9	3
148	inf49	22	24	16	0	8	0
157	inf30	23	27	14	0	15	8
166	inf46	22	28	12	0	16	11
171	inf48	23	31	25	0	11	2
177	inf13	24	30	20	0	10	2
177	inf28	24	28	12	0	14	2
179	inf12	19	36	14	2	7	4
199	inf09	24	32	20	0	7	5
202	inf07	22	29	16	0	14	3

Table E2. Pairwise comparisons for PROSODIC POSITION × LEXICAL STRESS: GLOTTALIZATION model, with estimates and 95 % Credible Intervals (CI). 'LW-ini' – *LW-initial* position; 'PW-ini' – *PW-initial* position; 'uns' – *unstressed*; 'str' – *stressed*. Pairwise comparisons for which 0 is not included in the CI are marked in bold.

Contrast	Estimate	Lower CI	Upper CI
LW-initial, uns – PW-initial, uns	**-1.61**	**-3.01**	**-0.24**
LW-initial, uns – LW-initial, str	**-3.49**	**-4.86**	**-2.21**
LW-initial, uns – PW-initial, str	**-4.78**	**-6.57**	**-3.10**
PW-initial, uns – LW-initial, str	**-1.89**	**-3.20**	**-0.48**
PW-initial, uns – PW-initial, str	**-3.17**	**-4.86**	**-1.52**
LW-initial, str – PW-initial, str	-1.28	-2.92	0.22

Table E3. Raw count number of types of glottalization ($N = 314$).

Type	Number
Prototypical creaky voice	166
Vocal fry	78
Multiply pulsed voice	22
Aperiodic voice	10
Glottal stop	13
Prototypical–vocal fry	13
Aperiodic–prototypical	9
Prototypical–multiply pulsed	3

Table E4. Selection of relevant pairwise comparisons for VOWEL TYPE × MEASURE: *H1*–H2** model, with estimates and 95 % Credible Intervals (CI). 0 is included within the upper and lower boundaries of the CI in all contrasts. 'm' – *modal* vowel,'g' – *glottalized* vowel. Percentages (e.g., *0 %*) refer to the levels of the variable MEASURE.

Contrast	**Estimate**	**Lower CI**	**Upper CI**
m × 0 % – g × 0 %	0.58	-1.23	2.36
m × 0 % – m × 17 %	0.34	-0.74	1.38
g × 0 % – g × 17 %	1.01	-0.86	3.01
m × 17 % – g × 17 %	1.26	-0.20	2.80
m × 17 % – m × 33 %	0.48	-0.57	1.53
g × 17 % – g × 33 %	-0.72	-2.30	0.76
m × 33 % – g × 33 %	0.07	-1.31	1.40
m × 33 % – m × 50 %	-0.25	-1.27	0.82
g × 33 % – g × 50 %	0.77	-0.55	2.11
m × 50 % – g × 50 %	1.09	-0.20	2.38
m × 50 % – m × 67 %	-0.50	-1.55	0.55
g × 50 % – g × 67 %	-1.10	-2.37	0.17
m × 67 % – g × 67 %	0.48	-0.85	1.73

Table E5. Selection of relevant pairwise comparisons for VOWEL TYPE × MEASURE: *H1*–A3** model, with estimates and 95 % Credible Intervals (CI). 'm' – *modal* vowel; 'g' – *glottalized* vowel. Percentages (e.g., *0 %*) refer to the levels of the variable MEASURE. Pairwise comparisons for which 0 is not included in the CI are marked in bold.

Contrast	Estimate	Lower CI	Upper CI
m × 0 % – g × 0 %	**3.41**	**0.94**	**5.92**
m × 0 % – m × 17 %	-0.52	-1.99	0.93
g × 0 % – g × 17 %	-1.02	-3.64	1.66
m × 17 % – g × 17 %	**2.93**	**0.87**	**5.04**
m × 17 % – m × 33 %	0.36	-1.13	1.82
g × 17 % – g × 33 %	-1.75	-3.89	0.34
m × 33 % – g × 33 %	0.81	-1.07	2.69
m × 33 % – m × 50 %	-0.38	-1.82	1.11
g × 33 % – g × 50 %	0.34	-1.44	2.19
m × 50 % – g × 50 %	1.52	-0.27	3.31
m × 50 % – m × 67 %	-0.32	-1.84	1.13
g × 50 % – g × 67 %	-0.24	-1.96	1.51
m × 67 % – g × 67 %	1.60	-0.16	3.41

Table E6. Selection of pairwise comparisons for VOWEL TYPE × MEASURE: *CPP* model, with estimates and 95 % Credible Intervals (CI). 'm' – *modal* vowel; 'g' – *glottalized* vowel. Percentages (e.g., *0 %*) refer to the levels of the variable MEASURE. Pairwise comparisons for which 0 is not included in the CI are marked in bold.

Contrast	Estimate	Lower CI	Upper CI
m × 0 % – g × 0 %	**4.64**	**3.47**	**5.82**
m × 0 % – m × 17 %	**-2.89**	**-3.57**	**-2.19**
g × 0 % – g × 17 %	**-3.77**	**-5.04**	**-2.52**
m × 17 % – g × 17 %	**3.75**	**2.74**	**4.71**
m × 17 % – m × 33 %	-0.17	-0.88	0.50
g × 17 % – g × 33 %	**-1.85**	**-2.84**	**-0.87**
m × 33 % – g × 33 %	**2.07**	**1.21**	**2.96**
m × 33 % – m × 50 %	-0.32	-1.01	0.38
g × 33 % – g × 50 %	**-1.34**	**-2.20**	**-0.46**
m × 50 % – g × 50 %	**1.05**	**0.22**	**1.91**
m × 50 % – m × 67 %	0.13	-0.56	0.84
g × 50 % – g × 67 %	-0.19	-1.04	0.63
m × 67 % – g × 67 %	0.72	-0.11	1.57

References

Aguilar, L. (2003). Effects of prosodic and segmental variables on vowel sequences pronunciation in Spanish. *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona, Spain, 2003*, 2111–2114. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/papers/p15_2111.pdf

Aguilar, L. (2005). Los enlaces vocálicos: ¿Cuestión de dominios prosódicos? *Revista Internacional de Lingüística Iberoamericana*, *3*(2), 29–48.

Aguilar, L., & Gutiérrez González, Y. M. (2018). Patterns of prominence, phrasing and tonal events in Spanish news reading: An illustrative case study. *Loquens*, *5*(1), 48. https://doi.org/10.3989/loquens.2018.048

Aguilar, L., de-la-Mota, C., & Prieto, P. (2009). *Sp_ToBI Training materials*. Sp_ToBI Training Materials. http://prosodia.upf.edu/sp_tobi/

Alameda, J. R., & Cuetos, F. (1995). *Diccionario de frecuencias de las unidades lingüísticas del castellano*. Universidad de Oviedo.

Alarcos Llorach, E. (1991). *Fonología española* (4th ed.). Gredos.

Alba, M. C. (2006). Accounting for variability in the production of Spanish vowel sequences. In N. Sagarra & A. J. Toribio (Eds.), *Selected Proceedings of the 9th Hispanic Linguistics Symposium* (pp. 273–285). Cascadilla Proceedings Project. http://www.lingref.com/cpp/hls/9/paper1385.pdf

Albalá, M. J., Battaner, E., Carranza, M., de-la-Mota, C., Gil, J., Llisterri, J., Machuca, M. J., Madrigal, N., Marquina, M., Marrero, V., Riera, M., & Ríos, A. (2008). VILE: Nuevos datos acústicos sobre vocales del español [Special issue 1]. *Language Design*.

Amengual, M., & Simonet, M. (2019). Language dominance does not always predict cross-linguistic interactions in bilingual speech production. *Linguistic Approaches to Bilingualism*, 1–26. https://doi.org/10.1075/lab.18042.ame

Amengual Watson, M. (2013). *An experimental approach to phonetic transfer in the production and perception of early Spanish-Catalan bilinguals* [University of Texas at Austin]. http://hdl.handle.net/2152/21732

Austin, L., Hernandez, A. E., & Schwieter, J. W. (2019). *Proficiency predictors in sequential bilinguals*. Cambridge University Press. https://doi.org/10.1017/9781108641395

Avelino, H. (2011). Intonational patterns of topic and focus constructions in Yucatec Maya. In H. Avelino (Ed.), *New perspectives in Mayan linguistics* (pp. 56–79). Cambridge Scholars Publishing.

Avelino, H. (2016). Phonetics in phonology: A cross-linguistic study of laryngeal contrast. In H. Avelino, M. Coler, & L. Wetzels (Eds.), *The phonetics and phonology of laryngeal features in Native American languages* (pp. 157–179). Brill. https://doi.org/10.1163/9789004303218_007

Avelino, H. (2018). Mexico City Spanish. *Journal of the International Phonetic Association*, *48*(2), 223–230. Cambridge Core. https://doi.org/10.1017/S0025100316000232

Avelino, H., Shin, E., & Tilsen, S. (2011). The phonetics of laryngealization in Yucatec Maya. In H. Avelino (Ed.), *New perspectives in Mayan linguistics* (pp. 1–20). Cambridge Scholars.

Barlow, J. A. (2003). The stop-spirant alternation in Spanish: Converging evidence for a fortition account. *Southwest Journal of Linguistics*, *22*(1), 51–86.

Barrera Vásquez, A. (1977). El idioma español en Yucatán. In L. H. Hoyos Villanueva, R. Irigoyen Rosado, R. Ruz Menéndez, & H. Lara y Lara (Eds.), *Enciclopedia yucatanense 6. Yucatán actual* (2nd ed., pp. 341–375). Gobierno de Yucatán.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using `lme4`. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bauer, L. (2008). Lenition revisited. *Journal of Linguistics*, *44*(3), 605–624. https://doi.org/10.1017/s0022226708005331

Baumann, S., & Winter, B. (2018). What makes a word prominent? Predicting untrained German listeners' perceptual judgements. *Journal of Phonetics*, *70*, 20–38. https://doi.org/10.1016/j.wocn.2018.05.004

Beckman, M. E. (1996). The parsing of prosody. *Language and Cognitive Processes*, *11*(1–2), 17–68. https://doi.org/10.1080/016909696387213

Beckman, M. E., & Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. *Phonology*, *3*, 255–309. https://doi.org/10.1017/S095267570000066X

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. J. Munró (Eds.), *Language experience in second language learning: In honor of James Emil Flege* (pp. 13–34). John Benjamins.

Birdsong, D. (2015). Dominance in bilingualism: Foundations of measurement, with insights from the study of handedness. In C. Silva-Corvalán & J. Treffers-Daller (Eds.), *Language dominance in bilinguals: Issues of measurement and operationalization* (pp. 85–105). Cambridge University Press; Cambridge Core.

Birdsong, D., Gertken, L. M., & Amengual, M. (2012). *Bilingual Language Profile: An easy-to-use instrument to assess bilingualism*. COERLL, University of Texas at Austin. https://sites.la.utexas.edu/bilingual/

Bissiri, M. P., Lecumberri, M. L., Cooke, M., & Volín, J. (2011). The role of word-initial glottal stops in recognizing English words. *Proceedings of Interspeech 2011, Florence*, 165–168. https://www.isca-speech.org/archive/archive_papers/interspeech_2011/i11_0165.pdf

Blankenship, B. (2002). The timing of nonmodal phonation in vowels. *Journal of Phonetics*, *30*(2), 163–191. https://doi.org/10.1006/jpho.2001.0155

Boersma, P., & Weenink, D. (2019). *Praat: Doing phonetics by computer* (6.0.43) [Computer software]. http://www.praat.org/

Briceño Chel, F., & Can Tec, G. R. (Eds.). (2017). *U nu'ukbesajil u ts'íibta'al maayat'an-Normas de escritura para la lengua maya*. Instituto Nacional de Lenguas Indígenas. http://site.inali.gob.mx/Micrositios/normas/pdf/Norma_Maya.pdf

Bürkner, P.-C. (2017). `brms`: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28 [2.12.0]. https://doi.org/10.18637/jss.v080.i01

Bye, P., & de Lacy, P. (2008). Metrical influences on Fortition and Lenition. In J. Brandão de Carvalho, T. Scheer, & P. Ségéral (Eds.), *Lenition and fortition* (pp. 173–204). Mouton de Gruyter.

Cano Aguilar, R. (1992). *El español a través de los tiempos*. Arco/Libros.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *1*(1), 1–32. https://www.doi.org/10.18637/jss.v076.i01

Carrasco, P., Hualde, J. I., & Simonet, M. (2012). Dialectal differences in Spanish voiced obstruent allophony: Costa Rican versus Iberian Spanish. *Phonetica*, *69*(3), 149–179. https://doi.org/10.1159/000345199

Catford, J. C. (1964). Phonation types: The classification of some laryngeal components of speech production. In D. Abercrombie, D. B. Fry, P. MacCarthy, & N. C. Scott (Eds.), *In honour of Daniel Jones: Papers contributed on the occasion of his eightieth birthday* (pp. 26–37). Longmans.

Catford, J. C. (1977). *Fundamental problems in phonetics*. Edinburgh University Press.

Chappell, W. (2013). *Social and linguistic factors conditioning the glottal stop in Nicaraguan Spanish* [Ohio State University]. http://rave.ohiolink.edu/etdc/view?acc_num=osu1366282634

Cho, T. (2016). Prosodic boundary strengthening in the phonetics–prosody interface. *Language and Linguistics Compass*, *10*(3), 120–141. https://doi.org/10.1111/lnc3.12178

Cho, T., & Keating, P. A. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics*, *29*(2), 155–190. https://doi.org/10.1006/jpho.2001.0131

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics*, *27*(2), 207–229. https://doi.org/10.1006/jpho.1999.0094

Cho, T., & McQueen, J. M. (2005). Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics*, *33*, 121–157. https://www.doi.org/10.1016/j.wocn.2005.01.001

Cho, T., McQueen, J. M., & Cox, E. A. (2007). Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics*, *35*, 210–243. https://doi.org/10.1016/j.wocn.2006.03.003

Clopper, C. G., & Turnbull, R. (2018). Exploring variation in phonetic reduction: Linguistic, social, and cognitive factors. In F. Cangemi, M. Clayards, O. Niebuhr, B. Schuppler, & M. Zellers (Eds.), *Rethinking reduction. Interdisciplinary perspectives on conditions, mechanisms, and domains for phonetic variation* (pp. 25–72). De Gruyter Mouton. https://www.degruyter.com/view/product/481060

Colantoni, L., & Marinescu, I. (2010). The scope of stop weakening in Argentine Spanish. In M. Ortega-Llebaria (Ed.), *Selected Proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology* (pp. 100–114). Cascadilla Proceedings Project.

Colazo-Simón, A. (2007). *Phenomenes glottaux en situation de contact linguistique: Maya et espagnol du Yucatán, Mexique* [PhD, Université Paris III]. http://tel.archives-ouvertes.fr/docs/00/67/76/35/PDF/A.C.S._thA_se.pdf

Cole, J., Hualde, J. I., & Iskarous, K. (1999). Effects of prosodic and segmental context on /g/-lenition in Spanish. In O. Fujimura, B. D. Joseph, & B. Palek (Eds.), *Proceedings of the Fourth International Linguistics and Phonetics Conference* (pp. 575–589). Karolinum.

Cole, J., Kim, H., Choi, H., & Hasegawa-Johnson, M. (2007). Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from Radio News speech. *Journal of Phonetics*, *35*(2), 180–209. https://doi.org/10.1016/j.wocn.2006.03.004

Colina, S. (2016). On onset clusters in Spanish: Voiced obstruent underspecification and /f/. In R. A. Núñez-Cedeño (Ed.), *The syllable and stress* (pp. 107–138). De Gruyter Mouton. https://doi.org/10.1515/9781614515975-006

Colina, S. (in press). Descripción fonológica de las paravocales. In J. Gil Fernández & J. Llisterri Boix (Eds.), *Fonética y fonología descriptivas de la lengua española*. Georgetown University Press.

Correa, J. A. (2012). El español hablado en el Pacífico de Colombia: Dos rasgos fonéticos de presunto sustrato africano. In R. J. File-Muriel & R. Orozco (Eds.), *Colombian varieties of Spanish* (pp. 43–66). Iberoamericana-Vervuert.

Crawley, M. J. (2013). Bayesian statistics. In *The R book. Second edition* (pp. 752–767). John Wiley & Sons.

De Krom, G. (1993). A cepstrum-based technique for determining a harmonics-to-noise ration in speech signals. *Journal of Speech, Language, and Hearing Research*, *36*(2), 254–266. https://doi.org/10.1044/jshr.3602.254

D'Imperio, M., Elordieta, G., Frota, S., Prieto, P., & Vigário, M. (2005). Intonational phrasing in Romance: The role of syntactic and prosodic structure. In S. Frota, M. Vigário, & M. J. Freitas (Eds.), *Prosodies. With special reference to Iberian languages* (pp. 59–98). Mouton de Gruyter.

Davies, M. (2016). *Corpus del español: Web/Dialects* [Online corpus]. https://www.corpusdelespanol.org/web-dial/

Delais-Roussarie, E., Post, B., Avanzi, M., Buthke, C., Di Cristo, A., Feldhausen, I., Jun, S.-A., Martin, P., Meisenburg, T., Rialland, A., Sichel-Bazin, R., & Yoo, H.-Y. (2015). Intonational phonology of French: Developing a ToBI system for French. In S. Frota & P. Prieto (Eds.), *Intonation in Romance*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199685332.003.0003

Di Napoli, J. (2018). *The phonetics and phonology of glottalization in Italian*. Peter Lang.

Dilley, L., Shattuck-Hufnagel, S., & Ostendorf, M. (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics*, *24*, 423–444. https://doi.org/10.1006/jpho.1996.0023

Dmitrieva, O., Llanos, F., Shultz, A. A., & Francis, A. L. (2015). Phonological status, not voice onset time, determines the acoustic realization of onset f0 as a secondary voicing cue in Spanish and English. *Journal of Phonetics*, *49*, 77–95. https://doi.org/10.1016/j.wocn.2014.12.005

Eddington, D. (2011). What are the contextual phonetic variants of /β, ð, γ/ in colloquial Spanish? *Probus*, *23*(1), 1–19. https://doi.org/10.1515/prbs.2011.001

Edmondson, J. A., & Esling, J. H. (2006). The valves of the throat and their functioning in tone, vocal register and stress: Laryngoscopic case studies. *Phonology*, *23*(2), 157–191. https://doi.org/10.1017/S095267570600087X

Edwards, J. (2006). Foundations of bilingualism. In T. K. Bhatia & W. C. Ritchie (Eds.), *Handbook of bilingualism* (pp. 7–31). Blackwell.

Elordieta, G. (2014). The word in phonology. In I. Ibarretxe-Antuñano & J.-L. Mendívil-Giró (Eds.), *To be or not to be a word: New reflections on the definition of word* (pp. 6–65). Cambridge Scholars.

Elordieta, G., Frota, S., & Vigário, M. (2005). Subjects, objects and intonational phrasing in Spanish and Portuguese. *Studia Linguistica*, *59*(2-3), 110–143. https://doi.org/10.1111/j.1467-9582.2005.00123.x

Elvin, J., Vasiliev, P., & Escudero, P. (2019). Production and perception in the acquisition of Spanish and Portuguese. In M. Gibson & J. Gil (Eds.), *Romance phonetics and phonology* (pp. 367–380). Oxford University Press. https://doi.org/10.1093/oso/9780198739401.003.0018

England, N. C., & Baird, B. O. (2017). Phonology and phonetics. In J. Aissen, N. C. England, & R. Zavala Maldonado (Eds.), *The Mayan languages* (pp. 175–200). Routledge.

Ennever, T., Meakins, F., & Round, E. R. (2017). A replicable acoustic measure of lenition and the nature of variability in Gurindji stops. *Journal of the Association for Laboratory Phonology*, *8*(1), 1–32. https://doi.org/10.5334/labphon.18

Escandell Vidal, M. V., & Leonetti, M. (2000). Categorías funcionales y semántica procedimental. In M. Martínez Hernández (et al.) (Ed.), *Cien años de investigación semántica, de Michel Breal a la actualidad: Actas del Congreso Internacional de Semántica* (Vol. 1, pp. 363–378). Clásicas. http://www3.uah.es/leonetti/papers/FuncyProc.pdf

Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization* [LOT Dissertation Series 113]. Utrecht University.

Esling, J. H., & Edmondson, J. A. (2011). Acoustical analysis of voice quality for sociophonetic purposes. In M. Di Paolo & M. Yaeger-Dror (Eds.), *Sociophonetics. A student's guide* (pp. 131–148). Routledge.

Estebas-Vilaplana, E., & Prieto, P. (2010). Castilian Spanish Intonation. In P. Prieto & P. Roseano (Eds.), *Transcription of Intonation of the Spanish language* (pp. 17–48). LINCOM Europa. http://prosodia.upf.edu/home/arxiu/activitats/4th_workshop/protegit/Castilian_final.pdf

Face, T.L. (2003). Intonation in Spanish declaratives: Differences between lab speech and spontaneous speech. *Catalan Journal of Linguistics*, *2*, 115–131.

Fant, G. (1960). *Acoustic theory of speech production* (2nd edition). Mouton.

Feldhausen, I., Gabriel, C., & Pešková, A. (2010). Prosodic phrasing in Argentinean Spanish: Buenos Aires and Neuquén. *Speech Prosody 2010*, Paper 111. https://www.isca-speech.org/archive/sp2010/papers/sp10_111.pdf

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. https://doi.org/10.1016/S0095-4470(19)30537-6

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* (pp. 233–277). York Press.

Flege, J. E. (1999). Age of learning and second language speech. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 101–121). Lawrence Erlbaum.

Flege, J. E. (2002). Interactions between the native and second-language phonetic systems. In P. Burmeister, T. Piske, & A. Rhodes (Eds.), *An integrated view of language development: Papers in Honor of Henning Wode* (pp. 217–244). Wissenschaftlicher Verlag.

Flege, J. E., & Eefting, W. (1986). Linguistic and developmental effects on the production and perception of stop consonants. *Phonetica*, *43*, 155–171. https://doi.org/10.1016/S0095-4470(19)30538-8

Flege, J. E., & Eefting, W. (1987). The production and perception of English stops by Spanish speakers of English. *Journal of Phonetics*, *15*(1), 67–83. https://doi.org/10.1016/S0095-4470(19)30538-8

Fletcher, J. (2010). The prosody of speech: Timing and rhythm. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (2nd edition, pp. 523–602). Wiley-Blackwell. https://doi.org/10.1002/9781444317251.ch15

Fougeron, C., & Keating, P. A. (1995). Demarcating prosodic groups with articulation. *The Journal of the Acoustical Society of America*, *97*, 3384. https://doi.org/10.1121/1.412582

Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *The Journal of the Acoustical Society of America*, *101*(6), 3728–3740. https://doi.org/10.1121/1.418332

Frazier, M. (2009). *The production and perception of pitch and glottalization in Yucatec Maya* [PhD]. University of North Carolina.

Frazier, M. (2013). The phonetics of Yucatec Maya and the typology of laryngeal complexity. *STUF - Language Typology and Universals*, *66*(1), 7–21.

Frota, S. (2014). The intonational phonology of European Portuguese. In S.-A. Jun (Ed.), *Prosodic Typology II*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199567300.003.0002

García Fajardo, J. (1984). *Fonética del español de Valladolid, Yucatán*. Universidad Nacional Autónoma de México, Instituto de Investigaciones Filológicas.

Garellek, M. (2012a). Glottal stops before word-initial vowels in American English: Distribution and acoustic characteristics. *Working Papers in Phonetics*, *110*, 1–23.

Garellek, M. (2012b). The timing and sequencing of coarticulated non-modal phonation in English and White Hmong. *Journal of Phonetics*, *40*, 152–161. https://www.doi.org/10.1016/j.wocn.2011.10.003

Garellek, M. (2013). *Production and perception of glottal stops* [University of California]. http://idiom.ucsd.edu/~mgarellek/files/Garellek_dissertation.pdf

Garellek, M. (2014). Voice quality strengthening and glottalization. *Journal of Phonetics*, *45*, 106–113. https://doi.org/10.1016/j.wocn.2014.04.001

Garellek, M. (2019). The phonetics of voice. In W. F. Katz & P. F. Assmann (Eds.), *The Routledge handbook of phonetics* (pp. 75–106). https://doi.org/10.4324/9780429056253

Georgeton, L., & Fougeron, C. (2014). Domain-initial strengthening on French vowels and phonological contrasts: Evidence from lip articulation and spectral variation. *Dynamics of Articulation and Prosodic Structure*, *44*, 83–95. https://doi.org/10.1016/j.wocn.2014.02.006

Gertken, L. M., Amengual, M., & Birdsong, D. (2014). Assessing language dominance with the Bilingual Language Profile. In P. Leclercq, A. Edmonds, & H. Hilton (Eds.), *Measuring L2 proficiency. Perspectives from SLA* (pp. 208–225). Multilingual Matters. https://doi.org/10.21832/9781783092291-014

Gobl, C., & Ní Chasaide, A. (2010). Voice source variation and its communicative functions. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (2nd edition, pp. 378–423). Wiley-Blackwell. https://doi.org/10.1002/9781444317251.ch11

Gordon, M. (2014). Disentangling stress and pitch accent: Toward a typology of prominence at different prosodic levels. In H. van der Hulst (Ed.), *Word stress: Theoretical and typological issues* (pp. 83–118). Cambridge University Press.

Gordon, M., & Ladefoged, P. (2001). Phonation types: A cross-linguistic overview. *Journal of Phonetics*, *29*, 383–406. https://doi.org/10.006/jpho.2001.0147

Granda, G. de (1982). Observaciones sobre la fonética del español en el Paraguay. *Anuario de Letras*, *20*, 145–194.

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, *36*(1), 3–15. https://doi.org/10.1016/0093-934X(89)90048-5

Grosjean, F. (2008). Methodological and conceptual issues. In *Studying bilinguals* (pp. 241–272). Oxford University Press.

Gussenhoven, C., & Teeuw, R. (2008). A moraic and a syllabic H-tone in Yucatec Maya. In E. Z. Herrera & P. Martín Butragueño (Eds.), *Fonología instrumental: Patrones fónicos y variación* (pp. 49–71). El Colegio de México. http://gep.ruhosting.nl/carlos/tonesyucatecmaya[1].pdf

Hall, T. A. (1999). The phonological word: A review. In T. A. Hall & U. Kleinhenz (Eds.), *Studies on the phonological word* (pp. 1–22). John Benjamins. https://doi.org/10.1075/cilt.174.02hal

Hall, T. A., & Kleinhenz, U. (Eds.). (1999). *Studies on the phonological word*. John Benjamins. https://doi.org/10.1075/cilt.174

Hanson, H. M., & Chuang, E. S. (1999). Glottal characteristics of male speakers: Acoustic correlates and comparison with female data. *Journal of the Acoustical Society of America*, *106*(2), 1064–1077. https://doi.org/10.1121/1.427116

Harris, J. W. (1969). *Spanish phonology*. The MIT Press.

Harris, J. W. (1983). *Syllable structure and stress in Spanish. A nonlinear analysis*. MIT Press.

Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. Chicago University Press.

Hildebrandt, K. A. (2015). The prosodic word. In John R. Taylor (Ed.), *The Oxford handbook of the word* (pp. 221–245). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199641604.013.035

Hillenbrand, J., Cleveland, R. A., & Erickson, R. L. (1994). Acoustic correlates of breathy vocal quality. *Journal of Speech and Hearing Research*, *37*, 769–778.

Hualde, J. I. (1989). Silabeo y estructura morfémica en español. *Hispania*, *72*(4), 821–831.

Hualde, J. I. (2003). El modelo métrico y autosegmental. In P. Prieto (Ed.), *Teorías de la entonación* (pp. 155–183). Ariel.

Hualde, J. I. (2007). Stress removal and stress addition in Spanish. *Journal of Portuguese Linguistics*, *6*(1), 59–89. https://doi.org/10.5334/jpl.145

Hualde, J. I. (2009). Unstressed words in Spanish. *Language Sciences*, *31*(2–3), 199–212. https://doi.org/10.1016/j.langsci.2008.12.003

Hualde, J. I. (2013). *Los sonidos del español*. Cambridge University Press. https://doi.org/10.1017/CBO9780511719943

Hualde, J. I. (in press). Descripción fonológica del acento. In J. Gil Fernández & J. Llisterri Boix (Eds.), *Fonética y fonología descriptivas de la lengua española*. Georgetown University Press.

Hualde, J. I., Beristain, A., Icardo Isasa, A., & Zhang, J. (2019). Lenition of word-final plosives in Basque. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019* (pp. 642–646). Australasian Speech Science and Technology Association Inc. http://www.assta.org/proceedings/ICPhS2019/papers/ICPhS_691.pdf

Hualde, J. I., & Prieto, P. (2014). Lenition of intervocalic alveolar fricatives in Catalan and Spanish. *Phonetica*, *71*(2), 109–127. https://doi.org/10.1159/000368197

Hualde, J. I., & Prieto, P. (2015). Intonational variation in Spanish European and American varieties. In S. Frota & P. Prieto (Eds.), *Intonation in Romance* (pp. 350–391). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199685332.003.0010

Hualde, J. I., Shosted, R., & Scarpace, D. (2011). *Acoustics and articulation of Spanish /d/ spirantization*, 906–909. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2011/OnlineProceedings/RegularSession/Hualde/Hualde.pdf

Hualde, J. I., Simonet, M., & Nadeu, M. (2011). Consonant lenition and phonological recategorization. *Journal of the Association for Laboratory Phonology*, *2*(2), 301–329. https://doi.org/10.1515/labphon.2011.011

Hualde, J. I., Simonet, M., & Torreira, F. (2008). Postlexical contraction of nonhigh vowels in Spanish. *Lingua*, *118*(12), 1906–1925. https://doi.org/10.1016/j.lingua.2007.10.004

Instituto Nacional de Estadística y Geografía-INEGI. (2015). *Encuesta Intercensal (2015). Panorama sociodemográfico de Quintana Roo*. http://coespo.qroo.gob.mx/Descargas/doc/PUBLICACCIONES%20DE%20INTERES/PANORAMA%20SOCIODEMOGRAFICO%20DE%20QUINTANA%20ROO%202015.pdf

Iseli, M., Shue, Y.-L., & Alwan, A. (2007). Age, sex, and vowel dependencies of acoustic measures related to the voice source. *Journal of the Acoustical Society of America*, *121*(4), 2283–2295. https://www.doi.org/10.1121/1.2697522

Ito, J., & Mester, A. (2009). The extended prosodic word. In J. Grijzenhout & B. Kabak (Eds.), *Phonological domains. Universals and deviations* (pp. 135–194). Mouton de Gruyter.

Jun, S.-A. (2005a). Korean intonational phonology and prosodic transcription. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 201–229). Oxford University Press.

Jun, S.-A. (2005b). Prosodic typology. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 430–458). Oxford University Press.

Jun, S.-A., & Fletcher, J. (2014). Methodology of studying intonation: From data collection to data analysis. In S.-A. Jun (Ed.), *Prosodic typology II: The phonology of intonation and phrasing*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199567300.003.0016

Kaland, C., & Himmelmann, N. (2019). Repetition reduction revisited: The prosody of repeated words in Papuan Malay. *Language and Speech*, 1–25. https://doi.org/10.1177/0023830918820044

Katz, J., & Pitzanti, G. (2019). The phonetics and phonology of lenition: A Campidanese Sardinian case study. *Journal of the Association for Laboratory Phonology*, *10*(1), 16. https://doi.org/10.5334/labphon.184

Keating, P. A. (2003). Phonetic encoding of prosodic structure. *Proceedings of the 6th International Seminar on Speech Production*, 119–124. http://linguistics.ucla.edu/people/keating/sydney.pdf

Keating, P., Cho, T., Fougeron, C., & Hsu, C.-S. (2004). Domain-initial articulatory strengthening in four languages. In J. Local, R. Ogden, & R. Temple (Eds.), *Phonetic interpretation: Papers in Laboratory Phonology VI* (pp. 145–163). Cambridge University Press. https://doi.org/10.1017/CBO9780511486425.009

Keating, P., & Esposito, C. (2007). Linguistic voice quality. *11th Australasian International Conference on Speech Science and Technology 2006*, 1–6.

Keating, P., Garellek, M., & Kreiman, J. (2015). *Acoustic properties of different kinds of creaky voice*. 18th International Congress of Phonetic Sciences, Glasgow, Scotland, UK. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0821.pdf

Keating, P. A., Wright, R., & Zhang, J. (2001). Word-level asymmetries in consonant articulation. *UCLA Working Papers in Phonetics*, *97*, 157–173.

Kidder, E. (2013). *Prominence in Yucatec Maya: The role of stress in Yucatec Maya words* [PhD thesis]. University of Arizona.

Kingston, J. (2008). Lenition. In L. Colantoni & J. Steele (Eds.), *Selected proceedings of the Third Conference on Laboratory Approaches to Spanish Phonology* (pp. 1–31). Cascadilla Proceedings Project. http://www.lingref.com/cpp/lasp/3/paper1711.pdf

Kirby, J. (2018–2019). *PraatSauce* [Computer software]. https://github.com/kirbyj/praatsauce

Kisler, T., Reichel, U., & Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech & Language*, *45*(C), 326–347 [2.22]. https://doi.org/10.1016/j.csl.2017.01.005

Kohler, K. J. (2001). Plosive-related glottalization phenomena in read and spontaneous speech. A stød in German? In N. Grønnum & J. Rischel (Eds.), *To honour Eli Fischer-Jørgensen: Festschrift on the occasion of her 90th birthday February 11th, 2001* (pp. 174–211). Reitzel.

Krämer, M. (2001). Yucatec Maya vowel alternations—Harmony as syntagmatic identity. *Zeitschrift Für Sprachwissenschaft*, *20*(2), 175–217. https://doi.org/10.1515/zfsw.2001.20.2.175

Kruschke, J. K. (2015). *Doing Bayesian data analysis. A tutorial with R, JAGS, and Stan* (2nd ed.). Academic Press.

Kruschke, J. K., Aguinis, H., & Joo, H. (2012). The time has come: Bayesian methods for data analysis in the organizational sciences. *Organizational Research Methods*, *15*(4), 722–752. https://doi.org/10.1177/1094428112457829

Kügler, F., Skopeteas, S., & Verhoeven, E. (2007). Encoding information structure in Yucatec Maya: On the interplay of prosody and syntax. *Interdisciplinary Studies on Information Structure*, *8*, 187–208.

Ladefoged, P. (1971). *Preliminaries to linguistic phonetics*. University of Chicago.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Blackwell.

Lahoz Bengoechea, J. M. (2015). *Fonética y fonología de los fenómenos de refuerzo consonántico en el seno de unidades léxicas en español* [Unpublished, Universidad Complutense de Madrid]. http://eprints.ucm.es/40122/

Lahoz-Bengoechea, J. M. (2018). *Word boundary perception in ambiguous resyllabification contexts in Spanish*. 16th Conference on Laboratory Phonology, Lisbon.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174. https://doi.org/10.2307/2529310

Laver, J. (1980). *The phonetic description of voice quality*. Cambridge University Press.

Laver, J. (1994). *Principles of phonetics*. Cambridge University Press.

Lavoie, L. M. (2001). *Consonant strength. Phonological patterns and phonetic manifestations*. Garland.

Lenth, R. (2020). *emmeans: Estimated marginal means, aka least-squares means* (1.4.5) [Computer software]. https://CRAN.R-project.org/package=emmeans

Lipski, J. M. (1987). Breves notas sobre el español filipino. *Anuario de Letras*, *25*, 209–219.

Lipski, J. M. (1996). *El español de América* (3rd ed.). Cátedra.

Lipski, J. M. (1997). Spanish word stress: The interaction of moras and minimality. In F. Martínez-Gil & A. Morales-Font (Eds.), *Issues in the phonology and morphology of the Major Iberian languages* (pp. 559–593). Georgetown University Press.

Llisterri, J. (in press). Descripción fonética del acento. In J. Gil Fernández & J. Llisterri Boix (Eds.), *Fonética y fonología descriptivas de la lengua española*. Georgetown University Press.

Lope Blanch, J. M. (1987). *Estudios sobre el español de Yucatán*. Universidad Nacional Autónoma de México.

Lozano, M. del C. (1979). *Stop and spirant alternations: Fortition and spirantization processes in Spanish phonology*. Indiana University Linguistics Club.

Malisz, Z., Żygis, M., & Pompino-Marschall, B. (2013). Rhythmic structure effects on glottalisation: A study of different speech styles in Polish and German. *Laboratory Phonology*, *4*(1), 119–158. https://doi.org/10.1515/lp-2013-0006

Martín Butragueño, P. (2014). *Fonología variable del español de México. Volumen I: procesos segmentales*. El Colegio de México.

Martín Butragueño, P. (2017). Contacto dialectal entonativo. Estudio exploratorio. In L. Orozco & A. Guerrero Galván (Eds.), *Estudios de variación geolingüística* (pp. 139–168). Instituto Nacional de Antropología e Historia.

Martínez Celdrán, E. (2013). Caracterización acústica de las aproximantes del español. *Estudios de Fonética Experimental*, *22*, 11–35.

Martínez Celdrán, E., & Fernández Planas, A. M. (2007). *Manual de fonética española: Articulaciones y sonidos del español*. Ariel.

Martínez-Celdrán, E., Fernández-Planas, A. M., & Carrera-Sabaté, J. (2003). Castilian Spanish. *Journal of the International Phonetic Association*, *33*(2), 255–259. Cambridge Core. https://doi.org/10.1017/S0025100303001373

Martínez García, N. (2017). /b d g/ en Quintana Roo (Yucatán). Estructura prosódica y lenguas en contacto. In V. Marrero Aguiar & E. Estebas Vilaplana (Eds.), *Tendencias actuales en fonética experimental. Current Trends in Experimental Phonetics* (pp. 165–170). UNED.

Martínez García, N., & Kaland, C. (2019). The prosody of repeated mentions in Yucatecan Spanish. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019* (pp. 2494–2498). Australasian Speech Science and Technology Association Inc. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2019/papers/ICPhS_2543.pdf

Martínez García, N., & Uth, M. (2019). Lack of syllable duration as a post-lexical acoustic cue in Spanish in contact with Maya. *Languages*, *4*(4), 84. https://doi.org/10.3390/languages4040084

McKinnon, S. (2018). *A sociophonetic analysis of word-initial vowel glottalization in monolingual and bilingual Guatemalan Spanish*. Hispanic Linguistics Symposium, University of Texas, Austin.

McKinnon, S. (2020). *Quantifying propagation in language contact: A variationist analysis of stops in bilingual (Spanish-Kaqchikel Maya) Guatemalan Spanish*. Indiana University.

Michnowicz, J. (2008). Final nasal variation in Merida, Yucatan. *Spanish in Context*, *5*(2), 278–303. https://doi.org/10.1075/sic.5.2.13mic

Michnowicz, J. (2009). Intervocalic voiced stops in Yucatan Spanish: A case of contact-induced language change? In M. Lacorte & J. Leeman (Eds.), *Español en Estados Unidos y otros contextos de contacto: Sociolingüística, ideología y pedagogía* (pp. 67–84). Iberoamericana/Vervuert.

Michnowicz, J. (2011). Dialect standardization in Merida, Yucatan: The case of (b d g). *Revista Internacional de Lingüística Iberoamericana*, *9*(2), 191–212. https://www.jstor.org/stable/41678477

Michnowicz, J., & Carpenter, L. (2013). Voiceless stop aspiration in Yucatan Spanish: A sociolinguistic analysis. *Spanish in Context*, *10*(3), 410–437.

Michnowicz, J., & Hyler, A. (2020). The acento pujado in Yucatan Spanish: Prosodic rhythm and the search for the "yucateco" accent. In L. A. Ortiz López, R. E. Guzzardo Tamargo, & M. González-Rivera (Eds.), *Hispanic contact linguistics. Theoretical, methodological and empirical perspectives* (pp. 116–136). John Benjamins. https://doi.org/10.1075/ihll.22.05mic

Michnowicz, J., & Kagan, L. (2016). On glottal stops in Yucatan Spanish. In S. Sessarego & F. Tejedo-Herrero (Eds.), *Spanish language and sociolinguistic analysis* (pp. 219–239). John Benjamins. https://doi.org/10.1075/ihll.8.09mic

Moreno Fernández, F. (2009). *La lengua española en su geografía*. Arco/Libros.

Mota, C. de-la, Martín Butragueño, P., & Prieto, P. (2010). Mexican Spanish Intonation. In P. Prieto & P. Roseano (Eds.), *Transcription of intonation of the Spanish language* (pp. 319–350). Lincom Europa.

Navarro Tomás, T. (1996). *Manual de pronunciación española* (26th ed.). Consejo Superior de Investigaciones Científicas.

Nespor, M., & Vogel, I. (2007). *Prosodic phonology. With a new foreword* (2nd ed.). Mouton de Gruyter.

Nicenboim, B., Logačev, P., Gattei, C., & Vasishth, S. (2016). When high-capacity readers slow down and low-capacity readers speed up: Working memory and locality effects. *Frontiers in Psychology*, *7*, article 280. https://doi.org/10.3389/fpsyg.2016.00280

Obediente Sosa, E., & Méndez Seijas, J. (in press). La sílaba. In J. Gil Fernández & J. Llisterri Boix (Eds.), *Fonética y fonología descriptivas de la lengua española*. Georgetown University Press.

Oh, E. (2011). Effect of speaker gender on voice onset time in Korean stops. *Journal of Phonetics*, *39*(1), 59–67. https://doi.org/10.1016/j.wocn.2010.11.002

Olson, D. J. (2013). Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in VOT production. *Journal of Phonetics*, *41*(6), 407–420. https://doi.org/10.1016/j.wocn.2013.07.005

Onaka, A. (2003). Domain-initial strengthening in Japanese: An acoustic and articulatory study. *Proceedings of 15th ICPhS Barcelona*, 2091–2094.

Orie, O. O., & Bricker, V. R. (2000). Placeless and historical laryngeals in Yucatec Maya. *International Journal of American Linguistics*, *66*(3), 283–317.

Ortega-Llebaria, M. (2004). Interplay between phonetic and inventory constraints in the degree of spirantization of voiced stops: Comparing intervocalic /b/ and intervocalic /g/ in Spanish and English. In T. L. Face (Ed.), *Laboratory approaches to Spanish phonology* (pp. 237–253). Mouton de Gruyter.

Ortega-Llebaria, M. (2006). Phonetic cues to stress and accent in Spanish. In M. Díaz-Campos (Ed.), *Selected Proceedings of the 2nd Conference on Laboratory Approaches to Spanish Phonetics and Phonology*. Cascadilla Proceedings Project.

Ortega-Llebaria, M., & Prieto, P. (2007). Disentangling stress from accent in Spanish: Production patterns of the stress contrast in deaccented syllables. In P. Prieto, J. Mascaró, & M.-J. Solé (Eds.), *Segmental and prosodic issues in Romance phonology* (pp. 155–197). John Benjamins. https://doi.org/10.1075/cilt.282.11ort

Ostendorf, M., Price, P., & Shattuck-Hufnagel, S. (1995). *The Boston University Radio News Corpus*. Linguistic Data Consortium. https://catalog.ldc.upenn.edu/LDC96S36

Parrell, B. (2011). Dynamical account of how /b, d, g/ differ from /p, t, k/ in Spanish: Evidence from labials. *Laboratory Phonology*, *22*(2), 423–449. https://doi.org/10.1515/labphon.2011.016

Parrell, B., & Narayanan, S. (2018). Explaining coronal reduction: Prosodic structure and articulatory posture. *Phonetica*, *75*(2), 151–181. https://doi.org/10.1159/000481099

Pentland, C., & Laughren, M. (2005). Distinguishing prosodic word and phonological word in Walpiri: Prosodic constituency in morphologically complex words. In I. Mushin (Ed.), *Proceedings of the 2004 Conference of the Australian Linguistic Society (ALS 2004)* (pp. 1–17).

Pérez Aguilar, R. A. (2002). *El habla de Chetumal: Fonética, gramática, léxico indígena y chiclero*. Fondo Estatal para la Cultura y las Artes de Quintana Roo and Instituto Quintanarroense de Cultura and Universidad de Quintana Roo.

Pierrehumbert, J. (1995). Prosodic effects on glottal allophones. In O. Fujimura & M. Hirano (Eds.), *Vocal fold physiology: Voice quality control. Proceedings of the 8th Vocal Fold Physiology* (pp. 39–60). Singular.

Pierrehumbert, J., & Talkin, D. (1992). Lenition of /h/ and glottal stop. In G. J. Docherty & D. R. Ladd (Eds.), *Papers in laboratory phonology 2. Gesture, segment, prosody* (pp. 90–117). Cambridge University Press.

Pike, K. L. (1946). Phonemic pitch in Maya. *International Journal of American Linguistics*, *12*(2), 82.

Piller, I., & Pavlenko, A. (2006). Bilingualism and gender. In T. K. Bhatia & W. C. Ritchie (Eds.), *Handbook of bilingualism* (pp. 489–511). Blackwell.

Podesva, R. J., & Kajino, S. (2014). Sociophonetics, gender, and sexuality. In S. Ehrlich, M. Meyerhoff, & J. Holmes (Eds.), *The handbook of language, gender, and sexuality. Second edition* (pp. 103–122). Wiley Blackwell.

Polo Cano, N. (2015). Aproximación preliminar al sintagma fonológico en español. *Loquens*, *2*(2), e020. https://doi.org/10.3989/loquens.2015.020

Polo Cano, N., & Elordieta, G. (2016). Evidencia segmental del sintagma fonológico en español. *Lingüística Española Actual*, *38*(1), 43–67.

Prieto, P. (2006). Phonological phrasing in Spanish. In F. Martínez-Gil & S. Colina (Eds.), *Optimality-Theoretic studies in Spanish phonology* (pp. 39–61). John Benjamins. https://doi.org/10.1075/la.99.03pri

Prieto, P., & Roseano, P. (2010). Introduction. In P. Prieto & P. Roseano (Eds.), *Transcription of Intonation of the Spanish language* (pp. 1–13). LINCOM Europa.

Quilis, A. (1999). *Tratado de fonología y fonética españolas* (2nd ed.). Gredos.

R Core Team. (2020). [Computer software] (3.6.3). R Foundation for Statistical Computing. https://www.R-project.org/

Rao, R. (2005). *The phonetics, phrasing, and pragmatics of spontaneous speech intonation in Mexican Spanish*. Universitat Autónoma de Barcelona, Phonetics and Phonology in Iberia (PaPI). http://webs2002.uab.es/filologiacatalana/papi/files/handout_Rajiv_Rao.pdf

Rao, R. (2007). Effects of deaccenting in two speech styles of Barcelona Spanish. In R. Colavin, K. Cooke, K. Davidson, S. Fukuda, & A. Del Guidice (Eds.), *Proceedings of the 37 Western Conference on Linguistics (WECOL 2007)* (pp. 202–216). Department of Linguistics, University of California, San Diego.

http://www.fresnostate.edu/artshum/linguistics/documents/WECOL%202007%20VOLUME%2018.pdf

Real Academia Española. (1997–2019). *Corpus de Referencia del Español Actual (CREA)*. http://corpus.rae.es/creanet.html

Real Academia Española & Asociación de Academias de la Lengua Española. (2009). *Nueva gramática de la lengua española: Morfología y sintaxis* (Vol. 1). Espasa.

Real Academia Española & Asociación de Academias de la Lengua Española. (2010). *Ortografía de la lengua española*. Espasa.

Redi, L., & Shattuck-Hufnagel, S. (2001). Variation in the realization of glottalization in normal speakers. *Journal of Phonetics*, *29*(4), 407–429. https://doi.org/10.1006/jpho.2001.0145

Roettger, T., & Gordon, M. (2017). Methodological issues in the study of word stress correlates. *Linguistics Vanguard*, *1*, 20170006. https://doi.org/10.1515/lingvan-2017-0006

Rojo, G. (1990). Frecuencia de fonemas en español actual. In M. Brea & F. Fernández Rei (Eds.), *Homenaxe ó profesor Constantino García* (Vol. 1, pp. 451–467). Universidade de Santiago de Compostela.

Rosado Robledo, L. (2011). Variación fónica. El caso de (b, d, g), (p, t, k) y (ɲ) en el español yucateco. In P. Martín Butragueño (Ed.), *Realismo en el análisis de corpus orales. Primer Coloquio de Cambio y Variación Lingüística* (pp. 147–167). El Colegio de México.

Santos, F. (2018). *`KappaGUI`: An R-Shiny application for calculating Cohen's and Fleiss' kappa* (2.0.2). https://cran.r-project.org/web/packages/KappaGUI/KappaGUI.pdf

Sebastián Gallés, N., Martí Antonín, M. A., Carreiras Valiña, M. F., & Cuetos Vega, F. (2000). *LEXESP. Léxico informatizado del español*. Universitat de Barcelona.

Ségéral, P., & Scheer, T. (2008). Positional factors in Lenition and Fortition. In J. Brandão de Carvalho, T. Scheer, & P. Ségéral (Eds.), *Lenition and fortition* (pp. 131–172). Mouton de Gruyter.

Selkirk, E. O. (1980). *On prosodic structure and its relation to syntactic structure*.

Selkirk, E. O. (1984). *Phonology and syntax: The relation between sound and structure*. The MIT Press.

Selkirk, E. O. (2004). The prosodic structure of function words. In J. J. McCarthy, L. Walsh Dickey, & S. Urbanczyk (Eds.), *Optimality Theory in phonology. A reader* (pp. 464–482). Blackwell. https://doi.org/10.1002/9780470756171.ch25

Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, *25*(2), 193–247. https://doi.org/10.1007/BF01708572

Shue, Y.-L. (2010). *The voice source in speech production: Data, analysis and models*. University of California at Los Angeles.

Shue, Y.-L., Keating, P., Vicenik, C., & Yu, K. (2011). VoiceSauce: A program for voice analysis. *Proceedings of ICPhS 17*, 1846–1849.

Silva-Corvalán, C., & Treffers-Daller, J. (2015a). Digging into dominance: A closer look at language dominance in bilinguals. In C. Silva-Corvalán & J. Treffers-Daller (Eds.), *Language dominance in bilinguals: Issues of measurement and operationalization* (pp. 1–14). Cambridge University Press; Cambridge Core. https://doi.org/10.1017/CBO9781107375345.001

Silva-Corvalán, C., & Treffers-Daller, J. (Eds.). (2015b). *Language dominance in bilinguals: Issues of measurement and operationalization*. Cambridge University Press; Cambridge Core. https://doi.org/10.1017/CBO9781107375345

Silverman, D. (1997). Laryngeal complexity in Otomanguean vowels. *Phonology*, *14*, 235–261.

Simonet, M. (2014). Phonetic consequences of dynamic cross-linguistic interference in proficient bilinguals. *Journal of Phonetics*, *43*, 26–37. https://doi.org/10.1016/j.wocn.2014.01.004

Simonet, M. (2016). The phonetics and phonology of bilingualism. In *Oxford handbooks online*. https://doi.org/10.1093/oxfordhb/9780199935345.013.72

Simonet, M., Hualde, J. I., & Nadeu, M. (2012). Lenition of /d/ in spontaneous Spanish and Catalan. *Proceedings of Interspeech 2012*, 1416–1419. https://www.isca-speech.org/archive/interspeech_2012/i12_1416.html

Skopeteas, S. (2010). Syntax-phonology interface and clitic placement in Mayan languages. In V. Torrens, A. Gavarró, & J. Gutiérrez (Eds.), *Movement and clitics: Adult and child grammar* (pp. 307–331). Cambridge Scholars.

Sobrino Gómez, C. M. (2007). El proceso fonológico de elisión de la segunda vocal en el maya yucateco. *Estudios de Cultura Maya*, *30*, 197–222.

Sobrino Gómez, C. M. (2010). Algunos rasgos tipológicos del idioma maya yucateco. *Escritos Revista Del Centro de Ciencias Del Lenguaje*, *42*, 29–41.

Sobrino Gómez, C. M. (2018). *El sistema fonológico protoyucatecano: Reconstrucción y desarrollo* [El Colegio de México]. https://repositorio.colmex.mx/concern/theses/6q182k57v?f%5Bsubject_sim%5D%5B%5D=Fonolog%C3%ADa&locale=en

Soler, A., & Romero, J. (1999). The role of duration in stop lenition in Spanish. *Proceedings of ICPhS 1999*, 483–486. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS1999/papers/p14_0483.pdf

Solís-Barroso, C., & Stefanich, S. (2019). Measuring language dominance in early Spanish/English bilinguals. *Languages*, *4*(3), 62. https://doi.org/10.3390/languages4030062

Sorensen, T., Hohenstein, S., & Vasishth, S. (2016). Bayesian linear mixed models using Stan: A tutorial for psychologists, linguists, and cognitive scientists. *The Quantitative Methods for Psychology*, *12*(3), 175–200. https://doi.org/10.20982/tqmp.12.3.p175

Sosa, J. M. (1999). *La entonación del español: Su estructura fónica, variabilidad y dialectología*. Cátedra.

Stan Development Team. (2020). *RStan: The R interface to Stan* (2.19.3) [Computer software].

Starr, R. L. (2015). Sweet voice: The role of voice quality in a Japanese feminine style. *Language in Society*, *44*, 1–34. https://www.doi.org/10.1017/S0047404514000724

Stevens, K. N. (1977). Physics of laryngeal behavior and larynx modes. *Phonetica*, *34*, 264–279. https://doi.org/10.1159/000259885

Strobl, C., Hothorn, T., & Zeileis, A. (2009). *Party on! A new, conditional variable importance measure for random forests available in the* `party` *package* (1.3-4) (Technical No. 50; pp. 1–4). Department of Statistics, University of Munich. https://epub.ub.uni-muenchen.de/9387/1/techreport.pdf

Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods*, *14*(4), 323–328.

Strycharczuk, P., & Kohlberger, M. (2016). Resyllabification reconsidered: On the durational properties of word-final /s/ in Spanish. *Journal of the Association for Laboratory Phonology*, *7*(1), 1–24. https://doi.org/10.5334/labphon.5

Stuart-Smith, J. (2007). Empirical evidence for gendered speech production: /S/ in Glaswegian. In J. Cole & J. I. Hualde (Eds.), *Change in phonology: Papers in laboratory phonology* (pp. 65–86). Mouton de Gruyter.

Stuart-Smith, J. (2020). Changing perspectives on /s/ and gender over time in Glasgow. *Linguistics Vanguard*, *6*(1), 1–13. https://doi.org/10.1515/lingvan-2018-0064

Suárez Molina, V. M. (1996). *El español que se habla en Yucatán: Apuntamientos filológicos* (3rd edition, revised and corrected by Miguel Güémez Pineda). Universidad Autónoma de Yucatán.

Tagliamonte, S. A., & Baayen, R. H. (2012). Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation and Change*, *24*(2), 135–178. Cambridge Core. https://doi.org/10.1017/S0954394512000129

Toledo, G. (2007). Fraseo en español peninsular y modelo autosegmental y métrico. *Estudios Filológicos*, *42*, 227–243.

Torreira, F., Simonet, M., & Hualde, J. I. (2014). Quasi-neutralization of stress contrasts in Spanish. *Proceedings of Speech Prosody 2014*, 197–201. https://pure.mpg.de/rest/items/item_1868210/component/file_2030729/content

Trawick, S., & Michnowicz, J. (2019). Glottal insertion before vowel-initial words in the Spanish of Asunción, Paraguay. In G. L. Thompson & S. M. Alvord (Eds.), *Contact, community, and connections. Current approaches to Spanish in multilingual populations* (pp. 147–174). Vernon.

Treffers-Daller, J. (2015). Conclusion. In C. Silva-Corvalán & J. Treffers-Daller (Eds.), *Language dominance in bilinguals: Issues of measurement and operationalization* (pp. 266–268). Cambridge University Press; Cambridge Core. https://doi.org/10.1017/CBO9781107375345.013

Turk, A., Nakai, S., & Sugahara, M. (2006). Acoustic segment durations in prosodic research: A practical guide. In *Methods in empirical prosody research* (pp. 1–28). Mouton de Gruyter.

Uth, M. (2016). Prosodic realization of contrastive focus in Yucatecan Spanish. Spanish dominant and monolingual speakers opposed to balanced bilinguals. *Romanistisches Jahrbuch*, *67*(1), 255–283. https://doi.org/10.1515/roja-2016-0017

Uth, M. (2018). *Language contact as a developmental co-determinant between systems and exemplars. Spanish and Yucatec Maya in Quintana Roo* [Unpublished habilitation treatise].

Vasishth, S., Nicenboim, B., Beckman, M. E., Li, F., & Kong, E. J. (2018). Bayesian data analysis in the phonetic sciences: A tutorial introduction. *Journal of Phonetics*, *71*, 147–161. https://doi.org/10.1016/j.wocn.2018.07.008

Veiga, A., & Arias, Á. (in press). Descripción fonológica de las oclusivas orales. In J. Gil Fernández & J. Llisterri Boix (Eds.), *Fonética y fonología descriptivas de la lengua española*. Georgetown University Press.

Verhoeven, E., & Skopeteas, S. (2015). Licensing focus constructions in Yucatec Maya. *International Journal of American Linguistics*, *81*(1), 1–40. https://doi.org/10.1086/679041

Vigário, M. (2010). Prosodic structure between the prosodic word and the phonological phrase: Recursive nodes or an independent domain? *The Linguistic Review*, *27*(4), 480–530. https://doi.org/10.1515/tlir.2010.017

Wiener, S., Speer, S. R., & Shank, C. (2012). Effects of frequency, repetition and prosodic location on ambiguous Mandarin word production. *Proceedings of Speech Prosody 2012*, 528–531. https://www.isca-speech.org/archive/sp2012/papers/sp12_528.pdf

Wright, M. N., & Ziegler, A. (2017). ranger: A fast implementation of random forests for high dimensional data in C++ and R (0.12.1). *Journal of Statistical Software*, *77*(1), 1–17. https://doi.org/10.18637/jss.v077.i01

Yager, K. (1989). La -m bilabial en posición final absoluta en el español hablado en Mérida, Yucatán (México). *Nueva Revista de Filología Hispánica*, *27*, 83–94. https://doi.org/10.24201/nrfh.v37i1.731

Zimmer, T. (2011). *El español hablado por los afrocostarricenses: Estudio lingüístico y sociolingüístico*. Kassel University Press. http://www.uni-kassel.de/upress/online/frei/978-3-86219-146-8.volltext.frei.pdf