Laryngealization in Upper Necaxa Totonac

by

Rebekka Puderbaugh

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Linguistics
University of Alberta

Examining committee:

Dr. Anja Arnhold, Supervisor
Dr. David Beck, Supervisor
Dr. Benjamin V. Tucker, Examiner
Dr. Stephanie Archer, Examiner
Dr. Ryan Shosted, External examiner
Dr. John Nychka, Pro Dean

Abstract

This dissertation examines laryngealization contrasts in vowels and fricatives in Upper Necaxa Totonac. In vowels the contrast is presumed to be realized as a form of non-modal phonation, while fricatives are supposed to differ according to their production mechanism. The goal of this dissertation is to provide evidence that will help to determine whether the phonetic characteristics of these sounds align with the impressionistic descriptions of their phonological categories.

Laryngealization categories were first examined via a corpus analysis in Chapter 3. The analysis revealed a highly frequent co-occurrence of laryngealized vowels and following glottal stops. No relationship was found between vowel laryngealization and ejective fricatives. In Chapter 4 an analysis of the difference in amplitude between the first and second harmonics (H1-H2) in laryngealized and non-laryngealized vowels showed that H1-H2 values were not influenced by vowel laryngealization categories, but were influenced the presence of a glottal stop following the vowel. This finding suggests that the laryngealization contrast neutralizes in vowels before glottal stops.

In order to consider the potentially glottalic nature of ejective fricatives in UNT, Chapter 5 compared durations of phonetic events that occur during fricative production, including oral closure and frication. Contrary to expectations, ejective fricatives were longer than pulmonic fricatives in overall duration due to longer silent intervals between the end of frication and the onset of vowel phonation. The closure intervals of the ejective fricatives fit nicely into a cross-linguistically attested continuum of decreasing closure duration at places of articulation nearer the back of the oral cavity, suggesting that ejective fricatives may be phonetic clusters in Upper Necaxa Totonac.

Preface

This thesis is an original work by Rebekka Puderbaugh. The research project, of which this thesis is a part, received research ethics approval from the University of Alberta Research Ethics Board, Project Name ”Phonetic documentation of Upper Necaxa Totonac”, No. Pro00032296, 2 July 2012, and Project Name ”Acoustic and Articulatory Documentation of Totonac-Tepehua”, No. Pro00043940, 13 November 2013.

Portions of this thesis have been previously published. Part of the analysis in Chapter 5 appears in Puderbaugh (2015). Parts of the analyses in Chapters 3 and 4 appear in Puderbaugh (2019).

Dedicated to the memory of

Sophia Jex-Blake

Isabel Thorne

Edith Pechey

Matilda Chaplin

Helen Evans

Mary Anderson

Emily Bovell

"[... Y]ou should never fully trust anyone else's description of the sounds of the language you are investigating. They may have been describing a different dialect, or the language might have changed. Or they might have been wrong."

Peter Ladefoged, *Phonetic Data Analysis*

Acknowledgements

The data upon which this thesis is based were collected on field trips supported by a travel grant from the Jacobs Research Fund at the Whatcom Museum in Bellingham, WA, and a grant from the Social Sciences and Humanities Research Council awarded to Dr. David Beck, who provided my first introduction to Upper Necaxa Totonac. This dissertation was only possible because of the generosity and friendship of the people of Patla and Chicontla. I am especially grateful to the Sampayo family, who treated me with kindness and care on both of my data collection field trips, and to all of the speakers who shared their voices with me. Preliminary analyses were reported at conferences where my attendance was supported by the Faculty of Graduate Studies and Research, the Graduate Students Association, and the Department of Linguistics at the University of Alberta, as well as the International Phonetic Association, the Acoustical Society of America, and the Canadian Acoustics Association.

Many people have contributed to the completion of this thesis. To anyone I may have been remiss in omitting, I am grateful for your contributions. I never would have been able to write this thesis were it not for the supervision of Dr. Anja Arnhold, who generously stepped in late in my graduate career and allowed me to pursue the questions that really interested me. I owe a deep debt of gratitude to Dr. Sarah Shulist, who provided many hours of advice and support, to Dr. Carrie Gillon, who coached me through the revision process, and to Chelsea Parlett who provided statistical consultation. Dr. Christina Bjorndahl read drafts of some revised text, gave feedback on typology of phonology systems, and shared her own experiences of spending many years working on her PhD. Dr. Tracy Connor, Dr. Lauren Ackerman, Dr. Michelle Sims and others also shared their own grad school experiences and provided insight into the process that enhanced my own perspectives in invaluable ways.

I will be forever indebted to my colleagues at the University of Edinburgh for giving me the opportunity to work with and among them. Their friendship, support, enthusiasm, and unwavering confidence in me have been miraculous and lifesaving forces over the past year. I am particularly grateful to Professor Nik Gisborne for the many compassionate and supportive pep talks. Thanks go also to Dr. Lauren Hall-Lew, who has been a great friend and source of advice. Dr. Josef Fruehwald was among the first to welcome me to Edinburgh, and the weekly knitting nights with him and Becky Mead and the other southside knitters were a great source of comfort to me. Thank you also to my end-of-grad school compatriots Colleen Patton and Mirjam Eiswirth for your friendship, commiseration, and shared snacks.

I also owe a debt of gratitude to the linguists and academics of Twitter, for being available at all hours of the global day, for laughing at and riffing on the bad jokes I sent into the void, for providing inspiration, solidarity, and advice, and above all for assuring me that I was not alone. There are too many of you to list, and I am certain that I would unintentionally leave someone out if I tried. I am particularly grateful to Kelly Nuttall for always seeming to be there when I needed someone to chat with.

Many others in Edmonton and elsewhere kept me afloat when I felt like all was lost. Dr. Caelan Marrville, my friend and grad school buddy, insisted I could finish when I was certain that I could not and provided unwavering support regardless of the final outcome. My partner, Dr. Nicholas Wakefield, and his family have been a source of strength when my own failed me. Jeff and Amy Nachtigall employed me and kept me well stocked with baked goods when my funding and teaching assignments ran dry. The Naming Commitee of Edmonton gave me a place to feel that my contributions were useful even when the research was at a standstill. Claire Wood and the Edinburgh Graduate Theatre Group provided community in a new country that has already led to many promising collaborations. Thank you all for being there.

Finally, I want to acknowledge the lasting impact that completing this PhD has caused to my physical, mental, and financial health. I hope that those who come after me will not have to endure the same experiences that I have.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

In this thesis, acoustic and statistical methods are used to describe two sets of contrasts in Upper Necaxa Totonac (UNT) (ISO [tku]). These contrasts, between ejective and pulmonic fricatives on the one hand, and laryngealized and modal vowels on the other, have not yet been described with instrumental phonetic methods. While contrastive phonation in vowels is widely attested in many languages of Mesoamerica, including Mixtec (Gerfen & Baker, 2005), Zapotec (Pickett et al., 2010; Silverman et al., 1995; Esposito, 2010a), Trique (DiCanio, 2010, 2011), and Cora (Kim & Valdovinos, 2014), as well as in the Totonacan family (Trechsel & Faber, 1992; Levy, 1987; MacKay, 1994; Mackay & Trechsel, 2015; Brown et al., 2011), ejective fricatives have never been known to occur in any system that does not otherwise make use of the glottalic airstream mechanism, as is reported in UNT (Beck, 2006).

Currently, the differences between pulmonic and ejective fricatives and laryngealized and non-laryngealized vowels in UNT are framed as indicating phonological contrasts between plain and modified segments. Pulmonic fricatives and modal vowels are produced without any modifications from the typical articulatory configurations and may therefore be characterized as simple; laryngealized vowels and ejective fricatives, on the other hand, are complex in that they are produced with the articulatory configurations of the simple segments with the addition of laryngeal tension or glottal closure. Each of these contrasts may alternatively be considered sequences of vowels or fricatives followed by glottal stops, which are highly frequent in UNT. The auditory differences between the plain and modified segments would then be the result of allophonic variation. The data presented in this thesis will be used to argue that the ejective fricatives are phonetic clusters of fricatives and glottal stops, and that at least some laryngealized vowels are the result of coarticulation with glottal stops. The thesis also touches upon the role of phonetics in documentary field linguistics, with particular emphasis on the interplay between phonetic descriptions and phonemic analysis.

Although it is impossible to do scientific work in a vacuum, I have tried as possible to remain as agnostic regarding the phonological identity of the segments and sounds that are investigated in this thesis. Unlike most research in phonetics and phonology, my goal is not to describe the acoustic correlates of a phonological contrast. Instead, I would like to build phonological categories out of collected and recorded phonetic materials. I have not succeeded in this task, or even anything really resembling it, but this position is very important for my interpretation of and conclusions about the data and what the analysis of it shows. This dissertation uses phonetic and computational methods to investigate contrasts that have heretofore largely been approached from impressionistic perspectives. As I am more of a phonetician than I am a Totonaquist, the text that follows is necessarily colored by my perspective as a phonetician. The majority of analyses are based in acoustic phonetics and are therefore of potential interest to phoneticians.

The thesis starts with materials that are currently available, specifically the rich resource that is *The Upper Necaxa Totonac Dictionary*, and proceeds through investigations into previously documented and newly documented patterns of laryngealization in vowels and fricatives with an eye to applying acoustic analysis to them. The data that is used in this study is necessarily specific and primarily of interest to those who are interested in Totonacan languages and their patterns and structures. However, the methods are intended to be examples of generalizable practices that may be applied to any language, whether it is widely analysed or as yet unknown to modern linguistic analysis. The ability to recognize patterns and structures that go unnoticed by many people is a key skill for a linguist, and a particularly difficult part of linguistic description is the recognition of phonetic detail. I would like to contribute in some small way to establishing methods for phonetic analysis that help linguists discover and subsequently make use of phonetic patterns to establish phonological contrasts and all of the larger linguistic structures conveyed through them. While linguistic analysis must begin with the linguist, a human mind, we can nevertheless make clever use of the analytical tools we have built to better manage and uncover the patterns that we seek to describe. It is in that spirit that this dissertation was undertaken. The findings are not earth shattering to phonetics as a discipline, and even within Totonacan linguistics, the contribution is small and potentially controversial. Nevertheless, I hope it is a step toward increasing transparency and reproducibility in language documentation efforts, toward treating languages on their own terms. Unlike phonological categories, phonetic patterns are physically measurable, and therefore definable independent of the people who hear and speak them. Those patterns, if carefully identified and described, will hopefully lead to phonological patterns that are themselves more transparent as well.

The remainder of this chapter first summarizes the sound system of UNT as typically described in the literature in Section 1.1, including information on consonants, vowels, syllables and prosody, then provides some background on ejective fricatives in Section 1.2 and laryngealized vowels in Section 1.3. Details pertaining to UNT are presented at the end of each these sections.

1.1 Phonology of Upper Necaxa Totonac

Upper Necaxa Totonac is a Totonacan language spoken in the villages of Patla, Chicontla, San Pedro Tlalontongo, and Cacahuatlan, which are situated along the banks of the Upper Necaxa River in the Sierra Norte of the Mexican state of Puebla. Approximately 3,400 people speak Upper Necaxa Totonac, a number that is fairly typical of Totonacan language communities, with most consisting of a few thousand speakers or fewer (McGraw, 2009; Lam, 2012). Most speakers of UNT are in their 40s or older, and few members of younger generations are learning the language at home.

Like other Totonacan languages (McQuown, 1990; Kung, 2007; McFarland, 2009), UNT is polysynthetic with nominative-accusative alignment. Constituent order is flexible and may be determined by information structure, but is usually verb initial (Levy & Beck, 2012). In contrast to its morphosyntactic complexity, UNT segmental phonology appears to be relatively straightforward with a typologically common inventory of consonants and vowels (with the exception of ejective fricatives), limited consonant

Table 1.1: UNT consonant inventory in IPA notation, borrowed/marginal phones in parentheses

	Bilabial	Alveolar	Post-alveolar	Lateral	Palatal	Velar	Glottal
Plosive	p	t				k	ʔ
Trill/Tap		(r,ɾ)					
Nasal	m	n				ŋ	
Affricate		ts	tʃ				
Fricative		s s’	ʃ ʃ’	ɬ ɬ’		x	(h)
Approximant	w			l	j		

clusters, and simple syllable structure. Further details on the segmental inventory, syllable structure and prosody are presented in Sections 1.1.1-1.1.3. Unless otherwise noted, descriptions of UNT phonology are drawn from Beck (2004).

1.1.1 Consonants

The consonant inventory of UNT as reported by Beck (2004) is shown in Table 1.1. Apart from the anomalous ejective fricatives, the consonant inventory of UNT resembles not only other Totonacan languages, but also the most common consonant inventories of the 317 languages surveyed in Maddieson (1984). Oral and nasal stops appear at bilabial, alveolar, and velar places of articulation, fricatives at alveolar, post-alveolar, and lateral places, and affricates at alveolar and post-alveolar places.[1] Voicing is not contrastive; all oral stops and fricative segments are phonologically voiceless. UNT differs from other Totonacan languages in the occurrence of ejective fricatives in parallel to the pulmonic fricatives. Alveolar trills and taps occur mainly in Spanish loanwords and are therefore considered marginal as contrastive segments. The velar fricative /x/ is often realized as a glottal fricative [h], with no apparent conditioning environment causing the variation. Glottal stops are considered to be part of the stop series, and a reflex of the Proto-Totonacan segment /*q/. As such, it retains some phonological properties

---

[1] All other Totonac varieties also have the lateral affricate /tɬ/, a distinction that has merged with /ɬ/ in UNT.

Table 1.2: UNT vowel inventory

	Front	Central	Back
High	i iː ḭ ḭː		u uː ṵ ṵː
Mid	e eː ḛ ḛː		o oː o̰ o̰ː
Low		a aː a̰ a̰ː	

of /q/ present in other varieties of Totonac, such as lowering preceding vowels, and triggering place assimilation in preceding nasals (Beck, 2004).

1.1.2 Vowels

The vowel inventory of UNT consists of a symmetrical 5-vowel system including the vowel qualities /a, e, i, o, u/ (Beck, 2004), shown in Table 1.2. The system likely began as a three-vowel system in Proto-Totonacan, where high vowels /i/ and /u/ were allophonically lowered to [e] and [o] in environments adjacent to /q,w,j,l,x/ (MacKay, 1994; Brown et al., 2011). Subsequent loss of some of the conditioning environments above resulted in the current five-vowel system. Each of the five vowel qualities may be contrastively short or long, laryngealized or non-laryngealized (also known as 'plain'), though short vowels are often laryngealized word-finally (Beck, 2004). The quantity distinction does not appear to be the result of any phonological processes and is believed to have been present in Proto-Totonacan (Brown et al., 2011). Garcia-Vega & Tucker (2019) report that long vowel quantity is expressed to increased vowel duration. They also find that short stressed vowels are also longer than short unstressed vowels. No limitations on combinations of vowel length, laryngealization, and stress have been noted in the documentary literature, though Garcia-Vega & Tucker (2019) note some asymmetries in their own data which may be indicative of more general asymmetries. In particular, no tokens of long laryngealized /ṵː,a̰ː,ḛː,o̰ː/ appeared in their data set, perhaps indicating that these particular combinations of qualities and lengths with laryngealization are uncommon in the language overall.

1.1.3 Syllables and prosody

Syllables in UNT, like other Totonacan languages, e.g. Huehuetla Tepehua (Kung, 2007), Filomeno Mata Totonac (McFarland, 2009), Tlachichilco Tepehua (Watters, 1980), and Misantla Totonac (MacKay, 1994), include obligatory onsets and optional codas. In vowel-initial roots and prefixes, the obligatory onset condition is satisfied through the insertion of a glottal stop word initially. Onsets may consist of a single consonant, including affricates and ejective fricatives, or clusters of up to two elements. Onset clusters may consist of a fricative followed by a stop, nasal, or approximant, or a stop followed by an approximant (Kirchner & Varelas, 2002).[2] Optional syllable codas may also contain clusters of up to two elements. Coda clusters may be made up of a stop or a nasal followed by a fricative.[3] Complex segments such as affricates and ejectives may not appear as elements of tautosyllabic clusters, regardless of syllable position. Ejective fricatives and fricative + stop clusters may not appear in syllable codas (Beck, 2004).

Phonemic stress often distinguishes nouns from verbs (Beck, 2004, 2008). In nouns, stress falls on the penultimate syllable, unless the word ends in a long vowel or a closed syllable, in which case the heavy weight draws stress to the final syllable. Verbs are always stress final, except suffixed verbs, which follow the same stress rules as nouns. While cues to stress in UNT are only sparsely documented, Beck (2008) indicates that stressed syllables are longer and louder, and have a “marked pitch contour” when the vowel is long (Beck, 2004, p. 15). Though their results do not speak to the pitch contour, Garcia-Vega & Tucker (2019) report that stressed vowels have higher F0 than unstressed vowels, as well as increased duration in short vowels.

Little is known about prosody and intonational contours in UNT. However, vowel laryngealization and glottal stops may be used to indicate phonological phrase boundaries, as reported in other varieties of Totonac and Tepehua, such as Coatepec Totonac (Levy, 2015; McQuown, 1940, 1990), Filomeno Mata Totonac (McFarland, 2009), and Tlachichilco Tepehua (Watters, 2010). In many cases, word final vowels are devoiced,

---

[2] In practice, only the velar and alveolar stops have been found in stop + approximant clusters. /k/ appears in /kl/, /kw/, and /kɾ/ clusters, while the only such cluster /t/ appears in is /tw/. Note that Kirchner & Varelas (2002) considers the alveolar tap to be a phonological approximant.

[3] Although this structure resembles the production sequence of affricates, the phonemic affricate segments /ts/ and /tʃ/ are notably absent from syllable codas, except in cases of ideophonic sound symbolism.

laryngealized, or dropped entirely. The presence of both phonemic and prosodic laryngealization could result in potentially complex interactions requiring extensive exploration and description to untangle.

1.1.4 Phonotactics

Descriptions of phonotactic patterns in UNT have focused primarily on consonant sequences and clusters. Table 1.3 summarizes the environments in which stops, fricatives and clusters have been reported to occur (Beck, 2006; Kirchner & Varelas, 2002). The table demonstrates the similarity between the distributions of stops and fricatives, generally. In comparison to pulmonic fricatives and stops, ejective fricatives occur in more restricted environments. They do not occur word finally, before pulmonic fricatives, or before stops. However, ejective fricatives occur in precisely the same phonotactic environments as fricative + stop clusters. A major difference between ejective fricatives and clusters is that clusters may resyllabify into sequences that cross syllable boundaries[4], while ejective fricatives occur only in syllable onsets (Beck, 2006). In situations where a fricative-final prefix attaches to a glottal initial stem, the sequence is maintained rather than fusing into a phonetic ejective fricative, as in *kilhhó'x'a'* /kiɬ.ʔó'.ʃ'a̰/ 'one's lips' where the prefix *kilh-* attaches to the stem *hó'x'a'*. In fact, the only cases of reported fricative + glottal stop clusters cross both syllable and morpheme boundaries (Beck, 2006) (c.f. section 1.1.3).

1.2 Ejective fricatives

Ejective fricatives are rarely reported in linguistic speech sound inventories, presumably due to the complex nature of their articulation and the difficulty in generating the airflow necessary for frication from the glottalic airstream mechanism (Maddieson et al., 2001; Shosted & Rose, 2011). They account for only a small proportion (1.52%) of segments reported in the UCLA Phonological Segment Inventory Database (UPSID) and appear in only ten languages, or 2.22%, in comparison to ejective stops in 15.08%

---

[4] Bauer (2015) differentiates between clusters and sequences in precisely this scenario, where the main difference between them is that a cluster occurs within a syllable, while a sequence occurs across the syllable boundary.

*Table 1.3: Phonotactic environments of ʔ and ejective fricatives. ✓ indicates that the segment occurs in the given environment. 'R' indicates that the occurrence is restricted in some way. *Sources differ about whether ejective fricatives occur in these environments; the dictionary data confirms that they do occur, but perhaps only across morpheme or syllable boundaries.*

Context	Oral stops (P)	ʔ	Pulm fricatives (F)	Ej fricatives	Fric + Stop clusters
#_V	✓	✓	✓	✓	✓
V_#	✓	✓	✓	-	-
V_V	✓	✓	✓	✓	✓
F_	✓	✓R	✓R	✓*	✓
_F	✓	✓	✓	-	-
N_V	✓	✓	✓	✓	✓
P_	✓	✓R	✓	✓	✓
_P	✓	✓R	✓	-	-

of languages and ejective affricates in 13.08% (Maddieson & Precoda, 1990). Shosted & Rose (2011) report an additional eight languages with ejective fricatives, including Upper Necaxa Totonac, and a further 13 languages with /s'/ appear in PHOIBLE Online (Moran et al., 2014). Other than UNT, none of these languages has ejective fricatives as the only glottalic segments.

Ejective speech is produced by means of the glottalic airstream; airflow is initiated not by the lungs, but by the manipulation of pressure in the air trapped above the glottis (Catford, 2010; Henton et al., 1992; Maddieson, 2013). In stops, the compression necessary to increase air pressure and allow for egressive airflow is achieved via simultaneous closures at the glottis and an oral place of articulation. While these closures are maintained, the larynx moves upward, increasing air pressure in the newly sealed chamber. After compression, the oral closure is released, followed by release of the glottal closure. Ejective fricatives are complicated by the need for sustained outward airflow in order to produce frication. As a result of these articulatory pressures, ejective fricatives are often preceded by silent intervals indicating glottal closure, as in Kabardian (Gordon & Applebaum, 2006) and Amharic (Demolin, 2002), or flanked by them as in Mehri (Ridouane et al., 2015). Ejective fricatives may also be affricated, as in Tigrinya (Shosted & Rose, 2011) where frication is preceded by an oral closure. Ejective fricatives have been differentiated from pulmonic fricatives by their shorter frication periods in Tlingit (Maddieson et al., 2001), Amharic (Demolin, 2015) and Kabardian (Gordon & Applebaum, 2006), presumably as a result of the limited air supply available for their production. Unlike ejective fricatives in other languages, frication in the ejective fricatives of UNT appeared to be longer than that of pulmonic fricatives in pre-vocalic environments, as well as in fricative + stop clusters (Beck, 2006). Frication in ejective fricatives was followed by a short silent interval; no appreciable preceding closures were noted.

Table 1.4: Syllabification of fricative + stop clusters and ejective fricatives. Examples are presented in practical UNT orthography (in italics) and phonemic notation using IPA symbols. Morpheme boundaries are indicated by '-', syllable boundaries by '.'. Fricative + stop clusters at syllable boundaries syllabify into coda and onset regardless of morphology, but ejective fricatives do not (Beck, 2006). Data were taken from (Beck, 2006, 2011a)

Clusters		**Ejectives**	
Morphology	**Syllabification**	**Morphology**	**Syllabification**
*mí**st**u′* /ˈmistṵ/ ‘cat’	/ˈmis.tṵ/	*hó′**x**ʼa′* /ˈʔo̰ʃʼa̰/ ‘skin’	/ˈʔo̰.ʃʼa̰/
*pá**lhk**a* /ˈpaɬka/ ‘griddle’	/ˈpaɬ.ka/	*pa:lh′á:* /paː-ɬʼaː/ ‘cut open sth′s belly’	/paː.ˈɬʼaː/
*ta**xt**ú* /ta-ʃtuː/ INCHOATIVE-out ‘leave’	/taʃ.ˈtuː/	*a′h**s**ʼawiní′* /a̰ʔ-sʼawi-ˈnḭ/ head-defeat-AGT ‘trickster, deceiver’	/a̰ʔ.sʼa.wi.ˈnḭ/
*ki**lhh**ó′xʼa′* /kiɬ-ʔo̰ʃʼa̰/ mouth-skin ‘one′s lips’	/kiɬ.ˈʔo̰.ʃʼa̰/	*ma:**x**ʼa′he:nín* /maː-ʃʼa̰ʔa-eː-nin/ CLS-shine-CLS-DTRN ‘be illuminated’	/maː.ʃʼa̰.ʔeː.ˈnin/

The phonology of ejective fricatives also plays a role in their identification in UNT. The synchronic distribution of ejective fricatives in UNT resembles those of fricative + stop clusters, which occur only in syllable onset position (Beck, 2006; Kirchner & Varelas, 2002). Clusters may also cross syllable boundaries such that the fricative appears in the coda of the preceding syllable, and the stop appears in the onset of the following syllable. In contrast, the ejective fricatives are reported not to resyllabify as fricative +

glottal stop clusters, remaining instead in the onset of the following syllable according to Beck (2006). Table 1.4 illustrates this syllabification scheme in both mono-morphemic and morphologically complex polysyllabic stems with fricative + consonant clusters and ejective fricatives at word internal syllable boundaries.

Ejective fricatives are believed to have their origins in historical fricative + /q/ clusters. In UNT, all instances of /q/ have shifted to become /ʔ/. Beck (2006) notes that a historical origin of ejective fricatives in fricative + /ʔ/ sequences would predict a parallel distribution between ejective fricatives and clusters, precluding the appearance of phonological evidence to differentiate between them. A very few instances of fricative + /ʔ/ are reported to arise in glottal-initial words preceded by a fricative-final prefix. All of the reported fricative + /ʔ/ clusters are produced by the addition of one of three fricative-final prefixes to stems beginning with /ʔ/: *ix-* /iʃ-/'his/her', *helh-* /ʔeɬ-/ 'mouth (interior)', and *kilh-* /kiɬ-/ 'mouth (exterior)'.

1.3 Laryngealized vowels

Laryngealization contrasts fall along a continuum of phonation from voiceless, produced with a widely spread glottis, through varying degrees of glottal constriction to complete glottal closure (Ladefoged, 1971; Gordon & Ladefoged, 2001; Blankenship, 2002). Vowel phonation contrasts are somewhat rare, occurring in less than 3% of languages in UPSID (Maddieson & Precoda, 1990). Laryngealized vowel contrasts occur in only 4 languages (0.89%) in UPSID. An additional 5 languages contrast breathy and modal vowels (1.11%), and a further 3 languages contrast voiced and voiceless vowels (0.67%). The PHOIBLE Online database includes 18 additional languages with /a̰/ in their vowel inventories (Moran et al., 2014).

Phonation types have highly variable acoustic profiles, with differences depending in part on whether the phonation is contrastive or allophonic, breathy, creaky, or otherwise modified. Spectral and acoustic measures have been found to differentiate between modal and non-modal phonation types in several languages including English, Korean, and Hmong (Garellek, 2010), Mazatec, Mpi, and Chong (Blankenship, 2002), Zapotec (Esposito, 2010b), and Gujarati (Esposito, 2006; Keating & Esposito, 2006), among others.

The difference in amplitude between the first and second harmonics, often reported as H1-H2, or H1*-H2* to indicate adjustment for values of the first formant, appears to be the most reliable measure for differentiating phonation types across languages (Keating et al., 2011; Slifka, 2006), though unadjusted H1-H2 may also be used (Keating & Esposito, 2007). Other measures, such as the difference in amplitude between the first harmonic and the first or higher formant (H1-A1, H1-A2, H1-A3, etc.), cepstral peak prominence, and others have also been used to measure voice quality differences (Keating & Esposito, 2007; Keating et al., 2011).

In addition to spectral differences, the timing of vowel phonation is affected by contrastiveness. Non-modal phonation "lasts longer and is more highly differentiated from modal phonation" when the phonation differences are contrastive than when they are non-contrastive (Blankenship, 2002; Garellek, 2010). In Hupa, laryngealization spreads from following consonants onto a portion of preceding vowels, but never laryngealizes the entire length of the vowel (Gordon, 2001).

Further details on acoustic measures of non-modal phonation and their perceptual relevance are presented in section 4.1.

1.3.1 Laryngealized vowels in Upper Necaxa Totonac

Not much is known about the phonetics of laryngealized vowels in UNT, in part because some speakers seem to laryngealize more clearly and reliably than others (Beck, personal communication). The most comprehensive source of information on the laryngealized vowels is likely the *Upper Necaxa Totonac Dictionary* (Beck, 2011b), which provides transcriptions, definitions, grammatical notes, and morphological, syntactic and dialectal information for 9,000 lexical entries. Laryngealized vowels in UNT are not generally to be considered the result of any phonological processes, though short vowels may laryngealize word finally, and the second person subject marker appears to involve laryngealization, perhaps spreading from the right edge of the word as in other Totonac varieties (Beck, 2004; Watters, 2010; Mackay & Trechsel, 2018).

While laryngealization has not been directly investigated in Upper Necaxa Totonac, Garcia-Vega & Tucker (2019) found some adventitious effects of laryngealization category on suggested acoustic indicators of stress. Their findings suggest that vowel space and fundamental frequency (F0) may be related to the laryngealization contrast. While their primary focus was on verifying potential acoustic indicators of stress (F1-F2 vowel space, duration, and F0) suggested in a grammatical sketch of UNT (Beck, 2004), Garcia-Vega & Tucker (2019) included vowel laryngealization as a categorical predictor in their statistical models. Their analysis suggested that F2 in laryngealized vowels occurring in the first syllable of a word may be higher than modal vowels in the same position, and that vowel duration may be shorter in stressed laryngealized vowels than in stressed non-laryngealized vowels. Garcia-Vega & Tucker (2019) also reported a simple main effect of vowel laryngealization on fundamental frequency, where the measured F0 was lower in laryngealized vowels than non-laryngealized vowels. However, the statistical model of F0 also indicated that vowel laryngealization participated in a significant interaction with the position of the syllable in which it occurs, revealing that vowel laryngealization only influenced F0 when the vowel appeared in initial or medial position, and not word finally. Their findings highlight the need for further phonetic research in this area, of which this dissertation is a part.

1.3.2 Laryngealized vowels in Totonacan

UNT is similar to other languages in the Totonacan family in maintaining a phonation contrast in vowels. In fact, laryngealization has long been of particular interest to linguists working with Totonacan languages and is prevalent in the descriptions of several languages of the family including Zapotitlán de Mendez (Aschmann, 1946), Papantla (Levy, 1987), Misantla (MacKay, 1994), and Filomeno Mata (McFarland, 2009), and Upper Necaxa Totonac (Beck, 2011b), among others. The related Tepehua languages reflect a similar contrast in glottalized consonants that correspond to laryngealized vowels in languages of the Totonac lineage (Watters, 1980, 1988, 2010; Kung, 2007; MacKay & Trechsel, 2013). Historical reconstructions variably support the reconstruction of laryngealized vowels (Brown et al., 2011) or glottalized consonants (Mackay & Trechsel, 2018; MacKay & Trechsel, 2011) in Proto-Totonacan based on these correspondences. The Tepehua and Totonac branches of the Totonacan family tree are differentiated in part based on the temporal location of laryngealization in the syllable: Tepehua languages generally manifest laryngealization in syllable onsets, and Totonac languages realize laryngealization in the syllable nucleus.

There is a strong relationship between laryngealized vowels and glottal stops in Totonacan languages, and they feature prominently in phonological descriptions, though their phonemic status is not always clear, and descriptions may even be unclear about distinctions between descriptions of contrastive laryngelization in vowels and phonetic glottal stops. In Coatepec Totonac, McQuown (1940) analyzed the glottal stop as an independent phoneme, distinct from the vowels adjacent to it. In the Totonac of Zapotitlán de Mendez, Aschmann (1946) instead described glottal stops as integrated with the vocalic nucleus rather than a distinct and separate phonemic category. Arana Osnaya (1953) subsequently pointed out in her reconstruction of Proto-Totonacan, the "special problem" of vowels in Totonacan, referring to inconsistencies in vowel quality, length, and laryngealization across cognate sets that led to difficulties in reconstruction by the comparative method. Glottal stops are themselves often portrayed as a bit problematic for research on Totonacan languages, especially as it relates to reconstruction efforts. (MacKay & Trechsel, 2013) report that synchronic glottal stops derived from glottal stops that were present in Proto-Totonac-Tepehua have a distribution unlike other stop phonemes. Typically, such glottal stops appear at the ends of words or, less commonly, at the beginnings of syllables. Aschmann (1946) also describes the unusual distribution of glottal stops in Zapotitlán de Mendez Totonac, citing the preponderance of consonant clusters that would result from interpreting it as a consonant.

The discussion of glottal stops and their relationship to vowel laryngealization continued with Levy (1987), who reported that in some cases, phonemic vowel laryngealization in Papantla Totonac was indicated entirely by the presence of a phonetic glottal stop without any observed phonetic laryngealization of the vowel at all. Levy also reported that words and phrases were ended with a phonetic glottal stop and suggested that at the end of a word the difference between a non-laryngealized vowel followed by a glottal stop and a laryngealized vowel followed by a glottal stop was audible, though difficult to identify in isolation. The manner of vocal fold vibration and its cessation was also reported to be an indicator of word-final vowel length contrasts: long vowels ended in a brief interval of devoicing, while short vowels ended abruptly, like a gentle glottal stop. The phonetic characteristics of these non-contrastive glottal stops are unclear from Levy's description.

Perhaps related to the close relationship between phonetic glottal stops and laryngealized vowels in Totonacan, descriptions of vowel laryngealization are somewhat varied. Aschmann (1946) notes that glottal stops give preceding vowels their laryngealized character in the Totonac of Zapotitlán de Mendez. Aschmann's portrayal of laryngealized vowels and their conditioning environments is among the most comprehensive in the Totonacan literature. He described three types of laryngealized vowels ranging from a modal vowel followed by a glottal stop to a "completely laryngealized" vowel flanked by glottal stops on either side. Based on a comparison of the distributions of other consonants, Aschmann concludes that interpreting them as a feature of adjacent vowels is preferable to the large number of consonant clusters that would result from interpreting them as part of the stop series. Aschmann finds instead that glottal stops are integral to the vowel nucleus that precedes it, resulting in what are called laryngealized vowels. These vowels are realized in one of three ways:

1. V + ʔ, where the vowel is produced with modal phonation
2. V + ʔ, where the vowel is produced with "a more or less rough glottal vibration" (completely laryngealized, preceded by a voiced continuant, where the consonant is also somewhat laryngealized.
3. ʔ + V + ʔ, with "complete laryngealization" of the vowel

Since Aschmann considered the glottal stop to be part of the vocalic nucleus, it is important to note here that the glottal stops described above are phonetic and do not denote a consonant contrast.[5] The above vowel phonation types were conditioned by surrounding speech, though the second type (a completely laryngealized vowel followed by a glottal stop) appeared to be the most frequent. Following a voiceless consonant, a laryngealized vowel appeared as either type 1, or a "mild form" of type 2, with only slight laryngealization at the end of the vowel. Following another laryngealized vowel, the

---

[5] Also important to note that while they may be phonetic, such a label does not clarify how they are produced or what they sound like. A growing body of research shows that glottal stops rarely involve closure and most often are realized as an interval of creaky voice or other non-modal phonation on surrounding sonorants. See, for example,Whalen (2016).

second vowel is completely laryngealized as in type 2. Following a voiced continuant that is preceded by another laryngealized vowel (i.e. $V_1\ C_{[+voice,+continuant]}\ V_2$), the second vowel is completely laryngealized as well as part of the preceding consonant. In word initial position, a glottal stop is inserted, and the following vowel is completely laryngealized. This does not occur if the word appears phrase medially following a consonant-final word. Modal vowels do not occur in word initial position. Aschmann acknowledged that variation in production is observed even within the given conditioning environments, and that variation between language communities was also common, with the same word being produced in alternate ways from one language to another.

Levy (1987) describes the laryngealized vowels in Papantla Totonac as having an auditorally sharper pitch and less variable vowel quality than non-laryngealized vowels, with non-modal vowel phonation either followed or indicated entirely by a non-phonemic glottal stop. They occur only after stops and affricates Mackay & Trechsel (2018). Levy also reports that laryngealization is also reported to indicate ends of utterances or even fatigue, in addition to its contrastive role. Word finally, Levy reports that both laryngealized and non-laryngealized vowels are followed by (non-phonemic) glottal stops, though the difference between such productions was reportedly difficult to identify and was not described in phonetic terms. Although the contrast is framed as belonging to the laryngealized vowels in Papantla, other sources note realizations of such laryngealization as glottalized or sometimes voiced stops (Alarcón Montero, 2008; García Ramos, 1979; Ramos, 2007). Indeed Garcia Ramos (2007) notes that vowels in Papantla are laryngealized as a result of proximity to glottalized consonants, with laryngealization appearing to move rightward from the consonant onto the vowel.

Other Totonacan languages realize laryngealized vowels in various ways and may not even maintain the distinction. Mackay & Trechsel (2018) report that laryngealization serves a morphological function in both the Totonac and Tepehua branches of the family even though Tepehua languages do not otherwise maintain a laryngealization distinction in vowels. Laryngealized vowels in languages of the Totonac branch may also be limited in the environments in which they occur, with fricatives and sonorants being frequent precursors to laryngealized vowels. In Misantla Totonac, glottal stops inserted at the beginning of otherwise vowel-initial words provoke early laryngealization on the following vowel (MacKay, 1994). A pattern opposite to that described by Aschmann in Zapotitlán de Mendez, where glottal stops induce laryngealization on the preceding vowel.

In Filomeno Mata Totonac the three original vowel qualities /a, i, u/ also bear a contrast in length, but although Filomeno Mata Totonac is reported to have laryngealized vowels, only a few lexical items are distinguished by the vowel laryngealization contrast (McFarland, 2009). Some other varieties of Totonac do not appear to maintain a distinction between laryngealized vowels at all. In Apapantilla Totonac (Reid 1974, Reid & Bishop 1991) vowels contrast in length and quality, but not phonation. In Huehuetla Totonac, vowel laryngealization appears to have been lost entirely (Troiani, 2004, 2007).

In Coatepec Totonac vowel laryngealization is not considered to be contrastive (McQuown, 1940, 1990), but /ʔ/ is phonemic and may also appear in /ʃʔ/ clusters. In this case the phonetic characteristics of /ʔ/ were left largely undescribed apart from the effect of lowering adjacent vowels, and a noted tendency for /ʔ/ to be followed by "an echo vowel" in absolute final position. While such a description does tell us something about the glottal stops, it does not constitute a phonetic description of them.

In languages of the Tepehua branch of the Totonacan language family, vowel laryngealization is not considered contrastive and instead appears allophonically in vowels that are preceded by glottal stops or glottalized stops (MacKay & Trechsel, 2013). This rightward conditioning of vowel laryngealization resembles that of Misantla Totonac, and contrasts with that of Zapotitlán de Mendez. Glottal stops are also reported to trigger the appearance of echo vowels when they are situated in the coda of stressed, phrase-final syllables. Instead of laryngealized vowels, Tepehua languages are reported to contrast between pulmonic and glottalized or glottalic stops (Kung, 2007; Watters, 1988, 1980). Since the Tepehua languages are rather more different from the Totonacan languages than the Totonacan languages are from each other, this dissertation does not touch on the Tepehua languages in any further detail, instead leaving them for future consideration and cross-linguistic comparison.

The variability and overlap of glottal stops with phonation is well known in languages outside of the Totonacan family as well, where they often appear as intervals of non-modal phonation in vowels rather than intervals of closure (Bao, 2009; Esling et al., 2005; Garellek, 2013; Quick, 2003; Whalen, 2016; Elías-Ulloa, 2016). Such variability likely contributes to the difficulty in identifying and classifying glottal stops according to traditional phonological categories or features.

The reported patterns of laryngealization in Upper Necaxa Totonac in some ways resemble those of other Totonacan languages, and in some ways differ. UNT maintains a phonation contrast in vowels, a characteristic that is strong but not universal among Totonac languages. UNT also makes extensive use of glottal stops, both phonemically and prosodically. Whereas some Totonac languages show patterns of allophonic laryngealization beginning with a glottal stop in the onset of a syllable, similar to Tepehua languages which allophonically laryngealize vowels following glottalic stops, others show the opposite pattern. Current descriptions of UNT do not explicitly state which pattern is exemplified by vowel productions.

Though the descriptions summarized here are highly detailed and nuanced, none of them define vowel laryngealization or glottal stops in phonetic terms. The impressionistic descriptions of glottal stops and vowel laryngealization suggest the need for greater clarity and detail in both analysis and terminology. To that end, some phonetic investigations have been performed. Illustrative studies of vowel laryngealization (Trechsel & Faber, 1992; Herrera Zendejas, 2014) and of vowel space more generally (Garcia-Vega & Tucker, 2019) are now available to those interested in Totonac phonetics, but there remains much to do.

1.4 Summary and Research Questions

UNT has been reported to maintain contrasts between two series of fricative and vowel phonemes. These contrasts both involve an altered laryngeal configuration, leading to ejective fricatives contrastive with pulmonic fricatives on the one hand and laryngealized vowels contrasting with modal vowels on the other. In addition to these contrasts, the glottal stops is posited as an independent phoneme belonging to the stop series. The laryngeal contrasts and glottal stops all share similar articulatory mechanisms in that they involve some degree of glottal constriction or closure, likely leading to similarities in acoustic output. Phonological patterns of co-occurrence between fricatives, vowels and glottal stops also play a role in the identification of the reported contrasts, often with appeals to diachronic origins.

There are two main questions to be addressed in this dissertation. The first is whether the ejective fricatives might be better interpreted as clusters of fricatives + glottal stops. The second is whether laryngealized vowels might be allophonically laryngealized as a result of proximity to glottal stops rather than phonemically laryngealized. These questions are addressed first through a corpus study of the *Upper Necaxa Totonac Dictionary*, which reveals patterns of co-occurrence across lexical forms, and then through analyses of certain acoustic aspects of each class of sound. The findings support the analysis of ejective fricatives as clusters and suggest that at least some laryngealized vowels may be the result of coarticulation with glottal stops. The thesis concludes with some general discussion about the usefulness of phonetic research to documentary linguistics generally and especially the establishment of reproducible phonological analyses.

The remainder of this thesis consists of distributional and acoustic descriptions of speech segments in UNT. Chapter 2 summarizes the methods of data collection and annotation that are used in this thesis, along with illustrative spectrograms of target segments. Further details of the methods are presented in each chapter as relevant. Chapter 3 examines patterns of segmental distributions throughout the *Upper Necaxa Totonac Dictionary* (Beck, 2011b), paying particular attention to the patterns of co-occurrence of glottal stops and ejective fricatives with laryngealized and non-laryngealized vowels. Chi-squared tests are used to determine the independence (or lack thereof) between laryngeal features of adjacent segments. Chapter 4 analyzes two acoustic measures known to be associated with vowel phonation cross-linguistically: the difference in amplitude between the first two spectral harmonics, and fundamental frequency. Each analysis is performed separately for two contexts, one where vowels are preceded by consonants, and the other where vowels are followed by consonants. Qualitative predictors reflecting laryngeal category of consonants and vowels, consonant manner, and syllable stress, as well as interactions between such factors are included in the models. Chapter 5 analyzes segmental and subsegmental duration of ejective and pulmonic fricatives in a variety of phonetic environments. Comparisons are made between fricatives in clusters and before vowels. Additional contextual factors such as word position, stress, and adjacent vowel laryngealization are also addressed. Chapter 6 summarizes the findings of the previous three chapters and provides some critical commentary and directions for future research.

Chapter 2

Materials

The majority of the analyses presented in this thesis are based on a collection of acoustic data obtained over the course of two field trips to Patla and Chicontla in September 2012 and January 2014. The present chapter describes how audio data were recorded and subsequently annotated for acoustic analysis based on the phonemic transcriptions available in the *Upper Necaxa Totonac Dictionary* (Beck, 2011b). Sample spectrograms are presented to illustrate the annotation conventions. Further background and methods particular to each chapter will be introduced in situ.

2.1 Segmental corpus analysis

Chapter 3 analyzes segmental collocates of laryngealization in both consonants and vowels. The analysis is based on words and affixes available in the *Upper Necaxa Totonac Dictionary* (Beck, 2011b). All entries were included in the analysis, after conversion to a phonemic transcription. Further details of the data extraction process are presented in Chapter 3.

2.2 Acoustic analysis

This section provides information about word list materials, the speakers who provided the audio data for the current studies, and the method of data collection.

2.2.1 Data collection

Four speakers of Upper Necaxa Totonac (two women, two men) provided the audio data included in the acoustic analyses. Speakers ranged in age from early 30s to about 60 years old. The speakers were all native to Patla and were bilingual in Spanish. All speakers had grown up speaking UNT, with younger speakers being exposed to more Spanish earlier in life. All of the speakers still speak UNT in the community on a daily basis, though Spanish is also very frequently used. Interactions with the author were undertaken in Spanish.

The word list for this study was initially compiled from orthographic forms found in the *Nuevo diccionario del idioma totonaco del Río Necaxa* (*New dictionary of the Upper Necaxa Totonac language*, Beck 2011a), a practical bilingual dictionary compiled for the use of speakers of UNT and derived from the more comprehensive *Upper Necaxa Totonac Dictionary* (Beck, 2011b). The original list of 66 words, collected in the field in 2012, was designed to capture potential variability of the three ejective fricative segments /s’, ʃ’, ɬ’/, including words where ejective fricatives appeared before vowels word initially, and between vowels word medially. In its initial form, the word list was balanced for laryngealization of the following vowel and syllable stress as much as possible. Other characteristics of vowels, such as quality and length were not included in the word list design due to the added complexity including such factors would entail. On a subsequent field trip in 2014, a supplemental list was collected that included words containing pulmonic fricatives in parallel environments to those of the ejective fricatives, in addition to affricates and pulmonic fricatives in fricative + stop clusters to allow for a comparison between these segment types and ejectives. The final word list consisted of 130 word forms, though some forms were not produced by all speakers. The complete wordlist from both field trips appears in Appendix A.

Recordings were made in speakers’ homes using a Marantz portable digital audio recorder (PMD 660) and a head-mounted ear set microphone.[1] By using a head-mounted setup, the distance from the speaker’s mouth to the microphone was kept fairly constant throughout the recording session and across speakers, ensuring reasonable consistency

---

[1] The microphone differed from one field trip to the next, but both microphones were high quality and produced audio that was not judged by the author to be substantially different from each other.

in the audio recordings. All recordings were made at a sample rate of 44.1 kHz with the exception of one, which was made at 96 kHz and subsequently down-sampled to 44 kHz in order to remain consistent with the other recordings.

The word list was not intentionally arranged in any particular order, and all speakers were presented with words in the same list order. The procedure for recording the word list was explained to speakers prior to beginning the recording. Speakers were asked to repeat each word three times within the frame sentence in *ixla wanli' ... chuwa* [ʃla wanlḭ ... tʃuwa] 'he said ... now'. During recording, speakers had visual access to the orthographic form of list items in UNT and written translations in Spanish, both of which were presented on the author's laptop screen. In addition to these written prompts, speakers were also orally prompted with a Spanish translation. No restrictions were imposed on speech rate or speed of moving through the word list. Speakers were encouraged to identify problems or points of confusion with any and all word list items and often took the opportunity to discuss each item as it was recorded. In several cases, alternative words were suggested by the speaker when the item on the word list was unfamiliar to them. Often, the suggested forms were related to the originally proposed item with a different morphology. Alternate word forms were accepted at the time of recording, but phonologically unrelated lexical forms were later excluded from the analysis.

Each elicitation session resulted in a single audio file in WAV (.wav) format. The recordings were allowed to run throughout the session without stopping unless the speaker requested that the recorder be turned off. This practice ensures minimal disruption during the recording sessions, making the speakers more likely to be at ease, as well as capturing any incidental discussion between the researcher, the speaker, and other consultants, where present (Bowern, 2008). Preserving the recording session from start to finish was also found to work as a memory aid for the author, as listening to the recordings provided a re-immersion in the surrounding environment of the recording.

2.2.2 Annotation

Audio files were annotated in their full, unedited forms. First, each file was segmented at the word level using the segmentation mode in ELAN (Wittenburg et al., 2006).

Segments were annotated with the orthographic form of each target word using ELAN's Transcription mode. This process of segmenting and transcribing allowed the original recording session to be maintained intact, while allowing for future avoidance of long gaps in the recording, whether they were full of conversation, background noise, or other material not relevant to the present analyses. The annotations were then exported in the Praat TextGrid format, which served as the basis for annotations of individual phonemic segments.

After the initial word-level segmentation in ELAN, audio data were further segmented and measured using Praat (Boersma & Weenink, 2018). Words were annotated from start to finish, allowing for the contribution of more than one relevant segment per word token to some analyses. In keeping with traditional analyses and to avoid the imposition of unnecessarily subjective judgments about the identity of segments, the audio files were annotated in accordance with the phonemic forms of words as transcribed in the dictionary (Beck, 2011b), rather than the audible impression of each segment. For example, many instances of laryngealized vowels were produced without audible non-modal phonation; nevertheless, all vowels that are orthographically transcribed as laryngealized were also labeled as laryngealized in the TextGrid. Phonemes were labeled with standard IPA notation, based on conventional orthographic transcriptions of UNT. Conversion from orthography to phonemic representation was straightforward as a result of the one-to-one relationship between orthographic and phonemic segments. Further details of the annotation conventions are presented in the subsections below.

The following sections provide descriptions of annotation conventions along with illustrations for each segment type by way of sample spectrograms. All figures in this section represent frequency on the x-axis and time on the y-axis. The frequency range is shown from 0-10000 Hz in order to show differences in the centers of spectral energy during release periods. Higher amplitudes of spectral energy are represented by darker shading in the spectrograms. All of the spectrograms were created from the recording of speaker GMM's speech unless otherwise noted.

Fricatives, oral stops, and affricates Oral stops, affricates, and all fricatives were annotated with a single set of conventions in order to reflect similarities in their production. Each of these segment types involve oral constriction, sometimes resulting in complete closure, and release of that constriction allowing air to flow. In some instances the release was then followed by a further interval of silence or near silence. Thus, fricatives, oral stops and affricates were annotated according to three possible events: closure, release, and lag. *Closure* and *lag* both refer to an interval of silence, where *closure* was defined as silent intervals preceding release, and *lag* was used to refer to silence following release. *Release* indicates the interval where air is flowing and noise is generated.

Stops and affricates typically involved a period of closure, followed by a release. The release of closure was often visible as a brief spike in amplitude in the acoustic waveform, corresponding with broad spectrum energy (a dark vertical line) in the spectrogram. In affricates the release burst was closely followed by a period of turbulent frication noise with no intervening silent period between them. Fricatives were most commonly produced without any preceding closure, though some instances of pre-frication lag were observed in both ejective and non-ejective fricatives. Onset of the fricative release interval typically did not involve a burst, beginning instead with the abrupt onset of sustained turbulent noise at a steady amplitude. After the release, fricatives and affricates were sometimes observed with a lag interval before the onset of vowel phonation.

Closure Figure 2.1 illustrates annotation of pre-release closure in one token of the lateral fricative /ɬ/. Closure intervals were defined as beginning at the end of the second formant in the preceding vowel or sonorant, or the abrupt end of frication noise where applicable, and ending with the onset of broad spectrum energy in the release (burst or frication), or the onset of vowel phonation in cases where the release burst was not apparent. Closure intervals were labeled separately from the release and encoded with the segment label followed by a ’c’ for closure (*e.g.* ‘tc’ indicates the closure of an alveolar stop).

Release Figure 2.2 shows the closure and release intervals of one token of the post-alveolar affricate /tʃ/. Release periods began with a burst, if present, or the onset

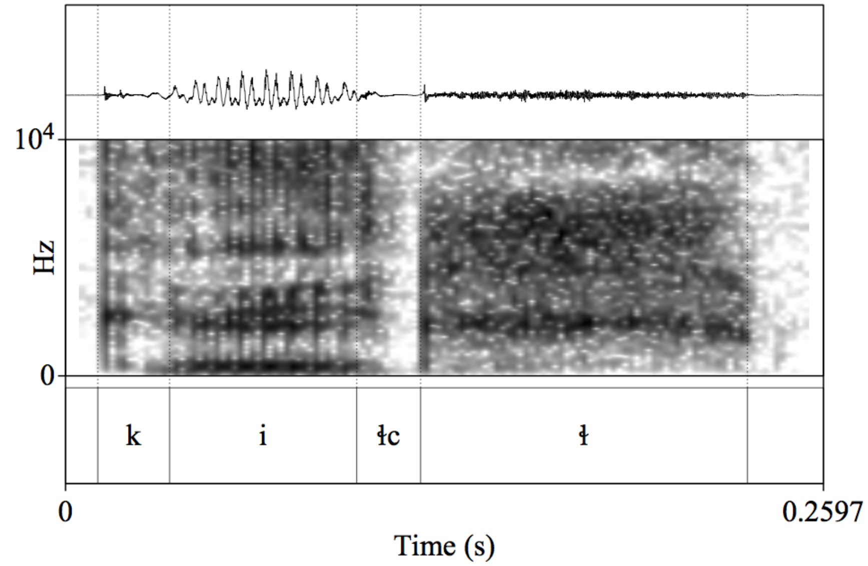


Figure 2.1: Pre-release silence in one token of /ɬ/ from speaker HFM. Detail taken from **kilh**pa'nlhúlu' /kiɬpa̰nˈɬulṵ/ *'jowly, with swollen cheeks'.*

of frication noise, and ended with the end of noise. Release intervals were labeled according to the segmental category, with no further modifications (*e.g.* 't' indicates the release burst of a voiceless alveolar stop). The release portion of affricates and some fricatives was made up of a burst followed by frication. Both the burst and any following frication were considered to be part of the release. As a result, the release was substantially longer in fricatives and affricates than in stops.

Lag Figure 2.3 illustrates an interval of post-release lag in one token of /tʃ/ including non-modal phonation in the following vowel. In many tokens of fricatives and affricates, the release period was followed by a period of silence before the onset of vowel phonation. These lag periods began at the end of frication and continued until the resumption of modal phonation in the following vowel. Lag intervals were often silent, but at times also included low amplitude noise, intermittent bursts of broad spectrum noise, or periods of non-modal phonation. This is addressed further in Section 2.2.2 which describes the annotation of glottal stops and relates them to the lag intervals described here. Lags were labeled with the symbol indicating the consonant phoneme followed by a '-'.

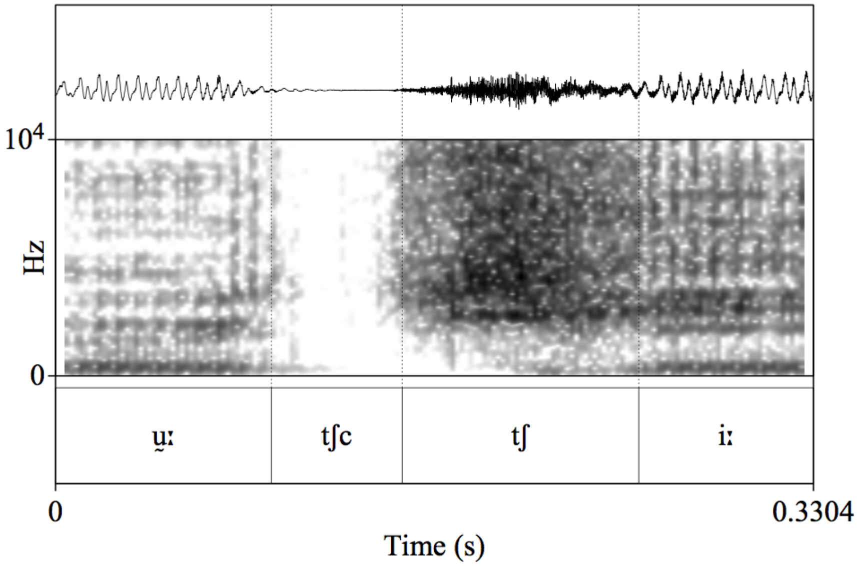


*Figure 2.2: /*tʃ*/ segment produced by speaker HFM. Detail from* a:'**tu:'chi:**yé:klh /a̰ːtuːtʃiːˈjeːkɬ/ *'mint'.*

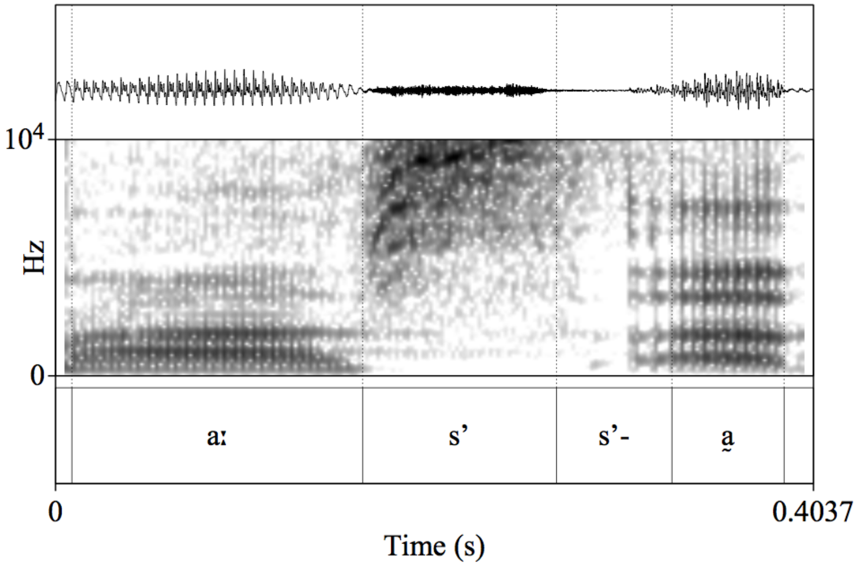
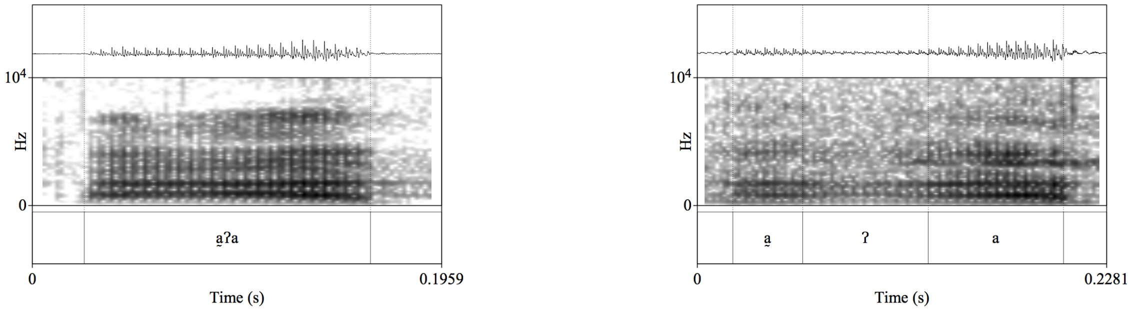


*Figure 2.3: Post-frication lag in the production of an /*s’*/ produced by speaker GMM. Detail taken from* **ma:s**’**a’**ta:nán /maːs’a̰taːˈnan/ *'raise children'.*

Glottal stops Glottal stops are highly variable in their production (see Section 1.3), and therefore do not present a predictable acoustic profile that lends itself easily to a single set of annotation conventions. In the present data, glottal stops were produced in several forms, a selection of which are presented here. There was no obvious pattern in the data that predicted which form a glottal stop would take. Segments were labeled as glottal stops only if they were represented in the orthography by the character <h>. In addition to orthographic glottal stops, the data include several cases of phonetic segments acoustically reminiscent of glottal stops that were not identified in transcriptions. Such segments have not been included in the present analysis.

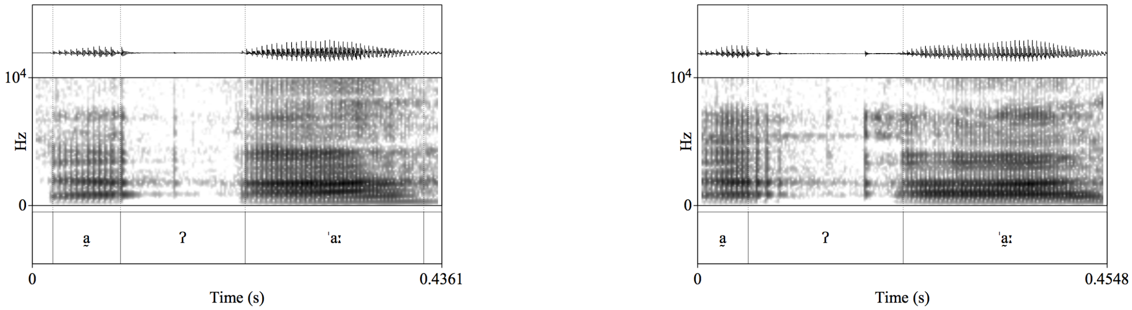
Figure 2.4 presents some of the possible acoustic realizations of glottal stops between vowels and sonorants. While these examples are not necessarily exhaustive, they represent some of the most frequent productions and illustrate the complicated nature of such highly variable segments. Figure 2.4a shows an example of a glottal stop produced as non-modal phonation that persists throughout the duration of surrounding vowels without any visible boundaries between them. Figure 2.4b shows an example of a glottal stop realized as a period of non-modal phonation with decreased amplitude relative to surrounding vowels of higher amplitude. Figure 2.4c shows a glottal stop produced as a silent period interrupted by a brief glottal pulse with little or no notable changes in phonation of surrounding vowels. Figure 2.4d shows a brief period of non-modal phonation at the end of the preceding vowel, followed by a period of relative silence and little or no non-modal phonation in the following vowel. Figure 2.4e shows a glottal stop produced as a silent period punctuated by multiple glottal bursts and both preceded and followed by non-modal phonation in surrounding sonorants. Figure 2.4f shows a period of relative silence with continuous low amplitude energy reminiscent of a voice bar corresponding to the transcribed location of a glottal stop.

Glottal stops sometimes also appear adjacent to fricatives across syllable and morpheme boundaries, as illustrated in Figures 2.5 and 2.6. Such instances of glottal stops were much more uniform in their appearance, always occurring with a substantial period of silence, and often being flanked at start and end with bursts representing the impact of glottal closure and the subsequent reopening of the glottis. Visual comparison with the spectrograms of ejective fricatives, such as those in Figure 2.3 reveals similari-



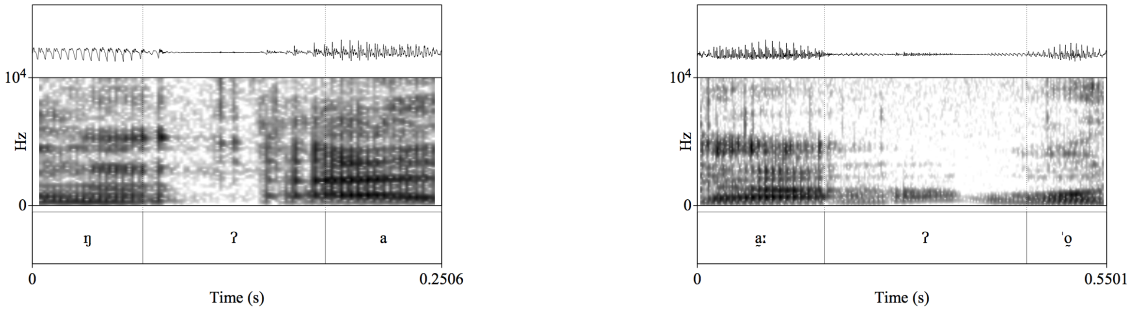
(a) Consistent non-modal phonation throughout a̰ʔa *sequence. Detail from* **a'h**atapu:x'á:ya' /a̰ʔatapuːˈʃ'aːja̰/ *'soot from cooking fires built-up on spiderwebs on the roof of a house'*

(b) Non-modal phonation with visibly reduced amplitude. Detail from **a'ha**tapu:x'á:ya' /a̰ʔatapuːˈʃ'aːja̰/ *'soot from cooking fires built-up on spiderwebs on the roof of a house'*



(c) Strong glottal closure at end of preceding vowel; modal onset of following vowel. Detail from tanhe:x'**a'há:** /tanʔeː-ʃ'a̰ˈʔaː/ *'begin to lighten on the horizon at dawn'*

(d) Brief non-modal phonation at end of preceding vowel, and modal onset of following vowel. Detail from ka:'x'**a'há:h** ka̰ːʃ'a̰ˈʔa̰ːʔ *'bright (place)'*



(e) Clearly delimited silent interval flanked by non-modal phonation. Detail from he**nha**li:s'óli' /henʔaliːˈs'o̰lḭ/ *'elephant'*

*(f) Sustained phonation, but visibly reduced amplitude. Detail from ch****a:'hó****x'a'* tʃa̰ːˈʔoʃ'a̰ *'tree bark'*

Figure 2.4: Variable realizations of glottal stop

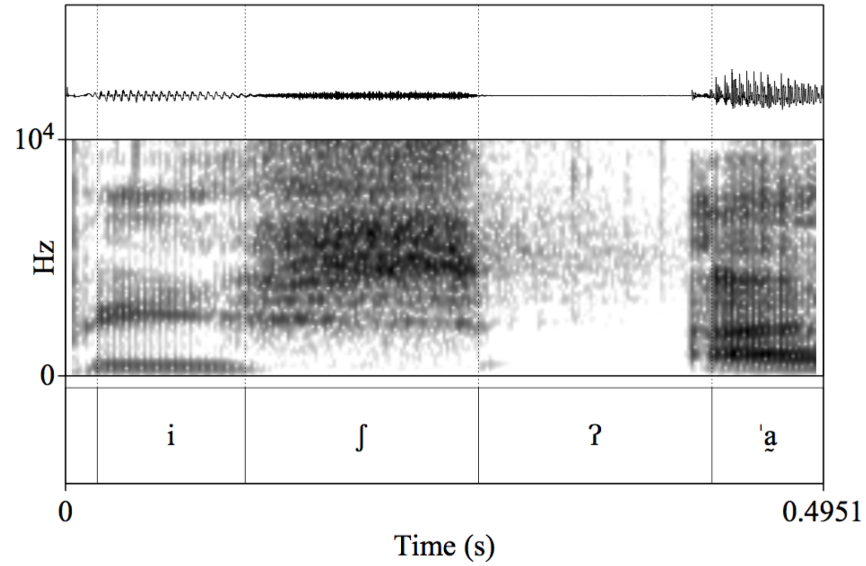


Figure 2.5: [ʃʔ] *sequence across a morpheme boundary. Detail from* pixhá'lha' [piʃ-ˈʔa̰ɬa̰] *'large (bunch or bouquet)'*

ties between these heterosyllabic/heteromorphemic fricative + glottal stop clusters and ejective fricatives.

Vowels and other sonorants Vowel segmentation relied on the cues of surrounding consonants. Generally vowels were considered to be portions of speech with clear formant structure apparent in the spectrogram. Vowel boundaries were sometimes determined by the presence or absence of the second formant, especially in relation to the onset of stop closure (see 2.2.2). In the case of vowels adjacent to glottal stops, shown in Figure 2.4, the segmentation boundary was placed at the onset or offset of more-or-less modal phonation in the vowel. Amplitude of the waveform was also used as an indication of the transition between vowel and glottal stop, with a higher amplitude indicating vowels rather than glottal stops.

Sonorant consonants do not play a role in the analyses presented in this dissertation, but they were nevertheless segmented in words where they occurred. Boundaries between vowel-sonorant sequences were placed at approximately the midway point of the transition from one oral configuration to the next. Vowel-nasal sequences were

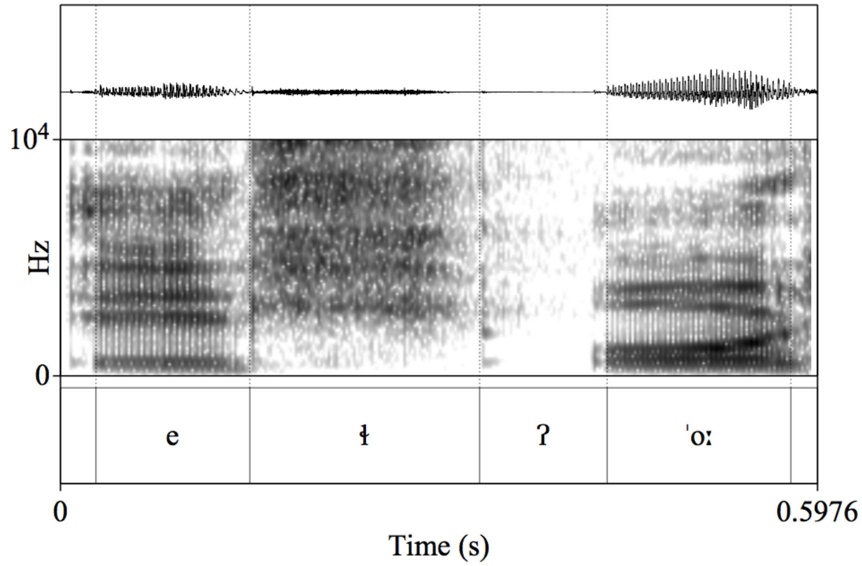


Figure 2.6: [ɬʔ] *sequence across a morpheme boundary. Detail from* helhhó:x'a' /ʔeɬ-ʔoːʃ'a̰/ *'one's lips'*

somewhat easier to segment due to the sharp transition into the period of damped energy and anti-formant structure of nasals when compared to surrounding vowels.

Chapter 3

Categorical laryngealization in adjacent segments

The present chapter provides a quantitative description and analysis of the segmental contexts in which laryngealization occurs in *The Upper Necaxa Totonac Dictionary* (Beck, 2011b). The analysis tests the hypothesis that laryngealized segments are more likely to occur adjacent to other laryngealized segments. Findings reveal interdependence between segmental laryngealization and laryngealization of preceding and following contexts, particularly between vowels and following glottal stops. The regular co-occurrence of laryngealization in adjacent segments suggests a redundancy in the phonemic representation that could be resolved by retaining the laryngeal feature in only one of the segments in the sequence and deriving the laryngealization of the other segment through traditional allophonic rules.

3.1 Introduction

Laryngealization is considered to be a contrastive characteristic of the vowel system of UNT (Beck, 2004). Consonants have not been described as maintaining laryngealization contrasts, but if we consider that glottal stops and ejective fricatives are both produced with laryngeal tension, perhaps we can consider these segments to be laryngealized as well. If laryngealization is contrastive in both consonants and vowels, there ought not be a strong correspondence between vowel laryngealization and laryngealization of surrounding consonants. On the other hand, a high correspondence between laryngealized vowels and laryngealized consonants might indicate an allophonic alternation that has heretofore been overlooked.

The main aim of this chapter is to explore the general patterns that arise between consonants and vowels with respect to laryngealization. Collocational patterns of laryngealization in both laryngealized and non-laryngealized environments are explored using methods inspired by corpus linguistics. Basic descriptions of other collocational patterns in the data are also presented. The discussion relates the findings discovered here to the following two chapters investigating acoustic dimensions of laryngealization in vowels in laryngeal and non-laryngeal contexts.

3.2 Methods

The materials for the present analysis were extracted from the digital version of the *Upper Necaxa Totonac Dictionary* (Beck, 2011a). All forms were included in the analysis, including head words, inflected forms, and affixes that received their own entries. For a first foray into quantitative phonotactic study of laryngealization, the *Dictionary* is good source material because it broadly represents the language as it has been encountered to date and serves as a transparent resource for future researchers to refer to. While it is possible that certain sequences may be over-represented in a corpus built from dictionary forms, the subject matter and vocabulary in the *Dictionary* are broad and not limited by style or topics as a corpus of stories or conversations would be. Potentially over-represented sequences are discussed where relevant below. The resulting corpus is the largest and most comprehensive dataset available for studying the segmental collocation patterns in UNT.

The data required little pre-processing before inclusion in the analysis because the orthography of UNT is transparent at the phonemic level. Dictionary forms were copied into a plaintext file format, then transliterated into IPA using grep and regex (regular expressions) find-and-replace methods. Following a method similar to Bauer (2015), transcriptions were accepted as given in the published dictionary. Because some segments were encoded by complex character sequences (i.e affricate digraphs, ejective fricatives, vowels with length, stress and laryngealization diacritics), phonemic symbols were delimited by inserted spaces on either side. Segment sequences such as clusters were therefore split into component segments by these spaces, while complex segments such as /tʃ/ were retained as a unit without a space between the two characters of the digraph. Word boundary markers (#) were inserted at the end of each word to ensure that segments were not interpreted as adjacent across word boundaries. The resulting list of prepared word data was then converted into a single text string from which immediately preceding and following segments were identified for all segments using the shift function from the data.table package (Dowle & Srinivasan, 2017) in R (R Core Team, 2017).

A single data frame was constructed from this collocational data, which was further classified according to the segment and the context in which it occurred. Each segment and its collocates were classified according to twenty binary identifiers loosely analogous to phonological features. These binary features allowed segments and their contexts to be separated into subsets on the basis of one or more features. Features were defined for each segment by way of a grid (see Appendix B), which was then merged with the segmental data sets in R. Three classes of segments (stops, fricatives, and vowels) and their contexts are analyzed here. Each segment class was divided into a set with laryngeal constriction and one without laryngeal constriction. In this way, fricatives were divided into ejective and pulmonic categories; vowels were divided into laryngealized and non-laryngealized, and stops were divided into oral (p,t,k) and glottal (ʔ). Cross-tabulations of these binary features resulted in the count data presented in the following section. Chi-squared tests were performed on these count data in order to determine whether laryngealization of context segments was related to laryngealization of target segments.

3.3 Results

This section presents the results of the analyses described above. First, distributions of preceding and following segmental collocates are described for stops, fricatives, and vowels. These distributions are then analyzed with chi-squared tests to determine whether there is a relationship between segmental and contextual laryngealization.

3.3.1 Segmental collocates

A general summary of the segmental collocates of stops, fricatives and vowels is reported here. Each segment class is further divided into member segments for illustrative purposes. Preceding and following contexts for all segments have been divided into the 6 categories: laryngealized vowels, non-laryngealized vowels, stops, fricatives, other consonants (including affricates, nasals, and other sonorants), and word boundaries. Laryngealization will be further addressed in the statistical analysis. The data are presented as raw counts as well as percentages of the total number of tokens per segment category.

Stops

A total of 15705 stop tokens occurred in the dictionary corpus. These tokens were distributed among the segment types as follows: /p/ segments 18%, /t/ segments 31.6%, /k/ segments 24.2%, and /ʔ/ segments 26.1%. Less than 0.04% of stops were represented by the marginal phonemic segments /b/ and /d/, which have been excluded due to their small numbers (6 tokens in total). The remaining 15699 tokens were included in the present analysis.

Table 3.1 presents a summary of stops and their preceding environments. The most striking finding here is that 72.94% of /ʔ/ tokens occur after laryngealized vowels. Closer inspection of glottal stop tokens revealed that some of these collocations were due to frequent meronymic prefixes made up of laryngealized vowels followed by glottal stops. Table 3.2 summarizes the count data for each of these prefixes in the dictionary. In total, these prefixes accounted for 1471, or 36% of instances of glottal stops after laryngealized vowels. In order to investigate whether these instances might bias the overall result, these 1471 items were removed, leaving 1523 instances of glottal stops preceded by laryngealized vowels out of 2634 instances of glottal stops in all environments, or approximately 58%. Removing the frequent prefix tokens to perform a more conservative analysis did not affect the overall results of the statistical tests in the present chapter. Therefore, the analyses presented below include all of the dictionary data.

Table 3.1: Counts of preceding environments of stop segments in Beck 2011b. 5 /b/ tokens and 1 /d/ token have been excluded. Percentages were calculated by segment (column).

Preceding context	p	t	k	ʔ	Total
v̰	268 (9.47%)	626 (12.61%)	715 (18.82%)	2994 (72.94%)	4603 (29.32%)
v	940 (33.22%)	1240 (24.97%)	1705 (44.88%)	406 (9.89%)	4291 (27.33%)
stops	226 (7.99%)	165 (3.32%)	46 (1.21%)	13 (0.32%)	450 (2.87%)
fricatives	327 (11.55%)	1326 (26.71%)	511 (13.45%)	12 (0.29%)	2176 (13.86%)
other consonants	131 (4.63%)	255 (5.14%)	228 (6%)	287 (6.99%)	901 (5.74%)
#_	938 (33.14%)	1353 (27.25%)	594 (15.64%)	393 (9.57%)	3278 (20.88%)
Total	2830	4965	3799	4105	15699

Table 3.2: Summary of highly frequent prefixes incorporating laryngealized vowels followed by glottal stops. (Note: ALN *= alienative,* STM *= stimulus Beck 2011b.)*

Orthography	IPA	Gloss	Count
a'h-	a̰ʔ-	'head'	493
a'ha-	a̰ʔa-	'ear'	714
la'ha-	la̰ʔa-	'face'	264
ma'h-	ma̰ʔ-	'ALN'	184
ma'ha-	ma̰ʔa-	'hand'/'STM'	112
		Total	1471

Table 3.3 summarizes following contexts of stop tokens. Here, both /p/ and /t/ occur before non-laryngealized vowels in more than half of their tokens, with /k/ and /ʔ/ also occurring more often before non-laryngealized vowels than before any other segment. In most cases, the remaining tokens occur mainly before laryngealized vowels,

resulting in the vast majority of stops occurring before vowels, reflecting the preferred CV structure of syllables.

Table 3.3: Counts of following environments of stop segments in Beck (2011b). 5 /b/ tokens and 1 /d/ token have been excluded. Percentages were calculated by column.

Following context	p	t	k	ʔ	Total
V̰	810 (28.62%)	1235 (24.87%)	837 (22.03%)	543 (13.23%)	3425 (21.82%)
V	1931 (68.23%)	3068 (61.79%)	1795 (47.25%)	1876 (45.7%)	8670 (55.23%)
stops	1 (0.04%)	20 (0.4%)	156 (4.11%)	273 (6.65%)	450 (2.87%)
fricatives	51 (1.8%)	25 (0.5%)	546 (14.37%)	686 (16.71%)	1308 (8.33%)
other consonants	8 (0.28%)	130 (2.62%)	334 (8.79%)	489 (11.91%)	961 (6.12%)
_#	29 (1.02%)	487 (9.81%)	131 (3.45%)	238 (5.8%)	885 (5.64%)
Total	2830	4965	3799	4105	15699

Fricatives

Fricative segments occurred 6292 times in the *UNT Dictionary*, about 93% of which were pulmonic. Approximately 30.21% were /s/ tokens, 31.04% were /ʃ/, and 31.8% were /ɬ/. Ejective fricatives make up 6.93% of fricative tokens in the dictionary, with /s'/ accounting for nearly half of these (2.97%). The post-alveolar /ʃ'/ made up 1.95% of fricative tokens, and /ɬ'/ accounted for the remaining 2%.

Table 3.4 summarizes preceding contexts for each of the 6 fricative categories. Ejective fricatives occur word initially in higher proportions than pulmonic fricatives. This is likely due to the high percentage of ejectives occurring word initially. Except for alveolar segments, ejective fricatives occur after modal vowels with lower frequency than pulmonic fricatives. Vowels are the most frequent preceding environment overall, though somewhat less frequent for ejective fricatives than pulmonic fricatives. A similar pattern is also apparent for laryngealized vowel contexts, with the exception of the post-alveolar segment. Ejective and pulmonic fricatives occur with similar frequencies after stops, with the exception of /ɬ/, which has a relatively low incidence of stops in its preceding environment. /ɬ/ also differs from other fricatives in occurring after modal vowels at a higher rate than any other fricative. It also has a reduced rate of occurrence word initially and following stops.

Table 3.4: Proportions of preceding environments for fricative segments in Beck (2011b). Percentages are calculated by segment (column).

Preceding Context	s	s’	ʃ	ʃ’	ɬ	ɬ’	**Total**
v̰	166 (8.73%)	3 (1.6%)	222 (11.37%)	16 (13.01%)	231 (11.54%)	10 (7.94%)	648 (10.30%)
v	680 (35.77%)	73 (39.04%)	825 (42.24%)	32 (26.02%)	1057 (52.8%)	33 (26.19%)	2700 (42.91%)
stop	442 (23.25%)	43 (22.99%)	409 (20.94%)	30 (24.39%)	348 (17.38%)	36 (28.57%)	1308 (20.79%)
fricative	59 (3.1%)	10 (5.35%)	46 (2.36%)	2 (1.63%)	4 (0.2%)	0	121 (1.92%)
other Cs	75 (3.95%)	1 (0.53%)	87 (4.45%)	1 (0.81%)	92 (4.6%)	1 (0.79%)	257 (4.08%)
#_	479 (25.2%)	57 (30.48%)	364 (18.64%)	42 (34.15%)	270 (13.49%)	46 (36.51%)	1258 (19.99%)
Total	1901	187	1953	123	2002	126	6292

Table 3.5 summarizes following contexts of fricative segments. Because ejective fricatives do not occur at the ends of words or as the first element in consonant clusters, their following environments only include vowels (cf. 1.1.4). Their occurrences are rather evenly distributed across laryngeal categories in vowels, though higher proportions of

all three ejective fricatives (all over 50%) occur before non-laryngealized vowels, which are of course more frequent than laryngealized vowels. In comparison, a relatively small proportion of pulmonic fricatives occurs before vowels (approximately 26-27% across vowel laryngealization categories in all three pulmonic fricatives), due in part to the greater number of possible contexts. A fairly high proportion of pulmonic fricatives occur before stops (/s/ 40.77%, /ʃ/ 41.07%, /ɬ/ 29.92%).

Table 3.5: Proportions of following environments for fricative segments in Beck (2011b).

Following Context	s	s’	ʃ	ʃ’	ɬ	ɬ’	**Total**
v̰	207 (10.89%)	85 (45.45%)	204 (10.45%)	55 (44.72%)	229 (11.44%)	53 (42.06%)	833 (13.24%)
v	311 (16.36%)	102 (54.55%)	342 (17.51%)	68 (55.28%)	310 (15.48%)	73 (57.94%)	1206 (19.17%)
stop	775 (40.77%)	0	802 (41.07%)	0	599 (29.92%)	0	2176 (34.58%)
fricative	0	0	5 (0.26%)	0	116 (5.79%)	0	121 (1.92%)
other consonants	442 (23.25%)	0	410 (20.99%)	0	383 (19.13%)	0	1235 (19.63%)
_#	166 (8.73%)	0	190 (9.73%)	0	365 (18.23%)	0	721 (11.46%)
Total	1901	187	1953	123	2002	126	6292

Vowels

28,966 vowel tokens were tallied in the *UNT Dictionary* corpus. Approximately 68% of these vowels were non-laryngealized, and 32% laryngealized. Table 3.6 summarizes the distributions of occurrences of laryngealized and non-laryngealized vowels across preceding contexts. Distributions of both laryngealized and non-laryngealized vowels

Table 3.6: Proportions of preceding environments for vocalic segments in Beck (2011b). Percentages are calculated by column.

Preceding context	v̰	v	**Total**
v̰	9 (0.1%)	6 (0.03%)	15 (0.5%)
v	210 (2.27%)	53 (0.27%)	263 (0.91%)
oral stop	2882 (31.14%)	6800 (34.50%)	9682 (33.43%)
glottal stop	543 (5.87%)	1876 (9.52%)	2419 (8.35%)
ejective fricative	193 (2.09%)	243 (1.23%)	436 (1.51%)
pulmonic fricative	640 (6.92%)	963 (4.89%)	1603 (5.53%)
other consonant	4197 (45.35%)	9718 (49.3%)	13915 (48.04%)
#_	580 (6.27%)	53 (0.27%)	633 (2.19%)
Total	9254	19712	28966

are fairly similar across context types. Nevertheless, laryngealized vowels tended to occur with somewhat lower frequency in all environments except word initially and immediately following another non-laryngealized vowel.

Table 3.7 summarizes the distributions of occurrences of laryngealized and non-laryngealized vowels across following contexts. Unlike the preceding contexts, the following contexts differ quite drastically from one another. Nearly a third of laryngealized vowels occurred before a glottal stop, compared to only 2% of non-laryngealized vowels. Laryngealized vowels also occurred more frequently in word final position compared to non-laryngealized vowels, although the proportional difference was only

Table 3.7: Proportions of following environments for vocalic segments in Beck (2011b). Percentages are calculated by column.

Following context	v̰	v	**Total**
v̰	9 (0.1%)	210 (1.07%)	219 (0.76%)
v	6 (0.06%)	53 (0.27%)	59 (0.20%)
oral stop	1609 (17.38%)	3889 (19.73%)	5498 (18.98%)
glottal stop	2994 (32.36%)	406 (2.06%)	3400 (11.74%)
ejective fricative	29 (0.31%)	138 (0.7%)	167 (0.58%)
pulmonic fricative	619 (6.69%)	2562 (13%)	3181 (10.98%)
other consonant	2085 (22.53%)	9498 (48.18%)	11583 (39.99%)
_#	1903 (20.56%)	2956 (15%)	4859 (16.77%)
Total	9254	19712	28966

about 5%. Only 22% of laryngealized vowels occurred before other consonants, far fewer than the 48% of non-laryngealized vowels.

3.3.2 Laryngealization collocates

The previous section provided an overview of segmental distributions with respect to their immediately preceding and following segmental collocates. The current section will investigate the same segmental categories by focusing only on laryngealization values in adjacent segments. Chi-squared tests of independence are used here to determine whether the occurrence of laryngealized segments is significantly related

Table 3.8: Chi-squared residuals of segment and preceding context laryngealization for stops, fricatives, and vowels, as well as all three segment classes together. '+' indicates counts greater than expected, '-' indicates counts smaller than expected. Residuals greater than +/-2 are generally considered to differ significantly from expected values. Compare to B.2.

Preceding Context		Non-laryng	Laryng	χ^2	df	p
Segment Class						
Stops	oral	+21.13	-26.29			
	glottal	-32.37	+40.29	3808.9	1	< 0.0001
Fricatives	pulmonic	+0.02	-0.03			
	ejective	-0.08	+0.13	0.03	1	0.87
Vowels	-laryng	-1.29	+3.89			
	+laryng	+1.88	-5.68	52.57	1	< 0.0001
All	-laryng	+6.23	-12.44			
	+laryng	-13.10	+26.16	1049.3	1	< 0.0001

to the laryngeal contexts in which they appear. Results are reported in terms of the χ^2 value as well as the size of the Pearson residuals of each cell, following (Arppe, 2009). The residuals represent how much the observed counts differed from what would be expected if there were no relationship between cells and were calculated using the *chisq.residuals* function from the questionr package (Barnier et al., 2017). Residuals greater than |2| indicate significant contributions to the overall χ^2 value. The raw count data are presented in Appendix B. Yates continuity correction is not used, due to the large number of observations (Yates, 1934). Tests are performed on the entire dictionary corpus, as well as separately on three subsets of the data: stops, fricatives and vowels. Subsequent tests were performed that excluded word boundaries, which were arbitrarily classified as "non-laryngealized", as well as consonant contexts of stops and fricatives. Removing these contexts resulted in a loss of significance in the chi-squared test of laryngealization between fricatives and following environments.

In preceding contexts, summarized in Table 3.8, stops whose laryngealization matched that of their context occurred with higher frequency than would be expected if there were no relationship between laryngealization of segments and preceding contexts. Stops that opposed their context's laryngealization were less frequent than expected. Fricative laryngealization was independent of context laryngealization. Vowels match-

Table 3.9: Chi-squared residuals of segment and following context laryngealization for stops, fricatives, and vowels, as well as all three segment classes together. '+' indicates counts greater than expected, '-' indicates counts smaller than expected. Residuals greater than +/-2 are generally considered to differ significantly from expected values. Compare to B.3.

Following Context		Non-laryng	Laryng	χ^2	df	p
Segment class						
Stops	oral	-3.06	+5.67			
	glottal	+5.15	-9.53	158.77	1	<0.0001
Fricatives	pulmonic	+1.88	-4.73			
	ejective	-6.89	+17.34	373.94	1	<0.0001
Vowels	-laryng	+13.92	-35.90			
	+laryng	-20.32	+52.40	4641.8	1	<0.0001
All	-laryng	+6.15	-13.22			
	+laryng	-13.22	+28.41	1194.3	1	<0.0001

ing the laryngealization of their preceding segment were less frequent than expected, and vowels that did not match were more frequent.

Table 3.9 summarizes the results of the analysis of following contexts. Here, stops that did not match the laryngealization of their following context were more frequent than expected, while stops that did match were less frequent. Fricatives, vowels, and all segments tallied together showed the opposite pattern, with higher rates of segments occurring before contexts that matched their own laryngealization, and lower rates of mismatched laryngealization than expected.

Table 3.10: Chi-squared residuals of segment and following context laryngealization for stops and fricatives after vowels (VC sequences).

Preceding Context		Non-laryng	Laryng	χ^2	df	p
Segment class						
Stops	oral	+23.98	-23.16			
	glottal	-30.49	+29.45	2908.3	1	<0.0001
Fricatives	pulmonic	-0.07	+0.13			
	ejective	+0.29	-0.58	0.45	1	0.50

The above analyses considered laryngealization in adjacent contexts for all segments in the dictionary. However, there are many instances in which stops and fricatives occur in clusters with oral stops, nasals and liquids that may have an effect on the outcome of the chi-squared tests above. In such clusters, the resulting sequence is made up of two non-laryngealized segments. These sequences of matching laryngealization will contribute to the relationship between laryngealization values in the complete dictionary data set. Additionally, contexts of ejective fricatives are limited in that they do not appear word-finally. In order to further investigate the relationship between laryngealization in consonants and vowels, the data set was limited to only CV or VC sequences where the consonants could be only stops or fricatives. After creating this smaller data set, chi-squared tests were again performed to test the relationships between laryngealization in adjacent segments.

Table 3.11: Chi-squared residuals of segment and following context laryngealization for stops and fricatives before vowels (CV sequences).

Following Context		Non-laryng	Laryng	χ^2	df	p
Segment class						
Stops	oral	-1.70	+2.71			
	glottal	+3.40	-5.41	51.1	1	<0.0001
Fricatives	pulmonic	+0.48	-0.58			
	ejective	-0.93	+1.11	2.67	1	0.10

Table 3.10 summarizes the results of the analysis of preceding environments for stops and fricatives appearing in VC sequences only. Here, oral stops were more likely to be preceded by non-laryngealized vowels, while glottal stops were more likely to be preceded by laryngealized vowels. In other words, segments were more likely to match in laryngealization values than would be expected in a case of true independence between segments. In the case of fricatives, no relationship was found between the laryngealization of preceding vowels and following fricatives. Table 3.11 summarizes the results of the analysis of following environments of stops and fricatives in CV sequences. Again, laryngealization between adjacent segments was significantly interrelated for stops, but not for fricatives. Unlike the VC analysis, CV

sequences were less likely to show a match between segment laryngealization and context laryngealization.

3.4 Discussion

The above analysis has shown that there is a relationship between laryngealization of segments and their contextual environments. This relationship depends partly on the relative position of the context, and partly on the class of segment being analyzed. Each of these contributing factors are summarized briefly below.

There is an overall tendency for segments to have different laryngealization values than their preceding environments. However, this relationship is not monolithic. Stops, and especially glottal stops, were likely to have the same laryngealization value as their preceding environments, i.e. glottal stops were likely to be preceded by laryngealized vowels. Fricative laryngealization was independent of the preceding environment. Vowels were likely to have different laryngealization categories than their preceding environments, though this effect was small compared to that seen in stops. Subsequent analysis strictly between consonants and preceding vowels showed the same relationships.

The relationship between segments and their following contexts was somewhat different. Here, the overall pattern showed that segments were likely to have different laryngealization values than their following contexts in the initial analyses of the complete data set. When the analyses were broken down by segment type, fricatives and vowels were both revealed to differ from this overall finding, tending to match the laryngealization of their following environments. Stops, however, were likely to have laryngealization values opposite that of their following contexts, in line with the overall finding. Subsequent exclusion of word boundaries from the analysis revealed that fricative laryngealization is independent from laryngealization of following vowels. This change in the effect of following environments is likely due to the relative infrequency of ejective fricatives overall, as well as the choice to code word boundaries as non-laryngealized. When word boundaries were included in the sample, pulmonic fricatives (which are not laryngealized) would have been far more likely to appear word-finally where ejective fricatives do not occur (cf. 1.1.3), adding to the number of sequences in which laryngealization was the same for both segments. Removing the word boundaries therefore revealed the independence of laryngealization between fricatives and following contexts. Limiting the data in this way confirmed that vowels were likely to match the laryngealization of the segment immediately following, while stops were likely to have the opposite laryngealization value from the segment immediately following. Fricatives did not show any pattern of laryngealization values in relation to their contexts.

By far the most striking finding is the relationship between glottal stops and their preceding vowels. The most likely explanation for this relationship is that the vowels are laryngealized as a result of coarticulation with the following glottal stop. Unlike oral stops, glottal stops lack cues to place, and even abrupt cessation of vocal fold vibration will involve some degree of non-modal phonation as the rate of vibration slows (see sections 1.3, 2.2.2). Given the nature of glottal stop as a highly variable segment produced at a place decoupled from the articulatory constraints that arise in adjacent oral articulations, some degree of temporal overlap is to be expected. Nevertheless, predictable variations in sound production are generally considered to be allophonic rather than contrastive. Without additional evidence to support the categorization of pre-glottal vowels as laryngealized, the current transcriptions are redundant and likely muddying the waters when it comes to describing the sound system of UNT. The present analysis has established that there is a correlation between laryngealized vowels and following glottal stops. However, the acoustic effects of glottal stops and ejective fricatives on neighboring vowels has yet to be addressed. In the next two chapters, the acoustics of laryngealized and non-laryngealized vowels in laryngeal and oral contexts, as well as ejective fricatives will be analyzed. The goal will be to determine whether categorical representations of laryngealization correlate to acoustic measures.

Further limitations on the conclusions that may be drawn from the present findings arise in part from the nature of the data. Morphological and syllabic structure, word position, and stress placement were not taken into account in the present analysis. Both freestanding lexical forms and affixes listed as separate entries in the dictionary were included in the analysis. The statistical analysis also does not take into account the sometimes extreme differences in phonotactic patterning between segments belonging to the same manner class. For example, within the set of oral stops, the velar stop /k/ occurs twice as often following laryngealized vowels than does /p/. In fact, /k/ more closely resembles /ʔ/ than either /p/ or /t/, perhaps suggesting that vowel laryngealization may be a secondary cue to the place of stop closure. In that case, again, vowels preceding /k/ would perhaps be better represented in the dictionary as non-laryngealized.

There is the possibility that these findings may be due in part to over- or under-representation of certain sequences as a result of highly frequent meronymic prefixes, as mentioned in section 3.3.1. However, removing these sequences from the data set did not alter the overall findings. There is also the possibility that word initial vowels, currently transcribed with laryngealization might more accurately be classified as /ʔv/ resulting in laryngealization on the surface (MacKay & Trechsel, 2013). Such sequences would only serve to strengthen the current finding that stops tend to differ from their following environments in terms of contrastive laryngealization. Prosodic processes that induce laryngealization at the ends of phonological phrases (cf. section 1.1.3) may explain the prevalance of laryngealized vowels in word final position. Given the prevalence of word- and phrase- final devoicing or laryngealization, one might expect to find a higher rate of occurrence of laryngealization the closer a segment is to the end of a word. Be that as it may, the basic assumption should be that vowels are phonemically non-laryngealized until they are acted upon by such processes, with dictionary forms representing lexical items unaffected by contextual factors.

Another important aspect of the present findings is the lack of relationship between laryngealization in fricatives and surrounding contexts. If the ejective fricatives were produced with the glottalic airstream, the glottal closure might be expected to begin before frication, potentially inducing some acoustic laryngealization on the preceding vowel in the same way that glottal stops do. No such contextual laryngealization is apparent in the dictionary forms. One possible explanation for the lack of pre-ejective laryngealization is that glottal closure does not begin until after frication onset, long after the vocal folds have stopped vibrating. If so, this would be evidence in favor of interpreting the ejective fricatives as clusters rather than complex glottalic segments.

With respect to following contexts, the cluster analysis is equally as consistent with the data as the ejective analysis. Whether they are clusters or ejectives, their behavior before vowels ought to resemble that of glottal stops. In fact this is the case, as neither glottal stops nor ejective fricatives were related to the laryngealization of their following contexts.

Chapter 4

Phonetics of laryngealized vowels in context

Chapter 3 established that glottal stops are preceded by laryngealized vowels far more frequently than would be expected if such segments were independent of each other. If some laryngealization occurs as the result of proximity to a following glottal stop, then it stands to reason that vowels occurring in that environment will have stronger acoustic cues to laryngealization than vowels occurring elsewhere. The present chapter analyzes a particular acoustic characteristic of vowels in different contexts in order to determine whether differences in phonation might be the result of consonant context. To this end, the difference in amplitude between the first and second harmonics (H1-H2) was extracted from laryngealized and non-laryngealized vowels adjacent to laryngealized and non-laryngealized consonants, as they are defined in Chapter 3. This measure was not found to be strongly affected by the laryngeal category of vowels themselves, however it did show strong effects of the laryngeal category of the following consonant. This result suggests that the consonant context may be the source of coarticulatory differences in phonation that therefore may not constitute a phonation contrast in vowels in UNT.

4.1 Background and hypotheses

This section provides background information on measures of phonation as well as previous studies on laryngealization in languages related to UNT and introduces the research questions and hypotheses of the chapter.

4.1.1 Phonation measures

Phonation contrasts and some possible acoustic characteristics of them were introduced in section 1.3. The present section provides a more detailed review of these characteristics, motivates the use of H1-H2 as a measure of phonation contrast, and provides an overview of other potential ways to capture phonation categories.

Phonetic characteristics of phonation differences may be described in terms of articulatory configuration, audible percept, or acoustic profiles. Ladefoged (1971) proposed a continuum of glottal constriction with full closure of the glottis at one end, and a fully spread glottis at the other. Various configurations in between these two extremes may be related to linguistic voice quality, but they may be largely grouped into three main categories: breathy, creaky and modal (Keating & Esposito, 2006). Breathy voice is typically associated with a more open glottis, creaky voice with a more closed one, and modal voice occurring in between. Beyond these simple glottal settings, overall muscle tension of the larynx and surrounding speech apparatus may contribute to the voice quality that is produced and perceived (Laver, 1980; Kreiman et al., 2014). Each of these linguistically relevant phonation types may be associated with a typical configuration and subsequent percept, though not all phonation types may be reliably associated with either a particular articulation setting nor acoustic cue (Gerratt & Kreiman, 2001). For example, the classic description of creaky voice is that it is produced with the arytenoid cartilages held together, while a small portion of the vocal folds continue to vibrate. The effect of creaky voice has been described as a "rapid series of taps, like a stick being run along a railing" (Catford, 1964, p. 32). What was once referred to simply as *creak* has since been subcategorized into several different varieties of voice quality according to articulatory, acoustic, and auditory properties. While Gordon (2001) refers to creaky voice interchangeably with "vocal fry", Keating et al. (2015) make a distinction between "prototypical creak" produced with a low and irregular F0 and constricted glottis, and several other types of creak including vocal fry, multiply pulsed voice, aperiodic voice, non-constricted creak, and tense or pressed voice. Each type of creak is associated with a particular suite of spectral and acoustic measures. The most common measure of creak is the difference between the amplitude of the first harmonic (H1) and the second harmonic (H2). Keating et al. (2015) (and many others) have found that most types of creaky voice had a low H1-H2 value with the exception of non-constricted creak, which was nevertheless produced with low and irregular F0.

In addition to characterizing linguistically relevant contrasts, much of the research on phonation and voice quality differences is motivated primarily by the goal of improving speech synthesis, describing speech produced with vocal pathologies for clinical purposes, and differentiating individual speakers from one another (Hanson, 1997; Keating et al., 2015). Perhaps due to the many different goals of voice research, terminology in the field is internally inconsistent in a number of ways, making it difficult to straightforwardly relate a physiological configuration with an acoustic output, on top of the already difficult problem of determining perceived qualities as auditorily similar. A number of measures have been proposed for quantifying different types of glottal configurations, phonation types, and voice qualities during speech. Although the number of possible measures for describing phonatory differences is quite large, many of the proposed measures are redundant or highly correlated with one another, as shown by Hanson (1997) and Kreiman et al. (2007), making them less likely to be independent enough to be perceptually relevant on their own.

Gerratt & Kreiman (2001) note that non-modal phonation is typically defined in comparison to what it is not, namely modal phonation, which itself does not have a precise phonetic definition. Indeed many variable terms have been applied to various types of non-modal phonation, often with the same terms applied to very different phonation types, and similar phonation types often referred to by different terms throughout the literature, making comparison and even reliable identification difficult. Gerratt & Kreiman (2001) report through a series of same-different tasks of carefully selected acoustic stimuli and subsequent investigation of related physiology that some types of non-modal phonation, namely period-doubled phonation, amplitude modulations, and vocal fry appear to have consistent acoustic and articulatory profiles, while other forms of non-modal phonation, such as breathiness and creak, vary continuously with respect to modal phonation, making characterization of them in articulatory or perceptual terms difficult.

Many of the reported measures of voice quality have their limitations. Some measures require the investigation of source characteristics via inverse filtering of the speech signal such that the influence of the vocal tract on the glottal source signal is removed (Gobl, 1992; Kreiman et al., 2007). However, inverse filtering is often labor intensive and otherwise difficult for a variety of reasons that lead to subjective and therefore potentially unreliable output (Hanson, 1997; Kreiman et al., 2007). Direct observation of the vocal apparatus during speech is another possibility for characterizing phonation contrasts (Hanson, 1997), but laryngoscopy is an invasive and expensive process. Spectral measures, on the other hand, are relatively easy to perform and accessible to most researchers thanks to freely available acoustic analysis software such as Praat (Boersma & Weenink, 2018). Acoustic analysis of the spectral domain is an affordable, non-invasive, reliable and accurate method for describing phonation and other voice and speech characteristics (Hanson, 1997).

The shape of the spectrum, known as spectral slope, is frequently cited for its likely relevance to perceivable differences in voice quality. Several measures of spectral slope have been put forth as a means of capturing the influence of spectral slope on voice characteristics. These measures include the difference in amplitude of the first and second harmonics (H1-H2) (Hanson, 1997; Gobl, 1992; Keating et al., 2015; Kreiman et al., 2007; Kreiman & Gerratt, 2010), the second and fourth harmonics (H2-H4) (Kreiman et al., 2007), the first harmonic and the first formant (H1-A1) (Hanson, 1997; Gobl, 1992), and the first harmonic and the third formant (H1-A3) (Hanson, 1997). Kreiman et al. (2007) also include overall spectral shape in their analyses, represented by the slope of a line fit to various numbers and combinations of harmonics. Hanson (1997) showed that of all of these, only H1-H2 was independent of other voice quality measures and therefore likely to be under speaker control, and potentially relevant to listeners. H1-H2 has been shown to be related to the degree of openness in the vocal folds (open quotient, OQ) when they vibrate (summarized in Garellek (2019)), though productions by individual speakers vary greatly to produce the same acoustic effect (Kreiman et al., 2012). While the acoustic profile of H1-H2 is reliable, its articulatory production may vary greatly on an individual basis. A low H1-H2 has been associated with creaky voice, but a high H1-H2 has been associated with breathy phonation, produced with widely spread vocal folds with low tension (summarized in Gordon & Ladefoged (2001)).

Measures of formant bandwidth have also been proposed as a means of characterizing differences in voice quality (Gobl, 1992; Hanson, 1997; Keating et al., 2015)). Theoretical models of the glottal source suggest that incomplete closure of the glottis, such as is presumed to occur during creaky voice, may result in a lower first formant bandwidth (B1) than complete closure (Hanson, 1997). B1 can be estimated reliably for vowels with high F1 and relatively wide separation between F1 and F2. Hanson (1995) further reports H1-A1 as another method for estimating F1 bandwidth in vowels with low F1 or first and second formants that are close together, which makes the direct estimation of B1 more difficult. While low B1 may indicate incomplete glottal closure, increased formant bandwidth may be related to breathiness in vowels (Gordon, 2001).

The fundamental frequency (F0) is related to the perceptual sensation of pitch and is reported to indicate certain types of creaky voice, including what Keating et al. (2015) call prototypical creak, vocal fry, and nonconstricted creak. Each of these types of creak are associated with a low F0, but prototypical and non-constricted creak are also reported to have irregular F0, which makes the measurement of it difficult. Irregular F0 can be described in terms of the variability from one pulse to the next, known as jitter, or various other means of determining the irregularity. Keating reports irregular F0 in terms of noise, addressed under Harmonic-to-Noise ratio, below. Furthermore, F0 has been observed to increase in the vicinity of ejective consonants (Gordon, 2016), which may lead to difficulties in the present data if the ejective fricatives are indeed produced with the glottalic mechanism.

Noise may also play a role in the characterization of non-modal voice. Noise is largely aperiodic and perceived as turbulence or a rushing sound. Noise may be measured in a number of ways including Harmonic-to-Noise ratio (Keating et al., 2015), or scales of noise ratings ranging from virtually no noise to completely aperiodic signals (Hanson, 1997). Both high and low Harmonic-to-Noise ratios have been reported to indicate different types of creaky voice (Keating et al., 2015), suggesting that they are perhaps not the most reliable measures when the nature of a contrast is not yet known. High frequency noise excitation emerged as one of the independent factors in the analysis in Kreiman et al. (2007), but was rejected as a likely cue to phonation characteristics due to its poor correlation with spectral shape and therefore low likelihood of relevance for perception. Further, Kreiman, et al. report that models of speech do not adequately account for sources of high frequency noise, entailing the unreliability of high frequency noise as a cue across voices.

Subharmonic-to-Harmonic ratio may be relevant for distinguishing multiply pulsed voice from other types of creak, which do not share any reported acoustic characteristics with each other (Keating et al., 2015). In this case, a high Subharmonic-to-Harmonic ratio may indicate the presence of a strong set of subharmonics, indicating a high degree of period doubling in the voice signal (Keating et al., 2015). This may be perceived as a rough or bitonal voice (Gerratt & Kreiman, 2001). Because this measure has only been reported as relevant to a single type of creaky voice, it is not the most robust measure for the present analysis, which deals with uncertain voice quality types.

Statistical reliability and consistency of a given acoustic measure are not enough to establish its usefulness as a characteristic of voice quality: listeners also need to perceive it, and speakers need to be able to manipulate the relevant dimensions with precision. Kreiman & Gerratt (2010) tested the just noticeable difference (JND) threshold of listeners with respect to H1-H2. The JND for each listener indicates the value of H1-H2 for which listeners could accurately distinguish between target and stimuli 70.7% of the time. The average JND value for all speakers was 3.18 dB, though the Mandarin speakers had consistently lower values than English speakers, perhaps suggesting an effect of language background on perceptual sensitivity. Because 3.18 dB is small relative to the variability across voices, Kreiman & Gerratt (2010) argue that it is likely to be perceptually useful to listeners. They propose the ratio of JND to the range of observed values as a measure of perceptual relevance. Reported values of H1-H2 differences in languages that are reported to make use of it are roughly twice as large as the JND values reported here, further supporting its utility as a linguistically relevant perceptual cue.

This section has presented a number of possible measures of voice quality differences. Because laryngealization in UNT has not been described in phonetic detail before, it is hard to know which measure would be best to characterize it. Many measures proposed for identifying different phonation types originate in theoretical acoustic models of the vocal tract and are intended to capture and explain individual differences in speaker voice characteristics (Hanson, 1997), or speaker attitude (Gobl, 1992) rather than linguistic contrast. Many are also useful for speech synthesis and description of various vocal pathologies in clinical settings (Hanson, 1997; Gobl, 1992; Kreiman et al., 2007). Application of these measures can also extend to uses of voice quality as part of the phonological system of contrast in a language, though some spectral characteristics are unlikely to fall under conscious speaker control due to their strong correlation and interdependence on each other (Hanson, 1997). Due to its cross-linguistic robustness and applicability, independence from other voice characteristics (and therefore likelihood of speaker control), and its perceptual salience, H1-H2 was chosen for analysis in the present study. Furthermore, using a well-established measure entails ease of comparison to previous literature and increased relevance of the present findings as a result. In some studies, the difference in amplitude between the first two harmonics is adjusted for the value of the first formant and reported as H1*-H2* (Garellek, 2010; Iseli & Alwan, 2004; Hanson et al., 2001). However, the benefit of adjusting for F1 and vocal tract size in this way is unclear and may also be compensated for by matching vowel quality to measured items (Keating & Esposito, 2007). In linear mixed effects regression modeling, random effects structure can be specified to control for the effects of factors such as vowel quality. In the present study, vowel quality and speaker sex were included as random effects in the statistical models, providing a statistical control for the influence of F1 and differences in vowel formants related to the size of the vocal tract. H1-H2 was selected as the measure to use in the analysis here based on its independence from other voice characteristics, its perceptual relevance, and its robustness for characterizing multiple types of non-modal phonation. The next section reviews the acoustic phonetic studies that are available Totonacan languages.

4.1.2 Phonetic studies of laryngealization in Totonacan languages

Previous descriptions of laryngealization in UNT have not specified particular acoustic or articulatory characteristics that might serve to differentiate laryngealized from non-laryngealized vowels. Some literature on other Totonacan languages suggest that spectral slope measures may be effective in capturing the laryngealization contrast in UNT as well.

One unpublished study of laryngealization measured the F1-F2 vowel space and the difference in amplitude between F0 and F1 in search of potential cues to vowel laryngealization in Misantla Totonac, a language related to UNT but distinct from it (Trechsel & Faber, 1992). The data set included acoustic measurements from two speakers who produced 43-44 words with varying vowel qualities, lengths, and phonation categories. Trechsel & Faber (1992) found that the F1-F2 vowel space was not reliably related to vowel laryngealization. Each speaker showed different patterns in their F1-F2 vowel space, and the authors concluded that F1-F2 was therefore not a reliable cue to the vowel laryngealization contrast. On the other hand, the difference in amplitude between F0 and F1, referred to by Trechsel & Faber (1992) as the Voice Quality Index (VQI), suggested a potential relationship between harmonic amplitudes and laryngealization category. Their findings showed that laryngealized vowels had lower F0-F1 values for all vowel qualities and both speakers. However, the difference in VQI between laryngealized and non-laryngealized vowels was usually only about 1-2 dB. The VQI measure used by Trechsel & Faber (1992) may also be expressed as H1-A1 as in Hanson (1997), among others. While this is a measure of spectral slope and therefore potentially captures certain aspects relevant to voice quality, Hanson (1997) showed that H1-A1 was highly correlated with H1-A3 as well as noise ratings, which suggests that speakers do not have precise control over it for the purposes of linguistic communication of linguistic contrasts. Trechsel & Faber (1992) also note that the perceptual relevance of differences in H1-A1 is unknown.

In a descriptive study of another related language, Herrera Zendejas (2014) looked at vowel laryngealization in Papantla Totonac, identifying and illustrating three forms of laryngealization: prototypical creak with irregular glottal pulsing throughout vowel duration, a sequence of creaky voice followed by modal phonation within the vowel, or a production with sustained vocal tension throughout, also referred to as stiff voice. While the prototypical and sequential productions were reportedly easy to identify in waveform and spectrographic representations, the third type of laryngealization was less straightforward. Herrera Zendejas (2014) suggested that these productions with stiff voice are conveyed less by the vowel phonation itself, and more by the spread of voicing onto stops that precede the laryngealized vowel. No similar process has been suggested in the sound system of Upper Necaxa Totonac, and I have not noted any apparent voicing effects in the data for this dissertation. Herrera Zendejas continued her description with a comparison of the amplitudes of the first and second harmonics (H1 and H2) to that of the harmonic nearest the first formant of the low central vowel /a/ (A1). The data set consisted of 36 tokens of laryngealized and non-laryngealized vowels. The harmonic amplitudes were averaged over all productions, and these average values were then used to calculate H1-H2 and H1-A1 differences. Non-laryngealized vowels had higher H1-H2 and H1-A1 values than laryngealized vowels.

Most recently, Garcia-Vega & Tucker (2019) reported some characteristics of vowels in UNT. In this acoustic study, several measures were investigated in relation to the stress contrast in UNT. These measures include F0, F1-F2 vowel space, formant values, and duration. Unlike the findings for Misantla Totonac (Trechsel & Faber, 1992), Garcia-Vega & Tucker (2019) did find an effect of laryngealization on the F1-F2 space in Upper Necaxa Totonac. In particular, they found that F2 was higher in laryngealized vowels than non-laryngealized vowels, but only when the vowel appeared in the first syllable of a word. They also reported lower F0 in laryngealized vowels than non-laryngealized vowels in initial and medial syllables.

Previous work on Totonacan languages has suggested that the location of laryngealization in the syllable is relevant to their classification as belonging to either the Totonac or Tepehua branch of the language family. Speakers Totonac languages are reported to produce laryngealization late in the syllable, i.e. in the nucleus, while speakers of Tepehua languages produce laryngealization early, in the onset (MacKay & Trechsel, 2013). This might suggest that measures of vowel laryngealization should be stronger at later time points in vowel production in UNT, which belongs to the Totonac branch of the family. Aschmann (1946) also indicates that the type and location of phonetic vowel laryngealization is conditioned by the surrounding environment. Most relevant to the present study are the vowels that are preceded by voiceless stops, which according to Aschmann (1946) were produced most often with modal phonation throughout the vowel followed by a glottal stop at the end, or a vowel with slight laryngealization near the end followed by a glottal stop. In a similar, though opposite, pattern, MacKay & Trechsel (2013) report that both glottal stops and glottalized stops induce laryngealization on vowels that follow them. Although no such conditioning environments have been explicitly identified in UNT, some form of allophony seems likely given the results of Chapter 3. Because the relationship was found especially between glottal stops and vowels that precede them, with no indication of any similar relationship between vowels and oral stops, the conditioning environment would seem to be unlike those described by either Aschmann (1946) or MacKay & Trechsel (2013).

4.1.3 Hypotheses

Based on findings in Chapter 3 and the previous literature on acoustic correlates of linguistic voice quality, two hypotheses will be tested. First, if vowel laryngealization is contrastive, then phonemically laryngealized and non-laryngealized vowels ought to show significantly different values for H1-H2. Specifically, laryngealized vowels are expected to have lower H1-H2 values than non-laryngealized vowels. Second, if vowel laryngealization is related to consonant context, then laryngealization of the immediately adjacent consonants ought also to affect H1-H2 values, with vowels in laryngeal contexts having lower H1-H2 than vowels in non-laryngeal contexts.

4.2 Methods

The present analysis was performed on audio recordings collected in the field and annotated as described in Chapter 2. The spectral measure of laryngealization H1-H2 was extracted from phonemically laryngealized and non-laryngealized vowels adjacent to stops and fricatives, which were classified as laryngealized or non-laryngealized as in Chapter 3 (i.e. glottal stops and ejective fricatives were classified as laryngealized, and oral stops and pulmonic fricatives as non-laryngealized). Measures were taken at three time points within the vowel, at one third of vowel duration, one half of vowel duration, and two thirds of vowel duration. In the statistical analyses that follow, the middle time point (at one half of the vowel duration) was omitted for simplicity and brevity. In order to assess whether the consonant context might have an effect on the phonation measures reported here, the data were subdivided into two conditions, one taking into account the consonant preceding the vowel (CV), and the second taking into account the consonant following the vowel (VC). In the CV analysis, any effect of the consonant would be expected to be apparent at the first measured time point, while in the VC analysis, the consonant effect would be expected to appear at the last measured time point.

H1-H2 measures were extracted from vowel tokens using a Praat script designed to imitate the measures taken by the VoiceSauce software developed at UCLA (Vicenik, 2009; Shue et al., 2011). Instead of correcting the H1 and H2 measures for F1, vowel quality was included in the linear models as a random intercept, allowing for the model to adjust the estimate for each vowel category. Since vowel quality is largely determined by the first formant (Peterson, 1961; Traunmüller, 1981), including vowel quality in the random structure of the model in this way can be seen as a means of approximating H1*-H2* measures reported in other voice quality literature.

4.3 Results

This section reports data summaries, as well as the results of linear mixed effects regression (lmer) models. Linear mixed modeling was chosen because of its flexibility in analyzing unbalanced data sets, and the possibility of controlling for interdependence between data points that have been collected from the same speaker (also known as a repeated measures design). Figure 4.1 illustrates H1-H2 values for all vowels across laryngealization categories of surrounding consonants and three relative time points (at one third, one half, and two thirds of vowel duration). Despite the noisy data indicated by the large error bars, an apparent pattern is visible: laryngealized vowels seem to have lower H1-H2 values than non-laryngealized vowels in all consonant laryngealization contexts. Over the time course of the vowel, H1-H2 values increase for non-laryngealized vowels, and decrease for laryngealized vowels, indicating that laryngealized vowels have stronger acoustic markers of non-modal phonation later in their production. The difference in H1-H2 values between vowel laryngealization categories is smaller in the context of laryngealized consonants, visible on the right of

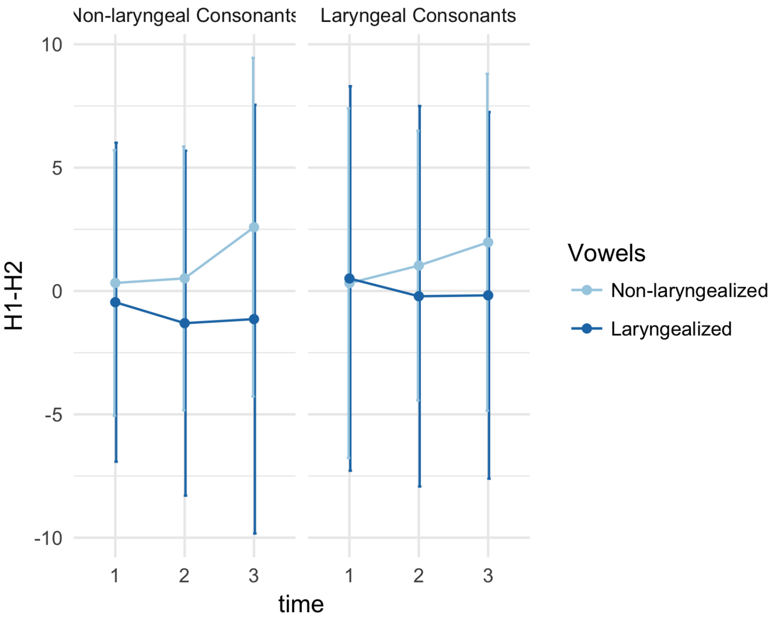


Figure 4.1: H1-H2 values at three time points. The vowels are identified according to their laryngeal category as well as that of their consonant context. Both preceding (CV) and following (VC) contexts are included here.

Figure 4.1. In the analyses that follow, the data were subsetted according to preceding (CV) and following (VC) consonant contexts. In the CV condition (n = 2626), vowels were included on the basis of being preceded by any consonant; in the VC condition (n = 1899), the data set was made up of vowels that were followed by any consonant. Vowels that were both preceded and followed by consonants were included in both analyses. In the CV condition, the early time point was expected to show greater influence from the consonant, while in the VC condition consonant effects would be expected to appear at the later time point.

The models were fitted using the *lme4* package in R through stepwise model comparisons (Gries, 2013, 2015). The *lmerTest* package (Kuznetsova et al., 2017) was used to calculate p-values based on Satterthwaite's approximation of degrees of freedom. Post-hoc comparisons were performed using the *lsmeans* package (Lenth, 2016). Model fitting for each analysis began with a maximal fixed effect structure, followed by building up the random structure. After the maximal random structure was attained, the fixed effect structure was backward fitted, removing non-significant interactions and main effects again in a stepwise manner, with analysis of variance (ANOVA) comparisons between each model step (Baayen et al., 2008; Matuschek et al., 2017). The maximal fixed effects structure included a four-way interaction among vowel laryngealization (**no**, yes), consonant laryngealization (**no**, yes), consonant manner (**fricative**, stop), and time point (**one third**, two thirds), reference levels indicated in bold. Main effects of vowel length category (**short**, long) and stress (**stressed**, unstressed) were also included as control variables. The initial models specified random intercepts for Word and Speaker. Random intercepts for vowel quality were then added to the models, significantly improving model fit in most cases. Random slopes by speaker were added to the models in the following order: Vowel Laryngealization, followed by Consonant Laryngealization, Stress, and Consonant Manner, if possible. In most cases, convergence issues kept the addition of random slopes to a minimum. The details of the random and fixed effects are presented for each model below. In order to reduce undue influence from outliers, data points with residuals outside ±2.5 standard deviations from the regression line were removed and the models refit to the trimmed data set, following Baayen et al. (2008).

The H1-H2 analysis was subdivided into CV and VC analyses. In both analyses, the initial model including time point as a predictor resulted in complex models with significant three-way interactions between time point, consonant laryngealization, and vowel laryngealization; time point, consonant laryngealization and consonant manner; and time point, vowel laryngealization and consonant manner. [1] In order to further investigate these effects, the models were subsequently divided into separate analyses of the the first and third time points in each of the CV and VC conditions, resulting in four models total. The original models with three-way interactions can be found in Appendix C.

4.3.1 CV condition

Figure 4.2 shows H1-H2 values for vowels preceded by consonants (CV condition). The values are shown at all three time points and divided by laryngealization of the preceding consonant. The following analysis further subdivides the data into two time points, at one-third (Time 1) and two-thirds (Time 3) of vowel duration.

---

[1] In the VC condition, all four of these factors interacted significantly.

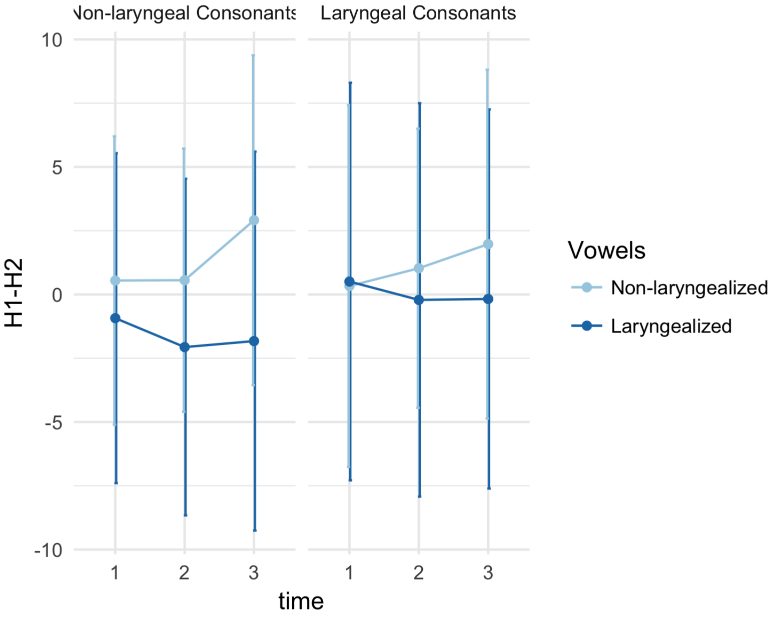


Figure 4.2: H1-H2 values at three time points during vowel production after consonants (CV condition).

Table 4.1: Summary of trimmed model of H1-H2 values for vowels following consonants (CV condition) at Time 1.

coef	Estimate	Std. Error	df	t value	Pr(>—t—)	
(Intercept)	−1.8324	1.9933	5.4235	−0.9193	0.3970	
stressunstressed	0.6894	0.3701	3.2312	1.8628	0.1528	
vlaryngyes	−0.2832	0.1752	751.6187	−1.6163	0.1065	
claryngyes	2.1843	0.5226	3.5866	4.1796	0.0174	*

Table 4.2: Summary of trimmed model of H1-H2 values for vowels following consonants (CV condition) at Time 3.

coef	Estimate	Std. Error	df	t value	Pr(>—t—)	
(Intercept)	1.8847	2.0828	4.8027	0.9049	0.4086	
vlaryngyes	−2.7319	0.6340	4.4226	−4.3087	0.0101	*
claryngyes	−0.6465	0.5970	24.8061	−1.0829	0.2893	
cmannerstop	−0.1008	0.4115	2202.8833	−0.2450	0.8065	
stressunstressed	−1.0412	0.6656	3.2234	−1.5643	0.2095	
vlaryngyes:claryngyes	0.9758	0.4794	2149.9218	2.0356	0.0419	*
claryngyes:cmannerstop	3.3446	0.5464	2242.7276	6.1213	0.0000	***

At Time 1 (see Table 4.1), the best linear mixed effects regression model of H1-H2 included random intercepts for Speaker, Word, and Vowel Quality, and random slopes for Consonant Laryngealization, and Stress by Speaker. Trimming removed 72 data points, or 2.74% of the total data set. The final model included only main effects with no interactions and revealed a significant main effect of consonant laryngealization. Main effects of stress and vowel laryngealization were not found to be significant. The results of the model indicated that vowels following laryngealized consonants (i.e. glottal stops and ejective fricatives) have higher H1-H2 values than vowels following non-laryngealized consonants (oral stops and non-ejective fricatives) overall.

At Time 3 (see Table 4.2), the best linear mixed effects regression model of H1-H2 included random intercepts for Speaker, Word, and Vowel Quality, and random slopes for Vowel Laryngealization, Consonant Laryngealization and Stress by Speaker. The final model included main effects of Vowel Laryngealization, Consonant Laryngealization, Consonant Manner, and Stress, with interactions between Vowel Laryngealization and Consonant Laryngealization, and Consonant Laryngealization and Consonant Manner. Model trimming removed 83 data points, or 3.16% of the total data set. Because of the significant interaction between vowel laryngealization and consonant laryngealization, the main effect of vowel laryngealization was recalculated using the *lsmeans* package in R. The resulting main effect was found to be marginally significant (df = 3.63, t = 3.72, $p < 0.05$), with laryngealized vowels having H1-H2 values about 2 dB lower than those of non-laryngealized vowels. This effect is in line with expectations of non-modal vowel phonation established by previous literature.

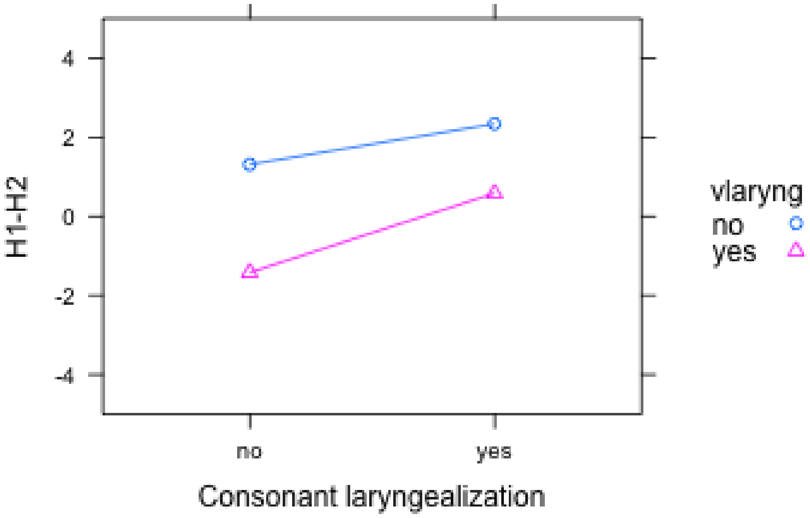


Figure 4.3: Illustration of consonant and vowel laryngealization effects in the CV condition at Time 3.

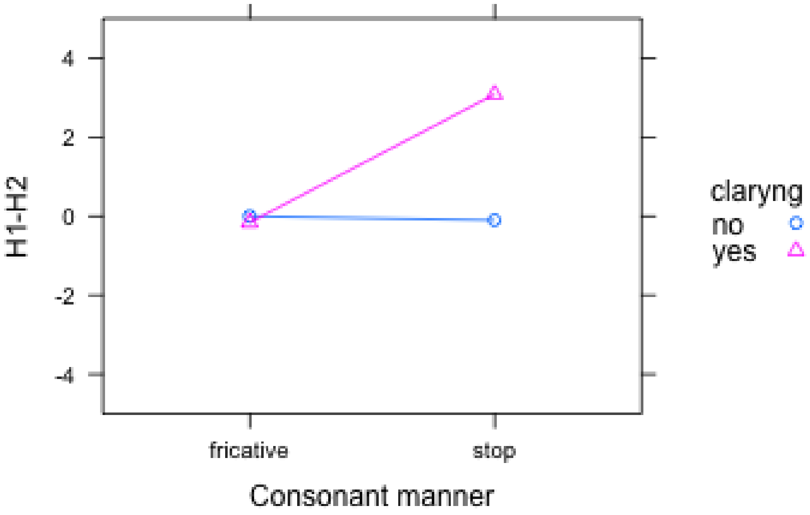


Figure 4.4: Illustration of interaction between consonant laryngealization and consonant manner effects at Time 3 in the CV condition.

The interaction between consonant laryngealization and vowel laryngealization at Time 3 is illustrated in Figure 4.3. The main effect of vowel laryngealization is visible in the overall tendency for laryngealized vowels to have lower H1-H2 values than non-laryngealized vowels. However, this effect was only significant when the preceding environment was not laryngealized. Pairwise comparison revealed that laryngealized vowels had significantly lower H1-H2 values than non-laryngealized vowels after non-laryngealized consonants, (est = 2.7318745, SE = 0.6340364 df = 4.42, t = 4.31, $p < 0.05$), but there was no significant difference between vowel laryngealization categories after laryngealized consonants. In other words, after glottal stops and ejective fricatives, there was no difference in H1-H2 between vowel laryngealization categories. Within the vowel laryngealization categories, consonant laryngealization had a significant effect for laryngealized vowels, but not for non-laryngealized vowels. That is, laryngealized vowels had significantly higher H1-H2 values after laryngealized consonants than non-laryngealized consonants (est = -2.0016506, SE = 0.5382438, df = 16.49, t = -3.72, $p < 0.01$). Non-laryngealized vowels, on the other hand, did not differ significantly based on consonant environment, despite appearing to show the same general pattern of increased H1-H2.

Figure 4.4 illustrates the significant interaction between consonant manner and consonant laryngealization. H1-H2 values differed significantly between oral and glottal stop environments (est = -3.19, SE = 0.4920047, df = 11.51, t = -6.48, $p < 0.0005$), but not between ejective and pulmonic fricatives. After glottal stops, vowels had H1-H2 values approximately 3 dB higher than vowels after any other consonant types. This effect is also visible in the raw data presented in Figure 4.3, above.

4.3.2 VC condition

Figure 4.5 illustrates H1-H2 values for vowels that were followed by consonants (VC condition). Here, vowels in non-laryngealized environments are highly similar to each other regardless of vowel laryngealization category. Both laryngealized and non-laryngealized vowels show H1-H2 values near 0 dB. On the other hand, vowels after laryngealized consonants show greater divergence between vowel laryngealization

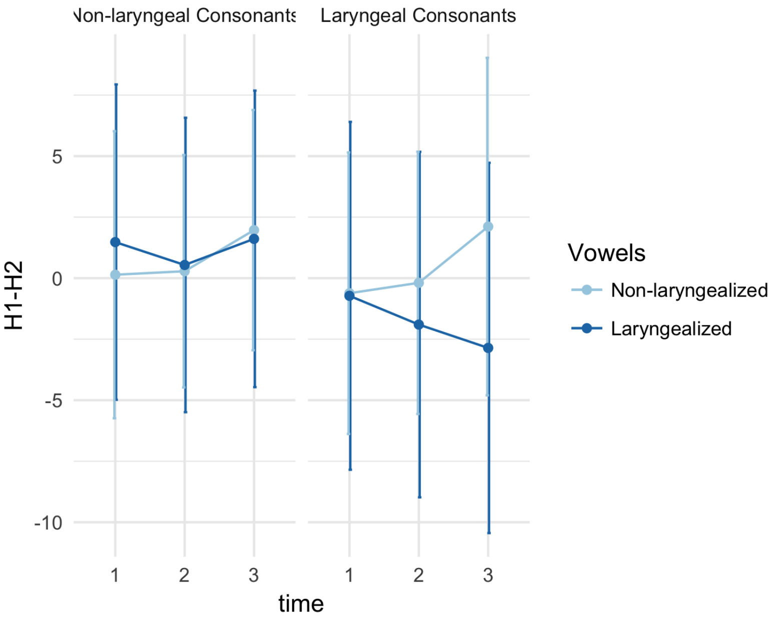


Figure 4.5: H1-H2 values at three time points during vowel production before consonants (VC condition).

categories than vowels after non-laryngealized consonants. This difference between the vowel categories increases at later measured time points.

At Time 1, the best model of H1-H2 included random intercepts for Speaker, Word, and Vowel Quality, as well as a random slope by Speaker for Consonant Laryngealization. Main effects of Vowel Laryngealization, Consonant Laryngealization, Consonant Manner, and Stress were also included, as well as an interaction between Consonant Manner and Consonant Laryngealization. Through model criticism, 52 data points, or 2.74%, were trimmed from the model. The output of the trimmed model is presented in Table 4.3. The model showed a significant main effect of Vowel Laryngealization, with a H1-H2 value about 1.3 dB higher on average for laryngealized vowels, an unexpected result given previous findings that non-modal phonation in vowels tends to result in lower H1-H2 values. The main effect of Stress was also highly significant, with higher H1-H2 values for stressed vowels than unstressed vowels. The interaction between Consonant Laryngealization and Consonant Manner, illustrated in Figure 4.6, was also significant. Pairwise comparison revealed significantly lower values before glottal stops than before ejective fricatives (est = 1.47, SE = 0.3497500, df = 1054.15, t = 4.208, $p < 0.0005$). Significantly lower values were also found before glottal stops than before oral

Table 4.3: Summary of trimmed model of H1-H2 values for vowels preceding consonants (VC condition), time = 1.

coef	Estimate	Std. Error	df	t value	Pr(>—t—)	
(Intercept)	−1.6213	2.0472	3.9583	−0.7919	0.4731	
vlaryngyes	1.3244	0.2492	934.2322	5.3136	0.0000	***
claryngyes	−0.3328	0.5469	7.0195	−0.6085	0.5620	
cmannerstop	0.1490	0.3009	780.2835	0.4952	0.6206	
stressunstressed	1.0877	0.2504	762.7265	4.3435	0.0000	***
claryngyes:cmannerstop	−1.6209	0.4516	835.8450	−3.5894	0.0004	***

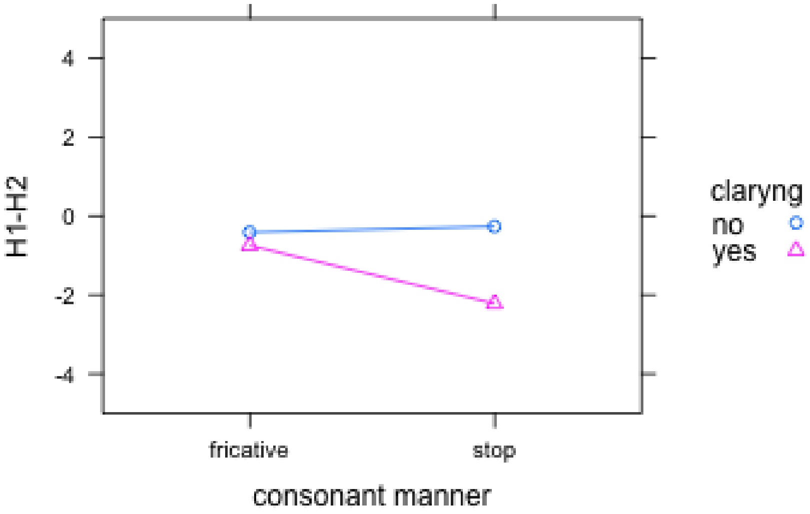


Figure 4.6: Illustration of interaction between consonant manner and consonant laryngealization for VC condition at Time 1.

stops (est = 1.9536577, SE = 0.5089788, df = 5.27, t = 3.838, $p < 0.05$). No difference was found between ejective and non-ejective fricative contexts.

At Time 3 of the VC condition, the best model of H1-H2 values included random intercepts for Speaker, Word, and Vowel Quality, and random slopes for Consonant and Vowel Laryngealization by Speaker. The fixed effects structure included main effects of Consonant and Vowel Laryngealization and Consonant Manner, as well as an interaction between Consonant Manner and Consonant Laryngealization. Model trimming removed 56 data points, or 2.95% of the data.

The model summary output is provided in Table 4.4. The main effect of Consonant Manner was significant, with stop contexts being associated with lower H1-H2 values

Table 4.4: Summary of trimmed model of H1-H2 values for vowels preceding consonants (VC condition), time = 3.

coef	Estimate	Std. Error	df	t value	Pr(>—t—)	
(Intercept)	1.4989	1.7825	4.6260	0.8409	0.4417	
claryngyes	1.3198	0.6328	5.4242	2.0856	0.0870	
cmannerstop	−1.9756	0.2943	921.7264	−6.7125	0.0000	***
vlaryngyes	−0.1001	0.7282	3.3200	−0.1375	0.8985	
claryngyes:cmannerstop	−4.9326	0.4468	945.9837	−11.0395	0.0000	***

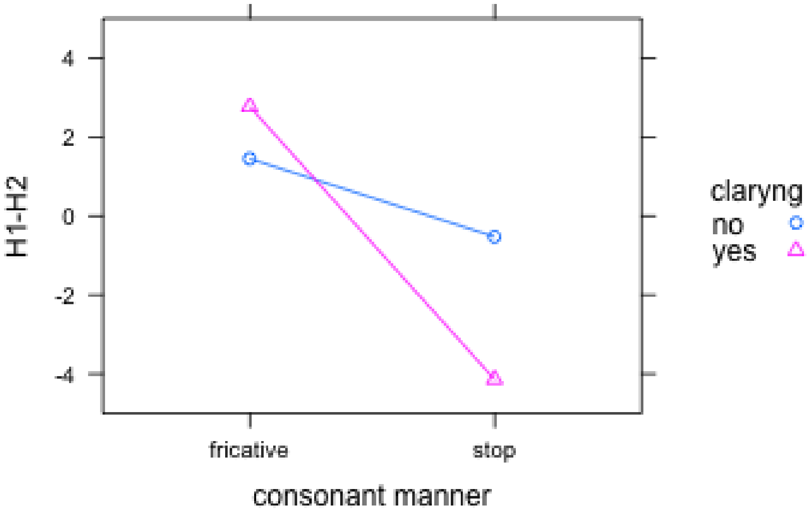


Figure 4.7: Illustration of interaction between consonant manner and consonant laryngealization for VC condition at Time 3.

than fricative contexts (est = -1.571507, SE = 0.2471497, df = 2404.65, t = -6.359, $p < 0.0001$). A highly significant interaction effect between Consonant Manner and Consonant Laryngealization, illustrated in Figure 4.7, was also found. Pairwise comparison showed that vowels before glottal stops had significantly lower H1-H2 values than vowels before oral stops (est = 3.612859, SE = 0.6042696, df = 4.51, t = 5.979, $p < 0.01$). Consonant laryngealization of fricatives did not have an effect on preceding vowels.

4.3.3 Summary

Table 4.5 summarizes the statistically significant findings of the H1-H2 analysis presented above including the size of the effect in decibels (dB) and the level of significance represented by *s.

In the CV condition at Time 1 (Table 4.1), vowel laryngealization did not affect H1-H2 values. Consonant laryngealization did have an effect: vowels preceded by laryngealized consonants had H1-H2 values about 2 dB higher than vowels following non-laryngealized consonants. At Time 3 (Table 4.2), the main effect of vowel laryngealization showed that laryngealized vowels had H1-H2 values about 2 dB lower that non-laryngealized vowels. The effect of consonant laryngealization showed that H1-H2 values that were 2.7 dB lower for laryngealized vowels than non-laryngealized vowels only when the preceding consonant was not also laryngealized (represented by the interaction between vowel laryngealization and consonant laryngealization in Table 4.5; see also Figure 4.3). Further analysis revealed that this effect was due to the presence of a glottal stop in the preceding environment: vowels preceded by a glottal stop had H1-H2 values 3.19 dB higher than vowels preceded by any other consonant, including ejective fricatives (represented by the interaction between consonant laryngealization and consonant manner in Table 4.5; see also Figure 4.4).

In the VC condition at Time 1 (Table 4.3), the model showed main effects of vowel laryngealization and stress. After controlling for interaction effects, laryngealized vowels were shown to have H1-H2 values that were 1.3 dB higher than non-laryngealized vowels, contrary to findings from other languages that non-modal phonation results in lower H1-H2 values. Stressed vowels had H1-H2 values 1.8 dB higher than unstressed vowels. The interaction between consonant manner and consonant laryngealization showed that vowels before glottal stops had H1-H2 values 1.47 dB lower than those before ejective fricatives, and 1.95 dB lower than those before oral stops (represented by the interaction between consonant laryngealization and consonant manner in Table 4.5; see also Figure 4.6). At Time 3 (Table 4.4) vowels preceding stops had H1-H2 values 1.57 dB lower than those preceding fricatives. Pairwise comparison of the significant interaction between consonant manner and consonant laryngealization revealed that vowels before glottal stops had H1-H2 values 3.6 dB lower than than vowels before oral stops, a larger effect than those seen in any other context (represented by the interaction between consonant laryngealization and consonant manner in Table 4.5; see also Figure 4.7). The increased magnitude of the effect of glottal stops on H1-H2 suggests increasing

Table 4.5: Summary of H1-H2 analysis fixed effects. Magnitude and direction of the effect is indicated in decibels. Stars indicate significance levels.

	CV		VC	
	T1	T3	T1	T3
Vlaryng		-2dB (*)	+1.3dB (***)	
Claryng	+2dB(*)			
Cmanner				-1.5dB(***)
Stress			+1.8dB(***)	
Claryng * Cmanner		+3dB (***)	+1.47-1.95 dB (***)	**-3.6 dB(***)**
Vlaryng * Claryng		-3dB(*)		

influence of the glottal stop with increasing proximity of the measurement point to the consonant.

4.4 Discussion

The above analysis has revealed some patterns in acoustic characteristics of laryngealized and non-laryngealized vowels in UNT with respect to their immediately preceding and following consonant environments. This section summarizes the analyses and relates the findings to the hypotheses stated at the beginning of the chapter. The analysis proceeded thus: A spectral measure indicating the difference in amplitude between the first and second harmonics of the spectrum (H1-H2) was extracted from vowel tokens at two time points during the course of vowel production, the first at one-third of vowel duration, and the second at two-thirds of vowel duration. The analyses were divided into two conditions, one where vowels followed consonants (CV), and the other where vowels preceded consonants (VC). Model predictors were all binary factors encoding information about segmental categories.

The first hypothesis was that categorical vowel laryngealization would result in lower H1-H2 values in laryngealized vowels than in non-laryngealized vowels. The present analyses show to the contrary that vowel laryngealization categories in UNT are not reliably related to this measure. Vowel laryngealization was found to have a significant main effect on acoustic measures in only a few cases: at Time 3 in the CV condition and Time 1 in the VC condition (see Table 4.5). At CV Time 3, vowel laryngealization had the effect of lowering H1-H2 values for laryngealized vowels, in line with findings in other languages. However, laryngealized vowels at VC Time 1 had higher H1-H2 values than non-laryngealized vowels, contrary to expectations and previous research.

The second hypothesis was that consonant laryngealization might have an effect on H1-H2 in vowels, with laryngealized consonants (i.e. glottal stops and ejective fricatives) expected to be related to lower H1-H2. This hypothesis was partially supported by the present analyses. Consonant laryngealization did in fact have a significant effect on H1-H2 values, particularly when the consonant was a glottal stop, but the direction of this effect varied. In the CV condition, vowels preceded by a glottal stop had significantly higher H1-H2 values, while in the VC condition, vowels followed by a glottal stop had lower H1-H2 values. These findings were largely independent of the laryngeal category of the vowels themselves (cf. Tables 4.2 and 4.5), with the exception of the CV analysis at Time 3 where laryngealized vowels after laryngealized consonants had significantly higher H1-H2 values than laryngealized vowels after non-laryngealized consonants (Figure 4.3); no effect was found within non-laryngealized vowels.

In general, the laryngeal category of the consonant was a stronger predictor of acoustically non-modal phonation than that of the vowel. Vowels before glottal stops showed lower H1-H2 values, regardless of whether the vowel itself was coded as laryngealized. These findings align well with the results in Chapter 3, which showed higher than expected counts of laryngealized vowels occurring before glottal stops and non-laryngealized vowels following glottal stops. In a further parallel with Chapter 3, no effect was found for laryngealization in fricatives; all of the consonant effects were due to the presence of glottal stops. Vowels *following* glottal stops, on the other hand, were produced with *higher* H1-H2 values.

The present findings demonstrate the strong influence of consonant context on H1-H2 on vowels in UNT. This influence was due entirely to the identity of the consonant as a glottal stop; no other consonant type was related to a change in H1-H2 values. The effect of the glottal stop was particularly strong when the vowel preceded the consonant, and increased in magnitude when the measurement was taken in closer proximity to the glottal stop. No effect was found relating fricative laryngealization to H1-H2. Previous results of similar studies on related languages showed small effects of laryngealization: Trechsel & Faber (1992) found that laryngealized vowels had F0-F1 values about 1-2 dB lower than non-laryngealized vowels. Herrera Zendejas (2014) reported H1-H2 values about 3 dB lower for laryngealized vowels in comparison to non-laryngealized vowels. The results are similar in magnitude to the effects found in the current chapter, but in the present analysis, that effect was not consistent across conditions. Trechsel & Faber (1992) similarly point out that although differences between laryngeal categories were consistent in that laryngealized vowels always had lower VQI than non-laryngealized vowels, the magnitude of the difference varied between speakers and vowel qualities. The consistency of the results in Trechsel & Faber (1992) suggest a more reliable indication of a phonation contrast than the results in the present study, despite the small magnitude of the effect. In the case of Herrera Zendejas (2014), no similar commentary or further comparison can be made regarding the consistency of the results, as they were reported as an average value across vowel qualities.

One limitation of the present analysis, and perhaps an avenue for future study, is that word position was not included in the model specifications. Findings in Garcia-Vega & Tucker (2019) indicate that laryngealized vowels in non-final position influenced the F0 of the vowel. In the present data, only vowels in the CV condition have the potential to appear in word final position. Seventy-four word items in the data set end in a vowel, representing about one third of the total number of vowels in the CV data. It is not yet known whether word position will have an effect on H1-H2 measures, but the interaction between vowel laryngealization and word position in the F0 analysis in Garcia-Vega & Tucker (2019) suggests that it could be playing a role here. It is furthermore possible that other acoustic measures might reliably indicate vowel laryngealization categories, such as fundamental frequency, F1-F2 vowel space, and vowel duration, as suggested in Garcia-Vega & Tucker (2019).

Chapter 5

Acoustic duration of ejective and pulmonic fricatives

This chapter provides an acoustic analysis of the contrast between ejective and pulmonic fricatives in UNT. Recordings of both classes of fricatives were collected in word initial and intervocalic contexts, preceding stressed and unstressed vowels, and preceding laryngealized and non-laryngealized vowels. In addition, pulmonic fricatives were collected in contexts where they preceded the oral stop consonants /p, t, k/ to allow for comparison between ejectives and fricatives in clusters. Duration data are analyzed from three intervals of consonant production: total duration from frication onset to vowel onset, duration of frication only, and duration of post-frication lag, where present. Statistical comparisons of duration reveal similarities between frication in ejective and pulmonic fricatives, and a durational pattern in post-frication closure relating to place of articulation following a cross-linguistic pattern of decreasing duration at places of articulation further back in the vocal tract. These findings suggest ejective fricatives may be phonetic clusters rather than true ejectives.

5.1 Background

Ejective consonants are produced by egressive airflow generated by raising the larynx to compress the air that is trapped in the supralaryngeal cavity. In stops, the necessary compression is achieved via simultaneous closures formed at the glottis and a second oral place of articulation. While these closures are maintained, the larynx moves upward, increasing air pressure in the newly sealed chamber. After compression, the oral closure is released, followed by release of the glottal closure. In the case of ejective fricatives, however, glottalic production is complicated by the need for sustained outward airflow in order to produce frication. Since the reservoir of available air trapped in the vocal tract is quite small, ejective fricatives are therefore expected to have a relatively short period of frication compared to pulmonic fricatives (Maddieson et al., 2001; Demolin, 2015; Gordon & Applebaum, 2006). Ejective fricatives have also been found to be preceded by silent intervals, either as a result of oral or glottal closure (Gordon & Applebaum, 2006; Demolin, 2002; Ridouane et al., 2015; Shosted & Rose, 2011). For further background on ejective fricatives, see Section 1.2.

Although expected to be short relative to pulmonic fricatives, the ejective fricatives of Tlingit, a Na-Dene language spoken in southern Alaska and western Canada, were found to have surprisingly long durations (Maddieson et al., 2001) when taking into account both frication and closure durations. Despite hypotheses that ejective fricatives ought to be shorter than pulmonic fricatives, the duration from onset of frication to onset of the following vowel was found to be quite similar across both airstream mechanisms: pulmonic fricatives had a mean duration of 223 ms, and ejective fricatives a mean duration of 194 ms. However, closer inspection of the acoustic data revealed that the frication portion of ejective fricatives was 74 ms shorter on average (148 ms) than that of pulmonic fricatives (222 ms). Whereas the total duration from onset of frication to onset of the following vowel was nearly identical to the frication duration for pulmonic fricatives, frication in ejective fricatives tended to be followed by a substantial period of (near) silence before the onset of vowel phonation. In word final position, this silence occurred before the onset of frication rather than after frication ended, an indication that glottal closure occurred before frication onset. Because glottal release was also evident after frication noise ceased, the authors concluded that glottal closure was maintained throughout fricative production.

In addition to duration data, Maddieson et al. (2001) collected intra-oral air pressure data for ejective and pulmonic fricatives at two places of articulation (s, s’, ɬ, ɬ’). Inspection of spectrograms and wave forms in relation to the air pressure data revealed very different pressure curves for ejective and pulmonic fricatives. In ejective fricatives, the peak in air pressure occurred later with respect to onset of frication, and lasted

Table 5.1: Summary of findings from Maddieson et al. (2001). Values represent average values across all places of articulation, rounded to the nearest millisecond. Peak air pressure data was reported for s, s′,ɬ′, ɬ only. Data have been adapted to maximize comparability with Beck (2006) (see Table 5.2).

	Frication Duration (ms)	Peak Intra-Oral Pressure (cm H20)	Post-frication Lag (ms)
Ejective	148	(later, shorter, parabolic peak)	46
ɬ’		13	
s’		16	
Pulmonic	222	(earlier, longer, flat peak)	1
ɬ		5	
s		8	

only briefly, ending at roughly the same time as frication (prior to the post-frication silence, where present). Peak air pressure in pulmonic fricatives, on the other hand, occurred nearly simultaneously with the onset of frication, and was maintained for longer, corresponding to the longer frication duration. These findings are summarized in Table 5.1. The air pressure and acoustic duration data of fricatives classified as ejective in Tlingit were ultimately determined to be consistent with expected properties of fricatives produced with the glottalic airstream.

While the overall durations of ejective fricatives in Tlingit were longer than initially expected, their ejective nature is uncontroversial. The large consonant inventory of Tlingit makes extensive use of the glottalic airstream across many places of articulation and includes six ejective fricatives and three ejective affricates. In contrast, ejective fricatives in UNT would entail a sound system like no other. Unlike Tlingit, UNT has no glottalic stops, suggesting that speakers somehow came to use a complex production mechanism to produce a rare sound type without also using it in simpler and more common ways.

Comparative reconstruction shows that the ejective fricatives originated from fricative + uvular stop clusters in Proto-Totonacan, which became fricative glottal stop clusters in Upper Necaxa as *q became ʔ in this language (Beck, 2006). After this shift, glottal closure began to overlap in time with preceding frication. In order to result in synchronic ejective fricatives, glottal closure would need to occur early enough that the frication was generated by the glottalic airstream. The timing of glottal closure in relation to frication production has not been demonstrated. Other potential sources of evidence to distinguish the ejective fricatives from fricative + stop clusters, such as phonotactics, are of little help due to identical behavior. The only posited difference between ejective fricatives and fricative + glottal stop clusters is that clusters may appear as two elements split into coda and onset of adjacent syllables, while ejective fricatives may only occur as a complex unit in syllable onset (cf. Section 1.1.4).

Using the findings in Maddieson et al. (2001) as benchmarks and following their methodology, Beck (2006) supported his historical account of ejective fricatives with a small study of aerodynamic and acoustic data elicited from two adult male speakers. The speakers produced seven repetitions of 10 words illustrating ejective fricatives at all three places of articulation. An additional 5 words containing /ʃʔ/ clusters that were the result of prefixation of the third person possessive prefix *ix-* /iʃ/ 'his/her' to stems beginning with /ʔ/ were also recorded (this being the only environment where fricative + glottal stop clusters are said to occur, cf. Section 1.1.4). Apart from the cluster items, no deliberate attempt was made to collect items containing pulmonic fricatives, although a single token of /ʃ/ followed by a laryngealized vowel was described as a basis of comparison. Beck (2006) describes the auditory impression of ejective fricatives as "sounding 'sharper' than ordinary fricatives" (p. 6) and reports visible glottal raising during the production of ejective fricatives. However, he also notes that instrumental means of distinguishing ejective and pulmonic fricatives in UNT were less reliable than impressionistic classifications.

The statistical analysis in Beck (2006) included only data from post-alveolar fricatives. Intra-oral pressure and airflow data were collected from a small sample produced by a single speaker and consisting of multiple repetitions of three lexical items: *mi**x**’á:m* /mi.ˈʃ’aːm/ 'your cornhusks' (intervocalic ejective fricative, *N* = 7), *i**x**hawá′cha′* /iʃ.ʔa.ˈwa̰.tʃa̰/ 'his/her son' (pulmonic fricative in fricative + glottal stop cluster, *N* = 6), and *i**x**a′hax’ólh* /i.ʃa̰.ʔa.ˈʃ’oɬ/ 'his/her ear' (intervocalic pulmonic fricative, *N* = 6). Only the first post-alveolar fricative in each word (in bold, above) was included in the analysis. Statistical analysis showed significant differences between ejective fricatives

*Table 5.2: Summary of findings from Beck (2006). Measures were reported for post-alveolar /ʃ/ tokens only. Values represent averages rounded to whole numbers. Significant findings are marked with * in the column heading. Time to peak intra-oral pressure was originally reported as the difference between time to peak airflow and time to peak intra-oral pressure but has been converted here for ease of comparison with other measures.*

	Frication Duration (ms)*	Time to Peak Airflow (ms)*	Time to Peak Intra-Oral Pressure (ms)*	Peak Intra-Oral Pressure (cm H20)	Post-frication Lag (ms)*
Ejective	143	25	20	8	9
Pulmonic (Pre-vocalic)	96	21	34	7	3
Pulmonic (ʃʔ Cluster)	101	27	33	9	–

on the one hand and both conditions of pulmonic fricatives on the other hand for all three measures: compared to pulmonic fricatives before vowels and in clusters, ejective segments were found to have a shorter time to peak intra-oral air pressure, a later relative time to peak airflow, and a longer overall duration. The duration of "hiatus", or lag time between the end of frication and the onset of the following vowel, was found to be longer in ejective fricatives than in pulmonic fricatives. Airflow and peak air pressure data were also collected, but no significant differences were found between ejective and pulmonic fricatives with respect to either measure.

The findings in Beck (2006), summarized in Table 5.2, are surprising for a few reasons. First, the average duration of pulmonic fricatives is remarkably short, at only half the duration of pulmonic fricatives in Tlingit. Second, ejective fricatives are longer than pulmonic fricatives by more than 50%. Given the limited air storage capacity in the glottalic airstream mechanism, there is no simple explanation as to how this increased duration could be produced by speakers. Third, the post-frication lag durations for ejective fricatives are surprisingly small, especially in comparison again to the lags found by Maddieson et al. (2001) in Tlingit (46 ms). These lags are important because the combination of frication + lag duration resulted in approximately equal durations

(~200 ms) from onset of frication to onset of vowel for ejectives and pulmonic fricatives in Tlingit, as well as being an indication of glottal closure during frication. In addition, the reported maxima of airflow and air pressure were more similar to the measurements of pulmonic fricatives reported in Maddieson (2001) than to those of ejective fricatives in the same study.

5.1.1 Duration

Duration is a robust and easily accessible acoustic measure that distinguishes many different linguistic contrasts including voicing (Ladefoged & Cho, 2001), place of articulation, manner, and airstream. It was chosen as the focus of this study for its simplicity, the relative ease of data collection, annotation and analysis, and because it is broadly comparable across languages and speakers. Duration also allows for direct comparison across all kinds of phones and the acoustic events that occur during their production. It is just as easy to compare the duration of a silent interval to the duration of a vowel as it is to compare to the duration of frication or any other speech sound. This is important in the present study because stop closures, frication intervals, and potentially epenthetic silences are all compared to each other.

Duration has been used to describe ejectives in many studies beyond the two in section 5.1 that inspired the present analysis. A selection of these studies are summarized here for further context. Wright et al. (2002) used voice onset time to compare unaspirated, aspirated and ejective alveolar stops in Witsuwit'en. They found that, ejective stops had longer voice onset time than unaspirated pulmonic stops, both within and across speakers, and that ejective bursts were followed by a silent interval before the onset of voicing began for the following vowel.

In a survey of Caucasian languages, Grawunder et al. (2010) reported comparisons of closure duration and voice onset time, with additional annotations for burst duration, post burst lag (or the time between the burst) and the onset of voicing occasionally also reported. They found that in Georgian ejectives tended to have short burst intervals frequently followed by a silence before the onset of vowel phonation. Closure duration did not differ significantly between pulmonic and ejective stops. Voice onset time was shorter for ejective stops than for aspirated stops. Ejectives also showed a shorter voice onset time when appearing in word medial position compared to word initial position. In Ingush, closure and burst duration did not differ across airstream mechanisms.

Ridouane et al. (2015) report the duration of vowels preceding fricatives, pre-release closure, frication interval, post-release lag duration, and the total duration of fricatives in ejective and pulmonic fricatives in Mehri. They found ejectivity had no effect on the duration of the preceding vowel. Similar to the analyses performed here, the presence of surrounding silences was observed in Mehri to sometimes distinguish ejective from pulmonic consonants, where pulmonic fricatives were never followed by lag silences. In Mehri, pulmonic fricatives were categorically not followed by a lag interval, while ejective fricatives varied across speakers and places of articulation (ranging from one speaker who almost never produced silences to another who always produced silences). Frication intervals were shorter for ejective fricatives than for pulmonic fricatives, but no difference in duration was observed when total duration including silent intervals was compared, or between airstream categories. They reported an average post-frication lag interval of about 25 ms between vowels, and 21 ms word initially. Intervocalic ejectives were also preceded by a silent interval of approximately 24 ms on average.

The ejective fricatives of Kabardian also had shorter frication intervals (130 ms, compared to 191 ms for pulmonic fricatives (Gordon & Applebaum, 2006)), and are reported to be followed frequently by a “gap”, though this silent interval is not quantified. A similar pattern is evident in Amharic (Demolin, 2002; Seidl et al., 2009) with shorter frication intervals in ejective fricatives following by post-frication lags of approximately 30 ms.

These studies, along with the results in Maddieson et al. (2001) show that duration is well established as a tool for comparing ejective to pulmonic consonants. The present study focuses on duration for comparability with Beck (2006) and Maddieson (2001) as well as other studies of ejective fricatives. While duration of frication and following silent intervals seems to identify particular characteristics of ejective fricatives, it is not the only possible measure that may be used to compare and contrast ejective to pulmonic fricatives. The discussion section of this chapter (5.5.2) addresses other non-temporal measures and their potential utility for describing the patterns in UNT fricative data. Both Beck (2006) and Maddieson (2001) also made use of both acoustic and aerodynamic measures to compare and contrast airstream mechanisms in fricatives. In the present study, I focus on acoustic duration and leave aerodynamic concerns for future research. The utility of these measures is also discussed in section 5.5.2.

5.2 Hypotheses

The data from Tlingit confirmed that fricatives of surprisingly long duration can be produced by the glottalic airstream. However, segmental duration was partly determined by periods of silence preceding or following frication. Removing these silences resulted in significantly shorter durations of frication intervals in pulmonic and ejective productions. Beck (2006) found the opposite, with longer frication in ejective fricatives than in pulmonic fricatives. Including following silences into the segmental duration exaggerates this difference rather than minimizing it. Given these unexpected findings, the present chapter presents analyses of frication and lag duration in UNT with a data set consisting of an extensive word list collected from multiple speakers. The word list includes fricative items from all three places of articulation and both laryngeal categories in multiple contexts. The materials and analyses have been designed to investigate the following six hypotheses based on previous findings in both Maddieson et al. (2001) and Beck (2006). These hypotheses address the overarching question of whether the ejective fricatives in UNT are phonetically ejective, or whether they are better described as phonetic clusters.

Frication duration - pulmonic vs. ejective Ejective fricatives will have shorter frication intervals than pulmonic fricatives (due to the limited supply of air available via the glottalic airstream production mechanism). This was the finding reported in Maddieson et al. (2001), even though the duration of frication was substantially longer than was predicted to be possible based on the presumed amount of available air trapped above the glottis during production of an ejective fricative. Beck (2006) found the opposite to be true: ejective fricatives had substantially longer frication intervals than pulmonic fricatives. Conversely, similar frication durations have been used to argue that “ejective” fricatives were in fact sequences of frication followed by glottal stops in Yapese (Maddieson, 1998).

Presence and duration of lags Pulmonic fricatives will have fewer instances of lag silences, and any such silences that do occur will be substantially shorter than those following ejective fricatives. Maddieson et al. (2001) found that ejective fricatives were produced with 46 ms of lag silence on average (with this lag being present "in general" (p. 157), although the exact number of silences is not reported), while pulmonic fricatives were only rarely followed by such a silence with an average duration of 1 ms. The data reported in (Beck, 2006) was insufficient to make any generalized descriptions of the presence or absence of lag silences.

Total duration Fricative segments from beginning of frication to beginning of the following vowel are of roughly equal length regardless of laryngeal type. This was the finding in Maddieson et al. (2001). When the durations of flanking silences were added to the length of frication, the overall durations of ejective and pulmonic fricatives were of comparable lengths. Combining durations of frication and silent periods in this way also accounted for the apparent lack of word final lengthening in ejective fricatives in comparison to pulmonic fricatives. Ejective fricatives were not found to undergo frication lengthening word finally, but they were flanked by audible glottal bursts (one closure burst preceding frication, and one release burst following frication) and periods of (near-)silence. Taking these component durations into account, ejective fricatives could also be interpreted as lengthened in word final position. In UNT, ejective fricatives may compensate for shorter frication intervals through the addition of flanking silent intervals, even though they do not occur word finally and therefore cannot undergo word final lengthening.

Frication duration - ejective vs. clusters Duration of frication in ejective fricatives will be distinct from pulmonic fricatives before vowels as well as from pulmonic fricatives in clusters (Beck 2006). Beck found that ejective fricatives were *longer* than pulmonic fricatives in both cluster and pre-vocalic contexts, a finding that contradicts Maddieson et al. (2001), who found that ejective frication was shorter than pulmonic frication.

Effect of place Frication and lag durations will not vary according to place of articulation of the fricative segment. Maddieson et al. (2001) found no effect of place on total duration, frication duration, or lag duration. No interaction was found between place and laryngeal type, either. Beck (2006) did not investigate such interactions in UNT. However, place of articulation has been shown to be correlated with varying durations in stop segments (Repp, 1984; Chao & Chen, 2008). On the other hand, place of articulation may have an effect on lag durations in clusters depending on the place of consonant closure.

Effect of external factors on frication and lag silence durations Word position may affect fricative duration (Maddieson et al., 2001). Maddieson et al. (2001) found duration differences between word initial and word final fricatives, with word final pulmonic fricatives being substantially longer than word initial pulmonic fricatives. Since ejective fricatives do not occur word finally (or in coda position, more generally) in UNT, the present study will look at word initial and word medial positions instead. Perhaps word initial tokens will be somewhat longer due to their prominent position. word final lengthening is known to occur in vowel-final stems in UNT (Beck, 2011b). Although this has not been reported for fricatives, and no other word position effects have yet been found, the potential for sensitivity to word position remains.

Factors external to the fricative segments, such as stress and following vowel laryngealization, may affect frication duration or the presence and duration of lag silences (Beck, 2006). Beck describes one instance of a pulmonic fricative followed by a laryngealized vowel. In this case, a brief lag of 7 ms is reported to follow frication [1], which is attributed to the following vowel rather than the fricative. Primary stress often occurs on the final syllable of a word. This fact in conjunction with word final lengthening may suggest that stress could have an effect on segmental duration, including fricatives.

---

[1] Inspection of the spectrogram in Beck (2006)'s Figure 3 calls this duration into question, however. The silent portion appears nearly as long as the frication itself, which is reported at 90 ms.

5.3 Methods

Analyses were performed on the acoustic data collected as specified in Chapter 2. Fricative and affricate segments were classified according to the frication condition in which they occurred, which are described in detail below. All conditions were limited to word initial or intervocalic positions. *Simplex* fricatives (hereafter simplex.fricative) are pulmonic fricatives that occur adjacent to vowels and/or word boundaries only and are produced with a single articulatory configuration. In contrast, the remaining frication conditions may be considered complex in that their production involves more than one articulatory configuration. *Fricative first* sequences (fricative.first) are those where frication occurs as the first part of a fricative + stop cluster and included tokens of frication before the oral stops /p, t, k/ as well as a few tokens before /ʔ/. *Laryngealized fricatives* (laryng.fricative) are segments that have previously been referred to as ejective fricatives (Beck, 2006); they were collected mainly in word initial position before vowels, or word medially between vowels[2]. Because the ejective nature of the fricatives in UNT is under investigation, the present naming convention was chosen in order to differentiate confirmed ejective fricatives from those currently under investigation here. *Affricates* (affricates), of which there are two in UNT ([tʃ] and [ts], are complex segments produced as sequences of complete oral closure followed by frication. Because they are considered to involve two articulatory phases within a single segment, they provide a basis of comparison between fricative.first clusters, which are considered sequences of two distinct segments, and laryng.fricatives, which are currently under investigation. The following frication conditions were excluded from the statistical analysis for simplicity: word final position, stop + fricative clusters (also called *Fricative second* (fricative.second) in Figure 5.4), all other fricative environments, such as before or after nasals, liquids, or glides (e.g. word initial [sl] in ***sl****a’hs’o’hó’jwa’* /sla̰ʔs’o̰ˈʔo̰hwa̰/ ‘a bit salty’), and clusters preceding or following anything other than vowels, as in the word medial sequence [kʃt] that occurs in *xta:’la**kxt**ím* /ʃta̰ːlakˈʃtim/ ‘one’s equivalent in age or size’ where the second fricative occurs between two stops (see section 1.2). Table 5.3 provides examples of each frication condition according to the control variables

---

[2] For a comparison to fricative + glottal stop clusters, see Figures 2.6 and 2.5.

word position (initial or intervocalic), *vowel laryngealization* (of the following vowel), and *vowel stress* (of the following vowel). For brevity (and because of the greater availability of forms), word initial items are shown crossed with laryngealization and frication condition, and word medial items are shown crossed with stress and frication condition.

Table 5.3: Examples of UNT words in various frication conditions. Stress and laryngealization conditions refer to the categorization of the following vowel. Target segments are shown in ***bold*** *in orthographic representations.*

Condition	Word Initial		Word Medial	
	stressed	**unstressed**	**laryng**	**non-laryng**
simplex. fricative	***s**á:sti’* [ˈsaːs.tḭ] ‘new’	***s**alún* [sa.ˈlun] ‘hoe’	*sé’h**s**i’* [ˈsḛʔ.sḭ] ‘sweet’	*ta**s**a:tanú:n* [ta.saː.ta.ˈnuːn] ‘stuck, fixed in place’
laryng. fricative	***s’**á’lhwa’* [ˈs’a̰ɬ.wa̰] ‘slow of movement or thought’	***x’**etím* ʃ’e.ˈtim] ‘seeded and deveined chili’	*li:**lh’**á:’n* [liː.ˈɬ’a̰ːn] ‘plough’	*ta**s’**awí* [ta.s’a.ˈwi] ‘lose, be defeated’
fricative. first	***x**ka’j* [ˈʃka̰h] ‘pineapple’	***lh**taká’la’* [ɬta.ˈka̰.la̰] ‘board’	*li:**x**pa’tán* [liːʃ.pa̰.ˈtan] ‘pestle of a *molcajete* (mortar)’	*tu:**s**púlh* [tuːs.ˈpuɬ] ‘one’s toes’
affricate	***ch**í’px* [tʃḭpʃ] ‘dense’	***tz**alá’j* [tsa.ˈla̰h] ‘brittle, fragile, thin (stick)’	*chu’**ch**ó’hx* [tʃṵ.ˈtʃo̰ʔʃ] ‘banana blossom’	*a:’tu:’**ch**i:yé:tlh* [a̰ː.tṵː.tʃiː.ˈjeːtɬ] ‘mint’

Forms that were not produced by all speakers, including speech errors and alternative word forms, have been excluded from the analyses, resulting in 121 common lexical items across the four speakers whose data made up the final data set. Further repetitions that were produced after the third one have also been excluded from this analysis. 1452 fricative tokens (121 words x 3 repetitions x 4 participants) were planned to be included in the analysis. Twelve words included 2 different fricative tokens, increasing the expected number of tokens to 1488. A number of words were excluded from the analyses because they did not occur in one of the following four conditions: *simplex.fricative, affricate, fricative.first* or *laryng.fricative*. Notably, approximately 20 word list items included instances of ejective fricatives occurring immediately after a glottal stop. Because the statistical analysis was limited to word initial and intervocalic environments, these forms were excluded. In some cases, speakers did not recognize the word that was intended by the word list, or produced fewer than 3 usable repetitions of some words, resulting in a final tally of 1430 tokens of frication and 1114 tokens of lag in the statistical models that follow.

5.4 Results

Results are presented in three parts. First, illustrative spectrograms of fricatives are presented in each of the four frication conditions defined above. Second, durations of acoustic events surrounding frication are summarized, along with indications of how frequently each event type occurred in the data. Finally, statistical analyses compare durations of three event intervals across segment types: frication, post-frication lag, and frication onset to vowel onset.

5.4.1 Spectrographic analysis of fricatives and affricates

This section provides spectrographic illustration of the four conditions defined above. Table 5.4 summarizes the duration data of frication, post-frication lag, and total fricative-to-vowel portions of the sample fricative spectrograms. See section 2.2.2 for details of the acoustic annotation.

Simplex fricatives

Each of the three places of articulation where simplex.fricatives are produced are illustrated below. The left-hand column in Figure 5.1 provides illustrations of pulmonic fricatives. Figure 5.1a shows a pulmonic alveolar between two /a/ vowels, with a brief (18 ms) intervening lag period between the end of frication and onset of the following

*Table 5.4: Summary of duration data taken from illustrative spectrograms in Figures 5.1-5.3. NB: pre-frication (lead) silences in affricates (marked with *) are reported in the post-frication lag column to reflect total segment duration.*

Segment	Figure	Frication (ms)	Post-frication lag (ms)	Fricative onset to vowel onset (ms)
s	5.1a	112	18	130
ʃ	5.1c	181	18	199
ɬ	5.1e	170	14	184
s’	5.1b	103	61	164
ʃ’	5.1d	141	66	207
ɬ’	5.1f	159	165	324
sp	5.2a	171	175	346
st	5.2b	131	130	260
sk	5.2c	119	83	203
ʃp	5.2d	146	107	253
ʃt	5.2e	185	130	316
ʃk	5.2f	233	168	402
ɬp	5.2g	114	139	252
ɬt	5.2h	134	151	284
ɬk	5.2i	177	150	326
ts	5.3a	74	63*	137
tʃ	5.3b	103	57*	160

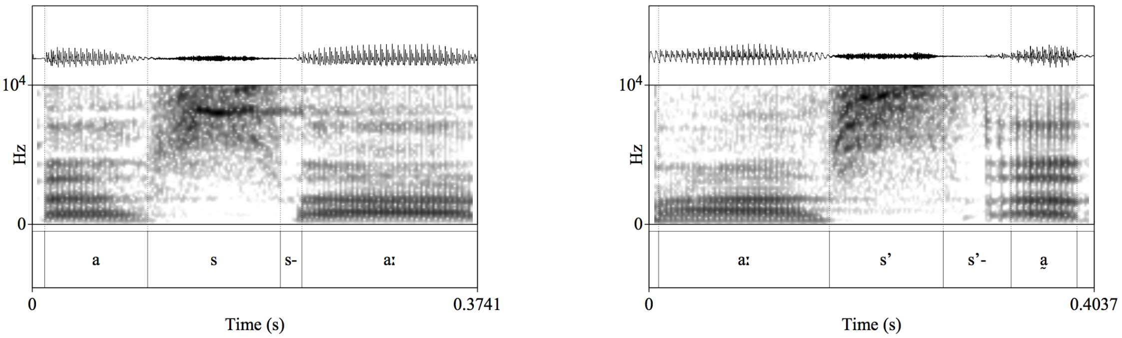
vowel. Frication lasts 112 ms. Figure 5.1c shows a pulmonic post-alveolar fricative separated from the following vowel by a brief lag period. Frication lasts approximately 181 ms and the lag 18 ms. Figure 5.1e illustrates a pulmonic lateral fricative also followed by a short lag before the onset of the following vowel. The frication lasts 170 ms and shows a formant-like structure with elevated spectral energy in several frequency bands.

Laryngealized Fricatives

The right-hand column of Figure 5.1 provides illustrations of laryngealized fricatives at each place of articulation. Figure 5.1b shows an example of /sʼ/ between two vowels, along with an intervening lag portion that includes some irregular pulsing of the vocal folds (represented by the vertical striations that are visible in the spectrogram). These segments were taken from the word item *ma:sʼa'ta:nán* /maːsʼa̰taːˈnan/ ‘raise children’. In this case, frication lasted approximately 103 ms. The post-frication lag lasted approximately 61 ms, with the silent portion of this lag making up 38 ms of that time. Figure 5.1d shows a laryngealized post-alveolar fricative between two laryngealized vowels. The frication period lasts around 141 ms, followed by a lag of 66 ms. Neither the preceding nor the following vowel shows an appreciable change in the temporal regularity of the glottal pulses. Figure 5.1f illustrates a token of the laryngealized lateral fricative. Frication lasts 159 ms and is followed by a lag of 165 ms (the last 29 ms of which involve erratic glottal pulsing).

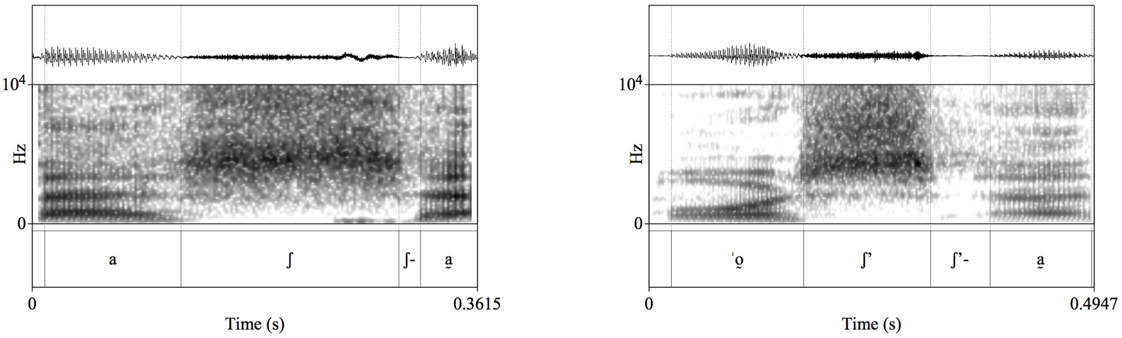
Fricative-first clusters

Figure 5.2 provides one example of each combination of fricative and stop places of articulation. Figure 5.2a shows a word medial /sp/ cluster before a stressed vowel. Frication lasts 171 ms, followed by a bilabial closure of 161 ms, and a release burst of 14 ms. Figure 5.2b shows a /st/ cluster in word medial position before a stressed vowel. Frication lasts for approximately 131 ms, followed by an alveolar closure of 122 ms, and a release burst of 8 ms. Figure 5.2c shows a word medial /sk/ cluster preceding an unstressed vowel. Frication lasts for 119 ms, followed by a 62 ms velar closure. The release burst lasts 21 ms. Figures 5.2d-5.2f illustrate the post-alveolar fricative before labial, alveolar, and velar stops. In the /ʃp/ sequence, frication lasts 146 ms, followed



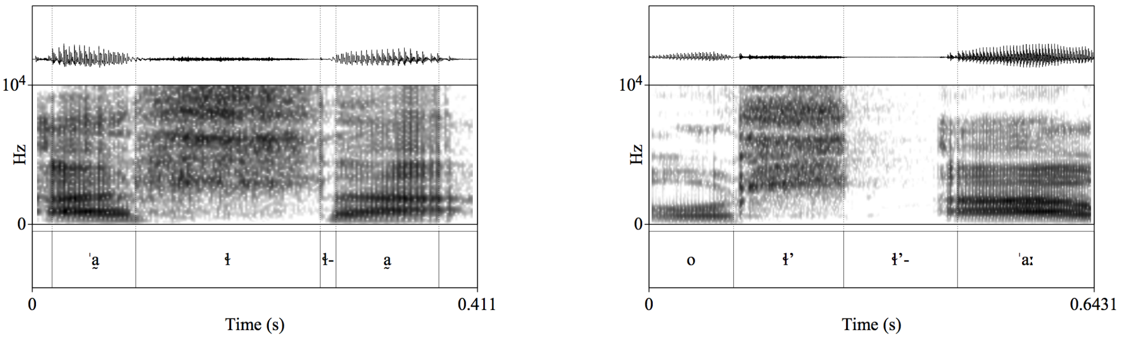
(a) /s/ *in* tasa:tanú:lh /tasaːtaˈnuːɬ/ *‘stuck, fixed in place’*

(b) /s’/ *in* ma:s’a'ta:nán /maːs’a̰taːˈnan/ *‘raise children’*.



(c) /ʃ/ *in* taxa'há /taʃa̰ˈʔa/ *‘scratch oneself’*

(d) /ʃ’/ *in* cha:’hó’x’a’ /tʃa̰ːˈʔo̰ʃ’a̰/ *‘tree bark’*



(e) /ɬ/ *in* pixhá’lha’ /piʃˈʔa̰ɬa̰/ *‘large (bunch or bouquet)’*

(f) /ɬ’/ *in* holh’á:’wa’ /ʔoɬ’ˈa̰ːwa̰/ *‘leathery, thick, stiff, not flexible’ (right)*.

Figure 5.1: Pulmonic and ejective fricatives at three places of articulation, produced by speaker GMM.

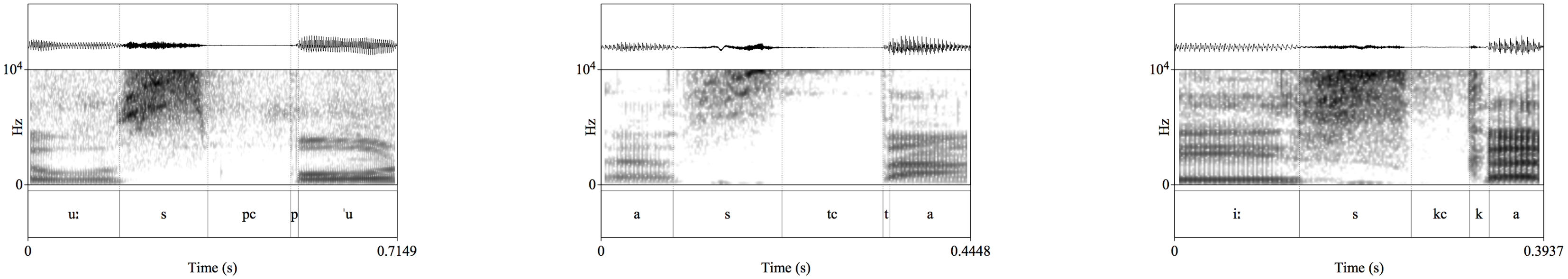
by a lag of 107 ms including closure burst. Frication in the /ʃt/ example lasts 185 ms, followed by a lag of 130 ms. Figure 5.2f shows 233 ms of frication, and 168 ms of lag. Figures 5.2g-5.2i illustrate lateral fricatives in clusters with oral stops. In the /ɬp/ sequence, frication lasts 114 ms followed by 139 ms of silence during the stop closure. Frication duration is 134 ms in the /ɬt/ cluster, followed by 151 ms of closure before the onset of vowel phonation. In the final configuration, /ɬk/, frication is even longer at 177 ms, followed by 150 ms for the /k/ closure. These sample spectrograms illustrate the similarities between the ejective fricatives and fricative + stop clusters, with long silent intervals between the offset of frication and the onset of vowel phonation.

Affricates

Figure 5.3 illustrates one token each of the alveolar affricate /ts/ (5.3a) and the post-alveolar affricate /tʃ/ (5.3b). The alveolar closure in /ts/ is approximately 63 ms, followed by 74 ms of frication. In /tʃ/ the closure lasts 103 ms, followed by 57 ms of frication. Note that in both cases the release of alveolar closure is visible as a broad spectrum band of energy across all visible frequencies, but that this burst has not been segmented separately from the frication that is contiguous with it.

5.4.2 Duration of acoustic events in frication production

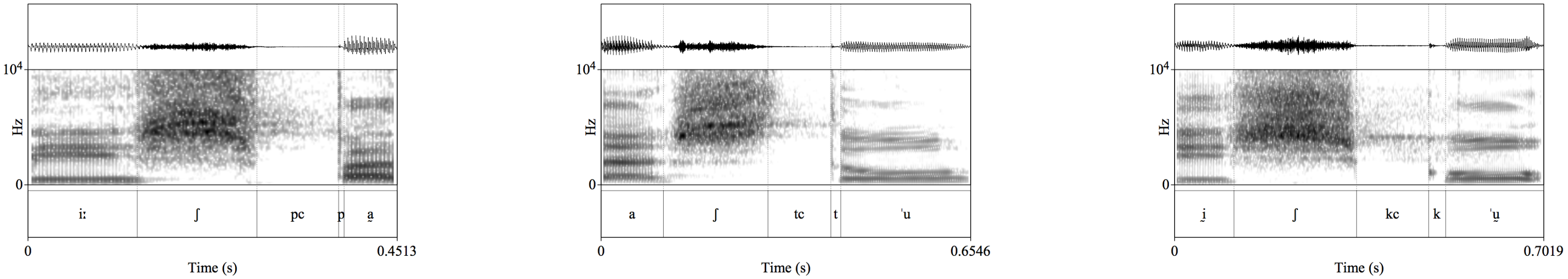
Figure 5.4 summarizes the mean durations of acoustic events that occurred in the vicinity of frication in each condition. Duration data for simplex stops occurring in environments analogous to those of frication (word initially before vowels, or word medially between vowels) have also been included as a point of reference for comparing the duration of affricates and clusters. All conditions were potentially made up of five types of events: *silence preceding frication* (lead), a *burst preceding frication* (lead.burst), *frication noise* (frication), *silence following frication* (lag), and a *burst following frication* (lag.burst). In the affricate condition, bursts indicating the release of oral closure were segmented as part of the frication intervals because they could not be differentiated from the frication noise, resulting in the apparent lack of release bursts. (See Chapter 2 for more information on segmentation and annotation procedures.) The colored



(a) tu:spúlh /tuːs.ˈpuɬ/ *'one's toes'*

(b) xastánku' /ʃas.ˈtan.kṵ/ *'youngest sibling'*

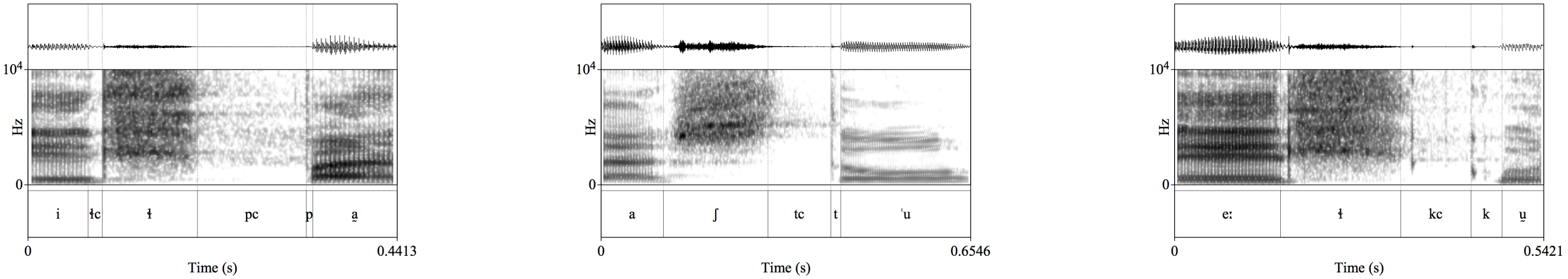
(c) li:s.ka.láj.wa' /liːskaˈlahwa̰/ *'danger'*



(d) li:x.pa'.tán /liːʃpa̰ˈtan/ *'pestle of a* molcajete'

(e) helhtaxtú /heɬ.taʃ.ˈtu/ *'fade (color)'*

(f) chi'xkú' [tʃḭʃ.ˈkṵ] *'man' (right)*

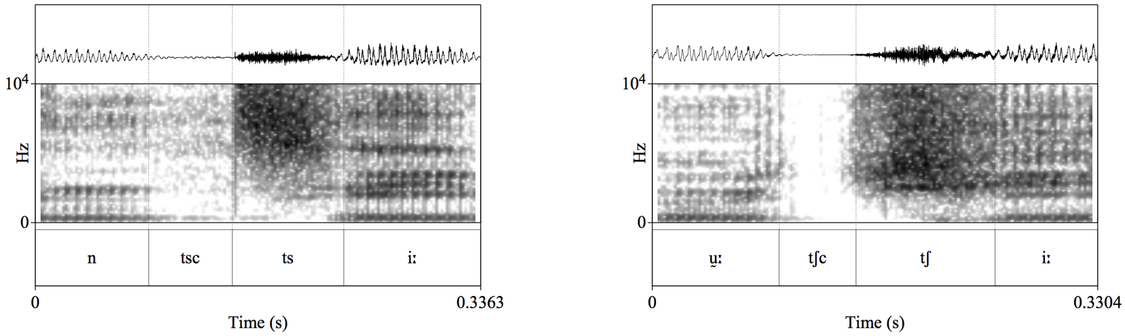


(g) kilhpa'nlhú'lu' [kiɬ.pa̰n.ˈɬṵ.lṵ] *'jowly, with swollen cheeks'*

(h) i'lhtí'n [ḭɬ.ˈtḭn] *'feces'*

(i) he:lhku'tán [ʔeːɬ.kṵ.ˈtan] *'smell of sweat'*

Figure 5.2: Fricative + stop clusters at three places of stop closure and frication constriction, produced by speaker GMM.



(a) tzi'ntsi:pá'hlhcha' /tsḭntsiːˈpa̰ʔɬtʃa̰/ *'cuatomate, small, tomato-like wild fruit'.*

(b) a:'tu:'chi:yé:klh /a̰ːtuːtʃiːˈjeːkɬ/ *'mint'.*

Figure 5.3: Affricates produced by speaker HFM

bars represent average durations of each event type with error bars indicating 95% confidence intervals.

By comparing the average durations of the events that did occur, some similarities and differences may already be observed between the conditions presented here. Laryng.fricative and fricative.first conditions bear a strong resemblance to each other with respect to both the overall durational profiles of their production, as well as the average durations of frication and lag events (see Figure 5.4). Both categories are produced with longer average durations than either the affricate or simplex.fricative conditions, as would be expected for sequences of two segments rather than a single complex segment such as an affricate, or a simplex segment such as a pre-vocalic fricative. Note also the similarity between overall length of affricate and simplex.fricative conditions despite the differences in their complexity; while affricate segments have longer lead silences than simplex tokens, their shorter frication intervals even out the total duration.

Frication was sometimes produced with lead and/or lag silence in all conditions, though the frequency of these silences varied by condition. The relative frequency of occurrence of each event type in each condition is summarized in Table 5.5. In addition to the frication conditions, simplex stops (that is, stops adjacent to vowels only) are also provided as a point of reference and basis of comparison. Laryng.fricatives and fricative.first clusters bear a strong resemblance to each other both in the types of events that occur during their production, as well as in terms of their (overall and component)

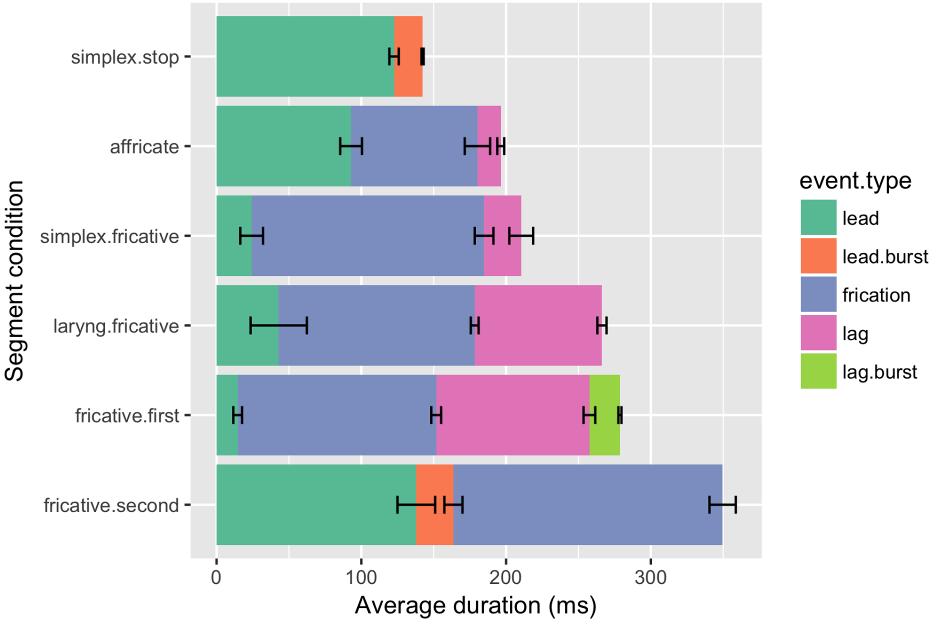


Figure 5.4: Mean duration of acoustic events during production of segments involving frication. Error bars represent 95% confidence intervals. Simplex stops (simplex.stop) are included here as a point of reference.

durations. These similarities contrast with the other remaining conditions. Lead silence occurred in less than 5% of simplex.fricatives, and less than 2% of laryng.fricatives and fricative.first clusters. Lead silence occurred in only just over 48% of affricate productions, a condition that would be expected to require complete closure prior to frication. This fairly low frequency is largely due to word initial tokens, in which the beginning of closure often could not be easily determined. Similarly, frequency of lag silence occurrence varied according to condition. Lag silences occurred in less than 15% of affricates, and less than 22% of simplex.fricatives. This is in stark opposition to the laryng.fricative condition, which was produced with lag silence 95% of the time, and fricative.first condition, which was produced with lag silence over 99% of the time. An additional difference between the laryng.fricative condition and the fricative.first condition is the lack of lag bursts in the former. This is due in part to the segmentation conventions, which did not indicate glottal bursts due to their variable and erratic appearance.

Table 5.5: Tabulation of event frequency in frication production. Numbers in parentheses represent counts of each event in the data set; percentages were calculated based on these condition-based counts.

	lead	lead.burst	frication	lag	lag.burst
simplex.stop	100% (947)	80.4% (762)	0	0	0
affricate	48.3% (73)	(–)[a]	100% (151)	14.6% (22)	0
simplex.fricative	4.1% (8)	0	100% (196)	21.9% (43)	0
laryng.fricative	1.5% (9)	0	100% (615)	95% (582)	0
fricative.first	1.9% (9)	0	100% (468)	99.8% (467)	90.0% (421)
fricative.second	100% (32)	18.8% (6)	96.9% (31)	0	0

[a] Bursts were not segmented separately from frication noise. See Section 2.2.2.

5.4.3 Statistics

In the statistical analysis, the data have been restricted from those reported in section 5.4.2. First, the fricative.second condition has been entirely excluded from the statistical analysis due to low numbers of tokens. Only 31 tokens of fricative.second frication occurred in non-word final (i.e. word medial) position throughout the entire data set. In addition, further conditions, such as sequences involving liquids or nasals, or sequences longer than two consonant segments in length were excluded from the analysis on similar grounds. Due to the low relative frequency and small overall numbers of occurrence, pre-frication (lead) silences have also been excluded (see Table 5.5).

Data from the remaining 1430 fricative tokens were analyzed using linear mixed effects regression analysis, as in Chapter 4. Separate models were fit to each of the following dependent measures: total (frication onset to vowel onset) duration, frication duration, and lag duration. Duration measures were log-transformed to improve the normality of their distributions. The model fitting procedure began with random intercepts specified for Word and Speaker to account for inherent differences between speakers and lexical items. The fixed effects structure included two-way interactions between all pairings of the independent variables (reference levels in bold): condition (**laryng.fricative**, fricative.first, simplex.fricative, or affricate), word position (**initial** or medial), place of articulation (**alveolar**, post-alveolar, or lateral), stress (of the following vowel; **unstressed** or stressed), and vowel laryngealization (of the following vowel; **no** or yes). Interactions between condition and place of articulation could not be assessed because there is no lateral affricate attested in the segmental inventory of UNT, resulting in an empty cell in the statistical model. The data set was insufficient in size to allow for reliable conclusions to be drawn with respect to three-way interactions between predictors, so such interactions were not included in the models. Random slopes by speaker, which allow the model to compensate for variance due to consistent patterns within a given speaker's output, were determined individually for each model by adding each predictor in the following order: condition, word position, laryngealization of the following vowel, place of articulation, and stress of the following vowel. After the addition of each random slope, the resulting model was compared to the previous, simpler model using ANOVA to determine the goodness of model fit (Baayen, 2006). The additional complexity in random structure was accepted into the model only if it resulted in significantly improved model fit. Interactions between factors were not tested in the random structure, again due to the limited size of the data set.

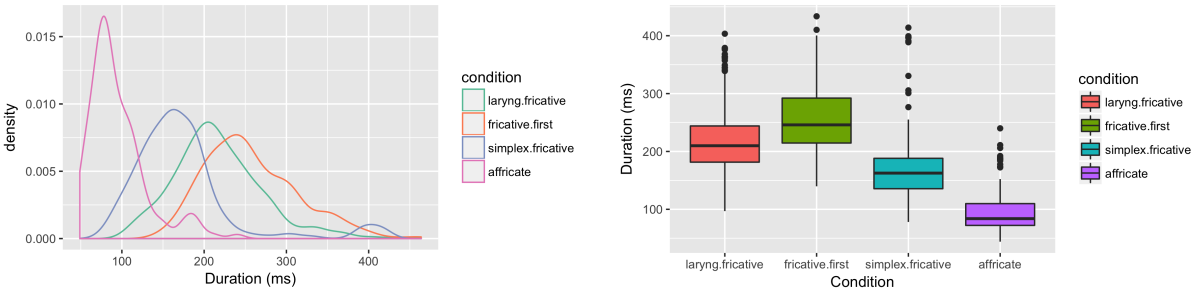
After constructing the maximal possible random structure that achieved the best model fit (within the limitations specified above), the fixed effects were then backward fitted by removing the least significant interactions and predictors, as long as doing so did not significantly decrease the goodness of model fit. Details of the fixed effect and random effect structures of each model are reported separately in the individual sections below. After arriving at the best fixed and random effects structures, each model was subjected to criticism in which residuals that fell outside ±2.5 standard deviations from the regression line were identified. The data points associated with these outlier residual values were then excluded, and the models refit to the remaining values to ensure that effects were not overly influenced by outliers, as well as to improve normality of the residual distribution.

Finally, after model fitting and model criticism were complete, multiple comparisons of conditional means were performed using the lsmeans package. Comparisons of least squares (LS) means allow for the partial effects of predictors to be evaluated against other partial effects without increasing the risk of Type I statistical errors.

Table 5.6: Total duration in four conditions. Summary statistics (means, medians, standard deviations, standard errors) are indicated for each condition as well as the overall distribution.

Total duration (ms)

	Mean	**Median**	**SD**	**SE**	***N***
affricate	98.36	83.97	58.29	4.74	151
simplex.fricative	179.15	163.62	73.33	5.24	196
laryng.fricative	218.09	210.29	59.76	2.41	615
fricative.first	259.72	247.45	66.49	3.08	465
Overall	213.63	209.59	79.63	2.11	1427



(a) Density plot of total durations by condition

(b) Total durations by condition. Differences between all pairs were significant.

Figure 5.5: Summary of total duration distributions by frication condition.

Total duration

The total durations from onset of frication to onset the following vowel across the condition types simplex.fricative, laryng.fricative, fricative.first and affricate are illustrated in Figure 5.5a. Affricates were found to have the shortest total durations. Within the remaining three conditions, fricative.first items were the longest, followed by laryng.fricative, and simplex.fricative, with considerable overlap between conditions. Descriptive statistics summarizing the distributions of durations in each condition are presented in Table 5.6. Three data points were excluded from the analysis because they were missing the lag duration data, bringing the total data set down to 1427 data points from 1430.

A linear mixed effects regression model was fit to the data as spelled out in section 5.4.3. The best model included interactions between condition and vowel laryngealization, condition and stress, condition and word position, and between vowel laryngealization and place. Although neither the effect of word position, nor the interaction between word position and condition were significant, removal of the interaction resulted in decreased goodness of fit. The interaction was therefore retained in the final model. The model specified random slopes by speaker for word position and place, meaning that each speaker might vary somewhat in the ways in which word position and place affect the total duration. Allowing for this by-speaker variation significantly improved the goodness of model fit. As a result of model criticism, 35 data points, or 2.45% of the data, were trimmed[3]. The trimmed model is summarized in Table 5.7.

The main finding of the model is a significant effect of condition. A pairwise comparison between least-squares means of all conditions revealed that all conditions differed significantly from each other ($p < 0.005$). Figure 5.5b illustrates the mean durations from onset of frication to onset of the following vowel in each condition. This figure, and all following figures, was generated from the trimmed data set, excluding data points that fell outside of ±2.5 standard deviations of the mean. The differences between the condition means is quite visibly apparent, with fricative.first items having longer average durations than laryng.fricatives, shorter durations in simplex.fricative and affricate conditions.

Significant main effects were also found for vowel laryngealization, stress, and place. These main effects participated in one or more two-way interactions, all of which were significant; these interactions are presented in Figure 5.6. Because of the presence of these interactions, interpretation of the main effects must be tempered by the comparison of conditional means performed with the *lsmeans* package. Rather than comparing the overall mean of a factor across conditions, we are instead able to compare means that have been averaged over the levels in factors that are not present in the interaction. This comparison allows us to see past some of the variability that is due to causes outside of the interactions of interest.

There was a significant interaction between vowel laryngealization and condition, illustrated in Figure 5.6a, with laryng.fricatives being significantly shorter than fricative.first items only when the following vowel was laryngealized (df = 250.17, t = -4.293,

---

[3] However, trimming the data did not result in normally distributed residuals, likely indicating that the model was missing information that would account for the variability.

$p < 0.001$). The effect of vowel laryngealization was significant only in simplex.fricative and affricate conditions. In the affricate condition, total duration was shorter when the following vowel was laryngealized (df = 728.51, t = 3.175, $p < 0.05$). In simplex fricatives, total duration was longer when the following vowel was laryngealized (df = 461.62, t = -3.167, $p < 0.05$).

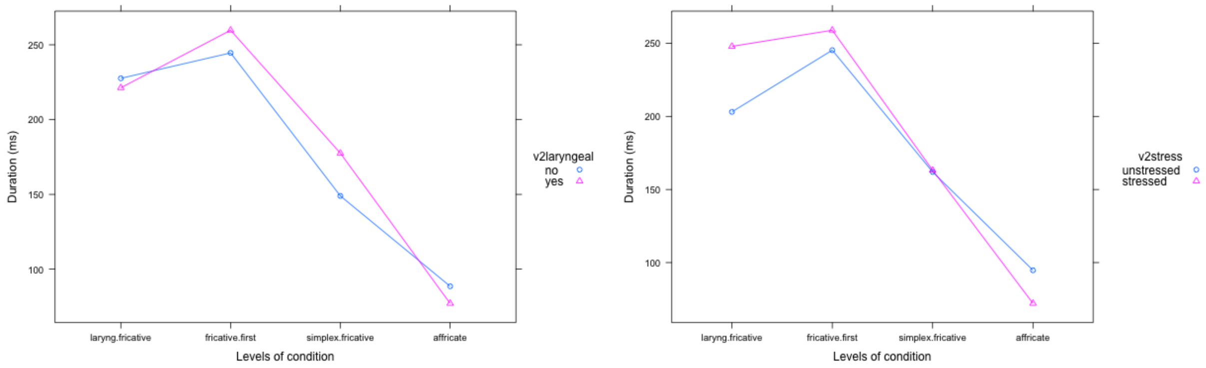
The interaction between condition and stress, illustrated in Figure 5.6b, revealed that the durational differences between the laryng.fricative and fricative.first conditions was only significant when the following vowel was unstressed (df = 139.43, t = -5.410, $p < 0.0001$). In addition, the effect of stress was significant only in the laryng.fricative (df = 133.05, t = -4.482, $p < 0.0005$) and affricate (df = 848.7, t = 5.024, $p < 0.0001$) conditions, where stress on the following vowel resulted in greater total duration, while in the affricate condition, a following stressed vowel decreased total duration.

The model also showed a significant interaction between vowel laryngealization and place of frication. The lsmeans comparison revealed that the effect of place was enhanced when the following vowel was not laryngealized. Despite this result, the effect was nevertheless quite small, as is indicated by the 30 ms range of the y-axis in Figure 5.6c, which illustrates the interaction between vowel laryngealization and place. The difference between alveolar and post-alveolar fricatives was greater before non-laryngealized vowels than before laryngealized vowels (df = 28.07, t = 3.585, $p < 0.05$). All other pairwise comparisons failed to reach significance.

The results of this model indicate that condition had a significant influence on total duration. Fricative + stop clusters had the longest durations, followed by ejective fricatives, simplex fricatives, and finally, affricates. Vowel laryngealization, stress, and place of frication also had significant effects. These predictors participated in interactions that showed that differences between the various factor levels only appeared significant in relatively few comparisons, with inconsistent results. The differences between laryng.fricative and fricative.first conditions were significant only before unstressed vowels and before laryngealized vowels. The main effect of stress was only significant in laryng.fricatives and affricates; vowel laryngealization was only significant in simplex.fricative and affricate conditions; the effect of place was enhanced between alveolar and post-alveolar fricatives when preceding a non-laryngealized vowel. Despite the interaction between place and vowel laryngealization, the effect was quite small, and no other place effect was significant.

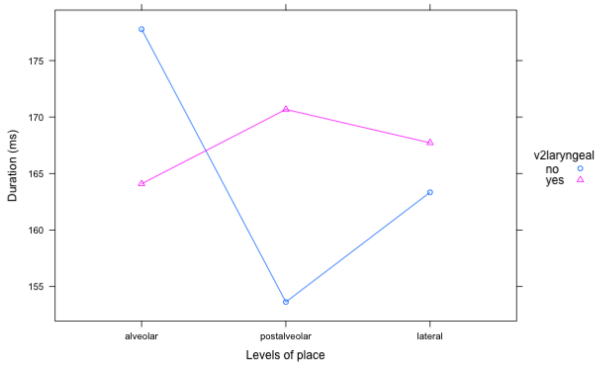
Table 5.7: Summary of linear mixed effects regression model of total duration from onset of frication to onset of following vowel (N = 1427).

coef	Estimate	Std. Error	df	t value	Pr(>—t—)	
(Intercept)	5.4677	0.0702	6.8185	77.8802	0.0000	***
conditionfricative.first	0.1438	0.0517	120.4509	2.7797	0.0063	**
conditionsimplex.fricative	−0.3024	0.0598	217.4712	−5.0608	0.0000	***
conditionaffricate	−0.7603	0.0632	245.0148	−12.0260	0.0000	***
v2laryngealyes	−0.1254	0.0381	527.3923	−3.2951	0.0010	**
v2stressstressed	0.1994	0.0430	136.0534	4.6328	0.0000	***
wordposmedial	−0.1259	0.0627	8.4812	−2.0085	0.0774	
placepostalveolar	−0.1460	0.0407	28.0098	−3.5879	0.0013	**
placelateral	−0.0847	0.0395	113.6976	−2.1435	0.0342	*
conditionfricative.first:v2laryngealyes	0.0882	0.0429	578.6095	2.0535	0.0405	*
conditionsimplex.fricative:v2laryngealyes	0.2035	0.0591	492.7427	3.4456	0.0006	***
conditionaffricate:v2laryngealyes	−0.1106	0.0612	742.6632	−1.8074	0.0711	
conditionfricative.first:v2stressstressed	−0.1453	0.0575	230.5044	−2.5291	0.0121	*
conditionsimplex.fricative:v2stressstressed	−0.1921	0.0870	143.5211	−2.2082	0.0288	*
conditionaffricate:v2stressstressed	−0.4729	0.0743	403.3724	−6.3671	0.0000	***
conditionfricative.first:wordposmedial	0.0017	0.0592	148.7945	0.0287	0.9771	
conditionsimplex.fricative:wordposmedial	−0.0507	0.0718	164.9282	−0.7063	0.4810	
conditionaffricate:wordposmedial	0.1060	0.0801	274.9668	1.3234	0.1868	
v2laryngealyes:placepostalveolar	0.1854	0.0455	474.7367	4.0715	0.0001	***
v2laryngealyes:placelateral	0.1066	0.0502	565.1744	2.1219	0.0343	*



(a) Interaction plot of condition x vowel laryngealization.

(b) Interaction plot of condition x stress.



(c) Interaction plot of place x vowel laryngealization.

Figure 5.6: Interaction effects from model of total duration

Table 5.8: Frication duration in four conditions. Summary statistics (means, medians, standard deviations, standard errors) are indicated for each condition.

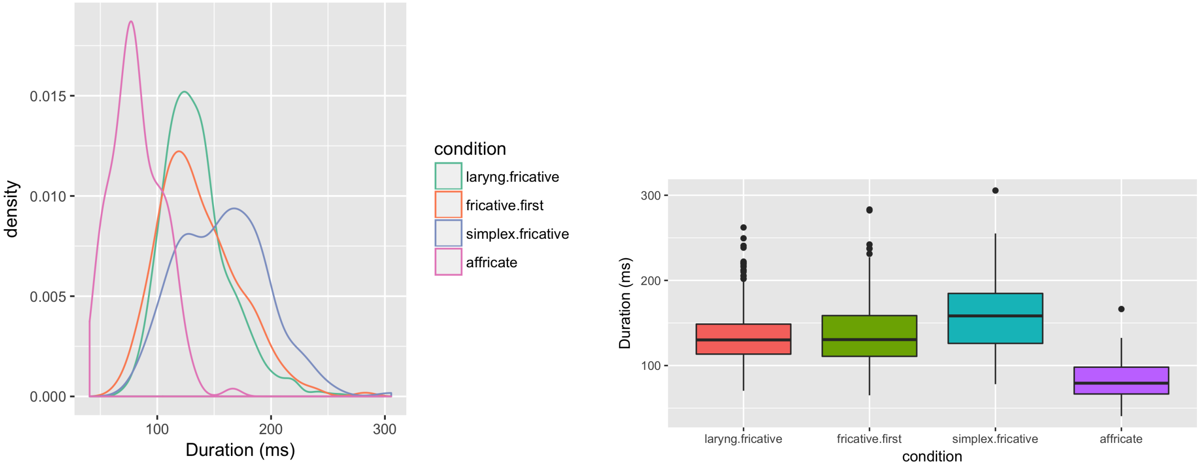
	Duration (ms)				
	Mean	**Median**	**SD**	**SE**	***N***
affricate	87.28	79.17	54.56	4.44	151
simplex.fricative	160.55	158.63	45.83	3.27	196
laryng.fricative	135.38	130.2	34.12	1.38	615
fricative.first	137.14	130.55	37.46	1.73	468
Overall	134.33	128.86	43.47	1.15	1430

Frication duration

The distributions of frication durations for each of the four conditions is illustrated in a density plot in Figure 5.7a. The data appear to be somewhat divided by condition, with affricates having the shortest frication intervals and simplex.fricatives the longest. Simplex.fricatives are also widely distributed in the range of durations, with the suggestion of a bimodal distribution. Laryng.fricatives and fricative.first segments appear to have frication intervals of comparable duration, continuing the pattern of similarities between the laryngealized fricatives and fricative-initial clusters. There is substantial overlap between all conditions except for affricates, which are markedly shorter. Table 5.8 provides summary statistics of duration data for frication in the four segment conditions.

The best linear mixed-effects model of frication duration included random intercepts by speaker and word, random slopes for vowel laryngealization and word position by speaker, and interactions among the fixed effects between condition and vowel laryngealization, condition and stress, condition and word position, and between vowel laryngealization and place of articulation. Trimming removed 27 data points, corresponding to 1.89% of the data, and resulted in normally distributed residuals. The trimmed model is summarized in Table 5.9.

The most important result of the model is the significant main effect of condition. A pairwise comparison of least-squares means between all four conditions showed significant differences between all conditions ($p < 0.005$) except between laryng.fricative



(a) Density plot of frication intervals.

(b) Frication durations by condition. Differences between all pairs were significant except between laryng.fricative and fricative.first conditions.

Figure 5.7: Summary of frication duration distributions by frication condition.

and fricative.first conditions (df = 148.17, t = 1.457, p = 0.47). Frication was longest in simplex.fricatives, and shortest in affricates, with laryng.fricatives and fricative.first clusters falling between the other two categories. Figure 5.7b illustrates the distributions of duration of frication in each of the four conditions. The similarity between the laryng.fricative and fricative.first categories is visually apparent, as are the differences between all other conditions.

All main effects were found to be significant, but except for condition the size of the main effects were relatively small. All of the main effects also participated in at least one significant interaction, illustrated in Figure 5.8, complicating the a priori interpretation of the main effects. In order to investigate the interactions, pairwise comparisons were performed between all combinations of factors participating in significant interactions.

The main effect of stress was significant (df = 305.07, t = 3.003, $p < 0.005$), with frication having longer duration before stressed vowels than unstressed vowels. The interaction between condition and stress was assessed in a paired lsmeans comparison that found that although all fricatives tended to show the effect of stress, the difference in duration before stressed and unstressed vowels was only significant in the laryng.fricative condition (df = 146.3, t = 4.314, $p < 0.001$). Figure 5.8a illustrates the interaction between condition and stress.

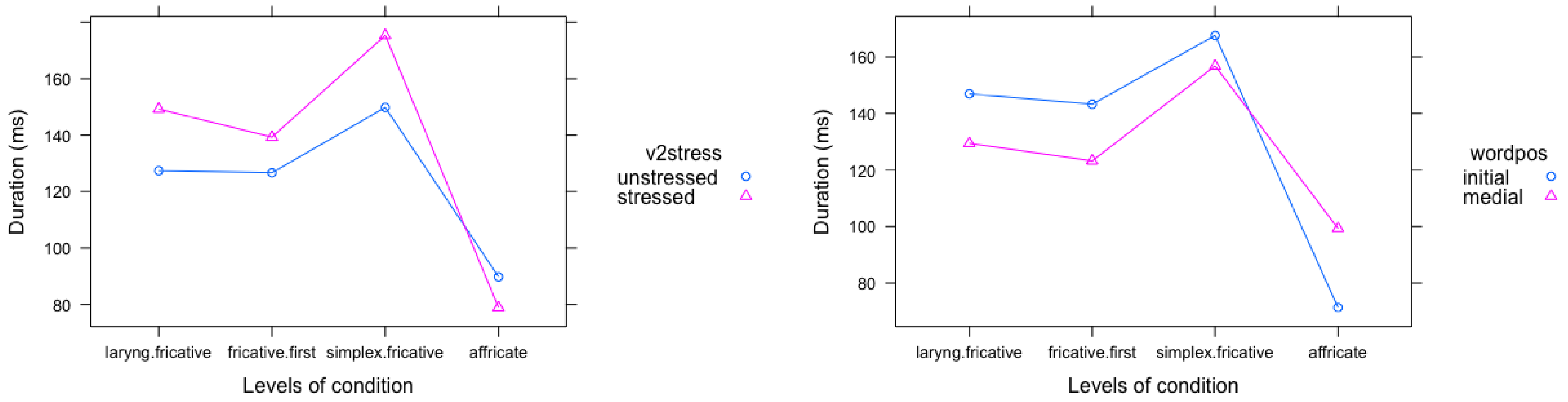
The main effect of word position approached but did not attain significance in post hoc tests, with shorter frication occurring in word medial position. Figure 5.8b shows the LS mean comparisons across word position and condition levels. Pairwise comparison of conditional means revealed that the effect of word position was only significant in the affricate condition, where it showed a pattern opposite that of the other conditions (df = 23.86, t = 4.356, $p < 0.005$). All other conditions differed from affricate durations in both initial and medial position, but notably the laryng.fricative and fricative.first conditions did not differ from each other in either word position.

The main effect of vowel laryngealization was not found to be significant within any of the four condition levels in post hoc comparisons. The interaction between vowel laryngealization and condition is shown in Figure 5.8c. None of the within-condition pairs differed significantly based on vowel laryngealization, although the effect of vowel laryngealization did approach significance in the simplex.fricative condition (df = 60.26, t = -3.047, $p = 0.0634$). Laryng.fricatives did not differ from fricative.first items regardless of of vowel laryngealization.

Figure 5.8d illustrates the subtle interaction between the factors place and vowel laryngealization. Post-alveolar and lateral fricatives were longer before laryngealized vowels, while alveolar fricatives were longer before non-laryngealized vowels. The differences between vowel laryngealization pairs were only significant in the post-alveolar place, however (df = 17.34, t = 3.399, $p < 0.05$), with a very small magnitude of difference between the vowel laryngealization conditions.

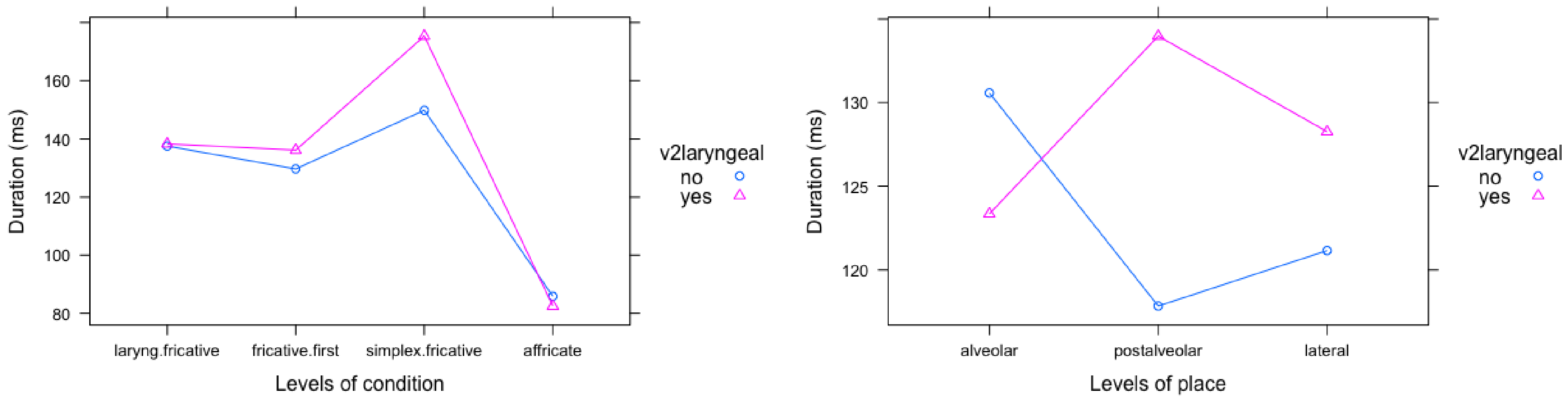
Table 5.9: Summary of linear mixed effects regression model of frication duration (N = 1430).

coef	Estimate	Std. Error	df	t value	Pr(>—t—)	
(Intercept)	4.9672	0.0740	4.9433	67.0861	0.0000	***
conditionfricative.first	−0.0152	0.0437	136.3694	−0.3481	0.7283	
conditionsimplex.fricative	0.0562	0.0522	225.2882	1.0752	0.2835	
conditionaffricate	−0.5560	0.0560	262.1392	−9.9350	0.0000	***
v2laryngealyes	−0.0941	0.0410	24.1915	−2.2921	0.0309	*
v2stressstressed	0.1582	0.0367	146.3012	4.3149	0.0000	***
wordposmedial	−0.1269	0.0552	7.2171	−2.3005	0.0539	*
placepostalveolar	−0.1026	0.0311	240.4922	−3.3037	0.0011	**
placelateral	−0.0749	0.0335	211.5120	−2.2353	0.0264	*
conditionfricative.first:v2laryngealyes	0.0428	0.0390	490.2559	1.0990	0.2723	
conditionsimplex.fricative:v2laryngealyes	0.1516	0.0534	438.6484	2.8383	0.0047	**
conditionaffricate:v2laryngealyes	−0.0458	0.0567	654.4760	−0.8077	0.4196	
conditionfricative.first:v2stressstressed	−0.0631	0.0503	244.9984	−1.2543	0.2109	
conditionsimplex.fricative:v2stressstressed	−0.0009	0.0741	154.8787	−0.0117	0.9907	
conditionaffricate:v2stressstressed	−0.2871	0.0677	452.4925	−4.2419	0.0000	***
conditionfricative.first:wordposmedial	−0.0239	0.0502	166.8000	−0.4748	0.6356	
conditionsimplex.fricative:wordposmedial	0.0603	0.0616	177.9500	0.9792	0.3288	
conditionaffricate:wordposmedial	0.4576	0.0713	311.3013	6.4141	0.0000	***
v2laryngealyes:placepostalveolar	0.1852	0.0412	411.2193	4.4943	0.0000	***
v2laryngealyes:placelateral	0.1139	0.0458	473.6831	2.4881	0.0132	*



(a) Interaction plot of condition and stress. Stress was only significant among laryngealized fricatives.

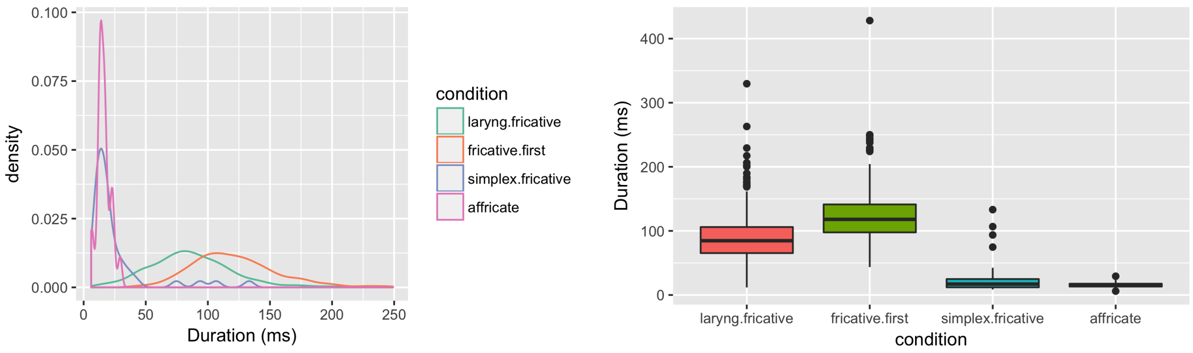
(b) Interaction effect of condition and word position. Word position was only significant among affricates.



(c) Interaction effect of condition and vowel laryngealization. No pairings within conditions were significant, but simplex.fricatives approached significance at $p = 0.06$.

(d) Interaction effect of vowel laryngealization and place on frication duration. Vowel laryngealization was only significant among post-alveolar fricatives.

Figure 5.8: Interaction effects from model of frication duration



(a) Density plot of lag durations by condition.

(b) Lag durations by condition.

Figure 5.9: Summary of lag duration distributions by frication condition.

The results of this model indicate that condition is again a significant predictor of duration. However, unlike total duration, frication duration did not differ between laryng.fricative and fricative.first conditions regardless of interactions with other predictors. The model further contained interactions and showed effects of stress of the following vowel, laryngealization of the following vowel and position in the word, but these effects did not produce consistent significant differences for all conditions.

Post-frication lag duration

A total of 1107, or roughly 77%, of frication tokens were produced with post-frication lag. The majority of lag productions occurred in the fricative.first (463) and laryng.fricative (582) conditions. The remaining tokens were produced across the simplex.fricative (43) and affricate (22) conditions. Lag periods were considered to include any periods of silence/closure as well as consonant release bursts or other erratic sounds, such as glottal bursts, that occurred before the onset of modal vowel phonation. Figure 5.9a illustrates the distributions of duration data across all four conditions. Table 5.10 provides summary statistics of lag duration in each of the four conditions. Lag durations in affricate and simplex.fricative conditions were quite short in addition to being less frequent overall. Fricative.first lags appear to be somewhat longer than laryng.fricative lags, but both distributions are very widely spread with a high degree of overlap between the two conditions.

A linear mixed-effects regression model was fit to the lag data from all four conditions. The best model included main effects of condition, frication place, and word

Table 5.10: Lag duration in four conditions. Summary statistics (means, medians, standard deviations, standard errors) are indicated for each condition as well as the overall distribution.

	Duration (ms)				
	Mean	**Median**	**SD**	**SE**	***N***
affricate	16.11	15.38	5.4	1.15	22
simplex.fricative	25.67	17.02	26.75	4.08	43
laryng.fricative	87.91	84.45	38.92	1.61	582
fricative.first	123.89	117.98	41.95	1.96	460
Overall	99.01	96.87	47.14	1.42	1107

position, as well as interactions between place of frication and word position, place of frication and stress, and word position and stress. Random slopes of condition, word position, and stress by speaker were also included. Interactions with condition were not included in the model due to the small numbers of tokens in affricate and simplex.fricative conditions. The model was trimmed during criticism, resulting in the removal of 28, or 2.53%, of the data points. The trimmed model is summarized in Table 5.11.

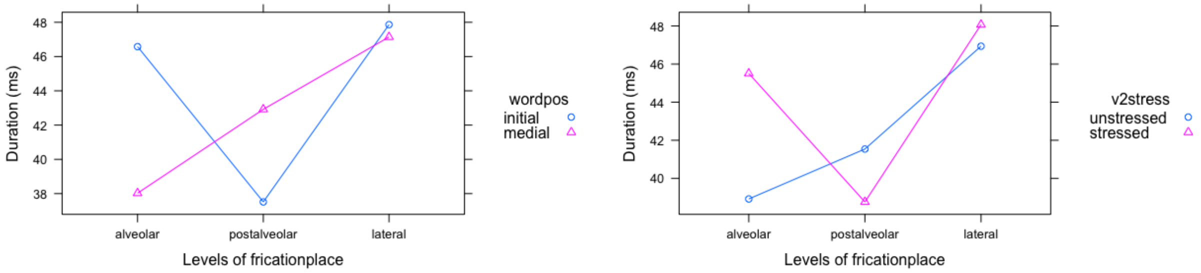
As with previous models, the main effect of condition was highly significant. A pairwise comparison revealed that all conditions differed significantly from each other except for the simplex.fricative and affricate conditions. The effect of condition is illustrated in Figure 5.9b, which shows that laryng.fricative lag periods are shorter than those of fricative.first condition. Both laryng.fricative and fricative.first items have longer lags than either of the simplex.fricative or affricate conditions, which are of comparable duration.

Table 5.11: Summary of linear mixed effects regression model of lag duration including all four conditions (N = 1107).

coef	Estimate	Std. Error	df	t value	Pr(>—t—)	
(Intercept)	4.5332	0.0878	10.9339	51.6322	0.0000	***
conditionfricative.first	0.3402	0.0805	6.5995	4.2252	0.0045	**
conditionsimplex.fricative	−1.2863	0.2167	3.8357	−5.9356	0.0046	**
conditionaffricate	−1.5736	0.1443	3.9165	−10.9087	0.0004	***
fricationplacepostalveolar	−0.1036	0.0738	163.9262	−1.4037	0.1623	
fricationplacelateral	0.0935	0.0827	106.3854	1.1303	0.2609	
wordposmedial	−0.3994	0.0846	69.9077	−4.7204	0.0000	***
v2stressstressed	−0.0400	0.0998	13.7680	−0.4012	0.6944	
v2laryngealyes	−0.0843	0.0351	402.9234	−2.4028	0.0167	*
fricationplacepostalveolar:wordposmedial	0.3375	0.1100	138.0113	3.0697	0.0026	**
fricationplacelateral:wordposmedial	0.1877	0.1162	128.6923	1.6151	0.1087	
fricationplacepostalveolar:v2stressstressed	−0.2256	0.1077	273.3722	−2.0947	0.0371	*
fricationplacelateral:v2stressstressed	−0.1326	0.1052	227.2147	−1.2605	0.2088	
wordposmedial:v2stressstressed	0.3929	0.1014	209.4187	3.8760	0.0001	***

Main effects of word position, stress and vowel laryngealization were also significant, with slightly longer durations occurring in word initial position, before stressed vowels, and before non-laryngealized vowels. The model also showed that all included interactions, illustrated in Figure 5.10, were also significant. None of the word position pairs differed from each other within frication place, but in post-alveolar fricatives, the direction of the effect was opposite to those in alveolar and lateral places. A similar pattern emerged in the interaction between place of frication and stress. This change in sign across places of frication is likely the source of the significant interaction effect in both cases. However, the magnitude of these effects was quite small, with the greatest duration differences only amounting to about 10 ms between factor levels. It is unlikely that speakers are able to make use of such small differences in duration (cf. Hawkins 1977, who reports just noticeable difference on the order of 25 ms), so while these results may be significant, they are of questionable importance. Likewise, the interaction between stress and word position indicates that initial fricatives had longer lags before unstressed vowels and shorter lags before stressed vowels, and vice versa for medial fricatives, but the effect is again so small (around 10 ms) as to be disregarded as unlikely to be useful in identifying these segments.

In sum, the results of the model in Table 5.11 indicates a strong effect of condition once again. Laryng.fricative and fricative.first conditions differ from one another as well as from the simplex.fricative and affricate conditions, which are not statistically different from each other. Although other effects were reported in the model as significant, post hoc tests show that pairs within factors rarely differ significantly. When they do, the effects are exceedingly small.



(a) Interaction effect of frication place and word position on lag duration across all four conditions.

(b) Interaction effect of frication place and stress on lag duration across all four conditions.

(c) Interaction effect of stress and word position on lag duration across all four conditions.

Figure 5.10: Interaction effects from lag duration model summarized in Table 5.11

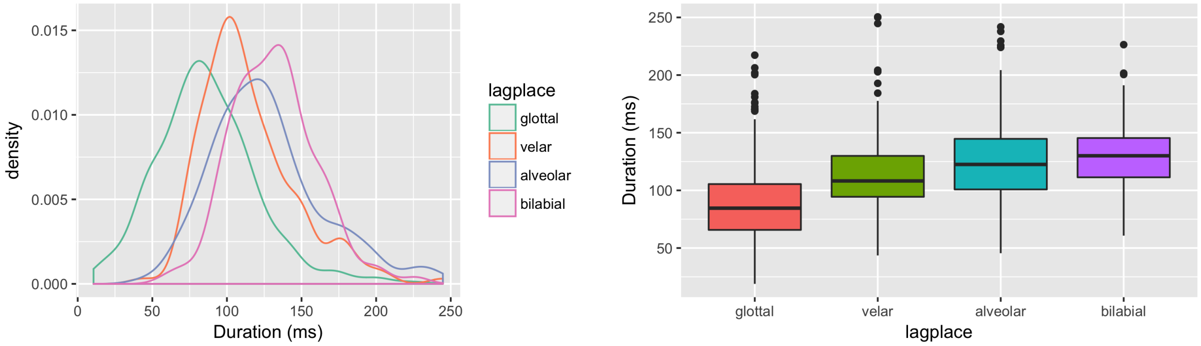
Post-frication lag in ejectives and clusters

In the preceding model of lag duration, the effect of condition was highly significant. However, the conditions appear to fall into two groups, with simplex.fricatives and affricates on one side, and laryng.fricatives and fricative.first clusters on the other. Despite finding significant differences between all conditions, the differences between these two groups is rather larger than the differences within them. In addition, the frequencies of lag occurrence in each condition were highly imbalanced, with far fewer affricate and simplex.fricatives items than laryng.fricative and fricative.first items, perhaps undermining the efficacy of the model in characterizing the data. Because of this imbalance, a further analysis was conducted on a data set restricted to only the two conditions with high numbers of lag periods: laryng.fricative and fricative.first (N = 1042).

In order to further investigate the potential differences among laryng.fricatives and fricative.first tokens, the lag durations were classified according to their place of articulation. In clusters, this place is identified according to the identity of the stop consonant; in laryng.fricatives, the place was identified as 'glottal'. A new factor, called *lagplace* was added to the restricted data set to encode these places (bilabial, alveolar, velar, glottal, with **glottal** serving as the reference level). This factor will allow for a comparison to previous findings relating place of articulation to stop closure duration. In addition to the laryng.fricatives item, 41 fricative.first items were also produced with a following glottal closure as a result of the cluster occurring across a morpheme boundary. In order to avoid conflating fricative + glottal stop clusters with potential ejective fricatives, these items were removed from the dataset.

Summary statistics for the remaining 1001 observations across lag place values are summarized in Table 5.12. The distributions of lag durations by place of closure are presented in Figure 5.11. Note that the place of frication was also included in the present model, as in previous models.

A linear mixed effects regression model was built for duration of post-frication lags in laryng.fricative and fricative.first items, coded for place of lag closure as described above. The duration data was log-transformed before model fitting began. The factors lag place, frication place, word position, stress, and vowel laryngealization were included



(a) Density plot of lag durations by place of closure.

(b) Lag durations by place of closure.

Figure 5.11: Summary of lag duration distributions included in model summarized in Table 5.13.

Table 5.12: Lag duration in four closure places. Summary statistics (means, medians, standard deviations, standard errors) are indicated for each closure place as well as the overall distribution.

	Duration (ms)				
	Mean	**Median**	**SD**	**SE**	***N***
bilabial	133.95	130.53	39.43	3.68	115
alveolar	129.77	122.87	47.24	3.79	155
velar	116.54	108.19	35.04	2.87	149
glottal	87.91	84.45	38.92	1.61	582
Overall	103.92	99.16	44.33	1.40	1001

as main effects in the best model. Random slopes of lag place and stress by speaker were also included, meaning that speakers might vary in the ways in which stress and lag place affect their outputs. The model also included interactions between lag place and stress, and frication place and word position. Model trimming removed 30 data points, or 3% of the data.[4] The trimmed model is summarized in Table 5.13.

The model showed a significant effect of lag place, with durations increasing as place of lag closure moved further forward in the vocal tract. A pairwise comparison within the levels of lag place showed that these durations differed significantly between the bilabial and glottal places (df = 13.30, t = -4.482, $p < 0.005$), and between glottal and alveolar places (df = 6.19, t = -3.45, $p < 0.05$). No other pairs differed significantly from one another. Main effects of stress, word position, and vowel laryngealization were also observed. Lags were longer before stressed vowels, shorter in word medial position, and shorter before laryngealized vowels.

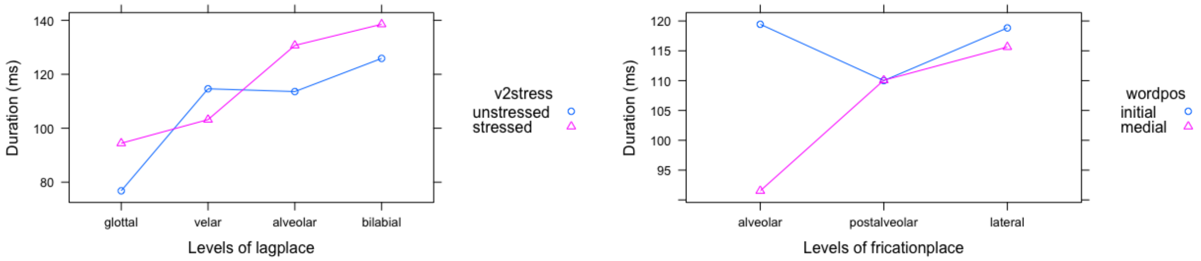
The model also showed two significant interactions, shown in Figure 5.12. The first occurred between lag place and stress. While there is a general trend of longer lag durations preceding stressed vowels, this pattern was reversed when the lag closure occurred at the velar place. Once again, however, this effect was very small (see Figure 5.12a). The second interaction took place between word position and frication place, illustrated in Figure 5.12b. This interaction revealed that the effect of word position on lag durations was only present when frication was produced at the alveolar place (df = 83.41, t = 3.858, $p < 0.005$). Word initial tokens with an alveolar place of frication have far longer lag durations (+30 ms) than medial tokens.

---

[4] Although normality was improved, trimming did not result in normal residuals.

Table 5.13: Summary of linear mixed effects regression model of lag duration of laryng.fricative and fricative.first conditions only. Fricative.first clusters with lags occurring at the glottal place have also been removed (N = 1001).

coef	Estimate	Std. Error	df	t value	Pr(>—t—)	
(Intercept)	4.4709	0.0817	9.2411	54.7242	0.0000	***
lagplacevelar	0.4004	0.1100	6.3525	3.6387	0.0098	**
lagplacealveolar	0.3914	0.1052	6.5248	3.7189	0.0085	**
lagplacebilabial	0.4941	0.0909	10.0492	5.4357	0.0003	***
v2stressstressed	0.2067	0.0698	7.4221	2.9625	0.0197	*
fricationplacepostalveolar	−0.0823	0.0656	106.8316	−1.2551	0.2122	
fricationplacelateral	−0.0052	0.0674	80.8360	−0.0764	0.9393	
wordposmedial	−0.2662	0.0690	83.4076	−3.8584	0.0002	***
v2laryngealyes	−0.1034	0.0312	317.8737	−3.3087	0.0010	**
lagplacevelar:v2stressstressed	−0.3118	0.0941	295.5765	−3.3147	0.0010	**
lagplacealveolar:v2stressstressed	−0.0665	0.0969	257.5402	−0.6861	0.4933	
lagplacebilabial:v2stressstressed	−0.1110	0.1323	275.0579	−0.8389	0.4022	
fricationplacepostalveolar:wordposmedial	0.2667	0.0925	92.8476	2.8837	0.0049	**
fricationplacelateral:wordposmedial	0.2389	0.1033	82.8800	2.3138	0.0232	*



(a) Interaction of lag place and stress in laryng.fricative and cluster conditions only.

(b) Interaction of frication place and word position in laryng.fricative and cluster conditions only.

Figure 5.12: Interaction effects from lag duration model summarized in Table 5.13.

In summary, this model took into account the place of closure in fricative.first clusters, and analyzed laryng.fricative items as though they were clusters with a place of closure at the glottis. This analysis found few differences between lag durations according to the place of closure associated with the lag period: laryngealized fricatives were significantly shorter than fricatives at bilabial and alveolar places only. These differences were part of a possible continuum of closure duration from front to back of the vocal tract. Vowel laryngealization had a small but significant effect on lag duration, with laryngealized vowels leading to shorter lag durations. The interaction between place of frication and stress of the following vowel was also significant, with slightly longer lag durations in alveolar and lateral fricatives before stressed vowels, and slightly shorter lag durations in post-alveolar fricatives before stressed vowels. The interaction between frication place and word position revealed longer lag durations word initially than word medially, but only for alveolar fricatives.

5.5 Discussion

This section summarizes the findings from previous sections and relates them to the hypotheses as they were set out at the beginning of the chapter. Table 5.14 recapitulates the present findings, as well as the data from Maddieson et al. (2001) and Beck (2006), for ease of comparison.

Overall, laryng.fricatives did not exhibit the expected durational characteristics of ejective fricatives. Frication duration was comparable to fricatives preceding oral

Table 5.14: Summary of findings from Maddieson et al. (2001), Beck (2006), and the present paper. Data represent mean durations in milliseconds.

Language	Source	Condition	Frication	Lag	Total
UNT	Beck (2006)	Ejective	143	9	152
		Simplex	96	3	99
		F.first	101	27	–
UNT	Puderbaugh (2019)	Ejective	135	88	218
		Simplex	161	26	179
		F.first	137	124	260
Tlingit	Maddieson et al. (2001)	Ejective	148	46	194
		Simplex	222	1	222
		F.first	–	–	–

stops in clusters. Frication was longer in the simplex.fricative condition than in the laryng.fricative condition, but the same was true when comparing simplex.fricatives before vowels to fricative.first clusters. Total duration differed across all conditions, largely due to differences in lag durations. Post-frication lag silences indeed might indicate a difference between simplex (pulmonic) and ejective fricatives, but the inclusion of fricatives in clusters to the analysis suggests an alternate explanation: segments previously described as ejective fricatives can be analysed as sequences of fricatives + stops, just as they were at their historical origins. The duration of glottal closures fits neatly into a cross-linguistically attested pattern of shorter durations at places of articulation that are further back in the vocal tract and would be consistent with an analysis of these sequences as clusters rather than ejective fricatives (Repp, 1984; Chao & Chen, 2008). This analysis would constitute an alternative to positing the existence of a separate and (cross-linguistically) unusual sound class in UNT, for which the present data do not provide empirical support.

5.5.1 Review of hypotheses with current findings

Frication duration Frication was expected to differ in duration between ejective and pulmonic fricative categories. Specifically, ejective fricatives were expected to have shorter frication intervals than pulmonic fricatives based on the constraints of the glottalic airstream mechanism and findings on other languages (Maddieson et al., 2001). This was not found to be the case. Ejective fricatives did not differ significantly in frication duration from pulmonic fricatives regardless of whether the pulmonic fricative occurred in fricative + stop clusters or pre-vocalically (as in simplex pulmonic fricatives). On the other hand, affricates, which are undisputed in their status as complex, doubly-articulated segments, were in fact produced with shorter frication intervals.

Presence and duration of lags Ejective fricatives are generally expected to be produced with lag silences more often than frication in simplex pulmonic conditions due to physiological constraints on the resumption of vocal fold vibration after release of glottal closure. This was also the expectation in the current study, and it was indeed borne out: laryng.fricatives were produced with lags 95% of the time, compared to only 21.9% of simplex.fricatives, and 14.6% of affricates (cf. Table 5.5). However, fricative.first clusters were produced with lag silence 99% of the time, a rate comparable to that of laryng.fricative tokens.

A further expectation was that lags would be longer for ejective fricatives than simplex fricatives. Again this expectation was met: lag periods were substantially longer in the laryng.fricative condition than in simplex.fricative, where they were both highly infrequent, and exceedingly brief even when present. However, there is another possible source of lag silences, namely closure that occurs during the production of fricative + stop clusters. In the case of laryng.fricatives, the closure place was coded as 'glottal', as indeed it would be regardless of whether they are truly ejective or in fact fricative + glottal stop clusters. Analyzing the lags in this way revealed that while the laryng.fricatives have the shortest lag durations, they were not significantly different from lags resulting from velar stop closure in clusters. In fact, no two adjacent places of articulation differed significantly in terms of lag duration, suggesting a gradual continuum with long closures occur at the bilabial place, and shorter durations further back in the vocal tract.

Total duration Based on prior findings in Maddieson et al. (2001), total duration was expected to be roughly equal between laryng.fricative and simplex.fricative segments when considering the total time from onset of frication to onset of the following vowel. These total durations were compared for simple pulmonic fricatives before vowels, pulmonic fricatives before stops as part of a cluster, and ejective fricatives before vowels. The analysis found that ejective fricatives were significantly longer than simple pulmonic fricatives, and similar in duration to fricative + stop clusters. Each of these three conditions was significantly different from the others, a question that was further addressed by way of separate analyses of frication and lag durations in each condition.

Frication duration Previous studies found that frication duration differed between ejective fricatives and pulmonic fricatives (Maddieson et al., 2001). Beck (2006) found frication duration in ejective fricatives to be distinct from both intervocalic pulmonic fricatives, and pulmonic fricatives preceding stops as the first element of a fricative + stop cluster. In the present chapter, duration measures of laryng.fricatives were found to overlap substantially with simplex.fricative or fricative.first conditions in most cases. In fact, none of the three conditions differed significantly from the others in the analysis of frication duration. These similarities suggest the same airstream production mechanism across all three frication conditions. The differences in total durations can be attributed almost entirely to the varying durations of post-frication lag silences.

Effect of place Place of frication was not expected to affect frication duration. Indeed, in line with previous findings in Tlingit, place of frication did not appear to have much of an effect. In the analysis of frication duration, place differences, although significant, were so small that they are unlikely to be relevant to speakers or linguistic descriptions.

Place of closure was also included in the analysis of lag durations. Although place of frication did not have an effect on lag duration, place of closure did. By analyzing the laryng.fricatives with a glottal place of closure, their lag intervals fit nicely into a continuum of stop closure durations, starting with longer closures for bilabial stops, and gradually shortening at each subsequent position further back in the vocal tract.

The establishment of this pattern allows for an interpretation of laryng.fricatives as cluster sequences rather than complex glottalic segments.

Effect of external factors on frication duration and lag silences Previous studies found an effect of word position in Tlingit, namely that word final pulmonic fricatives had longer frication intervals than non-final ones, and that word final ejective fricatives were preceded, rather than followed, by a period of silence that effectively increased their total durations as well. In the present chapter, word position did not significantly affect frication duration between word initial and word medial tokens, although a tendency for longer frication to occur in word initial position was observed. Lag duration varied according to place of frication and word position. While no word position differences were observed in the post-alveolar and lateral places, alveolar fricatives were significantly shorter in word medial than in word initial position.

Some differences in lag durations have previously been attributed to factors occurring outside the immediate segment, particularly in adjacent vowels. In the present chapter, vowel laryngealization and stress of the following vowel were included in all component analyses of duration in ejective and pulmonic fricatives. Although they improved the model fit and were sometimes found to be significant predictors, the effects were always small and often inconsistent across levels of other predictors.

The effect of vowel laryngealization appears to be limited to the lag period only and does not extend to the duration of preceding frication. Stress, on the other hand, appeared to have more uniform effects whereby fricatives before stressed vowels tended toward longer durations of both frication and lag intervals, although again the magnitude of the effect was very small.

5.5.2 Future research

While the present study has focused on duration as a potential distinguishing characteristic between proposed fricative categories in UNT, many acoustic, impressionistic, articulatory, and aerodynamic measures have been proposed and reported as potentially distinguishing ejective consonants from pulmonic consonants, and fricatives from each other.

Articulatory and aerodynamic data could contribute substantially to a clearer picture of the fricatives in UNT, whether they are ejective, pulmonic, or in clusters. These data may be related to air pressure and flow, and to observations of the state of the glottis during production. Some ejective fricatives and affricates are described qualitatively as having a scraping or pulsing quality (Maddieson, 2001; Grawunder et al., 2010), or with visible larynx raising (Grawunder et al., 2010; Beck, 2006). Articulatory measures such as laryngographs, electroglottographs, and measures of air pressure and airflow are also sometimes useful in distinguishing ejective from pulmonic consonants (Grawunder et al., 2010) by determining the state and position of the glottis during production.

Two examples of methods for observing glottal states are laryngoscopy, which provides a means of directly observing the glottis, and electroglottography, which is an indirect method. I briefly address the potential contributions of each of these measures here. Air pressure and flow are relevant to the description of fricatives because airflow is the source of their characteristic frication noise. During fricative production, air is forced through a narrow passage to generate turbulence. Naturally, this requires an increase in air pressure which is sustained in pulmonic fricatives by the careful control of a large reservoir of air in the lungs. The glottalic mechanism does not have access to that reservoir, yet must generate enough pressure to create turbulence. Since the volume of air is small, the speaker will not be able to sustain such high pressure for long. As a result, we expect to see differences in air pressure and flow during ejective fricatives vis-à-vis pulmonic fricatives. Whereas pulmonic fricatives will begin with high pressure and maintain a steady pressure throughout, ejective fricatives will begin with a peak in air pressure that quickly dies off. Maddieson et al. (2001) demonstrated that this pattern may hold even if the frication extends for a longer time interval than might be expected when generated by the glottalic mechanism. Aerodynamic data can therefore provide indirect evidence of either glottalic or pulmonic airflow. This in conjunction with information about the timing and duration of frication and adjacent silences can give a clearer picture of the production mechanism. In the Tlingit data, the long frication durations seemed to suggest that ejectivity was unlikely, but the airflow data provided overriding evidence in favor of ejectivity. The same could potentially be found in UNT.

Another observational method that may be potentially useful to the description of ejectives is laryngoscopy. Here, a small camera is passed into the pharyngeal cavity through the nose in order to provide a means of observing directly the usually hidden workings of the glottis during speech (Esling, 1999; Esling et al., 2005). By simultaneously recording audio and video, a precise determination of the timing of glottal closure with respect to frication production may be established. This could also illuminate the process that allows for sustained frication after the initial glottal pump such as occurs in Tlingit.

Finally, electroglottagraphy (EGG) measures glottal aperture by electrical impedence through the skin and fleshy tissues of the neck and throat. When the glottis is closed, impedance of a weak electrical signal is lowered, allowing the amplitude of the EGG waveform to peak. This is useful information for the study of ejectives, because the timing of glottal closure and subsequent movement of the larynx is key for the generation of the pressure necessary to initiate airflow. Simpson & Brandt (2019) have found that the presumed glottal pump, achieved by raising the larynx while the glottis and mouth are closed, may not be necessary to achieve the effect of ejective stops. The same may prove to be true for ejective fricatives.

In the acoustic domain there is also the potential for further study. Acoustic descriptions may focus on measures of the consonant itself such as intensity of either the release burst or frication, and center of gravity or spectral characteristics of frication, as well as measures of surrounding vowels such as H1-H2, Harmonic-to-Noise Ratio.

Spectral measures and parameters may also illuminate the comparison of ejective and pulmonic fricatives. In a discriminant analysis of obstruents in English, Forrest (1988) found that sibilant fricatives were well classified by a moments analysis of two spectral cross sections covering the first 20 ms of fricative noise. Moments are statistical parameters that measure characteristics of a probability density distribution. These moments, which include the mean, variance, skewness, and kurtosis of the probability density function, are useful for the classification of speech sounds because they are mathematically derived and do not require subjective human judgment. In a moment analysis of an acoustic spectrum, the moments act as indices of spectral shape. Forrest (1988) examined three of the first four moments of fricative spectra: mean, skewness and kurtosis. The mean is the average value of the distribution; skewness refers to the symmetry of the distribution, with long right tails resulting in positive skewness, long left tails negative skewness; and kurtosis is a measure of how "peaky" or how sharp or flat the distribution is, with positive values indicating a distribution with a sharper peak. Using these three moments, a discriminant analysis was able to classify 98% of sibilant fricative tokens in the sample when the analysis was performed on Bark transformed frequency data. Although similar analyses were quite successful in classifying place of articulation in voiceless stops for both linear and Bark transformed data, the fricative analysis was only successful when the sibilants were analyzed separately from non-sibilant fricatives.

Further spectral analyses of fricatives appear in Jesus & Shadle (2002), who investigated voiced and voiceless fricatives in European Portuguese using three spectral parameters, along with derivatives and combinations of them. The parameters capture the maximum amplitude, spectral slope and dynamic amplitude. These parameters were applied to fricatives in four different corpora including frication produced in various context and with different levels of effort. Jesus & Shadle (2002) found that the slope parameters were useful in distinguishing between effort levels but not stress (S_p), and between voicing categories within each place of articulation (S'_p). The dynamic amplitude parameter, A_d, separated the sibilant fricatives from non-sibilants; sibilant fricatives had higher dynamic amplitudes than non-sibilants in both of the sustained fricative corpora. There was also a tendency for voiced fricatives to have lower A_d than voiceless, though the result was inconsistent. Multiparametric spaces defined by $\bar{F}$ vs. S'_p vs. S_p and $\bar{F}$ vs. A_d vs. S_p, showed that fricatives formed clear groupings by place of articulation. It is unclear whether these findings would extend to ejective fricatives, but perhaps dynamic amplitude might differ between ejective and pulmonic fricatives due to their differences in production mechanisms. Another often-cited measure of fricative place, center of gravity, may also be of potential use. Ridouane et al. (2015) found that center of gravity was higher for ejective fricatives than pulmonic fricatives in both initial and intervocalic word position in Mehri.

Measures of voice quality in vowels following ejective consonants may also be cue to ejectivity of consonants. As summarized in Chapter 4, H1-H2 is a common measure of voice quality that is also applied to the task of identifying ejectivity (Grawunder et al., 2010). In the present UNT data, H1-H2 was not related to the laryngeal category of adjacent fricatives, but was affected by the laryngeal category of stops. Harmonic-to-noise ratio (HNR), another potential measure of voice quality, was found to be lower in the first third of vowel duration following glottalic stops in Georgian (Grawunder et al., 2010), indicating a noisier output at the start of the vowel while the glottis resets to modal configuration.

Many of the acoustic measures and parameters used to characterize fricatives capture differences in place of articulation or voicing. While the place parameters could be useful in describing the fricatives of UNT in a general sense, they may not directly help to characterize the difference between airstream mechanisms. It is possible, however, that slight differences in the place of frication could serve as a cue to the different airstream categories without being a direct consequence of the airstream, so an investigation into such characteristics could be justified in a more thorough and detailed study of acoustic parameters.

5.5.3 Conclusion

There are essentially two options in our interpretation of the 'ejective' fricatives of UNT: either they are clusters, or they are produced with the glottalic airstream, utilizing glottal closure to manipulate air pressure. In the first case, the lag closure would be associated with a distinct consonant segment, /ʔ/; in the second case, lag closure would be an epenthetic by-product of the articulatory configurations necessary for glottalic speech. If the closure were epenthetic, we might expect the lag time to be somewhat short, akin to the findings of Ohala (1997) regarding epenthetic [p] in English words such as 'hamster'. Note that although the lag periods reported for ejective fricatives in Tlingit were referred to as 'long', their mean duration was only 46 ms. The glottal lags measured in the present study were on average nearly twice that, at 88 ms.

On the other hand, we can instead interpret these glottal lags as part of the stop series. The resulting pattern follows a well-known pattern that occurs in many languages: as place of closure moves deeper into the vocal tract, closure duration becomes shorter (Cho & Ladefoged, 1999). Glottal stops are formed at the place of articulation at the furthest back point of the vocal tract, so their closure durations would be expected to be shorter than any others. This was indeed the pattern that we saw here: when the lags were categorized according to their place of closure (or presumed place, in the case of ejective fricatives), the lags of laryng.fricatives were the shortest of all. Comparisons between lag places showed that pairs of adjacent places did not differ significantly from one another, suggesting a gradual continuum of closure duration based on place of articulation rather than a categorical distinction based on different speech mechanisms. This pattern, coupled with the statistical model showing no significant differences between most places of lag closure, supports the analysis of what have heretofore been referred to as “ejective fricatives” in UNT as fricative + stop clusters.

Chapter 6

Discussion and Conclusions

This thesis began with a question about laryngealization in Upper Necaxa Totonac, which has been described as having contrastive laryngealization in vowels as well as a contrast between ejective and pulmonic fricatives. The goal of this thesis was to shed some light on these contrasts and provide insight into how the acoustics align with existing descriptions. In light of the results of the analyses, alternative hypotheses were also presented, namely, that vowels may be allophonically laryngealized as the result of coarticulation with glottal stops, and that ejective fricatives appear to be phonetic clusters of fricative + glottal stop. These questions were addressed in three chapters. The first of these chapters (Chapter 3) explored segmental collocates present in the UNT dictionary, including both vowels and consonants. The remaining two chapters focused on acoustic aspects of vowels (Chapter 4) and fricatives (Chapter 5) separately.

The present chapter proceeds first by summarizing and evaluating the findings in each of the three chapters, organized by segment type. The summary begins with the fricatives and continues with the vowels. Suggestions for possible follow-up research are included in each section. The chapter concludes with a summary of the contributions of this thesis.

6.1 Fricatives in Upper Necaxa Totonac

The previously supposed contrast between ejective and pulmonic fricatives in UNT was investigated from two perspectives. The first, presented in Chapter 3, looked at segmental laryngealization in context with the hypothesis that laryngealized vowels might be more likely to occur adjacent to other segments involving glottal closure, such as ejective fricatives. Chapter 5 considered acoustic duration as evidence for or against ejective production. Results of the contextual analysis showed that fricative laryngealization is independent of laryngealization in surrounding segments, meaning that laryngealized vowels were no more likely to occur adjacent to ejective fricatives than pulmonic fricatives. The acoustic analysis showed that UNT ejective fricatives are produced with overall longer durations than pulmonic fricatives, due to the combination of long frication intervals followed by substantial glottal closure.

The results of Chapter 5 showed that ejective fricatives differed from pre-vocalic pulmonic fricatives both in terms of frication duration and the duration of the silent periods that appeared following frication. These differences were explained by comparing the same intervals to those of fricatives in clusters before stops. Like the vowel analysis, it is possible that a larger sample would reveal a different pattern from the one reported in this dissertation. However, the analysis focused only on a single acoustic measurement: duration. Fricatives may also be differentiated from each other on the basis of center of gravity, spectral tilt, and relative intensity to name a few. It is possible, therefore, that another measure would support the contrastive categories that have previously been reported in Beck (2006). The contrast could turn out to be a matter of airstream, or it could turn out to be the result of some other phonetic process that has yet to be identified. A further possibility is that speakers interpret the acoustic sequence of fricative + silence as a single unit regardless of its similarity to the fricative + stop clusters. Such an eventuality would likely need to be determined through phonological analysis rather than phonetic, although it may be possible to design a perceptual experiment to test how speakers interpret these phonetic sequences as well.

In the most uncontroversial scenario of canonical glottalic egressive airflow, the glottis must be closed in order to facilitate the compression and subsequent release of oral air pressure. If the ejective fricatives were produced by the glottalic airstream mechanism, the glottis would have to be closed before the onset of frication in order to allow for the compression of ambient air in the vocal tract before release. In such a scenario, we might expect a phonetically ejective fricative to have an effect on preceding vowels similar to that of the glottal stops, which were found to be highly correlated with preceding laryngealized vowels (see Section 6.2 for further details). No such relationship was observed, suggesting that glottal closure does not occur at or before the onset of frication. On the other hand, if the ejective fricatives are phonetically clusters, glottal closure would not be expected to occur until after frication (or at least, not until some time after frication onset). In the cluster scenario, we should expect the fricatives to have the same pattern of co-occurrence with vowel laryngealization as pulmonic fricatives, that is, no relationship at all. In fact, this is what we do see. Likewise, we would also expect fricative + glottal stop clusters to behave like glottal stops with respect to their following environment. Again, the results of the collocational analysis are consistent with such expectations in that no discernible pattern was observed. The ejective fricatives of UNT show collocational patterns similar to those of pulmonic fricatives in their preceding environments, and patterns similar to those of glottal stops in their following environments, exactly as would be expected for fricative + glottal stop sequences.

The analysis of phonetic characteristics of the contrast between ejective and pulmonic fricatives focused on the duration of the component acoustic events on either side of frication in three conditions: simplex fricatives, ejective fricatives and fricative + stop clusters. The analysis was limited to instances of each condition that occurred before vowels word initially or between vowels word medially. First, each condition was segmented into its component events based on the acoustic signal. Due to their presumed production mechanism, ejective fricatives would be expected to be preceded by an interval of silence more often than pulmonic fricatives. While some tokens were produced with this lead silence, it was highly infrequent overall, regardless of the frication condition. The bulk of the analysis focused instead on the durations of frication and following (lag) silent intervals. Some differences were found between ejective fricatives and pulmonic fricatives, especially regarding the duration of lag silences and overall duration from onset of frication to onset of the following vowel (an interval that allows for the comparison of total duration across pulmonic, ejective, and cluster conditions). Contrary to expectations, ejective fricatives had frication intervals equal in duration to pulmonic fricatives, followed by substantially longer periods of silence. These silences, in conjunction with frication, resulted in longer overall durations from the onset of frication to the onset of the following vowel. This is again counter to expectations of ejective fricatives, which have been found to have duration roughly equal to that of pulmonic fricatives when measured from frication onset to vowel onset, as a result of shorter frication followed by a silent interval (cf. Section 5.1).

The overall longer duration of ejective fricatives in comparison to pulmonic fricatives is puzzling: a long lag silence could easily be integrated into an account based on the glottalic mechanism, but the long frication interval is inexplicable given the limited reservoir of air available to the glottalic airstream. Luckily, an alternative hypothesis is at our disposal. Rather than positing a rather unusual sound in an otherwise simple segmental system, we can instead consider the possibility that they are in fact clusters of fricatives followed by glottal stops. Analyzing the ejective fricatives as clusters would assume a pulmonic airstream to initiate and sustain frication, and provide an explanation for the lengthy silence that follows as necessary to indicate a stop closure in the cluster. Taken as a unit from onset of frication to onset of the following vowel, the ejective fricatives appeared inexplicably longer than pulmonic fricatives, but when compared to fricative + oral stop clusters they were of roughly equal total duration.

Despite the overall similarity in duration between ejective fricatives and clusters, there were differences with respect to the duration of lag silences following frication. In order to investigate these differences further, lag periods were annotated according to their place of articulation, with ejective fricatives being labeled as "glottal". Comparisons of lag duration across these place of articulation categories revealed a continuum of closure duration that parallels findings from numerous languages of increasing duration as the place of stop closure approaches the front of the mouth. The ejective fricative closures fit nicely into this pattern due to their closure intervals being the shortest. This pattern, in conjunction with frication intervals of equal duration in simplex, ejective and cluster conditions, supports the hypothesis that the ejective fricatives would be better described as fricative + glottal stop clusters.

The acoustic evidence notwithstanding, a final pronouncement on the phonetic nature of UNT ejective fricatives also requires articulatory measurements and perceptual studies. Ejective speech is produced by way of the glottalic airstream, which requires glottal closure and has repercussions on oral airflow. Further analysis of such articulatory factors might still favor the analysis of ejective fricatives as complex segmental units rather than clusters. A limited set of airflow data has been used to argue for the ejective analysis in the past (Beck, 2006), but the findings were substantially different from airflow measures of ejective fricatives in other languages. A larger airflow study with multiple speakers and a balanced word list would improve the strength of claims that could be made based on such findings. Electroglottography could be used to verify glottal closure during fricative production, if present. Together, airflow and EGG measurements would provide definitive evidence for the articulatory mechanism used in the production of ejective fricatives in UNT. Articulatory analysis would also provide information about the variability of production in these segments across speakers. It is possible that some speakers do produce ejective forms while others do not, or that speakers produce ejective forms in certain linguistic or conversational contexts but not in others.

Phonological evidence for the analysis of ejective fricatives as complex units is hard to come by, since they occur in precisely the same phonotactic environments as fricative + stop clusters. In fact, since UNT allows fricative + stop clusters at all places of articulations for stops except glottal, analyzing the ejectives instead as clusters would fill an apparent gap. Beck (2006) has suggested that syllable structure might be a potential avenue for distinguishing ejective fricatives from clusters. Fricative + stop clusters have been described as splitting into coda and onset segments across syllable boundaries, while ejective fricatives, on the other hand, appear only in syllable onset position (c.f. Table 1.4, Beck 2006), but this analysis has not been empirically tested. In order to do so, speakers could be trained on presumably unambiguous CV syllable types, then tested on clusters and ejective fricatives in a variety of word positions as a means of testing the syllabification argument. Although somewhat unnatural, a simple approach to eliciting speaker judgments of syllable boundaries might be asking them to separate words into syllables (or "beats") by speaking slowly and clapping on each beat. The task would require no special equipment, though audio recording of the task would be advisable for transparency, and the data could be collected fairly quickly. Linguist judgments notwithstanding, speaker judgments are necessary for establishing distinct patterns in syllabification between clusters and ejective fricatives. In lieu of such evidence of syllable structure, and in light of the current findings presented in this thesis, a cluster analysis of the ejective fricatives would be simpler and more straightforward from a phonological point of view. It would also resolve the typological oddity of a system with ejective fricatives, but no ejective stops.

6.2 Vowels in Upper Necaxa Totonac

Laryngealized vowels were found in the segmental corpus analysis of Chapter 3 to be followed by glottal stops far more frequently than would be expected if the two segment types were independent of each other. More than 70% of glottal stops were preceded by laryngealized vowels in the corpus, and nearly a third of laryngealized vowels occur before glottal stops. These correlations were stronger than any other segment pairs, which generally did not show relationships between laryngeal categories.

Chapter 4 investigated this pattern of co-occurrence by comparing measured H1-H2 values between laryngealized and non-laryngealized vowels in a variety of segmental contexts. Lower values of H1-H2 have been related to creaky voice in several languages, while higher H1-H2 sometimes indicates breathy phonation. Most languages with proposed phonation contrasts exceed the just noticeable difference for H1-H2 by a ratio of nearly 2:1. In the present analysis, small effects of categorical vowel laryngealization were found to significantly impact H1-H2 in limited conditions, but these effects were all very close to the threshold for just noticeable differences reported in Kreiman & Gerratt (2010), who note that in order for acoustic characteristics to be relevant to linguistic contrasts, listeners must be perceptually sensitive to them. If an effect does not surpass the threshold for just noticeable difference, listeners are unlikely to be able to use that characteristic in any meaningful way. The effect of glottal stops on their immediately preceding vowels was somewhat stronger, however, nearing 4 dB. While this is still close to the roughly 3 dB average of just noticeable differences Kreiman (2010) also report differences in sensitivity to H1-H2 depending on the listener's first and dominant language. If speakers of UNT make use of phonation to contrast between vowel categories, they may be more sensitive to changes in H1-H2, lowering their threshold for just noticeable difference, and therefore also lowering the size of the change necessary to maintain a reliably perceptible contrast.

On the other hand, if vowel phonation is not contrastive, the effects of glottal stops on the vowels that precede them would still be near the threshold of perceptibility. These changes in the low frequency region of the spectrum are audible, therefore, but need not be accessed be speakers to maintain a phonation contrast. Instead, they could simply be the result of proximity to the glottal stop in an example of coarticulatory assimilation, a classic source of allophonic variation. Another potential bit of evidence in favor of the coarticulatory influence of glottal stop as the source of non-modal phonation in UNT vowels is the temporal location of the lower H1-H2 values in the time course of the vowels. Unlike previous studies of contrastive laryngealization that showed effects in the first two-thirds of vowel duration (Blankenship, 2002; Garellek, 2010; Gordon, 2001), the present study showed that H1-H2 values were lower later in the vowel, particularly in cases when the vowels were followed by glottal stops. This increased influence of the glottal stop at the edge of the vowel nearer to it strengthens the notion that the phonation difference is conditioned by the vowel.

Conversely, vowels that followed glottal stops had H1-H2 values roughly 2 dB higher than vowels following fricatives from either laryngeal category or oral stops, again regardless of the reported laryngealization category of the vowel. This effect was not predicted by previous findings. One potential explanation for this effect is that the vowels are produced with a non-constricted type of creak (Keating et al., 2015), or with breathy phonation, perhaps as a result of increased airflow that occurs when the glottal closure is released.

The results of Chapter 4 showed that H1-H2 was not reliably associated with laryngealization categories of vowels in the present data sample. There are a number of possibilities for why this might be the case. In the present analysis, the data suggested that vowels before glottal stops might be allophonically non-modal due to the conditioning environment. Another possibility is that the sample, which consisted of speech from only four speakers, is simply too small to show differential effects of categorical vowel and conditioned vowel laryngealization. This would be especially true in the case that speakers use a variety of phonetic strategies to indicate the laryngeal contrast. Although

H1-H2 is known to be perceptually available to listeners from languages such as English, Mandarin, Thai and Gujarati (Kreiman & Gerratt, 2010; Kreiman, 2010) and has been related to phonation categories in many languages (Garellek, 2010; Keating et al., 2015; Blankenship, 2002; Esposito, 2010a, 2006; Keating & Esposito, 2006), there are many other acoustic characteristics that may instead carry the contrast, such those summarized in section 4.1. Because there are many approaches to measuring laryngealization, and different measures are better suited to capturing specific types of non-modal phonation, it may be wise to approach vowel laryngealization more agnostically than I have done here. A future study may extract multiple measures of phonation characteristics and compare each measure to the existing categories identified in UNT. Comparisons may also be made between measures, to provide insight into how and why the laryngealized vowels might differ from each other. These measures could be extracted at multiple time points throughout the vowel duration to observe how laryngealization may change over time. Such a study would serve not only to support the contrast as described, but would also provide a high degree of detail, adding immensely to our understanding of phonation in UNT, and providing new avenues of comparison among other Totonacan languages as well.

Due to the complex nature of voice characteristics, it is sometimes (and perhaps often) the case that a single acoustic parameter is insufficient to capture a phonation contrast. Similar to the F1-F2 vowel space, plotting more than one measure against each other can sometimes reveal patterns that were otherwise overlooked. Because there are a great many potential measures that can describe voice quality, computational and statistical approaches may be used to determine which of the available measures are most relevant, such as in multidimensional scaling analyses. Keating et al. (2011) used such an approach to compare phonation categories within and across languages. They found that a three dimensional space differentiated between languages, but that phonation categories, though they might share a label, occupied different regions of the space. Future studies may measure such acoustic characteristics of vowels in either the present sample or an expanded data set, and compare each of them individually or in various combinations with other measures across the vowel categories to look for potential correlations with the existing laryngealization categories.

The present findings show that, on the one hand, H1-H2 does not appear to be associated with the transcribed vowel laryngealization in UNT. On the other hand, this measure does show a strong relationship to vowels followed by glottal stops. Unlike in the Totonac of Zapotitlán de Mendez, UNT is reported to use /ʔ/ as a robust phoneme in its own right, not simply as a phonetic manifestation of phonation contrast in vowels. It stands to reason, therefore, that glottal stops might influence surrounding speech in a manner similar to any other pattern of phonetic processes between adjacent speech segments. The frequent appearance of glottal stops as intervals of non-modal or creaky phonation in the acoustic signal without clearly defined silent or closure intervals supports this hypothesis (see Figure 2.4 for some examples of glottal stop production). Consequently, differentiating sequences of laryngealized vowels followed by glottal stops from non-laryngealized vowels followed by glottal stops is difficult, if not impossible. Without an acoustic distinction between laryngealized and non-laryngealized vowels, transcription of laryngealized vowels in some cases and non-laryngealized vowels in others is misleading, indicating a contrast where it may be impossible to maintain or identify. Merely stipulating a contrast is not enough to establish its existing. At a minimum, a phonetic description of the contrast between laryngealized and non-laryngealized vowels before glottal stops is necessary, followed by a phonemic analysis to verify that the distributions of these sounds cannot be explained by predictable phonetic processes. Whether or not the contrast is justified, a corpus analysis such as the one performed in Chapter 3 can serve to identify redundant transcriptions of predictable allophonic variation for further investigation, or as a means to replace them in a straightforward and systematic way.

There may very well be a phonation contrast among the vowels of UNT, but the present results have shown that vowels are more likely to be labeled as laryngealized in an environment that is likely to condition non-modal phonation as an allophonic variation: before glottal stops. Further study into the nature of the contrast is certainly warranted. One approach might be annotate a data set according to both categorical laryngealization (as indicated by transcription) and audible laryngealization identified by both trained linguist listeners and native speakers of Upper Necaxa Totonac. Once the data is annotated, multiple potential measures of phonation contrasts could then be extracted and analyzed from each of these data sets, with comparisons between the most fruitful measures in each case revealing the acoustic characteristics that are likely to be most salient in each analysis. This would be beneficial on at least two fronts. First, it would highlight similarities and differences between linguist judgments and native speaker judgments, which could be used to improve the accuracy of linguistic annotations. A high correspondence across the data sets would be an indication that linguist judgments are in line with speaker perceptions, while variable outcomes would demonstrate the influence that training and experience have on classification judgments. Second, it would result in a collection of detailed information about the phonetic profile of laryngealization, which could lead to better descriptions, which improves the identification of phonological patterns and potentially affects higher order linguistic analyses as a result.

In order for a contrast to be maintained between two phonological categories, there must be information about that contrast contained in the phonetic signal. If there is a contrast between laryngealization categories in vowels, then a detailed and many-pronged acoustic study ought to be able to describe some characteristic or cluster(s) of characteristics that they share in common. Although contrast cannot be demonstrated on the basis of phonetics alone, the acoustic signal is nevertheless the first evidence of phonological categories linguists have access to, whether via their ears or through acoustic analysis. Future work could endeavor to collect data that would first confirm that conditioning environments reported by Aschmann and Levy are also relevant to the sound system of UNT and provide further acoustic detail relating to the reported differences in production and auditory impression. Such work could be undertaken from a top-down perspective, assuming the currently accepted categories are correct and looking for acoustic information to back them up, or it could proceed from the bottom up, beginning with the acoustics and then building categories based on careful, detailed phonetic annotation and systematic phonemic analysis.

6.3 Variability and alternative explanations

Previous treatments of laryngealization in related Totonacan languages have noted some patterns of variability in the production of laryngealized vowels, such as Aschmann (1946) who described three main types of laryngealized vowels in the Totonac spoken in Zapotitlán de Mendez (summarized in 1.3), conditioned to some degree by the voicing of the previous consonant, or the presence of another laryngealized vowel. Levy (1987) noted environments that seem to condition the appearance of a glottal stop or "hiatus" (p. 61) before laryngealized vowels, particularly the position of the vowel in a phrase, whether post-pausal or phrase-medial. This intermittent glottal stop sometimes resulted in laryngealization of the vowel that followed it, and sometimes did not. The glottal stop also affected syllabification by preventing the following vowel from becoming a nucleus for the consonant that preceded the glottal stop. Levy also noted "fluctuation" (p. 61) in the transcription of laryngealized vowels of Papantla Totonac, which she attributed to phonetic variation in production. Herrera-Zendejas (2014) provided some more detail on the variability of vowel phonation in Papantla, noting three types of non-modal phonation: prototypical, sequential, and tense. In the first two types, laryngealization appears as irregular or slow pulsing of the vocal folds, visible in the spectrogram and apparently similar to what Keating et al. (2015) refer to as prototypical creaky voice. In the productions with tense voice, preceding stop consonants were often voiced; this voicing was a more reliable cue to the laryngeal category of the vowel than any identified aspect of the vowel (Herrera Zendejas, 2014).

As noted in Section 1.3 laryngealized vowels have been a topic of interest among linguistics working on Totonac since the earliest descriptions in the middle of the twentieth century. Even now, more than seven decades after these initial impressionistic transcriptions were produced, they continue to influence scholarship in in Totonacan linguistics. Aschmann (1946) described laryngealized vowels in a variety of Totonac spoken in Zapotitl/'an de Mendez, Puebla, in the southwestern portion of the Totonac speaking region of Mexico. Unlike McQuown (1940, 1990), Aschmann identified ʔ not as a separate and distinct phone participating in the contrastive inventory of consonants, but rather as an integral part of the vowel nucleus of laryngealized vowels.

Aschmann (1946) reported that laryngealized vowels appear in three forms, all of which involved one or more glottal stops and varying degrees of non-modal phonation throughout vowel production. Aschmann also reports particular conditioning environments where each of the types of laryngealized vowels are supposed to occur. No such pattern of glottal stops in relation to laryngealized vowels has been reported for UNT, whether based on linguistic impressions or quantitative phonetic annotation and analysis. Levy (1987) also noted the relationship between laryngealized vowels and glottal stops in Papantla Totonac, along with several other functions of glottal stops including as markers of phrase and perhaps morpheme boundaries. Glottal stops also appear in many Totonacan languages as a reflex of /*q/, and appear to behave phonologically as part of the stop series in such instances (MacKay & Trechsel, 2013).

Herrera Zendejas (2014) illustrated patterns of laryngealization in Papantla Totonac with acoustic spectrograms and F1-F2 vowel plots without reference to this relationship between glottal stops and laryngealized vowels. According to Herrera-Zendejas, laryngealized vowels may present in three ways, ranging from completely creaky (characterized by irregular glottal pulses) throughout the vowel, creaky at the start of the vowel only, and a third "stiff voice" production that was reportedly difficult to identify both auditorally and by visual inspection of spectrograms. In fact, this third type of laryngealized vowel was primarily identified by the presence of voicing in the preceding stop consonant.

Both Aschmann and Herrera-Zendejas describe variable productions of laryngealized vowels. Perhaps coincidentally they both propose three different profiles, but these three types of laryngealized vowels are not the same across both languages. In fact, the descriptions are not at all similar to one another. This lack of similarity in description could mean that the languages differ in the patterns speakers use to represent and produce a particular contrast. On the other hand, it could also indicate simply that each linguist chose different terminology to describe acoustic patterns that are in fact very similar to each other. Herrera-Zendejas provides illustrations of "prototypical" examples of laryngealized vowels, but unfortunately such illustrations were not possible at the time of Aschmann's description. It is therefore impossible to know precisely what he meant by "completely laryngealized" or how it relates to modern presentations of related yet distinct languages. Likewise, even terms that appear to be quite technical are not adequately defined to allow for accurate identification of speech. For example, although Herrera-Zendejas reports one type of laryngealized vowel in Papantla Totonac as being characterized by irregular glottal pulses, there is no indication of how "irregularity" is defined or identified. In fact, some supposed examples of irregularity instead appear to show widely spaced glottal pulses (Herrera-Zendejas, figure 9). These widely spaced pulses indicate a low F0, a characteristic of several types of creak that does not necessarily entail irregularity (Keating et al., 2015).

The dissimilarities between these two closely related languages raises an important point about variability in language and the importance of detailed acoustic descriptions in addition to illustrations of "prototypical" speech segments: every language has potentially different form from other languages related to it. The system of one language need not resemble that of another language, even if they are geographically close and historically related. Language is continuously being learned and created by new generations of speakers who deduce some patterns from their linguistic environment and sometimes infer or create new ones at the same time.

The extent of variability present in productions of laryngealized vowels and ejective fricatives have not yet been reported in detail for Upper Necaxa Totonac. This dissertation provides some evidence for that variability in Figure 2.4, which shows various configurations of vowels adjacent to glottal stops, as well as in the distributions of measured acoustic data in Figures 4.1, 5.4. An important step in any future work would be to first establish that the same kinds of variability are present in UNT, and then to determine how much of that variability is due to individual differences in speakers, speech styles, utterance types, and similar factors. Determining the patterns of variability within the UNT data in detail will greatly bolster future investigations into this particular language of the Totonacan family.

6.4 Implications for phonetics, phonology and typology

Although both the ejective fricatives and the laryngealized vowels appear to be related to glottal stops, the implications of these relationships are not the same for each segment type. I will first discuss the implications for the vowels, and then follow up with a discussion of the fricatives.

6.4.1 Vowels

The results of the investigations in this dissertation suggest that laryngealized vowels and glottal stops are redundantly transcribed when they co-occur. This redundancy may be viewed from two perspectives: either glottal stops are conditioned by the laryngealized vowels that precede them, or the laryngealized vowels are conditioned by the glottal stops that follow them. I have suggested elsewhere in this thesis that the vowels are likely to be laryngealized by the presence of the glottal stops. If this were true, then there would be no need to transcribe vowel laryngealization before glottal stops. However, the overall frequency of occurrence of glottal stops might suggest the opposite to be true. Perhaps the glottal stops are themselves an allophone of vowel laryngealization. In that case, the transcription of the glottal stops would be redundant.

The H1-H2 analysis showed that laryngealized vowels were not significantly different from non-laryngealized vowels before glottal stops. Low H1-H2 is associated with creaky voice, and H1-H2 was lower in vowels before glottal stops. Therefore, a narrow phonetic transcription should indicate creakiness. In order to argue that some vowels before glottal stops are nevertheless non-laryngealized, we need to look to higher order evidence. Phonetically they look the same, and phonologically they are occurring in the same conditioning environment, so we could posit that they are underlyingly modal but become laryngealized before glottal stops. In the transcribed dictionary forms, the only cases of modal vowels followed by glottal stops occurred across morpheme boundaries. In these cases, the vowels might appear non-laryngealized when followed by morphemes beginning with segments other than glottal stop, and so a non-laryngealized transcription could be justified. However, in most of the forms included in the acoustic analysis here, vowels followed by glottal stops occurred within an individual morpheme and usually within the same syllable, so morphology is of little help in determining the underlying laryngeal category. Without a surface distinction between contrastively and contextually laryngealized vowels, establishing an underlying contrast is nontrivial.

Table 6.1: Possible analyses of present data

1	No contrast - all allophony	
2	Contrast & allophony	Overlapping distribution (neutralization) Complementary distribution
3	Incorrectly labeled contrast	

In the case that vowels are allophonically laryngealized before glottal stops, there are a few analytical possibilities, summarized in Table 6.1. The first possibility is that there is no contrast between non-laryngealized and laryngealized vowels anywhere in UNT. This scenario is potentially consistent with the analyses in this dissertation as well as descriptions of some other varieties of Totonac that do not have contrastive vowel laryngealization such as Coatepec or Apapantilla. The second possibility is that there is a laryngealization contrast that may or may not also be accompanied by allophonic laryngealization. In this scenario, the environments in which contrastive and non-contrastive laryngealization occur may either overlap or complement one another.

If there is no contrast, and phonetically laryngealized vowels are always allophonic (or prosodic) in UNT, we can appeal not only to the patterns within UNT but also to descriptions of similar sounds across languages for comparative reference points. In particular we might expect phonation contrasts to be maintained by similar parameters across languages. For example, some results have shown that phonation contrasts tend to begin early and persist throughout vowel duration, while allophonically conditioned differences in phonation tend to be shorter in duration (and appear later in the vowel). My results suggest that creak, as indicated by lower H1-H2 values, occurs later in the vowel, especially when the following context is a glottal stop, suggesting that the laryngealization is allophonic rather than phonemic. Outside of this conditioning environment, there was no reliable effect of laryngealization on H1-H2.

If laryngealization is contrastive, it may also appear in overlapping distributions between contrastive and allophonic laryngealization, the contrast may neutralize in a certain environment. In this case, that environment is before a glottal stop. If neutralization occurs, then it would be impossible to determine on the basis of phonological evidence whether the vowel is underlyingly laryngealized or not. In the case of no overlap in phonotactic distributions between laryngealized and non-laryngealized vowels, we may expect little difference between the appearance of allophonic and contrastive laryngealization, and the differences we do observe could always be attributed to external conditioning factors such as differences in phonotactic environments. No direct comparison between the two types of laryngealization would be possible. A third possibility is that there is a laryngealization contrast, but the laryngealized vowels have not yet been correctly labeled. Here is where careful phonetic documentation and description can do the most good for the phonological analysis. References to descriptive categories without relation to acoustic (or other phonetic) facts can only lead to further confusion.

6.4.2 Fricatives

Regarding the fricatives in UNT, a different issue arises. One of the main uses of phonological databases such as UPSID (Maddieson, 1984), Phoible (Moran & McCloy, 2019), P-Base (Mielke, 2008) is the investigation of so-called phonological universals, and the implicational relations that can be determined to hold in sound inventories. Prior to Beck (2006), no language had ever been described with ejective fricatives unless the sound inventory also included ejective stops. The existence of such a language would be a violation of the universal that states that the presence of ejective fricatives entails the presence of ejective stops. In a general sense, contrasts involving secondary articulations should not be present unless a more basic sound type is also present in the language.

There are at least two options for analysis here: ejective fricatives are ejective, or ejective fricatives are clusters. If ejective fricatives are ejective, the sound system of UNT is a typological oddity. The appearance of ejective fricatives complicates the segment inventory of UNT by adding a second series of fricatives, which ordinarily occur in relatively low numbers. On the other hand, if ejective fricatives are clusters, this implicational relation is not violated. This scenario would align with the phonetic data and analyses presented in this dissertation. It is furthermore not out of line with the general/overall phonotactic patterns of the language, which allows for fricative + stop clusters in every environment where ejective fricatives also occur. Interpreting the ejective fricatives as clusters would therefore simply the sound inventory of UNT and avoid violating an otherwise robust implicational relation in phonological systems. The relative simplicity of the cluster analysis and the subsequent resolution of what is otherwise a troublesome data point for linguistic typology, lends support to the phonetic analysis that also supports the cluster analysis.

6.5 Contributions of the thesis

This thesis has provided analysis of understudied contrasts related to laryngealization in Upper Necaxa Totonac. The sounds that have been referred to as ejective fricatives are very similar to fricative + stop clusters. There is at present no indication that they are produced with a non-pulmonic airstream, and therefore describing them as an exotic sound type is not necessary. Despite initial appearances in visual inspection of the data, the laryngealized vowels did not differ significantly overall compared non-laryngealized vowels in terms of H1-H2. Rather, the proximity of a following glottal stop appeared to influence vowel production, resulting in characteristics of creaky voice later in the vowel. In other words, vowels that preceded glottal stops had lower H1-H2 values than vowels in other environments, regardless off the laryngeal category of the vowel itself. The findings presented here may be useful in producing transcriptions that are more consistent with the acoustic data, which in turn could affect phonological and other higher order linguistic analyses. The thesis has also identified future research that could further confirm or question these conclusions. Similar analyses could be performed for other Totonac languages to shed light on the disparate analyses of laryngealization across the family and the consequences for comparative reconstruction. The findings of such research have the potential to inform our understanding of the relationship between phonetics, phonology and fieldwork, as well as reconstruction of historical developments in the language family.

References

Alarcón Montero, R. (2008). Indicios acústicos de las vocales rechinadas del totonaco.

Arana Osnaya, E. (1953). Reconstruccion del protototonaco. *Revista mexicana de estudios antropologicos*, (pp. 123–130).

Arppe, A. (2009). *Univariate, bivariate, and multivariate methods in corpus-based lexicography – a study of synonymy*. PhD thesis, University of Helsinki.

Aschmann, H. P. (1946). Totonaco phonemes. *International Journal of American Linguistics*, 12(1), 34–43.

Baayen, R. H. (2006). *Analyzing Linguistic Data*. Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412.

Bao, M. (2009). Phonetic realization of glottal stop in Shugni. *The Journal of the Acoustical Society of America*, 126, 2181.

Barnier, J., Briatte, F., & Larmarange, J. (2017). questionr: Functions to Make Surveys Processing Easier.

Bauer, L. (2015). English phonotactics. *English Language and Linguistics*, 19(3), 437–475.

Beck, D. (2004). *Upper Necaxa Totonac*. Muenchen: Lincom Europe.

Beck, D. (2006). The emergence of ejective fricatives in Upper Necaxa Totonac. *University of Alberta Working Papers in Linguistics*, 1, 1–18.

Beck, D. (2008). Ideophones, adverbs, and predicate qualification in Upper Necaxa Totonac. *International Journal of American Linguistics*, 74(1), 1–46.

Beck, D. (2011a). *Nuevo diccionario del idioma totonaco del Río Necaxa*. Self-published.

Beck, D. (2011b). *Upper Necaxa Totonac Dictionary*. Berlin: Mouton DeGruyter.

Blankenship, B. (2002). The timing of nonmodal phonation in vowels. *Journal of Phonetics*, 30, 163–191.

Boersma, P. & Weenink, D. (2018). Praat: doing phonetics by computer [Computer program].

Bowern, C. (2008). *Linguistic Fieldwork: A practical guide*. New York: Palgrave MacMillan, 1st edition.

Brown, C. H., Beck, D., Kondrak, G., Watters, J. K., & Wichmann, S. (2011). Totozoquean. *International Journal of American Linguistics*, 77(3), 323–372.

Catford, J. C. (1964). Phonation types. In D. Abercrombie, D. Fry, M. P.A.D., N. Scott, & J. Trim (Eds.), *In honour of Daniel Jones* (pp. 26–37). University of Chicago.

Catford, J. C. (2010). On the classification of stop consonants (1939). *Journal of the International Phonetic Association*, 40(3), 287–291.

Chao, K.-y. & Chen, L.-m. (2008). A Cross-Linguistic Study of Voice Onset Time in Stop Consonant Productions. *Computational Linguistics and Chinese Language Processing*, 13(2), 215–232.

Cho, T. & Ladefoged, P. (1999). Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics*, 27(2), 207–229.

Demolin, D. (2002). The search for primitives in phonology and the explanation of sound patterns: The contribution of fieldwork studies. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology 7* (pp. 455–514). The Hague: Walter de Gruyter.

Demolin, D. (2015). Dynamics and articulatory control in Amharic ejectives. Presented at LabPhon15, Cornell University.

DiCanio, C. T. (2010). Itunyoso Trique. *Journal of the International Phonetic Association*, 40(2), 227–238.

DiCanio, C. T. (2011). Perceptual cues of laryngeal contrasts in Trique. Linguistics Society of America.

Dowle, M. & Srinivasan, A. (2017). *data.table: Extension of 'data.frame'*. R package version 1.10.4.

Elías-Ulloa, J. (2016). The role of prominent prosodic positions in governing laryngealization in vowels: A case study of two panoan languages. In H. Avelino, M. Coler, & L. Wetzels (Eds.), *The Phonetics and Phonology of Laryngeal Features in Native American Languages*. Brill.

Esling, J. H. (1999). Language and Speech Laryngoscopic Observations of Pharyngeal Articulations and Larynx Height. *Language and Speech*, 42, 349–72.

Esling, J. H., Fraser, K. E., & Harris, J. G. (2005). Glottal stop, glottalized resonants, and pharyngeals: A reinterpretation with evidence from a laryngoscopic study of Nuuchahnulth (Nootka). *Journal of Phonetics*, 33, 383–410.

Esposito, C. M. (2006). *The Effects of Linguistic Experience on the Perception of Phonation*. PhD thesis, University of California Los Angeles.

Esposito, C. M. (2010a). The effects of linguistic experience on the perception of phonation. *Journal of Phonetics*, 38(2), 306–316.

Esposito, C. M. (2010b). Variation in contrastive phonation in Santa Ana Del Valle Zapotec. *Journal of the International Phonetic Association*, 40(2), 181–198.

Forrest, K., W. G. M. P. D. R. (1988). Statistical analysis of word-initial voiceless obstruents: Preliminary data. *Journal of the Acoustical Society of America*, 84(1), 115–123.

García Ramos, C. (1979). Fonología del totonaco del Tajín, Veracruz. *Cuadernos Antropológicos (Veracruz)*, 2, 133–176.

Garcia-Vega, M. & Tucker, B. V. (2019). Acoustic properties of vowels in Upper Necaxa Totonac. *Journal of the International Phonetic Association*, to appear.

Garellek, M. (2010). The acoustics of coarticulated non-modal phonation. *UCLA Working Papers in Phonetics*, 108, 66–112.

Garellek, M. (2013). *Production and perception of glottal stops*. PhD thesis, University of California, Los Angeles.

Garellek, M. (2019). Acoustic discriminability of the complex phonation system in !xóõ. *Phonetica*, to appear.

Gerfen, C. & Baker, K. (2005). The production and perception of laryngealized vowels in Coatzospan Mixtec. *Journal of Phonetics*, 33(3), 311–334.

Gerratt, B. & Kreiman, J. (2001). Toward a taxonomy of nonmodal phonation. *Journal of Phonetics*, 29, 365–381.

Gobl, C., N. C. A. (1992). Acoustic characteristics of voice quality. *Speech Communication*, 11, 481–490.

Gordon, M. (2001). Laryngeal timing and correspondence in Hupa. *UCLA Working Papers in Linguistics*, 7.

Gordon, M. (2016). Consonant-tone interactions: A phonetic study of four indigenous languages of the Americas. In *Phonetics and Phonology of Laryngeal Features in Native American Languages* (pp. 129–156). Leiden: Brill.

Gordon, M. & Applebaum, A. (2006). Phonetic structures of Turkish Kabardian. *Journal of the International Phonetic Association*, 36(02), 159.

Gordon, M. & Ladefoged, P. (2001). Phonation types: A cross-linguistic overview. *Journal of Phonetics*, 29, 383–406.

Grawunder, S., Simpson, A., & Khalilov, M. (2010). Phonetic characteristics of ejectives – samples from caucasian languages. In T. M. Z. M. Fuchs, S. (Ed.), *Turbulent Sounds: An Interdisciplinary Guide* (pp. 209–244). Berlin: Mouton de Gruyter.

Gries, S. T. (2013). *Statistics for Linguistics with R: A Practical Introduction*. Berlin and New York: Mouton DeGruyter, 2 edition.

Gries, S. T. (2015). The most under-used statistical method in corpus linguistics: Multilevel (and mixed-effects) models. *Corpora*, 10(1), 95–125.

Hanson, H. M. (1997). Glottal characteristics of female speakers: acoustic correlates. *The Journal of the Acoustical Society of America*, 101(1), 466–481.

Hanson, H. M., Stevens, K. N., Kuo, H.-K. J., Hill, M., Chen, M. Y., & Slifka, J. (2001). Towards models of phonation. *Journal of Phonetics*, 29, 451–480.

Henton, C., Ladefoged, P., & Maddieson, I. (1992). Stops in the world's languages. *Phonetica*, 49, 65–101.

Herrera Zendejas, E. (2014). *Mapa fónico de las lenguas mexicanas: formas sonoras 1 y 2, 2nd edition*. Mexico, D.F.: El Colégio de México.

Iseli, M. & Alwan, A. (2004). An improved correction formula for the estimation of harmonic magnitudes and its application to Open Quotient Estimation. In *Proceedings of ICASSP* (pp. 669–672). Montreal.

Jesus, L. M. T. & Shadle, C. H. (2002). A parametric study of the spectral characteristics of European Portuguese fricatives. *Journal of Phonetics*, (pp. 437–464).

Keating, P. & Esposito, C. (2006). Linguistic Voice Quality. In *SST 2006 Proceedings*.

Keating, P. & Esposito, C. (2007). Linguistic Voice Quality. *UCLA Working Papers in Phonetics*, 105(105), 85–91.

Keating, P., Esposito, C. M., Garellek, M., Khan, S., & Kuang, J. (2011). Phonation contrasts across languages. In *Proceedings of the XVII International Congress of Phonetic Sciences* (pp. 1046–1049).

Keating, P., Garellek, M., & Kreiman, J. (2015). Acoustic properties of different kinds of creaky voice. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)* (pp. Paper number 821, 1–5).

Kim, Y. & Valdovinos, M. (2014). The interaction of laryngealized vowels, stress, and falling pitch in Mariteco Cora. In R. Bennett, R. Dockum, E. Gasser, D. Goldenberg, R. Kasak, & P. Patterson (Eds.), *Proceedings of the Workshop on the Sound Systems of Mexico and Central America* New Haven.

Kirchner, R. & Varelas, E. (2002). A cue-based approach to the phonotactics of Upper Necaxa Totonac. In *7th Workshop on Structure and Constituency in the Languages of the Americas*.

Kreiman, J., G. B. u. D. K. S. (2010). Effects of native language on the perception of voice quality. *Journal of Phonetics*, 38, 588–593.

Kreiman, J., B., G., & Antonanzas-Barroso (2007). Measures of the glottal source spectrum. *Journal of Speech, Language, and Hearing Research*, 50, 595–610.

Kreiman, J. & Gerratt, B. (2010). Perceptual sensitivity to first harmonic amplitude in the voice source. *Journal of the Acoustical Society of America*, 128, 2085–2089.

Kreiman, J., Gerratt, B. R., Garellek, M., Samlan, R., & Zhang, Z. (2014). Toward a unified theory of voice production and perception. *Loquens*, 1(1), e009.

Kreiman, J., Shue, Y.-L., Chen, G., Iseli, M., Gerratt, B., Neubauer, J., & A., A. (2012). Variability in the relationships among voice quality, harmonic amplitudes, open quotient, and glottal area waveform shape in sustained phonation. *Journal of the Acoustical Society of America*, 132(4), 2625—-2632.

Kung, S. S. (2007). *A Descriptive Grammar of Huehuetla Tepehua*. PhD thesis, University of Texas at Austin.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26.

Ladefoged, P. (1971). *Preliminaries to Linguistic Phonetics*. Chicago: University of Chicago.

Ladefoged, P. & Cho, T. (2001). Linking linguistic contrasts to reality: The case of VOT. In *Travaux du cercle linguistique de Copenhague*, volume XXXI (pp. 212–225).

Lam, Y. (2012). Oportunidad, ideologia y la perdida del totonaco del rio Necaxa [Opportunity, ideology and the loss of Upper Necaxa Totonac]. In P. Levy & D. Beck (Eds.), *Las lenguas totonacas y tepehuas: Textos y otros materiales para su estudio [The Totonac and Tepehua languages: Texts and other reference materials]* (pp. 519–543). Mexico City, Mexico: Universidad Nacional Autonoma de Mexico.

Laver, J. (1980). *The phonetic description of voice quality*. Cambridge University Press.

Lenth, R. V. (2016). Least-squares means: The R package lsmeans. *Journal of Statistical Software*, 69(1), 1–33.

Levy, P. (1987). *Fonología del totonaco de Papantla, Veracruz*. Mexico City, Mexico: Instituto de Investigaciones Filológicas, Universidad Nacional Autonoma de México.

Levy, P. (2015). La fonología prosódica del totonaco de Coatepec: Los textos totonacos de N. A. McQuown (1938-1940). In *Memorias del VII Congreso de Idiomas Indígenas de Latinoamérica, 29–31 de octubre de 2015, Universidad de Texas en Austin*.

Levy, P. & Beck, D., Eds. (2012). *Las lenguas totonacas y tepehuas: Textos y otros materiales para su estudio [The Totonac and Tepehua languages: Texts and other reference materials]*. Mexico City, Mexico: Universidad Nacional Autonoma de Mexico.

MacKay, C. J. (1994). A Sketch of Misantla Totonac Phonology. *International Journal of American Linguistics*, 60(4), 369–419.

MacKay, C. J. & Trechsel, F. R. (2011). Relaciones internas de las lenguas totonaco-tepehuas. In *Memorias del V Congreso de Idiomas Indígenas de Latinoamérica, 6-8 de octubre de 2011, Universidad de Texas en Austin* (pp. 2–26).

MacKay, C. J. & Trechsel, F. R. (2013). A Sketch of Pisaflores Tepehua Phonology. *International journal of American Linguistics*, 79(2), 189–218.

Mackay, C. J. & Trechsel, F. R. (2015). Totonac-Tepehua Genetic Relationships. *Amerindia*, 37(2), 121–158.

Mackay, C. J. & Trechsel, F. R. (2018). An alternative reconstructino of Proto-Totonac-Tepehua. *International Journal of American Linguistics*, 84(1), 51–92.

Maddieson, I. (1984). *Patterns of Sounds*. Cambridge, United Kingdom: Cambridge University Press.

Maddieson, I. (1998). Why make life hard? - resolutions to problems of rare and difficult sounds types. In *Proceedings of the Twenty-Fourth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Phonetics and Phonological Universals* (pp. 367–380).

Maddieson, I. (2001). Good timing : Place-dependent voice onset time in ejective stops. In *Eurospeech-2001* (pp. 823–826).

Maddieson, I. (2013). Glottalized Consonants. In M. S. Dryer & M. Haspelmath (Eds.), *World Atlas of Linguistic Structures*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Maddieson, I. & Precoda, K. (1990). Updating UPSID. *UCLA Working Papers in Phonetics*.

Maddieson, I., Smith, C. L., & Bessell, N. J. (2001). Aspects of the phonetics of Tlingit. *Anthropological Linguistics*, 43(2), 135–176.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, 94, 305–315.

McFarland, T. A. (2009). *The phonology and morphology of Filomeno Mata Totonac*. PhD thesis, University of California, Berkeley.

McGraw, R. (2009). *Language Attitudes and Opportunities for Speaking a Minority Language What Lies Ahead for Ozelonacaxtla Totonac?* PhD thesis, University of Alberta.

McQuown, N. (1940). *A Grammar of the Totonac Language*. PhD thesis, Yale University.

McQuown, N. A. (1990). *Grámatica de la lengua totonaca (Coatepec, Sierra Norte de Puebla)*. México: UNAM.

Mielke, J. (2008). *The Emergence of Distinctive Features*. Oxford University Press.

Moran, S. & McCloy, D., Eds. (2019). *PHOIBLE 2.0.* Jena: Max Planck Institute for the Science of Human History.

Moran, S., McCloy, D., & Wright, R., Eds. (2014). *PHOIBLE Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Ohala, J. J. (1997). Emergent Stops. In *Proc. 4th Seoul International Conf. on Linguistics [SICOL]* (pp. 84–91).: Linguistic Society of Korea, Seoul.

Peterson, G. E. (1961). Parameters of vowel quality. *Journal of Speech and Hearing Research*, 4, 10–29.

Pickett, V. B., Villalobos, M. V., & Marlett, S. a. (2010). Isthmus (Juchitán) Zapotec. *Journal of the International Phonetic Association*, 40(03), 365–372.

Puderbaugh, R. (2015). Contextual effects on the duration of ejective fricatives in Upper Necaxa Totonac. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)* (pp. Paper number 931, 1–5). Glasgow, UK: University of Glasgow.

Puderbaugh, R. (2019). Phonetics and phonotactics of vowel laryngealization in Upper Necaxa Totonac. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019* (pp. 2085–2089). Canberra, Australia: Australasian Speech Science and Technology Association Inc.

Quick, P. (2003). *A grammar of the Pendau language of central Sulawesi, Indonesia*. PhD thesis, Australian National University.

R Core Team (2017). R: A language and environment for statistical computing.

Ramos, C. G. (2007). Diccionario - Totonaco - Español.

Repp, B. H. (1984). Closure duration and release burst amplitude cues to stop consonant manner and place of articulation. *Language and Speech*, 27(3), 245–254.

Ridouane, R., Gendrot, C., & Khatiwada, R. (2015). Mehri ejective fricatives: an acoustic study. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)* (pp. Paper number 995, 1–5). Glasgow, UK: University of Glasgow.

Seidl, H., Rajendran, S., & Yegnanarayana, B. (2009). Acoustic characteristics of ejectives in amharic. In *INTERSPEECH 2009 - 10th Annual Conference of the International Speech Communication Association, September 6-10, 2009* Brighton, UK: ISCA Archive.

Shosted, R. K. & Rose, S. (2011). Affricating ejective fricatives: The case of Tigrinya. *Journal of the International Phonetic Association*, 41(1), 41–65.

Shue, Y.-L., Keating, P., Vicenik, C., & Yu, K. (2011). Voicesauce: A program for voice analysis. In W. S. Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 1846–1849). Hong Kong: City University of Hong Kong.

Silverman, D., Blankenship, B., Kirk, P., & Ladefoged, P. (1995). Phonetic structures in Jalapa Mazatec. *Anthropological Linguistics*, 37(1), 70–88.

Simpson, A. & Brandt, E. (2019). Detecting larynx movement in non-pulmonic consonants using dual-channel electroglottography. In *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS 2019)*.

Slifka, J. (2006). Some physiological correlates to regular and irregular phonation at the end of an utterance. *Journal of Voice: Official Journal of the Voice Foundation*, 20(2), 171–86.

Traunmüller, H. (1981). Perceptual dimension of openness in vowels. *The Journal of the Acoustical Society of America*, 69(5), 1465–1475.

Trechsel, F. R. & Faber, A. (1992). Acoustic properties of plain and laryngealized vowels in the Misantla dialect of Totonac. Ms.

Vicenik, C. (2009). Praat voicesauce imitator.

Watters, J. K. (1980). Aspects of Tlachichilco Tepehua (Totonacan) Phonology. *Summer Institute of Linguistics. SIL-Mexico Workpapers*, 4, 85–129.

Watters, J. K. (1988). *Topics in Tepehua Grammar*. PhD thesis, University of California, Berkeley.

Watters, J. K. (2010). Phrase-final glottals in Tlachichilco Tepehua. In *Annual Meeting of the Society for the Study of the Indigenous Languages of the Americas, Baltimore, Maryland, January 2010* (pp. 1–17).

Whalen, D. H. (2016). Acoustic realization of a distinctive, frequent glottal stop: The Arapaho example. *The Journal of the Acoustical Society of America*, 139(4), 2212–2213.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: a Professional Framework for Multimodality Research. In *Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation*.

Wright, R., Hargus, S., & Davis, K. (2002). On the categorization of ejectives: data from Witsuwit'en. *Journal of the International Phonetic Association*, 32(01), 43–77.

Yates, F. (1934). Contingency Tables Involving Small Numbers and the $\chi 2$ Test. *Supplement to the Journal of the Royal Statistical Society*, 1(2), 217–235.

Appendix A

Supplement to Chapter 2

The following wordlist was compiled from the *Upper Necaxa Totonac Dictionary* (Beck, 2011b), its derivative practical dictionary *Nuevo diccionario del idioma totonaco del Río Necaxa* (Beck, 2011a), and a paper on UNT phonotactics by Kirchner & Varelas (2002). IPA transcriptions are derived from the orthographic representations in the *Dictionary*, shown in the second column here (UNT), and intended to indicate phonemic contrast. Syllabification is based on simple CV(C) construction rules with as described in Chapter 2. The wordlist is the basis of the analyses in Chapters 4 and 5. See Appendix C for information on vowel and context laryngealization encoding used in the analysis for Chapter 4, and Appendix D for information on frication condition and position used for the analysis in Chapter 5.

Table A.1: Upper Necaxa Totonac wordlist.

IPA	UNT	English
a̰ː.ˈtuːs	a:'tú:s	in awhile
a̰.ʔa.ta.puː.ˈʃ’aː.ja̰	a'hatapu:x’á:ya'	soot from cooking fires built up on spiderwebs on the roof of a house
a̰ʔ.ˈɬ’oː.ma̰	a'hlh’ó:ma'	water and lime mixture for making nixcomel
a̰ʔ.s’a.wi.ˈnḭ	a'hs’awiní'	trickster, deceiver
a.ka.kuː.ˈluːkɬ	akaku:lú:klh	scorpion
a̰ː.tṵː.tʃiː.ˈjeːtɬ, a̰ː.tṵː.tʃiː.ˈjeːkɬ	a:'tu:'chi:yé:tlh, a:'tu:'chi:yé:klh	mint

Table A.1: Upper Necaxa Totonac wordlist.

IPA	UNT	English
tʃa̰ː.ˈʔo̰.ʃ’a̰	cha:'hó'x’a'	treebark
tʃa̰.ˈʔaːn	cha'há:n	be washed roof beam
tʃa.ˈnaː	chaná:	sweat (pot, bottle), be beaded with condensation
tʃi.ḭ.ˈḭkʃ	chili'lí'kx	calcite (calcium carbonate crystals), used for lime
tʃḭpʃ	chi'px	dense
tʃḭʃ.ˈkṵ	chi'xkú'	man
tʃṵ.ˈtʃo̰ʔʃ	chu'chó'hx	banana blossom
ʔeː.tʃi.ˈwiʃ	he:chiwíx	Ocomantla (village)
ʔeː.ˈʔa̰ːʃ	he:há:'x	its shell (turtle, armadillo, snail)
ʔeː.mi.ˈjaːɬ	he:miyá:lh	be standing facing the other way
ʔeː.ta.wa.ˈka̰ɬ	he:tawaká'lh	1) be on the back of sth 2) have sth on one's back
ʔeɬ.ˈʔoː.ʃ’a̰	helhhó:x’a'	one’s lips
ʔeːɬ.kṵ.ˈtan	he:lhku'tán	smell of sweat
ʔeɬ.taʃ.ˈtu	helhtaxtú	fade (color)
ʔen.ʔa.liː.ˈs’o.ḭ	henhali:s’óli'	1) Northern Tamandu 2) elephant
ˈʔo̰.ʃ’a̰	hó'x’a'	its skin, its hide
ʔo̰.ʃ’a̰.ʔaː.ˈɬwa̰ːʔ	ho'x’a'ha:lhwá:'h	eggs with soft, flexible or leathery shells (lizards, turtles, etc.
ʔo.ˈɬ’aː.wa̰	holh’á:wa'	leathery, stiff, not flexible
ḭɬ.ˈtḭn	i'lhtí'n	1) faeces; 2) eggs of flies, mosquitoes, e
ḭʃ.ʔeːn.ḭʃ.ma.ˈkan	i'xhe:n i'xmakán	back of one’s hand
ka̰ː.s’e.ˈwi.wḭ	ka:'s’ewíwi'	cool (climate)
ka̰ː.ʃ’a̰.ˈʔa̰ːʔ	ka:'x’a'há:'h	bright (place)
kiɬ.ˈʔo̰.ʃ’a̰	kilhhó'x’a'	one's lips

Table A.1: Upper Necaxa Totonac wordlist.

IPA	UNT	English
kiɬ.pa̰n.ɬṵ.lṵ	kilhpa'nlhú'lu'	1) jowly, with swollen cheeks; 2) toothless
la̰ʔ.ˈɬ’aː	la'hlh’á:	cut sth into fine strips (meat)
la̰ʔ.ˈɬ’o.nḭ	la'hlh’óni'	flames
la.ka.man.saː.ˈnas	lakamansa:nás	a girl who blushes (term of endearment)
liː.ˈɬ’a̰ːn	li:lh’á:'n	plough
liː.ɬ’o.ˈluː	li:lh’olú:	snore because of something
liːs.ka.ˈlah.wa̰	li:skalájwa'	danger
liːʃ.pa̰.ˈtan	li:xpa'tán	pestle of a molcajete
lu.ˈlo̰ʔʃ	luló'hx	1) stringy endocarpal material that contains seeds of squash or melons; 2) rotten material inside of an old gourd
ɬa.ˈpa	lhapá	cover sth (with cloth, sheet, mulch, etc.
ɬapɬ	lhaplh	completely wet
ɬḛ.ʔeː.ˈnin	lhe'he:nín	cut sugarcane
ɬkaːk.ˈnan	lhka:knán	be hot (weather)
ɬpa̰.ˈma.ma̰	lhpa'máma'	cuddly, soft and furry
ˈɬkah.wa̰	lhkájwa'	1) disgusting; 2) mean, aggressive, vulgar; 3) immoral
ɬka.ˈwat	lhkawát	dry, stiff, hard
ɬpa.ˈpa̰ʔ	lhpapá'h	wrinkled, old and bent out of shape
ɬta̰.ˈʔa̰h	lhta'há'j	1) thick and flat; 2) lying flattened or face down; 3) flat up against a surface; 4) having a single-piece roof (house)
ɬta.ˈka̰.la̰	lhtaká'la'	board
ɬta.ˈyat	lhtayát	looking out of the corner of one's eye
ɬ’a̰ː.ˈna̰	lh’a:'ná'	1) (Pt. pu:tayán) fish net; 2) material for making nets
ɬ’ḛ.ˈʔeː.lḭ	lh’e'hé:'li'	unfinished, rough, pock-marked

Table A.1: Upper Necaxa Totonac wordlist.

IPA	UNT	English
ɬ’e.ˈli	lh’elí	wheeze heavily
ɬ’ḭ.wi.ˈliː	lh’i'wilí:	1) flatten sth out by pressing down on it; 2) set down sth heavy
ɬ’oː.ˈma.wa̰	lh’o:máwa'	half-cooked
ɬ’o̰.ˈʔo.lṵ	lh’o'hólu'	pocked, rough, bumpy
ˈɬ’o.lṵ	lh’ólu'	unfinished, rough, pock-marked
maː.s’a̰.ʔa̰ː.ˈniː	ma:s’a'ha:'ní:	make sby sweat
maː.s’a̰.taːˈnan	ma:s’a'ta:nán	raise children
maː.s’ḛʔ.ˈniː	ma:s’e'hní:	1) provide shade for sth, protect sth from the sun; 2) provide shelter for sth from the rain
maː.s’o̰.ˈʔoː	ma:s’o'hó:	salt sth
maːs.kaː.ˈkiː	ma:ska:kí:	dry sth
maː.ʃ’a̰.ʔeː.ˈnin	ma:x’a'he:nín	1) be illuminated; 2) shine; 3) light up, glow (of fireflies, glow worms); 4) be reflective; 5) be clear
ma̰ʔ.ˈʃuː	ma'hxú:	peel a thick slice of skin off sth (banana, mango, avocado)
paːʃ.ˈkat ka̰.ˈtsiːɬ	pa:xkát ka'tzí:lh	thank you
piː.ˈs’aːm	pi:s’á:m	splinter
pḭ.ˈpḛʔs	pi'pé'hs	its scale (eg fish -rp)
piʃ.ˈʔa̰.ɬa̰	pixhá'lha'	large (bunch or bouquet)
piʃ.ˈpam.wa̰	pixpámwa'	having long hair, having puffy hair
po̰ʔɬ	po'hlh	1) dark, lightless; 2) having clouded vision, seeing spots before one's eyes, feeling faint; 3) feeling fed up, irritated or lethargic
puː.ɬḛ.ˈʔe	pu:lhe'hé	count sth
ˈpuː.ʃ’a̰	pú:x’a'	stream
puks	puks	overcast, gloomy

Table A.1: Upper Necaxa Totonac wordlist.

IPA	UNT	English
sa̰.ˈsa	sa'sá	turn grey (hair)
sa.ˈlun	salún	hoe
ˈsḛʔ.sḭ	sé'hsi'	sweet
ˈska̰t.lḭ	ska'tli'	1) learn sth; 2) learn to do sth
ska.ˈma.ma̰	skamáma'	quiet, serious, staid
sla̰ʔ.s’o̰.ˈʔo̰j.wa̰	sla'hs’o'hó'jwa'	a bit salty
spa̰ɬ.ta.ˈtiː	spa'lhtatí:	caress sby to help them sleep
spa.ˈma.ma̰	spamáma'	velvety
sta̰h.ˈnḭ	sta'jní'	1) wet; 2) bloody
sta.ˈwaː	stawá:	weave sth out of roots, reeds or cords (net, basket, etc.)
s’a̰.ʔa̰ː.ˈna̰	s’a'ha:'ná'	sweat
ˈs’a̰.ta̰	s’á'ta'	1) child, baby, young of any animal; 2) doll; 3) (ni) one's infant
s’aɬ	s’alh	loosely, not well-placed
ˈs’a̰ɬ.wa̰	s’á'lhwa'	1) slow of movement or thought (of people and animals); 2) calm; 3) serious, quiet
s’ḛː.ˈʔeː	s’e:’hé:	etch sth, cut a groove in sth
ˈs’ḛʔ.tḭ	s’é'hti'	unidentified vine with solid, heart-shaped leaves, used by shaman to wrap chicks in before burying them alive
s’e.ti.ˈna̰ʔ	s’etiná'h	Chigger, Harvest Mite
s’o̰.ʔa.ˈna̰	s’o'haná'	person who hugs
s’oʔ.ta.ˈma̰ʔ	s’ohtamá'h	unidentified plant that ejects its seeds when they are mature
s’o.li.ˈnḭ	s’oliní'	whistler, musician
ʃ’aː.ˈna̰	x’a:ná'	corn-shucker

Table A.1: Upper Necaxa Totonac wordlist.

IPA	UNT	English
ʃ’a̰.ˈʔa̰ːʔ	x’a'há:'h	1) daylight; 2) moonlight; 3) (ni) its light, its shine
ʃ’a̰.ˈtan.ʔa̰	x’a'tánha'	reddish, color of ripe fruits such as guavas
ʃ’ḛ.ˈʔḛ.lḭ	x’e'hé'li'	rough, pock-marked
ʃ’e.ˈtim	x’etím	seeded and de-veined chili
ʃ’oː.ˈlu.lṵ	x’o:lúlu'	unidentified species of large green lizard
ʃ’o̰p.ˈliː	x’o'plí:	slim down, get thin
ʃ’o.nun.ˈʔo̰ː	x’onunhó:’	unidentified blood-sucking arthropod that lives in houses, lays eggs in clothing
ʃa̰ː.ˈʔan	xa:'hán	take a steam bath
ˈʃaː.ʃlḭ	xá:xli'	tepache, a drink made from fruit juices and fermented pineapple rinds
ʃas.tan.kṵ	xastánku'	1) youngest sibling; 2) socially inferior person
ʃka̰h	xka'j	pineapple
ʃka.ˈnan	xkanán	bite, sting
ʃliː.ma̰.ˈʔa̰ːs	xli:ma'há:'s	a long time ago, at a definite or experienced time in the past [a long time (Kirchner)]
ʃpa̰.ˈta	xpa'tá	grind something with a molcajete
ʃpḭ.pḭ.ˈleːʔ	xpi'pi'lé:h	butterfly
ʃta̰.ˈpu	xta'pú	1) dam sth (water) to make a pool; 2) dredge sth out, remove stones from sth (stream) to make it deeper; 3) trap or fish or shrimp by creating a pool on the edge of sth (river) and letting it dry up
ʃtaː.lak.ˈʃtim	xta:lakxtím	the same, evenly
ʃuː.ˈwiɬ	xu:wílh	dried minnows

Table A.1: Upper Necaxa Totonac wordlist.

IPA	UNT	English
ta̰ː.la̰.ʔa.s’a.ˈwi	ta:'la'has’awí	1) win sth away from sby; 2) win a thing that sby doesn’t want to give up using sth
ta̰ː.lak.s’a̰.ˈtan	ta:'laks’a'tán	infant of one's own age
ta̰ː.sa̰.ˈnḭn	ta:'sa'ní'n	1) gossip; 2) story, fiction
ta̰ːˈʃ’aː	ta:'x’á:	shuck sth (corn) with sby
ta.ʔeɬ’a.ˈmaːn	tahelh’amá:n	joke
ta.ˈlaː.ʃ’a̰	talá:x’a'	unidentified tree with a long, wide seedpod
tam.pṵˈ.ɬuː	tampu'lhú:	pull sth up by the roots
tan.ʔeː.ʃ’a̰.ˈʔaː	tanhe:x’a'há:	begin to lighten on the horizon at dawn
ta.piː.s’aː.na̰ː.ˈpa̰ʔ.ɬma̰	tapi:s’a:na:'pá'hlhma'	unidentified species of plant, used to treat wounds
ta.s’a.ˈwi	tas’awí	1) lose, be defeated (fight, bet, game); 2) break a promise, not do what one says
ta.saː.ta.ˈnuːn	tasa:tanú:n	stuck, fixed in place
ta.ˈʃ’aːn	tax’á:n	shucked (corn)
ta.ʃ’a̰.ˈʔeːt	tax’a'hé:t	1) daylight, sunlight; 2) aurora; 3) (ni) its light, its inherent brightness (light source), its sparkle
ta.ʃa̰.ˈʔa	taxa'há	scratch oneself
te.heːʃ.ˈkaːn	teje:xká:n	water running in the street
tsa.ˈla̰h	tzalá'j	brittle, fragile, thin (stick)
tsḭn.tziː.ˈpa̰ʔɬ.tʃa̰	tzi'ntzi:pá'hlhcha'	cuatomate, small, tomato-like fruit (possible Solanum glaucescens)
tuːs.ˈpuɬ	tu:spúlh	one's toes
wah.taʃ.tu.ˈtʃa	wajtaxtuchá	1) leave, get out; 2) turn out; 3) come up (sun, moon)

Appendix B

Supplement to Chapter 3

B.1 Annotation tools

The following table shows the values for each segment according to 19 feature-like dimensions that were used to code the collocate segments for corpus analysis.

Table B.1: Segment ID grid by which segments and collocates were coded for analysis

	a	e	i	o	u	lary	stress	long	fric	stop	nas	approx	vowel	alv	postalv	lat	velar	bilab	seg
ˈa̰ː	T	F	F	F	F	T	T	T	F	F	F	F	T	F	F	F	F	F	T
ˈḛː	F	T	F	F	F	T	T	T	F	F	F	F	T	F	F	F	F	F	T
ˈḭː	F	F	T	F	F	T	T	T	F	F	F	F	T	F	F	F	F	F	T
ˈo̰ː	F	F	F	T	F	T	T	T	F	F	F	F	T	F	F	F	F	F	T
ˈṵː	F	F	F	F	T	T	T	T	F	F	F	F	T	F	F	F	F	F	T
a̰ː	T	F	F	F	F	T	F	T	F	F	F	F	T	F	F	F	F	F	T
ḛː	F	T	F	F	F	T	F	T	F	F	F	F	T	F	F	F	F	F	T
ḭː	F	F	T	F	F	T	F	T	F	F	F	F	T	F	F	F	F	F	T
o̰ː	F	F	F	T	F	T	F	T	F	F	F	F	T	F	F	F	F	F	T
ṵː	F	F	F	F	T	T	F	T	F	F	F	F	T	F	F	F	F	F	T
ˈa̰	T	F	F	F	F	T	T	F	F	F	F	F	T	F	F	F	F	F	T
ˈḛ	F	T	F	F	F	T	T	F	F	F	F	F	T	F	F	F	F	F	T
ˈḭ	F	F	T	F	F	T	T	F	F	F	F	F	T	F	F	F	F	F	T
ˈo̰	F	F	F	T	F	T	T	F	F	F	F	F	T	F	F	F	F	F	T
ˈṵ	F	F	F	F	T	T	T	F	F	F	F	F	T	F	F	F	F	F	T
a̰	T	F	F	F	F	T	F	F	F	F	F	F	T	F	F	F	F	F	T
ḛ	F	T	F	F	F	T	F	F	F	F	F	F	T	F	F	F	F	F	T
ḭ	F	F	T	F	F	T	F	F	F	F	F	F	T	F	F	F	F	F	T
o̰	F	F	F	T	F	T	F	F	F	F	F	F	T	F	F	F	F	F	T
ṵ	F	F	F	F	T	T	F	F	F	F	F	F	T	F	F	F	F	F	T
ɬ’	F	F	F	F	F	T	F	F	T	F	F	F	F	F	F	T	F	F	T
s’	F	F	F	F	F	T	F	F	T	F	F	F	F	T	F	F	F	F	T
ʃ’	F	F	F	F	F	T	F	F	T	F	F	F	F	F	T	F	F	F	T
ʔ	F	F	F	F	F	T	F	F	F	T	F	F	F	F	F	F	F	F	T
ˈa	T	F	F	F	F	F	T	F	F	F	F	F	T	F	F	F	F	F	T
ˈaː	T	F	F	F	F	F	T	T	F	F	F	F	T	F	F	F	F	F	T
ˈeː	F	T	F	F	F	F	T	T	F	F	F	F	T	F	F	F	F	F	T
ˈiː	F	F	T	F	F	F	T	T	F	F	F	F	T	F	F	F	F	F	T

Table B.1: Segment ID grid by which segments and collocates were coded for analysis

	a	e	i	o	u	lary	stress	long	fric	stop	nas	approx	vowel	alv	postalv	lat	velar	bilab	seg
ˈoː	F	F	F	T	F	F	T	T	F	F	F	F	T	F	F	F	F	F	T
ˈuː	F	F	F	F	T	F	T	T	F	F	F	F	T	F	F	F	F	F	T
aː	T	F	F	F	F	F	F	T	F	F	F	F	T	F	F	F	F	F	T
eː	F	T	F	F	F	F	F	T	F	F	F	F	T	F	F	F	F	F	T
iː	F	F	T	F	F	F	F	T	F	F	F	F	T	F	F	F	F	F	T
oː	F	F	F	T	F	F	F	T	F	F	F	F	T	F	F	F	F	F	T
uː	F	F	F	F	T	F	F	T	F	F	F	F	T	F	F	F	F	F	T
ˈe	F	T	F	F	F	F	T	F	F	F	F	F	T	F	F	F	F	F	T
ˈi	F	F	T	F	F	F	T	F	F	F	F	F	T	F	F	F	F	F	T
ˈo	F	F	F	T	F	F	T	F	F	F	F	F	T	F	F	F	F	F	T
ˈu	F	F	F	F	T	F	T	F	F	F	F	F	T	F	F	F	F	F	T
a	T	F	F	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	T
a	F	T	F	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	T
i	F	F	T	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	T
o	F	F	F	T	F	F	F	F	F	F	F	F	T	F	F	F	F	F	T
u	F	F	F	F	T	F	F	F	F	F	F	F	T	F	F	F	F	F	T
j	F	F	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	T
l	F	F	F	F	F	F	F	F	F	F	F	T	F	F	F	T	F	F	T
ɾ	F	F	F	F	F	F	F	F	F	F	F	T	F	T	F	F	F	F	T
w	F	F	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	T	T
b	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	F	T	T
d	F	F	F	F	F	F	F	F	F	T	F	F	F	T	F	F	F	F	T
k	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	T	F	T
ɬ	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	T	F	F	T
m	F	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	T	T
n	F	F	F	F	F	F	F	F	F	F	T	F	F	T	F	F	F	F	T
ŋ	F	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	T	F	T
p	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	F	T	T

Table B.1: Segment ID grid by which segments and collocates were coded for analysis

	a	e	i	o	u	lary	stress	long	fric	stop	nas	approx	vowel	alv	postalv	lat	velar	bilab	seg
s	F	F	F	F	F	F	F	F	T	F	F	F	F	T	F	F	F	F	T
ʃ	F	F	F	F	F	F	F	F	T	F	F	F	F	F	T	F	F	F	T
t	F	F	F	F	F	F	F	F	F	T	F	F	F	T	F	F	F	F	T
ts	F	F	F	F	F	F	F	F	F	F	F	F	F	T	F	F	F	F	T
tʃ	F	F	F	F	F	F	F	F	F	F	F	F	F	F	T	F	F	F	T
x	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	T	F	T
#	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F

B.2 Chi-squared test results

Table B.2: Chi-squared tests of preceding contexts for stops, fricatives, and vowels individually, and all three segment classes together. Percentages are calculated from the matrix total. Counts in bold contributed most to the Chi-squared value. Compare to 3.8.

		Preceding Context					
		Non-laryng	Laryng		χ^2	df	p
Segment Class				**Total**			
Stops	oral	6833 (54.99%)	1882 (15.14%)	8715			
	glottal	718 (5.78%)	**2994 (24.09%)**	3712			
	Total	7551	4876	12427	3808.9	1	< 0.0001
Fricatives	pulmonic	4645 (72.15%)	1793 (27.85%)	6438			
	ejective	313 (4.86%)	**123 (1.91%)**	436			
	Total	4958	1916	6874	0.03	1	0.87
Vowels	-laryng	17587 (60.72%)	2125 (7.34%)	19712			
	+laryng	8509 (29.38%)	**745 (2.57%)**	9254			
	Total	26096	2870	28966	52.57	1	< 0.0001
All	-laryng	46155 (67.12%)	9932 (14.44%)	56087			
	+laryng	8815 (12.82%)	**3862 (5.62%)**	12677			
	Total	54970	13794	68764	1049.3	1	< 0.0001

Table B.3: Chi-squared tests of following contexts for stops, fricatives, and vowels individually, and all three segment classes together. Compare to Table 3.9.

		Following Context Non-laryng	Laryng		χ^2	df	p
Segment class				**Total**			
Stops	oral	8690 (55.33%)	2910 (18.53%)	11600			
	glottal	3468 (22.08%)	**637 (4.06%)**	4105			
	Total	12158	3547	15705	158.77	1	<0.0001
Fricatives	pulmonic	5192 (82.52%)	664 (10.55%)	5856			
	ejective	243 (3.86%)	**193 (3.07%)**	436			
	Total	5435	857	6292	373.94	1	<0.0001
Vowels	-laryng	18958 (65.45%)	754 (2.60%)	19712			
	+laryng	6222 (21.48%)	**3032 (10.47%)**	9254			
	Total	25180	3786	28966	4641.8	1	<0.0001
All	-laryng	53804 (69.40%)	9933 (12.81%)	54797			
	+laryng	9933 (12.81%)	**3862 (4.8%)**	13795			
	Total	63737	13795	77532	1194.3	1	<0.0001

Table B.4: Chi-squared tests of independence between laryngealization in consonants (stops and fricatives only) and preceding vowels. Percentages are calculated based on the total number of observations in each 2 x 2 matrix.

		Preceding Context					
		Non-laryng	Laryng		χ^2	df	p
Segment class				**Total**			
Stops	oral	3889 (43.71%)	1609 (18.08 %)	5498			
	glottal	**406 (4.56%)**	2994 (33.65%)	3400			
	Total	4295	4603	8898	2908.3	1	<0.0001
Fricatives	pulmonic	2562 (76.52%)	619 (18.48%)	3181			
	ejective	138 (4.12%)	**29 (0.87%)**	167			
	Total	2700	648	3348	0.45	1	0.50

Table B.5: Chi-squared tests of independence between laryngealization in consonants (stops and fricatives only) and following vowels. Percentages are calculated based on the total number of observations in each 2 x 2 matrix.

		Following Context					
		Non-laryng	Laryng		χ^2	df	p
Segment class				**Total**			
Stops	oral	6800 (56.19%)	2882 (23.82%)	9682			
	glottal	1876 (15.15%)	**543 (4.49%)**	2419			
	Total	8676	3425	12101	51.1	1	<0.0001
Fricatives	pulmonic	963 (47.23%)	640 (31.38%)	1603			
	ejective	243 (11.92%)	**193 (9.47%)**	436			
	Total	1206	833	2039	2.67	1	0.10

Appendix C

Supplement to Chapter 4

C.1 Materials

Orthographic forms and glosses of the word list are provided in Appendix A. The word list materials presented in Appendix A were further classified for use in the analysis in 4. The table below provides information on every vowel of all word list items. Each vowel appears in the table according to the word it appears in and its order of appearance. The table also identifies each vowel for every word list item according to its inherent laryngealization category, as well as that of its preceding and following segmental environments. A '+' indicates that the segment or environment is laryngealized, while a '-' indicates that it is not. The '#' symbol indicates a word boundary and entails the omission of the vowel from the revelant analytical grouping. For example, vowels appearing at the start of a word, indicated by a '#' in the **Preceding** column of the table, were omitted from the analysis of laryngealization in preceding contexts. Words in parentheses indicate a continuation from the previous page.

Table C.1: Word list according to vowel condition. Columns indicate laryngealization with either '+' or '-'; '#' indicates a word boundary. See Chapter 4 for further details.

Word (IPA)	Segment	Vowel	Preceding	Following
a̰ː.ˈtuːs	a̰ː	+	#	-
	uː	-	-	-
a̰.ʔa.ta.puː.ˈʃ’aː.ja̰	a̰	+	#	+
	a	-	+	-
	a	-	-	-
	uː	-	-	+
	aː	-	+	-
	a̰	+	-	#
a̰ʔ.ˈɬ’oː.ma̰	a̰	+	#	+
	oː	-	+	-
	a̰	+	-	#
a̰ʔ.s’a.wi.ˈnḭ	a̰	+	#	+
	a	-	+	-
	i	-	-	-
	ḭ	+	-	#
a.ka.kuː.ˈluːkɬ	a	-	#	-
	a	-	-	-
	uː	-	-	-
	uː	-	-	-
a̰ː.tṵː.tʃiː.ˈjeːtɬ, a̰ː.tṵː.tʃiː.ˈjeːkɬ	a̰ː	+	#	-
	ṵː	+	-	-
	iː	-	-	-
	eː	-	-	-
tʃa̰ː.ˈʔo̰.ʃ’a̰	a̰ː	+	-	+
	o̰	+	+	+
	a̰	+	+	#
tʃa̰.ˈʔaːn	a̰	+	-	+
	aː	-	+	-
tʃa.ˈnaː	a	-	-	-
	aː	-	-	#
tʃi.lḭ.ˈlḭkʃ	i	-	-	-
	ḭ	+	-	-
	ḭ	+	-	-
tʃḭpʃ	ḭ	+	-	-
tʃḭʃ.ˈkṵ	ḭ	+	-	-
	ṵ	+	-	#
tʃṵ.ˈtʃo̰ʔʃ	ṵ	+	-	-
	o̰	+	-	+
ʔeː.tʃi.ˈwiʃ	eː	-	+	-
	i	-	-	-

Table C.1: Word list with experimental conditions from Chapter 4.

Word (IPA)	Segment	Vowel	Preceding	Following
(ʔeː.tʃi.ˈwiʃ)	i	-	-	-
ʔeː.ˈʔa̰ːʃ	eː	-	+	+
	a̰	+	+	-
ʔeː.mi.ˈjaːɬ	eː	-	+	-
	i	-	-	-
	aː	-	-	-
ʔeː.ta.wa.ˈka̰ɬ	eː	-	+	-
	a	-	-	-
	a	-	-	-
	a̰	+	-	-
ʔeɬ.ˈʔoː.ʃ’a̰	e	-	+	-
	oː	-	+	+
	a̰	+	+	#
ʔeːɬ.kṵ.ˈtan	eː	-	+	-
	ṵ	+	-	#
	a	-	-	-
ʔeɬ.taʃ.ˈtu	e	-	+	-
	a	-	-	-
	u	-	-	#
ʔen.ʔa.liː.ˈs’o.lḭ	e	-	+	-
	a	-	+	-
	iː	-	-	+
	o	-	+	-
	ḭ	+	-	#
ˈʔo̰.ʃ’a̰	o̰	+	+	+
	a̰	+	+	#
ʔo̰.ʃ’a̰.ʔaː.ˈɬwa̰ːʔ	o̰	+	+	+
	a̰	+	+	+
	aː	-	+	-
	a̰	+	-	+
ʔo.ˈɬ’aː.wa̰	o	-	+	+
	aː	-	+	-
	a̰	+	-	#
ḭɬ.ˈtḭn	ḭ	+	#	-
	ḭ	+	-	-
ḭʃ.ʔeːn.ḭʃ.ma.ˈkan	ḭ	+	#	-
	eː	-	+	-
	ḭ	+	-	-
	a	-	-	-
	a	-	-	-
ka̰ː.s’e.ˈwi.wḭ	a̰ː	+	-	+

Table C.1: *Word list with experimental conditions from Chapter 4.*

Word (IPA)	Segment	Vowel	Preceding	Following
(ka̰ː.s’e.ˈwi.wḭ)	e	-	+	-
	i	-	-	-
	ḭ	+	-	#
ka̰ː.ʃ’a̰.ˈʔa̰ːʔ	a̰	+	-	+
	a̰	+	+	+
	a̰	+	+	+
kiɬ.ˈʔo̰.ʃ’a̰	i	-	-	-
	o̰	+	+	+
	a̰	+	+	#
kiɬ.pa̰n.ɬṵ.lṵ	i	-	-	-
	a̰	+	-	-
	ṵ	+	-	-
	ṵ	+	-	#
la̰ʔ.ˈƛ’aː	a̰	+	-	+
	aː	-	+	#
la̰ʔ.ˈƛ’o.nḭ	a̰	+	-	+
	o	-	+	-
	ḭ	+	-	#
la.ka.man.saː.ˈnas	a	-	-	-
	a	-	-	-
	a	-	-	-
	aː	-	-	-
	a	-	-	-
liː.ˈƛ’a̰ːn	iː	-	-	+
	a̰	+	+	-
liː.ɬ’o.ˈluː	iː	-	-	+
	o	-	+	-
	uː	-	-	#
liːs.ka.ˈlah.wa̰	iː	-	-	-
	a	-	-	-
	a	-	-	-
	a̰	+	-	#
liːʃ.pa̰.ˈtan	iː	-	-	-
	a̰	+	-	-
	a	-	-	-
lu.ˈlo̰ʔʃ	u	-	-	-
	o̰	+	-	+
ɬa.ˈpa	a	-	-	-
	a	-	-	#
ɬapɬ	a	-	-	-
ɬḛ.ʔeː.ˈnin	ḛ	+	-	+

Table C.1: Word list with experimental conditions from Chapter 4.

Word (IPA)	Segment	Vowel	Preceding	Following
(ɬḛ.ʔeː.ˈnin)	eː	-	+	-
	i	-	-	-
ɬkaːk.ˈnan	aː	-	-	-
	a	-	-	-
ɬpa̰.ˈma.ma̰	a̰	+	-	-
	a	-	-	-
	a̰	+	-	#
ˈɬkah.wa̰	a	-	-	-
	a̰	+	-	#
ɬka.ˈwat	a	-	-	-
	a	-	-	-
ɬpa.ˈpa̰ʔ	a	-	-	-
	a̰	+	-	+
ɬta̰.ˈʔa̰h	a̰	+	-	+
	a̰	+	+	-
ɬta.ˈka̰.la̰	a	-	-	-
	a̰	+	-	-
	a̰	+	-	#
ɬta.ˈyat	a	-	-	-
	a	-	-	-
ɬ’a̰ː.ˈna̰	a̰	+	+	-
	a̰	+	-	#
ɬ’ḛ.ˈʔeː.lḭ	ḛ	+	+	+
	eː	-	+	-
	ḭ	+	-	#
ɬ’e.ˈli	e	-	+	-
	i	-	-	#
ɬ’ḭ.wi.ˈliː	ḭ	+	+	-
	i	-	-	-
	iː	-	-	#
ɬ’oː.ˈma.wa̰	oː	-	+	-
	a	-	-	-
	a̰	+	-	#
ɬ’o̰.ˈʔo.lṵ	o̰	+	+	+
	o	-	+	-
	ṵ	+	-	#
ˈɬ’o.lṵ	o	-	+	-
	ṵ	+	-	#
maː.s’a̰.ʔa̰ː.ˈniː	aː	-	-	+
	a̰	+	+	+
	a̰	+	+	-

Table C.1: Word list with experimental conditions from Chapter 4.

Word (IPA)	Segment	Vowel	Preceding	Following
(maː.s’a̰.ʔa̰ː.ˈniː)	iː	-	-	#
maː.s’a̰.taːˈnan	aː	-	-	+
	a̰	+	+	-
	aː	-	-	-
	a	-	-	-
maː.s’ḛʔ.ˈniː	aː	-	-	+
	ḛ	+	+	+
	iː	-	-	#
maː.s’o̰.ˈʔoː	aː	-	-	+
	o̰	+	+	+
	oː	-	+	#
maːs.kaː.ˈkiː	aː	-	-	-
	aː	-	-	-
	iː	-	-	#
maː.ʃ’a̰.ʔeː.ˈnin	aː	-	-	+
	a̰	+	+	+
	eː	-	+	-
	i	-	-	-
ma̰ʔ.ˈʃuː	a̰	+	-	+
	uː	-	-	#
paːʃ.ˈkat ka̰.ˈtsiːɬ	aː	-	-	-
	a	-	-	-
	a̰	+	-	-
	iː	-	-	-
piː.ˈs’aːm	iː	-	-	+
	aː	-	+	-
pḭ.ˈpḛʔs	ḭ	+	-	-
	ḛ	+	-	+
piʃ.ˈʔa̰.ɬa̰	i	-	-	-
	a̰	+	+	-
	a̰	+	-	#
piʃ.ˈpam.wa̰	i	-	-	-
	a	-	-	-
	a̰	+	-	#
po̰ʔɬ	o̰	+	-	+
puː.ɬḛ.ˈʔe	uː	-	-	-
	ḛ	+	-	+
	e	-	+	#
ˈpuː.ʃ’a̰	uː	-	-	+
	a̰	+	+	#
puks	u	-	-	-

Table C.1: Word list with experimental conditions from Chapter 4.

Word (IPA)	Segment	Vowel	Preceding	Following
sa̰.ˈsa	a̰	+	-	-
	a	-	-	#
sa.ˈlun	a	-	-	-
	u	-	-	-
ˈsḛʔ.sḭ	ḛ	+	-	+
	ḭ	+	-	#
ˈska̰t.lḭ	a̰	+	-	-
	ḭ	+	-	#
ska.ˈma.ma̰	a	-	-	-
	a	-	-	-
	a̰	+	-	#
sla̰ʔ.s’o̰.ˈʔo̰h.wa̰	a̰	+	-	+
	o̰	+	+	+
	o̰	+	+	-
	a̰	+	-	#
spa̰ɬ.ta.ˈtiː	a̰	+	-	-
	a	-	-	-
	iː	-	-	#
spa.ˈma.ma̰	a	-	-	-
	a	-	-	-
	a̰	+	-	#
sta̰h.ˈnḭ	a̰	+	-	-
	ḭ	+	-	#
sta.ˈwaː	a	-	-	-
	aː	-	-	#
s’a̰.ʔa̰ː.ˈna̰	a̰	+	+	+
	a̰	+	+	-
	a̰	+	-	#
ˈs’a̰.ta̰	a̰	+	+	-
	a̰	+	-	#
s’aɬ	a	-	+	-
ˈs’a̰ɬ.wa̰	a̰	+	+	-
	a̰	+	-	#
s’ḛː.ˈʔeː	ḛː	+	+	+
	eː	-	-	#
ˈs’ḛʔ.tḭ	ḛ	+	+	+
	ḭ	+	-	#
s’e.ti.ˈna̰ʔ	e	-	+	-
	i	-	-	-
	a̰	+	-	+
s’o̰.ʔa.ˈna̰	o̰	+	+	+

Table C.1: Word list with experimental conditions from Chapter 4.

Word (IPA)	Segment	Vowel	Preceding	Following
(s’o̰.ʔa.ˈna̰)	a	-	+	-
	a̰	+	-	#
s’oʔ.ta.ˈma̰ʔ	o	-	+	+
	a	-	-	-
	a̰	+	-	+
s’o.li.ˈnḭ	o	-	+	-
	i	-	-	-
	ḭ	+	-	#
ʃ’aː.ˈna̰	aː	-	+	-
	a̰	+	-	#
ʃ’a̰.ˈʔa̰ːʔ	a̰	+	+	+
	a̰ː	+	+	+
ʃ’a̰.ˈtan.ʔa̰	a̰	+	+	-
	a	-	-	-
	a̰	+	+	#
ʃ’ḛ.ˈʔḛ.lḭ	ḛ	+	+	+
	ḛ	+	+	-
	ḭ	+	-	#
ʃ’e.ˈtim	e	-	+	-
	i	-	-	-
ʃ’oː.ˈlu.lṵ	o	-	+	-
	u	-	-	-
	ṵ	+	-	#
ʃ’o̰p.ˈliː	o̰	+	+	-
	iː	-	-	#
ʃ’o.nun.ˈʔo̰ː	o	-	+	-
	u	-	-	-
	o̰	+	+	#
ʃa̰ː.ˈʔan	a̰	+	-	+
	a	-	+	-
ˈʃaː.ʃlḭ	aː	-	-	-
	ḭ	+	-	#
ʃas.tan.kṵ	a	-	-	-
	a	-	-	-
	ṵ	+	-	#
ʃka̰h	a̰	+	-	-
ʃka.ˈnan	a	-	-	-
	a	-	-	-
ʃliː.ma̰.ˈʔa̰ːs	iː	-	-	-
	a̰	+	-	+
	a̰ː	+	+	-

Table C.1: Word list with experimental conditions from Chapter 4.

Word (IPA)	Segment	Vowel	Preceding	Following
ʃpa̰.ˈta	a̰	+	-	-
	a	-	-	#
ʃpḭ.pḭ.ˈleːʔ	ḭ	+	-	-
	ḭ	+	-	-
	eː	-	-	+
ʃta̰.ˈpu	a̰	+	-	-
	u	-	-	#
ʃtaː.lak.ˈʃtim	aː	-	-	-
	a	-	-	-
	i	-	-	-
ʃuː.ˈwiɬ	uː	-	-	-
	i	-	-	-
ta̰ː.la̰.ʔa.s’a.ˈwi	a̰	+	-	-
	a̰	+	-	+
	a	-	+	+
	a	-	+	-
	i	-	-	#
ta̰ː.lak.s’a̰.ˈtan	a̰ː	+	-	-
	a	-	-	-
	a̰	+	+	-
	a	-	-	-
ta̰ː.sa̰.ˈnḭn	a̰ː	+	-	-
	a̰	+	-	-
	ḭ	+	-	-
ta̰ːˈʃ’aː	a̰ː	+	-	+
	aː	-	+	#
ta.ʔeɬ’a.ˈmaːn	a	-	-	+
	e	-	+	-
	a	-	+	-
	aː	-	-	-
ta.ˈlaː.ʃ’a̰	a	-	-	-
	aː	-	-	+
	a̰	+	+	#
tam.pṵˈ.ɬuː	a	-	-	-
	ṵ	+	-	-
	uː	-	-	#
tan.ʔeː.ʃ’a̰.ˈʔaː	a	-	-	-
	eː	-	+	+
	a̰	+	+	+
	aː	-	+	#
ta.piː.s’aː.na̰ː.ˈpa̰ʔ.ɬma̰	a	-	-	-

Table C.1: Word list with experimental conditions from Chapter 4.

Word (IPA)	Segment	Vowel	Preceding	Following
(ta.piː.s’aː.na̰ː.ˈpa̰ʔ.ɬma̰)	iː	-	-	+
	aː	-	+	-
	a̰	+	-	-
	a̰	+	-	+
	a̰	+	-	-
ta.s’a.ˈwi	a	-	-	-
	a	-	+	-
	i	-	-	#
ta.saː.ta.ˈnuːn	a	-	-	-
	a	-	-	-
	a	-	-	-
	uː	-	-	-
ta.ˈʃ’aːn	a	-	-	+
	aː	-	+	-
ta.ʃ’a̰.ˈʔeːt	a	-	-	+
	a̰	+	+	+
	eː	-	+	-
ta.ʃa̰.ˈʔa	a	-	-	-
	a̰	+	-	+
	a	-	+	#
te.heːʃ.ˈkaːn	e	-	-	-
	eː	-	-	-
	aː	-	-	-
tsa.ˈla̰h	a	-	-	-
	a̰	+	-	-
tsḭn.tziː.ˈpa̰ʔɬ.tʃa̰	ḭ	+	-	-
	iː	-	-	-
	a̰	+	-	+
	a̰	+	-	#
tuːs.ˈpuɬ	uː	-	-	-
	u	-	-	-
wah.taʃ.tu.ˈtʃa	a	-	-	-
	a	-	-	-
	u	-	-	-
	a	-	-	#

C.2 Supplemental models

The models presented here were built with data from Time points 1 and 3. Due to multiple interactions in the data, subsequent analyses in Chapter 4 subdivided the data into smaller groups based on time point, simplifying the models and allowing for more straightforward interpretation of the effects and interactions, where present. Models of the complete data sets for each measure and CV/VC condition are presented here.

The best linear mixed effects regression model of H1-H2 included random intercepts for Speaker, Word, and Vowel Quality, and random slopes by for Vowel Laryngealization, Consonant Laryngealization, and Stress by Speaker. Main effects of Vowel Laryngealization, Consonant Laryngealization, Consonant Manner, Stress, and Time were included in the model, resulting in significant three-way interactions between all fixed effects terms except for stress. Stress was excluded from interactions due to missing data in some conditions. The model output summary

Table C.2: Summary of trimmed model of H1-H2 values for vowels following consonants (CV condition.) Highly significant three-way interactions warranted further analyses with time split into separate models.

coef	Estimate	Std. Error	df	t value	Pr(>—t—)	
(Intercept)	−1.0463	1.9509	5.7196	−0.5363	0.6119	
vlaryngyes	1.4120	0.5495	24.4747	2.5698	0.0167	*
claryngyes	1.4943	0.5705	12.6677	2.6191	0.0216	*
time3	2.7102	0.4255	4912.4777	6.3703	0.0000	***
cmannerstop	−0.4646	0.3827	4036.4073	−1.2140	0.2248	
stressunstressed	−0.0556	0.4614	3.2300	−0.1205	0.9112	
vlaryngyes:claryngyes	−1.1580	0.4489	4018.9641	−2.5795	0.0099	**
vlaryngyes:time3	−5.4144	0.5311	4913.2138	−10.1942	0.0000	***
claryngyes:time3	−2.0824	0.4762	4912.5359	−4.3725	0.0000	***
vlaryngyes:cmannerstop	−1.1137	0.4509	4428.0903	−2.4699	0.0136	*
time3:cmannerstop	−0.4513	0.4496	4912.4512	−1.0038	0.3155	
claryngyes:cmannerstop	1.4468	0.4592	3895.0249	3.1506	0.0016	**
vlaryngyes:claryngyes:time3	2.2403	0.5131	4912.5038	4.3661	0.0000	***
vlaryngyes:time3:cmannerstop	2.6963	0.5266	4913.5071	5.1198	0.0000	***
claryngyes:time3:cmannerstop	2.0125	0.5240	4913.0020	3.8402	0.0001	***

Table C.3: Summary of trimmed model of H1-H2 values for vowels preceding consonants (VC condition.) Significant three- and four-way interactions warranted further analyses with time split into separate models.

coef	Estimate	Std. Error	df	t value	Pr(>—t—)	
(Intercept)	−1.4921	1.8841	4.2065	−0.7919	0.4707	
vlaryngyes	1.1324	0.4947	1442.4843	2.2891	0.0222	*
claryngyes	0.6192	0.4865	17.8723	1.2728	0.2194	
cmannerstop	−0.1636	0.3410	2148.0665	−0.4798	0.6314	
time3	2.8181	0.2766	3529.3637	10.1881	0.0000	***
stressunstressed	0.5192	0.2283	11.3347	2.2742	0.0433	*
vlaryngyes:claryngyes	−2.1791	0.7583	1713.2899	−2.8737	0.0041	**
vlaryngyes:cmannerstop	1.1599	0.6489	1481.9870	1.7875	0.0741	
claryngyes:cmannerstop	−2.5461	0.6260	2847.3002	−4.0671	0.0000	***
vlaryngyes:time3	−1.4471	0.5515	3529.8918	−2.6239	0.0087	**
claryngyes:time3	0.8364	0.4312	3531.2087	1.9396	0.0525	
cmannerstop:time3	−2.1404	0.3885	3529.5896	−5.5101	0.0000	***
vlaryngyes:claryngyes:cmannerstop	1.6690	0.9839	2140.3561	1.6963	0.0900	
vlaryngyes:claryngyes:time3	1.9340	0.8579	3531.6078	2.2545	0.0242	*
vlaryngyes:cmannerstop:time3	−0.2100	0.7197	3530.4575	−0.2918	0.7704	
claryngyes:cmannerstop:time3	−1.0639	0.7659	3530.6530	−1.3890	0.1649	
vlaryngyes:claryngyes:cmannerstop:time3	−3.7098	1.1579	3531.3639	−3.2040	0.0014	**

Appendix D

Supplement to Chapter 5

D.1 Materials

Orthographic forms and glosses of the word list are provided in Appendix A. The table below provides information regarding the frication conditions and word position of frication for each of the word list items. Where there are multiple phones involving frication, each segment is presented on its own line, in order of its appearance in the word. Certain tokens were excluded according to criteria spelled out in the main text of Chapter 5.

Table D.1: Word list according to frication conditions. See Chapter 5 for further details.

IPA	Segment	Condition	Position
a̰ː.ˈtuːs	s	simplex	final
a̰.ʔa.ta.puː.ˈʃ’aː.ja̰	ʃ’	ejective	medial
a̰ʔ.ˈɬ’oː.ma̰	ɬ’	ejective	medial
a̰ʔ.s’a.wi.ˈnḭ	s’	ejective	medial
a.ka.kuː.ˈluːkɬ	ɬ	cluster	final
a̰ː.tṵː.tʃiː.ˈjeːtɬ	tʃ	affricate	medial
	ɬ	cluster	final
tʃa̰ː.ˈʔo̰.ʃ’a̰	tʃ	affricate	initial
	ʃ’	ejective	medial
tʃa̰.ˈʔaːn	tʃ	affricate	initial
tʃa.ˈnaː	tʃ	affricate	initial

Table D.1: Word list with experimental conditions from Chapter 5.

IPA	Segment	Condition	Position
tʃi.lḭ.ˈlḭkʃ	tʃ	affricate	initial
	ʃ	cluster	final
tʃḭpʃ	tʃ	affricate	initial
	ʃ	cluster	final
tʃḭʃ.ˈkṵ	tʃ	affricate	initial
	ʃ	cluster	medial
tʃṵ.ˈtʃo̰ʔʃ	tʃ	affricate	initial
	tʃ	affricate	medial
	ʃ	cluster	final
ʔeː.tʃi.ˈwiʃ	tʃ	affricate	medial
	ʃ	simplex	final
ʔeː.ˈʔa̰ːʃ	ʃ	simplex	final
ʔeː.mi.ˈjaːɬ	ɬ	simplex	final
ʔeː.ta.wa.ˈka̰ɬ	ɬ	simplex	final
ʔeɬ.ˈʔoː.ʃ’a̰	ɬ	cluster	initial
	ʃ’	ejective	medial
ʔeːɬ.kṵ.ˈtan	ɬ	cluster	medial
ʔeɬ.taʃ.ˈtu	ɬ	cluster	medial
	ʃ	cluster	medial
ʔen.ʔa.liː.ˈs’o.lḭ	s’	ejective	medial
ˈʔo̰.ʃ’a̰	ʃ’	ejective	medial
ʔo̰.ʃ’a̰.ʔaː.ˈɬwa̰ːʔ	ʃ’	ejective	medial
	ɬ	cluster	medial
ʔo.ˈɬ’aː.wa̰	ɬ’	ejective	medial
ḭɬ.ˈtḭn	ɬ	cluster	medial
ḭʃ.ʔeːn.ḭʃ.ma.ˈkan	ʃ	cluster	medial
	ʃ	cluster	medial

Table D.1: Word list with experimental conditions from Chapter 5.

IPA	Segment	Condition	Position
ka̰ː.s’e.ˈwi.wḭ	s’	ejective	medial
ka̰ː.ʃ’a̰.ˈʔa̰ːʔ	ʃ’	ejective	medial
kiɬ.ˈʔo̰.ʃ’a̰	ɬ	cluster	medial
	ʃ’	ejective	medial
kiɬ.pa̰n.ɬṵ.lṵ	ɬ	cluster	medial
	ɬ	cluster	medial
la̰ʔ.ˈɬ’aː	ɬ’	ejective	initial
la̰ʔ.ˈɬ’o.nḭ	ɬ’	ejective	initial
la.ka.man.saː.ˈnas	s	cluster	medial
	s	simplex	final
liː.ˈɬ’a̰ːn	ɬ’	ejective	medial
liː.ɬ’o.ˈluː	ɬ’	ejective	medial
liːs.ka.ˈlah.wa̰	s	cluster	medial
liːʃ.pa̰.ˈtan	ʃ	cluster	medial
lu.ˈlo̰ʔʃ	ʃ	cluster	final
ɬa.ˈpa	ɬ	simplex	initial
ɬapɬ	ɬ	simplex	initial
	ɬ	cluster	final
ɬḛ.ʔeː.ˈnin	ɬ	simplex	initial
ɬkaːk.ˈnan	ɬ	cluster	initial
ɬpa̰.ˈma.ma̰	ɬ	cluster	initial
ˈɬkah.wa̰	ɬ	cluster	initial
ɬka.ˈwat	ɬ	cluster	initial
ɬpa.ˈpa̰ʔ	ɬ	cluster	initial
ɬta̰.ˈʔa̰h	ɬ	cluster	initial
ɬta.ˈka̰.la̰	ɬ	cluster	initial
ɬta.ˈyat	ɬ	cluster	initial

Table D.1: Word list with experimental conditions from Chapter 5.

IPA	Segment	Condition	Position
ɬ’a̰ː.ˈna̰	ɬ’	ejective	initial
ɬ’ḛ.ˈʔeː.lḭ	ɬ’	ejective	initial
ɬ’e.ˈli	ɬ’	ejective	initial
ɬ’ḭ.wi.ˈliː	ɬ’	ejective	initial
ɬ’oː.ˈma.wa̰	ɬ’	ejective	initial
ɬ’o̰.ˈʔo.lṵ	ɬ’	ejective	initial
ˈɬ’o.lṵ	ɬ’	ejective	initial
maː.s’a̰.ʔa̰ː.ˈniː	s’	ejective	medial
maː.s’a̰.taːˈnan	s’	ejective	medial
maː.s’ḛʔ.ˈniː	s’	ejective	medial
maː.s’o̰.ˈʔoː	s’	ejective	medial
maːs.kaː.ˈkiː	s	cluster	medial
maː.ʃ’a̰.ʔeː.ˈnin	ʃ’	ejective	medial
ma̰ʔ.ˈʃuː	ʃ	cluster	medial
paːʃ.ˈkat ka̰.ˈtsiːɬ	ʃ	cluster	medial
	ts	affricate	medial
	ɬ	simplex	final
piː.ˈs’aːm	s’	ejective	medial
pḭ.ˈpḛʔs	s	cluster	final
piʃ.ˈʔa̰.ɬa̰	ʃ	cluster	medial
	ɬ	simplex	medial
piʃ.ˈpam.wa̰	ʃ	cluster	medial
po̰ʔɬ	ɬ	cluster	final
puː.ɬḛ.ˈʔe	ɬ	simplex	medial
ˈpuː.ʃ’a̰	ʃ’	ejective	medial
puks	s	cluster	final
sa̰.ˈsa	s	simplex	initial

Table D.1: Word list with experimental conditions from Chapter 5.

IPA	Segment	Condition	Position
	s	simplex	medial
sa.ˈlun	s	simplex	initial
ˈsḛʔ.sḭ	s	simplex	initial
	s	cluster	medial
ˈska̰t.lḭ	s	cluster	initial
ska.ˈma.ma̰	s	cluster	initial
sla̰ʔ.s’o̰.ˈʔo̰h.wa̰	s	cluster	initial
	s’	ejective	medial
spa̰ɬ.ta.ˈtiː	s	cluster	initial
	ɬ	cluster	medial
spa.ˈma.ma̰	s	cluster	initial
sta̰h.ˈnḭ	s	cluster	initial
sta.ˈwaː	s	cluster	initial
s’a̰.ʔa̰ː.ˈna̰	s’	ejective	initial
ˈs’a̰.ta̰	s’	ejective	initial
s’aɬ	s’	ejective	initial
	ɬ	simplex	final
ˈs’a̰ɬ.wa̰	s’	ejective	initial
	ɬ	cluster	medial
s’ḛː.ˈʔeː	s’	ejective	initial
ˈs’ḛʔ.tḭ	s’	ejective	initial
s’e.ti.ˈna̰ʔ	s’	ejective	initial
s’o̰.ʔa.ˈna̰	s’	ejective	initial
s’oʔ.ta.ˈma̰ʔ	s’	ejective	initial
s’o.li.ˈnḭ	s’	ejective	initial
ʃ’aː.ˈna̰	ʃ’	ejective	initial
ʃ’a̰.ˈʔa̰ːʔ	ʃ’	ejective	initial

Table D.1: Word list with experimental conditions from Chapter 5.

IPA	Segment	Condition	Position
ʃ’a̰.ˈtan.ʔa̰	ʃ’	ejective	initial
ʃ’ḛ.ˈʔḛ.lḭ	ʃ’	ejective	initial
ʃ’e.ˈtim	ʃ’	ejective	initial
ʃ’oː.ˈlu.lṵ	ʃ’	ejective	initial
ʃ’o̰p.ˈliː	ʃ’	ejective	initial
ʃ’o.nun.ˈʔo̰ː	ʃ’	ejective	initial
ʃa̰ː.ˈʔan	ʃ	simplex	initial
ˈʃaː.ʃlḭ	ʃ	simplex	initial
	ʃ	cluster	medial
ʃas.tan.kṵ	ʃ	simplex	initial
	s	cluster	medial
ʃka̰h	ʃ	cluster	initial
ʃka.ˈnan	ʃ	cluster	initial
ʃliː.ma̰.ˈʔa̰ːs	ʃ	cluster	initial
	s	simplex	final
ʃpa̰.ˈta	ʃ	cluster	initial
ʃpḭ.pḭ.ˈleːʔ	ʃ	cluster	initial
ʃta̰.ˈpu	ʃ	cluster	initial
ʃtaː.lak.ˈʃtim	ʃ	cluster	medial
	ʃ	cluster	initial
ʃuː.ˈwiɬ	ʃ	simplex	initial
	ɬ	simplex	final
ta̰ː.la̰.ʔa.s’a.ˈwi	s’	ejective	medial
ta̰ː.lak.s’a̰.ˈtan	s’	ejective	medial
ta̰ː.sa̰.ˈnḭn	s	simplex	medial
ta̰ːˈʃ’aː	ʃ’	ejective	medial
ta.ʔeɬ’a.ˈmaːn	ɬ’	ejective	medial

Table D.1: Word list with experimental conditions from Chapter 5.

IPA	Segment	Condition	Position
ta.ˈlaː.ʃ’a̰	ʃ’	ejective	medial
tam.pṵˈ.ɬuː	ʃ’	ejective	medial
tan.ʔeː.ʃ’a̰.ˈʔaː	ʃ’	ejective	medial
ta.piː.s’aː.na̰ː.ˈpa̰ʔ.ɬma̰	s’	ejective	medial
	ɬ	cluster	medial
ta.s’a.ˈwi	s’	ejective	medial
ta.saː.ta.ˈnuːn	s	simplex	medial
ta.ˈʃ’aːn	ʃ’	ejective	medial
ta.ʃ’a̰.ˈʔeːt	ʃ’	ejective	medial
ta.ʃa̰.ˈʔa	ʃ	simplex	medial
te.heːʃ.ˈkaːn	ʃ	cluster	medial
tsa.ˈla̰h	ts	affricate	initial
tsḭn.tsiː.ˈpa̰ʔɬ.tʃa̰	ts	affricate	initial
	ts	affricate	medial
	ɬ	cluster	medial
	tʃ	affricate	medial
tuːs.ˈpuɬ	s	cluster	medial
	ɬ	simplex	final
wah.taʃ.tu.ˈtʃa	ʃ	cluster	medial
	tʃ	affricate	medial