Vowels and Stress in Chichicastenango K'iche'

by

Elizabeth Anne Wood

Report

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF ARTS

THE UNIVERSITY OF TEXAS AT AUSTIN

August, 2020

Vowels and Stress in Chichicastenango K'iche'

APPROVED BY

SUPERVISING COMMITTEE:

Nora C. England, Supervisor

Scott P. Myers, Supervisor

Acknowledgments

Primero de todo me gustaría agradecer a todos los hablantes de k'iche' de los municipios de Chichicastenango y Nahualá que me han apoyado en esta investigación: Sebastián Tiniguar, Manuel Tahay y su familia, Miguel Ignacio y su familia, Alicia Algua, Manuel Chicoj, Wilson Conoz, Sebastián González, Sonia González, Magdalena Jerónimo, Teresa Jerónimo, Mildred Mejía, Celso Pérez, Rafaela Sen, Juan Tiriquiz, Juana Tol, María Xirum, y muchos más. Gracias por compartir la riqueza de su idioma y cultura conmigo. Maltyox chech alaq, maltyox chiwe!

I am also grateful for all of the support that I have received from my professors and fellow students in the Linguistics department at UT Austin, and especially from my supervisors Nora England and Scott Myers. The phonetics and phonology research group and the documentary and descriptive linguistics research group both provided me with valuable ideas and feedback. And my wonderful cohort and the west cubicles were a source of both academic and emotional support.

Finally, thank you to my family for encouraging me through all my adventures in grad school, and for catching so many typos in earlier drafts of this paper.

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. 000392968. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

Vowels and Stress in Chichicastenango K'iche'

Elizabeth Anne Wood, M.A.
The University of Texas at Austin, 2020

Supervisors: Nora C. England
Scott P. Myers

This papers presents an exploration of the vowel inventory and patterns of stress and vowel deletion in the phonologically innovative variety of K'iche' (Eastern Mayan) spoken in Chichicastenango, Guatemala. The vowel inventory (vowel quality and duration) is investigated through an acoustic production study. Results suggest that this dialect has five peripheral vowels /a e i o u/ and four or five centralized vowels /ə ɛ (ɪ) ɔ ʊ/. The vowel /ɪ/ may be merging with surrounding phonemes. There are not consistent duration differences between the peripheral and centralized vowels. The stress pattern is investigated through a database of polysyllabic forms collected from elicitation and texts, where the most prominent syllable in each word is determined impressionistically by the author. The preliminary results discussed here suggest that stress falls on the final syllable in non-verbs, but falls according to a hierarchy of syllable weights in verbs (CVC > CV > C$\ddot{\text{V}}$C > C$\ddot{\text{V}}$). Vowel deletion is investigated through the same database of polysyllabic forms, as well as additional textual materials. Results indicate that deletion in content words appears to be restricted to unstressed centralized vowels in non-final CV syllables. Deletion in function words appears to be less restricted. This study significantly expands the understanding of vowels and stress in Chichicastenango K'iche', building on previous descriptions and raising new questions for future research.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

This paper provides an investigation into the vowel system of the variety of K'iche' spoken in Chichicastenango, Guatemala. This dialect has an innovative and non-standard vowel system for K'iche' that has only been superficially described in the past, leaving many questions. Specifically, this dialect is characterized by replacing the length contrast found elsewhere with a quality contrast dubbed "tense" vs. "lax", by extensive vowel deletion, and by a shift in the stress pattern from fixed final to in some cases quantity-sensitive. This paper explores the following three topics:

1. Identity of the vowels: What is the phonetic identity of each vowel? Does this dialect maintain 10 phonemic vowels distinguished purely by quality, or are there duration differences as well? What distinguishes tense and lax vowels?

2. Stress: What is the stress pattern of this dialect? What factors are relevant to stress placement? Is stress the same for all types of words?

3. Vowel deletion: Under what environments are vowels deleted? Are only short unstressed vowels subject to deletion? Are there different types of deletion present in the language?

These three topics, though conceptually distinct, interact in a variety of ways. For instance, vowel quality appears to be a factor in the stress pattern of verbs, and both vowel quality and stress can be shown to be relevant to the patterns of vowel deletion.

The structure of the paper is the following. I begin with some background information on the K'iche' language, followed by a note on K'iche' orthography and an exploration of previous literature on the topics at hand. Chapter 2 details methodology and results of

a phonetic experiment investigating the duration and quality of the vowels, showing centralization as the main difference between tense and lax vowels and inconsistent duration differences. Chapter 3 explores the stress pattern of Chichicastenango K'iche', showing two distinct patterns which appear to be found in the language (fixed final stress for non-verbs, quantity/quality-sensitive stress for verbs). Chapter 4 addresses vowel deletion, showing some restrictions which it appears to obey and distinguishing between the patterns found in content words and function words and between diachronic vowel loss, synchronic vowel deletion, and optional vowels. Chapter 5 concludes, pointing out the highlights of this investigation and indicating future directions this research may take. A list of abbreviations and sources of data used throughout this document can be found in the appendices.

1.1 Language Background

K'iche' is a Mayan language belonging to the Eastern branch of the family (Campbell 2017). As of the 2001 census, it was spoken by around 900,000 speakers, located primarily in the northwestern highlands of Guatemala, in the departments of Quiché, Huehuetenango, Quetzaltenango, Retalhuleu, Sololá, Suchitepéquez, Totonicapán, San Marcos and Chimaltenango (Richards 2003). This makes it one of the most spoken and most well-known Mayan languages.

Following an initial division made by Kaufman, K'iche' is frequently classified into five broad dialect areas: West, East, Central, North and South (Par Sapón and Can Pixabaj 2000). However, even within each of these dialect areas there may be large differences between the varieties spoken in different towns, which are readily observed by speakers and linguists alike; this high degree of regional variation is found across the Mayan family (Romero 2017). Therefore, linguists usually identify varieties of K'iche' by the town (*municipio*) in which they are spoken (Velleman 2014); e.g., Chichicastenango K'iche', Nahualá K'iche', Totonicapán K'iche', etc. Speakers usually also identify most closely with their town (Can Pixabaj 2017). The two varieties of K'iche' most frequently mentioned in this document are those of Chichicastenango (abbreviated Chi) and Nahualá (abbreviated Nah); these belong to the Central and West dialect areas, respectively. Other towns included in the Central di-

alect area with Chichicastenango include Santa María Chiquimula, San Antonio Ilotenango and Santa Cruz de El Quiché (Par Sapón and Can Pixabaj 2000). Chichicastenango also borders with areas where the predominant language is Kaqchikel (Cojtí Ren 2019), a closely related Mayan language also in the K'ichean family (Campbell 2017). Figure 1.1 shows the dialect areas of K'iche'; Chichicastenango is represented in olive brown and Nahualá in blue.

In my observation, most K'iche' speakers refer to their language as *qach'ab'al* or *qatzij*, meaning 'our language'. They do not usually use the name K'iche' to refer to the language; this term is more often used to refer to the ethnic group or their historic nation. The language/ethnic group is sometimes also spelled *Quiché*, as historically spelled by the Spaniards; however, this is increasingly avoided by linguists. The word Quiché may also refer to the El Quiché Department or its capital city Santa Cruz del Quiché.

1.2 Orthographic Note

A K'iche' orthography was standardized by the *Academia de Lenguas Mayas de Guatemala* ("Academy of Mayan Languages of Guatemala") in 1987, by Decree 1046-87 (López Ixcoy 1997). This orthography corresponds to that listed in (1b) below. However, this orthography was never widely adopted since it was not the most suited for the majority of dialects. In practice, multiple orthographies are used by different individuals and organizations. These orthographies differ primarily in the spelling of vowels, and include the following:

1. Orthographies which distinguish 10 vowels, the maximum number found in K'iche' dialects. These may be of two types:
   (a) Orthographies which use single letters <a e i o u> for short/lax vowels and double letter <aa ee ii oo uu> for long/tense vowels
   (b) Orthographies which use single letters <a e i o u> for long/tense vowels and vowels with dieresis <ä ë ï ö ü> for short/lax vowels
2. Orthographies which distinguish 5 vowels, written <a e i o u>, leaving the distinction between short/lax and long/tense unmarked in writing

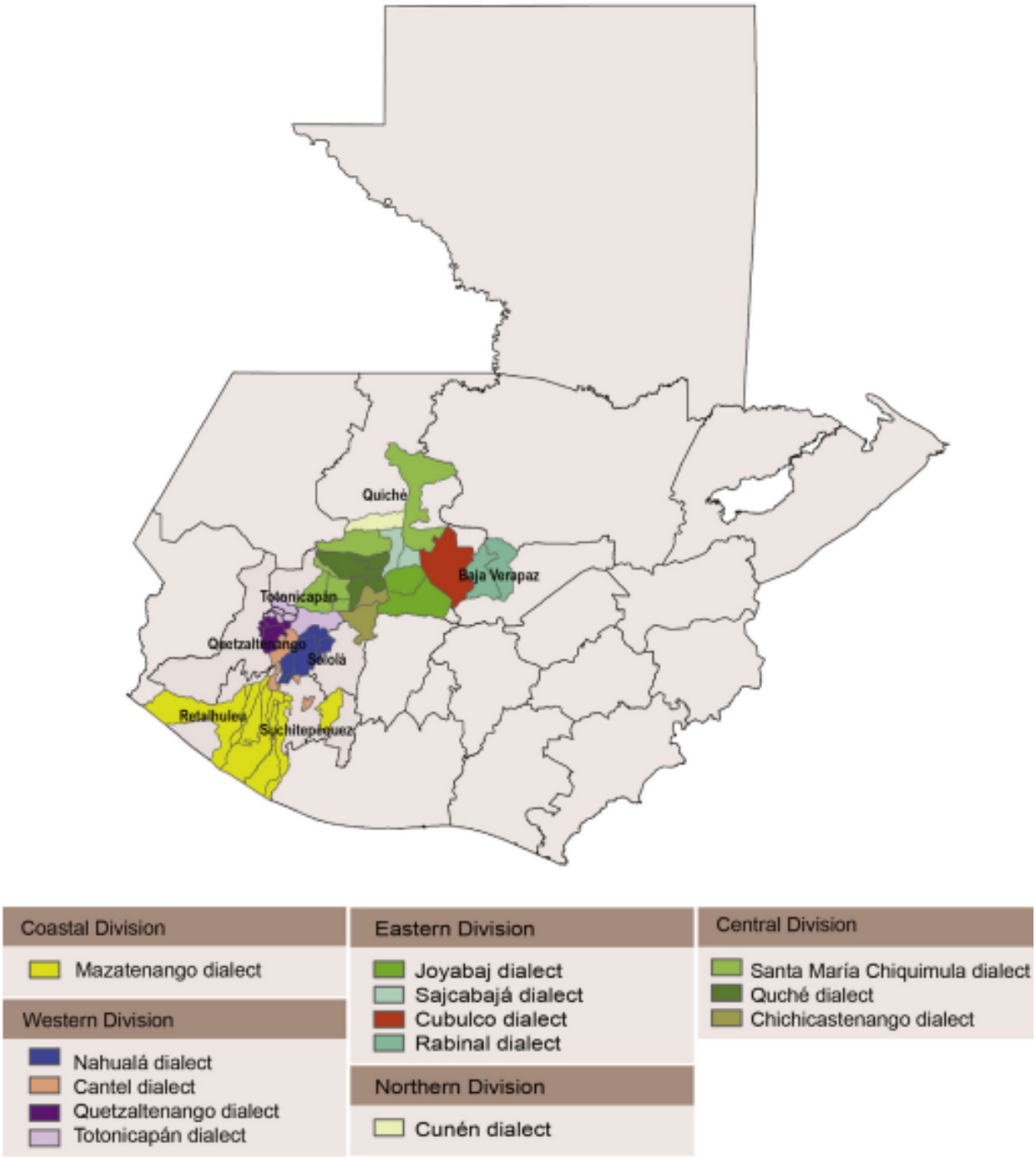

Figure 1.1: K'iche' dialect areas (Romero 2016)

As can be observed, this has the potential to create quite a lot of confusion, as the same symbol (single letter) may be used to mark short/lax vowels in some orthographies but

long/tense vowels in others.

This diversity of orthographies reflects differences in the vowel systems of different dialects of K'iche'. For dialects such as that of Nahualá, which have 10 contrastive vowels, 5 short and 5 long, the most practical system is the first described above (López Ixcoy 1994). However, for dialects such as Chichicastenango, in which the distinction is (at least described as) one of tenseness rather than length, the second is more practical (López Ixcoy 1994). Finally, for dialects which have lost some of the contrasts between vowel pairs, the third may be best.

This paper will primarily make reference to forms in the dialects of Chichicastenango and Nahualá. When Chichicastenango forms are referenced, they will be written in the tense/lax orthography, with single letters representing tense (long) vowels and dieresis marking lax (short) vowels. When forms from Nahualá are referenced, they will be written in the long/short orthography, with single letters representing short vowels and double letters long vowels. IPA transcriptions will additionally be provided in all cases.

Modern K'iche' orthographies all write consonants in a way that is quite similar to the IPA. Relevant differences are the following:

Orthography	b'	ch	j	tz	x	'	y
IPA	ɓ	t͡ʃ	χ	t͡s	ʃ	ʔ	j

Table 1.1: K'iche' orthography correspondence to IPA

Finally, a note on glossing conventions is in order. Due to the need to mark both morphological and syllable boundaries in glossed examples throughout this text, a multi-tier structure will be used. Each example will consist of the following 5 tiers: 1) orthography; 2) stress and syllable structure (IPA); 3) morpheme boundaries (IPA); 4) morpheme-by-morpheme glosses; 5) free translation and code to data source. Additionally, a recording of each example can be heard by clicking on the orthography tier. This structure can be seen in the following example, in which the tiers are numbered for explanatory purposes:

(1) 1. *xïnkchpa'*
2. ʃɪn.k[t͡ʃ(ə).ˈpaʔ]
3. ʃ-ɪn-k-t͡ʃ(ə)p-aʔ
4. CPL-B.1S-A.3P-hold-SS
5. 'They caught me.' [elic;MCX3]

In example glosses, the following two conventions are also used: vowels which are deleted in the surface form but posited to exist underlyingly are indicated in parentheses (as in (ə) above), and the stress domain is indicated in square brackets when some part of the word is not included in it (as in [t͡ʃ(ə).ˈpaʔ] above).

1.3 Previous Literature

As a phonologically innovative dialect, the Chichicastenango variety of K'iche' has received a fair amount of attention in the description of K'iche' vowels. However, there are many remaining questions about each of these topics. This section outlines what has been said before about Chichicastenango vowels, pointing out some areas where further work is necessary, several of which are treated in the following chapters.

1.3.1 Identity of the Vowels

López Ixcoy (1994) states that Chichicastenango K'iche', rather than contrasting five long and five short vowels, has a ten-vowel system which contrasts five vowels that are "tense" (*tensas*) and five that are "lax" (*relajadas*). This description is repeated in later works, such as Par Sapón and Can Pixabaj (2000), England and Baird (2017) and Can Pixabaj (2017). Although there has been no discussion of the precise meaning of the terms "tense" and "lax" in this context, it is likely that the difference the terms were intended to capture is one of centralization. This is inferred from the IPA labels give to these vowels, where the lax variants of the high and mid vowels /i u e o/ are the lower /ɪ ʊ ɛ ɔ/ but the lax variant of the low vowel /a/ is the higher /ə/ (López Ixcoy 1994; Par Sapón and Can Pixabaj 2000; Can Pixabaj 2017).

It is worth noting that this is not what is often understood by the terms "tense" and "lax" in other languages outside of the Mayan family, where these terms may be used to describe pure height differences or advanced/retracted tongue root systems like those found in many African languages (e.g. Fulop et al. 1998). However, these terms when used in reference to the neighboring Mayan language Kaqchikel also describe a centralization contrast (R. Bennett 2016b).

Although these various descriptions of Chichicastenango vowels agree that there is a quality difference between tense and lax vowels, there is disagreement about whether any duration difference can be found between the tense and lax sets. While Par Sapón and Can Pixabaj (2000) indicate that tense vowels are regularly long, although this is more clear in some cases than others, López Ixcoy (1994) states that there is no duration difference between the two.

Minimal (or near-minimal, in the case of /e ɛ/) pairs for the Chichicastenango vowels are shown in Table 1.2.

Lax Vowels				Tense Vowels			
/ə/	*chäj*	t͡ʃəχ	'pine tree'	/a/	*chaj*	t͡ʃaχ	'ashes'
/ɛ/	*xpëq*	ʃpɛq	'toad'	/e/	*xpetïk*	ʃpetɪk	's/he came'
/ɪ/	*k'ïx*	k'ɪʃ	'shame'	/i/	*k'ix*	k'iʃ	'thorn'
/ɔ/	*öj*	ɔχ	'we/us'	/o/	*oj*	oχ	'avocado'
/ʊ/	*tüx*	tʊʃ	'sprout'	/u/	*tux*	tuʃ	'kind of mange'

Table 1.2: Minimal pairs for Chichicastenango vowels (data from López Ixcoy (1994) and Can Pixabaj (2017), IPA added)

Although there is some degree of consensus in this description of the vowel inventory of Chichicastenango K'iche', all of these descriptions rely on auditory judgements of researchers. There may be differences between the vowels that are not easily noticed by non-native speakers of the language or of this dialect specifically, and furthermore there may be differences that even native speakers are not consciously aware of. For example, the quality difference between the two sets of vowels is quite salient, but this does not mean that there is no duration difference between them. For this reason, quantitative measurements of the vowels in acoustic and articulatory studies would greatly contribute to their description.

1.3.2 Cross-dialectal Correspondences and Neutralization

Generally speaking, lax vowels in Chichicastenango correspond to short vowels in other dialects, and tense vowels to long vowels in other dialects (Can Pixabaj 2017). However, this correspondence is not believed to hold in all syllables. López Ixcoy (1994), López Ixcoy (1997) and Par Sapón and Can Pixabaj (2000) state that the correspondence holds fully in monosyllabic words and in final syllables of polysyllabic words. It is apparently assumed in these works that stress is final in this dialect, like in most other dialects of K'iche'. Some examples from López Ixcoy (1997) are given in Table 1.3, showing the correspondence between lax and short, and tense and long, in final syllables.

Tense/lax varieties		Long/short varieties		Gloss
ch'äm	t͡ʃ'əm	*ch'am*	t͡ʃ'am	'sour'
k'ïx	k'ɪʃ	*k'ix*	k'iʃ	'shame'
choyob'äl	t͡ʃojoβəl	*choyob'al*	t͡ʃojoβal	'cutting instrument'
chun	t͡ʃun	*chuun*	t͡ʃuːn	'limestone'
k'oj	k'oχ	*k'ooj*	k'oːχ	'mask'
ab'ix	aβiʃ	*ab'iix*	aβiːʃ	'cornfield'

Table 1.3: Vowel correspondence in final syllables (data from López Ixcoy (1997), IPA added)

However, there is disagreement about what happens in non-final syllables. While López Ixcoy (1997) states that all non-final vowels are tense in Chichicastenango K'iche' (and short in other varieties), Par Sapón and Can Pixabaj (2000) mention that there are exceptions to this rule, and some lax vowels may be found in non-final syllables although as a general rule they are restricted to final syllables. Furthermore, this means that just like adding suffixes in long-short dialects results in long vowels becoming short, adding suffixes in Chichicastenango K'iche' results in lax vowels becoming tense (López Ixcoy 1997).

This characterization of neutralization, like the description of the vowel inventory, has necessarily relied on auditory impressions. However, if it is the case that there is neutralization of the contrast in non-final syllables or in unstressed syllables (which are not always non-final), then the realization of these vowels could have a larger range than that of vowels in the syllables where all 10 vowels are contrastive. Furthermore, my auditory judgements

do not agree with those in the literature. I observe a large number of lax vowels in non-final and unstressed syllables, not just a few exceptions, although I do also observe cases where short vowels in other dialects do correspond to tense-sounding vowels in Chichicastenango. For example, consider the following example of a word identified by López Ixcoy (1994) as having a non-final short vowel in other dialects which corresponds to a tense vowel in Chichicastenango:

(2) *ab'ix*
a.ˈɓiʃ
aɓiʃ
cornfield
'cornfield' [txt;planting]

The first vowel in this word corresponds to a short vowel in other dialects and does indeed sound very tense-like in the Chichicastenango data: it sounds closer to [a] than to [ə]. However, consider also the following examples which have very lax-sounding vowels in non-final unstressed syllables:

(3) a. *päjb'äl*
pəχ.ˈɓəl
pəχ-ɓəl
weigh-INSTR
'scale' [elic;MXM2]

b. *uk'äslmal*
u.k'əs.l(ɛ).ˈmal
u-k'əs-lɛmal
A.3S-live-?
'his/her life' [txt;church]

c. *säqwäch*
səq.ˈwət͡ʃ
səq-wət͡ʃ
white-face
'potato' [txt;healing]

The vowel in the unstressed first syllable of each of these words sounds much more like lax [ə] than tense [a]. A quantitative study of this suggested neutralization pattern would be very helpful to better understand this.

1.3.3 Deletion

The most detailed account of vowel deletion in Chichicastenango K'iche' is provided by Campbell (1977), who describes deletion as affecting all short vowels which are preceded by a consonant and followed by a stressed syllable. Some of his examples are shown in Table 4.1.

Surface form	Source	Gloss
ɾɓaːx	ɾ-**a**ɓaːx	'his rock'
t͡ʃkop	t͡ʃ**i**kop	'animal, bird'
t͡sˀnuːn	t͡sˀ**u**nun	'hummingbird'
kmat͡s	k**u**mat͡s	'snake'

Table 1.4: Examples of Chichicastenango vowel deletion according to Campbell (1977) (modified to match modern IPA; deleted vowels indicated in bold)

Others have not attempted comprehensive descriptions of deletion, but have remarked on the the fact that is is common in non-final syllables (López Ixcoy 1994), or in unstressed syllables (England and Baird 2017).

However, even if this description is accurate, many questions remain. All of the examples given by Campbell (1977) are of nouns with the structure CVCVC, where the first (short/lax) vowel is deleted and stress is on the second syllable. This raises the question of whether the same occurs in other word classes, or in words with different syllable structure or

different stress patterns. Additionally, there are many other factors which could be relevant to deletion which are not mentioned, such as register and speech rate.

1.3.4 Origins of the Tense/Lax Vowels

Proto-K'ichean (like Proto-Mayan) is reconstructed with long and short vowels (Campbell 2017), and therefore López Ixcoy (1994) and Par Sapón and Can Pixabaj (2000) hold the tense/lax vowel system found in Chichicasteango and Chiché to be an innovation. It is commonly mentioned in the literature that this vowel system is likely a result of contact with the neighboring language Kaqchikel, the only other Mayan language with 10 vowels divided into tense and lax sets (Par Sapón and Can Pixabaj 2000; Baird 2018). The municipality of Chichicastenango borders with modern-day Kaqchikel towns Tecpán (on the west) and Sololá and Concepción (on the south) (Cojtí Ren 2019), indicating that this contact and linguistic influence could be recent. However, as is pointed out by Cojtí Ren (2019), the tense and lax vowels are found homogeneously in all 90 villages that make up the municipality, and not just in the border region. Furthermore, England 1992 (in Baird 2010) points out that other K'iche' dialects outside of the Chichicastenango area have six vowels, with one being lax /ə/; these dialects are not in contact with Kaqchikel and therefore recent contact is an implausible source for this vowel.

An alternative is proposed by Cojtí Ren (2019), who places the contact in the pre-Columbian period. She argues from archeological evidence and historical documents that the Chichicastenango area was occupied by the Kaqchikel in the 15th century during their alliance with the K'iche'. During this period, the official language of the area was K'iche', which meant that the Kaqchikel in the area would have been bilingual. Due to their dominance in Kaqchikel over K'iche', they transferred certain elements of Kaqchikel into K'iche'. The most notable transfer would be the tense/lax vowels, but another possible example might be the absence of the honorific second person pronouns that are found in most other dialects of K'iche'. These honorific pronouns are also not used in three other varieties of K'iche' which have some lax vowels: Joyabaj, Zacualpa and Cunén (Par Sapón and Can Pixabaj 2000); these are also varieties believed to have had pre-Columbian contact with Kaqchikel. Cojtí Ren (2019) further points out that the influence of Kaqchikel in the Chichicastenango

area can also be seen in certain toponyms coming from Kaqchikel, such as *Saqiya'* (from Kaqchikel [səq] 'white' + [jaʔ] 'water') and *Chokoya'* (from Kaqchikel [t͡ʃokoj] 'parrot' + [jaʔ] 'water'). Crucially, these two toponyms both have the Kaqchikel term for water, [jaʔ], rather than the K'iche' term [χaʔ].

Chapter 2

Vowel Quality and Duration: An Experimental Investigation

2.1 Introduction

The vowel inventory of Chichicastenango K'iche' is described as consisting of five tense vowels /a e i o u/ and five lax vowels /ə ɛ ɪ ɔ ʊ/ (López Ixcoy 1994; López Ixcoy 1997; Par Sapón and Can Pixabaj 2000; Can Pixabaj 2017). Given the IPA labels, I deduce that tense and lax vowels differ in centralization: lax variants of the high and mid vowels are lower, but the lax variant of the low vowel is higher. López Ixcoy (1994) states that there is no duration difference between these two sets of vowels, while Par Sapón and Can Pixabaj (2000) indicate that tense vowels are usually long.

These generalizations about vowel quality and duration, though widely accepted, are purely impressionistic and are not supported by any available acoustic or articulatory evidence. This chapter details a production experiment which provides some preliminary acoustic data to address the questions of quality and duration of tense and lax vowels in Chichicastenango K'iche'. This experiment measured F1, F2 and duration of each of the 10 proposed Chichicastenango vowel phonemes, with the goal of determining the acoustic difference between "tense" and "lax" sets with respect to quality and duration. Results suggest that there is a difference that can be described as centralization between the tense and lax vowels at each place of articulation, while duration differences between tense and lax sets, though present in some cases, are not consistently found. However, there is also a lot of variability across speakers, lexical items, vowel pairs, and in some cases gender of the speaker. This variability points to the need for further research on vowels in this dialect.

2.2 Methods

2.2.1 Subjects

Eleven adult K'iche'-Spanish bilingual speakers were recruited for this experiment, all from the Chichicastenango area (city center or the surrounding communities Chij Tinimit, Pachoj, Chontalá, Chipacá, Chicua and Paquixic). The majority of the Chichicastenango community is bilingual with Spanish, and as such bilingual speakers are representative of the community. Reference codes for each participant, as well as their age group, sex and the community they belong to, are shown in Table 2.1.

Subject code	Age	Sex	Community
s1	young adult	female	Chij Tinimit
s2	young adult	male	city center
s3	young adult	female	Chij Tinimit
s4	young adult	female	Pachoj
s5	young adult	female	city center
s6	older adult	male	Chontalá
s7	older adult	male	Chipacá
s8	young adult	male	Chicuá
s9	young adult	female	Chicuá
s10	older adult	male	Chipacá
s11	older adult	female	Chij Tinimit

Table 2.1: Experiment subjects

2.2.2 Materials

Twenty items were used, two for each vowel phoneme. These are listed in Table 2.2.

Word Orthography	IPA	Gloss	Vowel	Word Orthography	IPA	Gloss	Vowel
k'ïx	k'iʃ	thorn	i	*chïq'*	t͡ʃɪq'	ripe, cooked	ɪ
q'ïj	q'iχ	sun		*pïx*	pɪʃ	tomato	
ch'ek	t͡ʃ'ek	knee	e	*q'ëq*	q'ɛq	black	ɛ
xex	ʃeʃ	stinky fish smell		*xpëq*	ʃpɛq	toad	
q'aq'	q'aq'	fire	a	*k'äx*	k'əʃ	difficult	ə
pas	pas	sash		*säq*	səq	white	
pop	pop	mat	o	*ch'ök*	t͡ʃ'ɔk	grackle	ɔ
q'oq'	q'oq'	*chilacayote* (squash)		*köj*	kɔχ	lion	
sutz'	sut͡s'	cloud	u	*t'üq*	t'ʊq	brood-hen	ʊ
tuj	tuχ	*temascal* (sweat bath)		*tüx*	tʊʃ	sprout	

Table 2.2: Experiment target words

All target words are monosyllabic, of the form CVC (as in /pas/ 'sash') or CCVC (as in /ʃpɛq/ 'toad'), where all consonants are voiceless obstruents /p t k q t͡ʃ t' k' q' t͡s' t͡ʃ' s ʃ χ/. When possible, common words were selected (e.g. /q'aq'/ 'fire', /səq/ 'white'). However, especially for the more uncommon mid vowels, this was not always possible, and less common words were used (e.g. /kɔχ/ 'lion').

The classification of the vowel in each word was checked in dictionaries of a related dialect (Ajpacajá Tum 2001; Ajpacajá Tum et al. 2005) through cross-dialect correspondence (long-tense, short-lax) since there is no dictionary of Chichicastenango K'iche'. One word, transcribed as /t͡ʃɪq'/, was not found in the dictionary, but it was included in the experiment due to the difficulty of finding suitable words with the phoneme /ɪ/. Classification of this vowel as /ɪ/ was done by analogy to the vowel in the word /t͡ʃɪk/ (*chïk* 'again, already'), which is known to correspond to short /i/ in related dialects. For reference, a spectrogram of these two words is shown in Figure 2.1, which shows that the vowels in the two words have very similar formant values. However, it is likely that this vowel was miscategorized, since a word transcribed as *chäq'* /t͡ʃaq'/ was later found in another dictionary of a related

variety (Christenson n.d.). Throughout the remainder of this chapter, the vowel in this word is treated as a lax high front vowel, but it is worth keeping in mind its possible miscategorization.

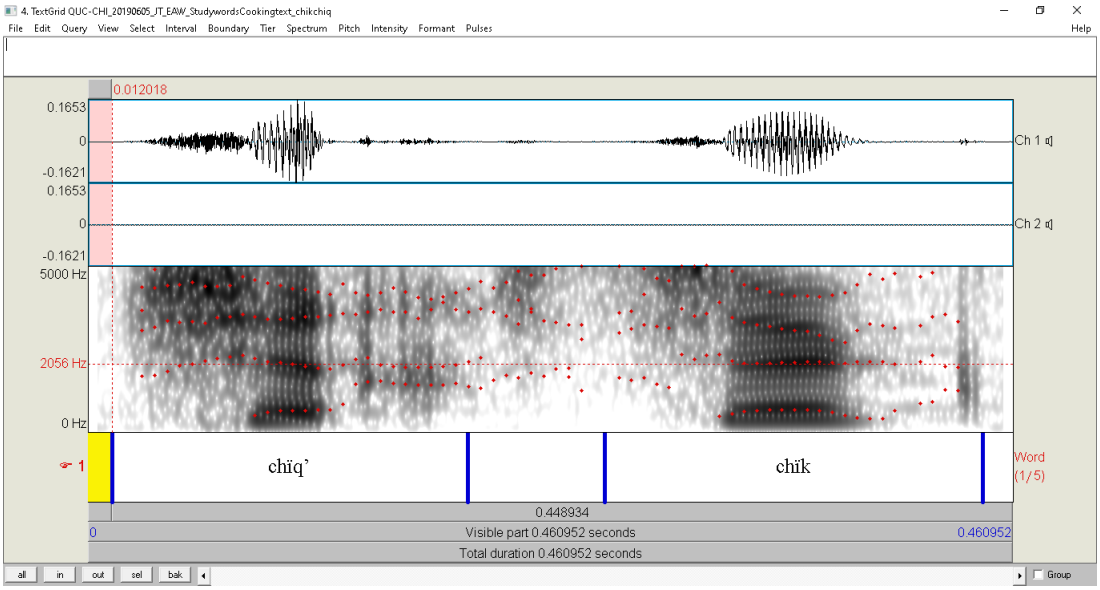

Figure 2.1: Spectrogram of the words *chïq'* and *chïk* from a free narration

2.2.3 Procedure

The experiment was conducted in a quiet room in the homes of the participants or at an office in the center of Chichicastenango. The experiment itself consisted of a translation task from Spanish. Participants memorized a carrier phrase *Ka'ij _ ju belt chïk* 'Say _ again', and inserted the relevant word from a Spanish prompt. They were asked to repeat this sentence with each word three times, creating 6 tokens for each vowel. Due to very low literacy rates in K'iche', no reading was used in this experiment. However, one participant who struggled to remember the carrier phrase did write it down, approximating the spelling from Spanish, and read it each time. Additionally, some participants produced slight variants of the carrier phrase, including changing *ka'ij* 'you say' to *kin'ij* 'I say' or *kq-ij* 'we say' and changing *ju belt chïk* 'again' for *jmul chïk* 'again'.

At times participants were unable to think of the target word directly from the Spanish translation, possible due to the use of some less common words or to the novelty of the task.

In this situation, I tried prompting speakers by very hesitantly pronouncing the target word, saying something like "I think I've heard people say a word that sounds a bit like X. Is that a word that could mean this?". They usually recognized the word, sometimes agreeing that it meant what I was asking about, and sometimes giving a slightly different meaning than the one I had originally proposed. In a few cases, even with prompting, participants did not recognize the word; in this case, I skipped that word and moved on to the next one. Occasionally participants produced a word that was not the target word, usually a word with a related meaning.

The entire experiment procedure was recorded with a Shure SM10A headset microphone connected to a Zoom H4n Pro Handy recorder, at a sampling rate of 44.1 kHz.

2.2.4 Measurements

Vowels were segmented in Praat (version 6.0.49, Boersma and Weenik 2019) with reference to the waveform by the following rules:

- The beginning of the vowel was marked at the point where amplitude begins to increase.
- The end of the vowel was marked at the point where amplitude is no longer decreasing.

A sample window can be seen in each of the following images. Figure 2.2 shows segmentation of a vowel between two stops (ʃpɛq 'toad'). The beginning of the vowel, where amplitude begins to increase, corresponds to the beginning of the stop burst for /p/ after the closure. The end of the vowel, where amplitude is no longer decreasing, corresponds to the beginning of the closure for /q/.

Figure 2.3 shows the segmentation of a vowel between two fricatives (ʃeʃ 'stinky fish smell'). The beginning of the vowel is marked where the amplitude begins to increase after the fairly steady amplitude during the preceding fricative, and the end of the vowel is marked where the amplitude no longer is decreasing and stays steady for the following fricative.

After segmentation, F1, F2 (at the midpoint) and duration of each vowel were extracted with a Praat script. For female speakers, the maximum formant value was set to 5500

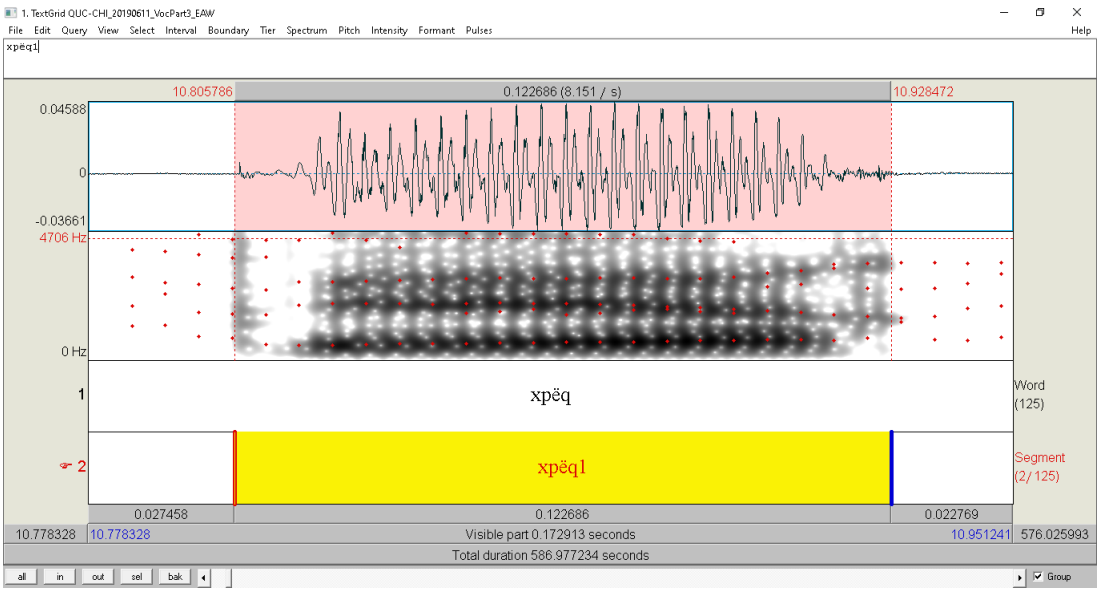

Figure 2.2: Segmentation of the vowel /ɛ/ between two stops

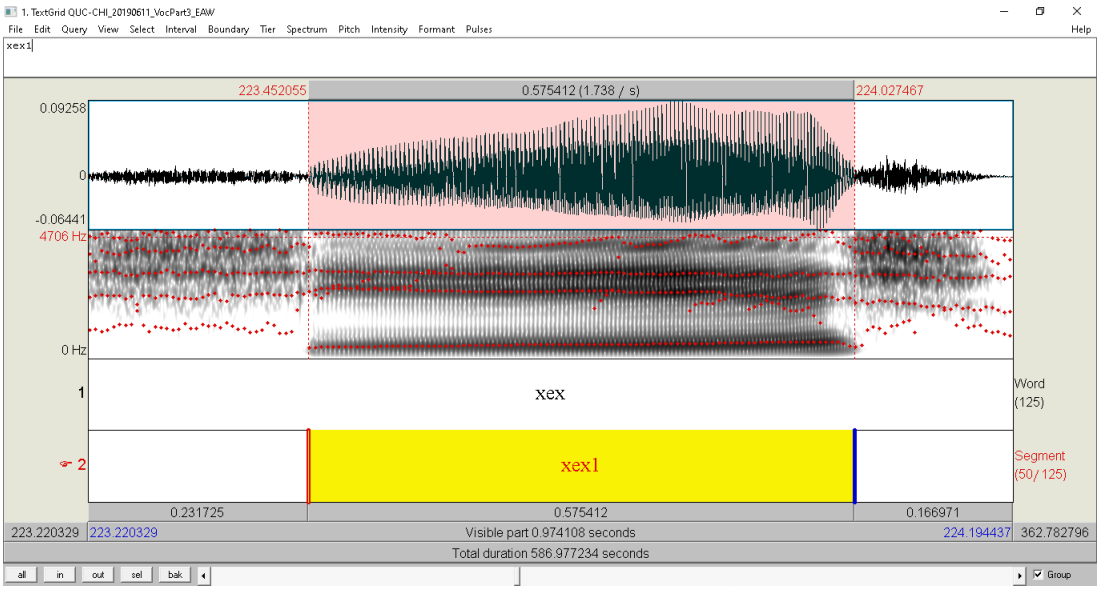

Figure 2.3: Segmentation of the vowel /e/ between two fricatives

Hz, and for male speakers, to 5000 Hz. Each participant's data was graphed individually in R (version 3.5.3, R Core Team 2019) using the package ggplot2 (Wickham 2016) in order to find outliers created by formant tracking errors; i.e. where Praat did not correctly identify the formants' locations. A variety of these were found and corrected. Once this was done, various graphs of the data or subsets of the data were made and observed.

2.2.5 Hypotheses

From previous descriptions of this dialect (López Ixcoy 1994; López Ixcoy 1997; Par Sapón and Can Pixabaj 2000), it was expected that the results would show 10 contrastive vowel qualities: two high front vowels, two high back vowels, two mid front vowels, two mid back vowels, and two low central vowels, with a tense and lax vowel at each place of articulation. High vowels were expected to have lower F1 values than mid vowels, and mid vowels to have lower F1 values than low vowels. The exception of this is the lax low vowel which, since it has been labeled /ə/, was expected to have F1 values similar to those of mid vowels. Front vowels were expected to have higher F2 values than central vowels, and central vowels higher F2 values than back vowels. Since the IPA labels previously given to the lax vowels suggest that they are more centralized than the tense vowels (lower for high and mid vowels but higher for the low vowel pair), the lax vowel of each pair was expected to be closer to the center of the vowel space than the tense vowel.

In keeping with cross-linguistically observed inherent duration differences (e.g. Lindblom 1968), low vowels were expected to have longer duration than mid vowels, and mid vowels longer than high vowels. Additionally, given the historic length contrast in this language (Campbell 1977), it was expected that there would be a duration difference between tense and lax vowels at each place of articulation, with the tense vowel being longer than the lax vowel in each case.

Significant differences between speakers, between the speech of men and women or between the contrast between tense and lax for each of the five different vowel pairs were not expected.

2.3 Results

During data collection, one participant (s7) produced almost all sentences with the target word in an unintended position, by shifting the target word to the beginning or end of the phrase. Therefore, all of this participant's data was excluded from the analysis and no measurements were taken. The same was true of close to half of the tokens of participant 11, as well as occasional tokens from other participants; these individual tokens were also

excluded. Finally, some participants failed to produce the correct target word from the Spanish prompt, resulting in some missing data. Excluded and missing data is summarized in Table 2.3.

Participant	Missing/excluded data
s1	all tokens of pop and kɔχ
s2	all tokens of kɔχ and tʊʃ
s3	1 token of t͡ʃ'ek
s4	-
s5	all tokens of kɔχ
s6	all tokens of kɔχ
s7	all data
s8	1 token of q'aq'
s9	all tokens of pɪʃ and t'ʊq
s10	-
s11	all tokens of k'iʃ, q'iχ, ʃeʃ, q'ɛq, k'əʃ, t͡ʃ'ɔk, kɔχ and sut͡s', 2 tokens of pop, q'oq' and tuχ, 1 token of q'aq'

Table 2.3: Data excluded or missing for each participant

The following sections detail the duration and formant results from the remaining data. In each section, results are shown for the whole group and then broken down by vowel pair, gender, speaker and item. As noted below, observing the differences in each of these dimensions reveals a great deal of variability in the data, which does not clearly follow the simple pattern expected from previous descriptions. Because of this variability observed in the data, the sample is too small for reliable generalizations, and therefore no statistical analysis is presented here beyond simple calculations of some means.[1] Rather, these results may be interpreted as a useful first glance at the variability in the system, which motivates a further study with more participants and more items for each vowel category.

2.3.1 Duration Results

The average duration of all vowels for each individual experiment participant is shown in Table 2.4.

[1] Averages calculated in R (R Core Team 2019)

Participant	Duration (ms)
s1	139
s2	124
s3	320
s4	119
s5	139
s6	113
s8	148
s9	131
s10	127
s11	132

Table 2.4: Average duration of all vowels for each experiment participant

This table shows that the average duration across all vowels is between 110 and 150 ms for every individual participant except for participant 3 who has an average of 320 ms, more than twice that of the next highest.

The duration of each vowel in the data is shown in Figure 2.4. This graph excludes data from subject 3, who had much longer durations on all vowels than the other participants, in order to more easily see the pattern. Tense vowels are in blue and lax in red.

The average values for the duration of each vowel are shown in Table 2.5, again excluding the data from participant 3.

Tense Vowels		Lax Vowels	
Vowel	Duration (ms)	Vowel	Duration (ms)
i	151	ɪ	103
e	148	ɛ	122
a	162	ə	91
o	136	ɔ	134
u	128	ʊ	130

Table 2.5: Average durations of each vowel, excluding data from participant 3

Comparing the average duration of the tense and lax variant in each vowel pair across all speakers, it can be seen that there is a difference for some pairs but not for others. There is a large difference between the average durations of the tense and lax central vowels, a

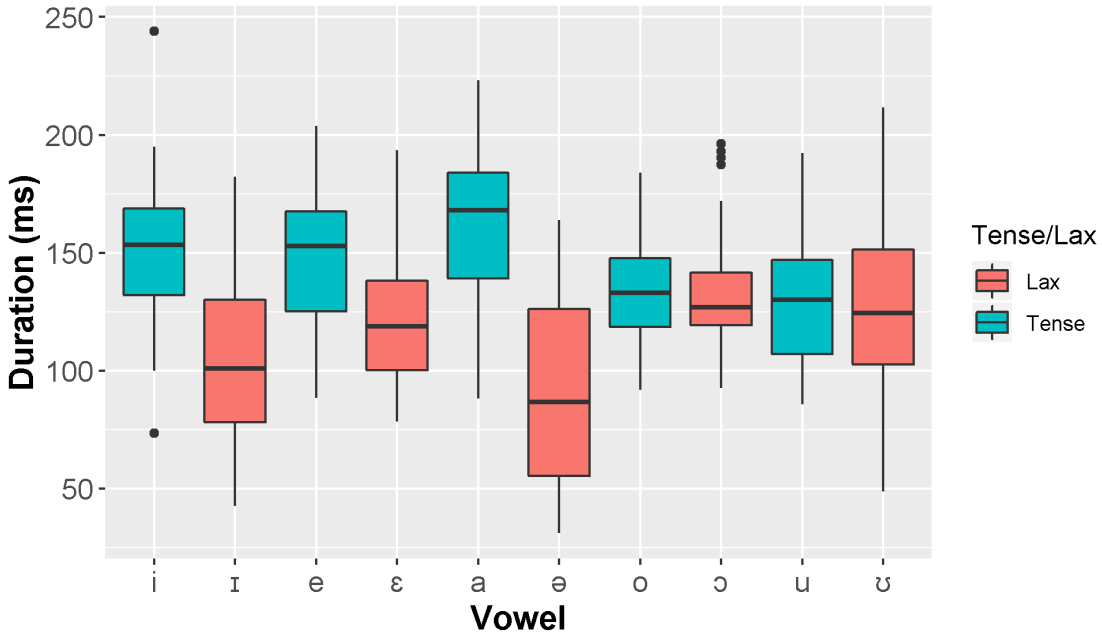

Figure 2.4: Duration results by vowel

smaller difference between the averages of the tense and lax front vowels, but almost no difference between the averages of the tense and lax back vowels. Excluding the data from participant 3, the vowel /a/ is about 70 ms longer on average than the vowel /ə/, the vowel /i/ is about 50 ms longer on average than the vowel /ɪ/ and the vowel /e/ is about 25 ms longer on average than the vowel /ɛ/. For the high and mid back vowel pairs in the data there is a tiny average difference of less than 5 ms between tense and lax vowels in each pair. Overall, front tense vowels have longer average durations than all back vowels (tense and lax), which have longer average durations than front lax vowels.

Comparing duration of each vowel according to vowel height, it can be seen that there is not a direct correlation between vowel height and duration. The vowel with the longest average duration is the low vowel /a/. However, mid vowels do not consistently have average durations longer than high vowels in this data, even taking into account the tense/lax contrast. Among lax vowels, the front high vowel /ɪ/ has a shorter average duration than the mid vowel /ɛ/, but high and mid back vowels have approximately the same average duration. Among tense vowels, the high and mid front vowels have approximately the same average durations, and so do the high and mid back vowels. The vowel with the shortest average

duration is /ə/, the lax counterpart of the low vowel /a/ (and modern reflux of historic short /a/) but itself a central vowel.

Individual graphs of the duration results of each experiment participant can be seen in Figure 2.5 (again excluding participant 3, whose averages are much larger than everyone else's, obscuring the pattern).

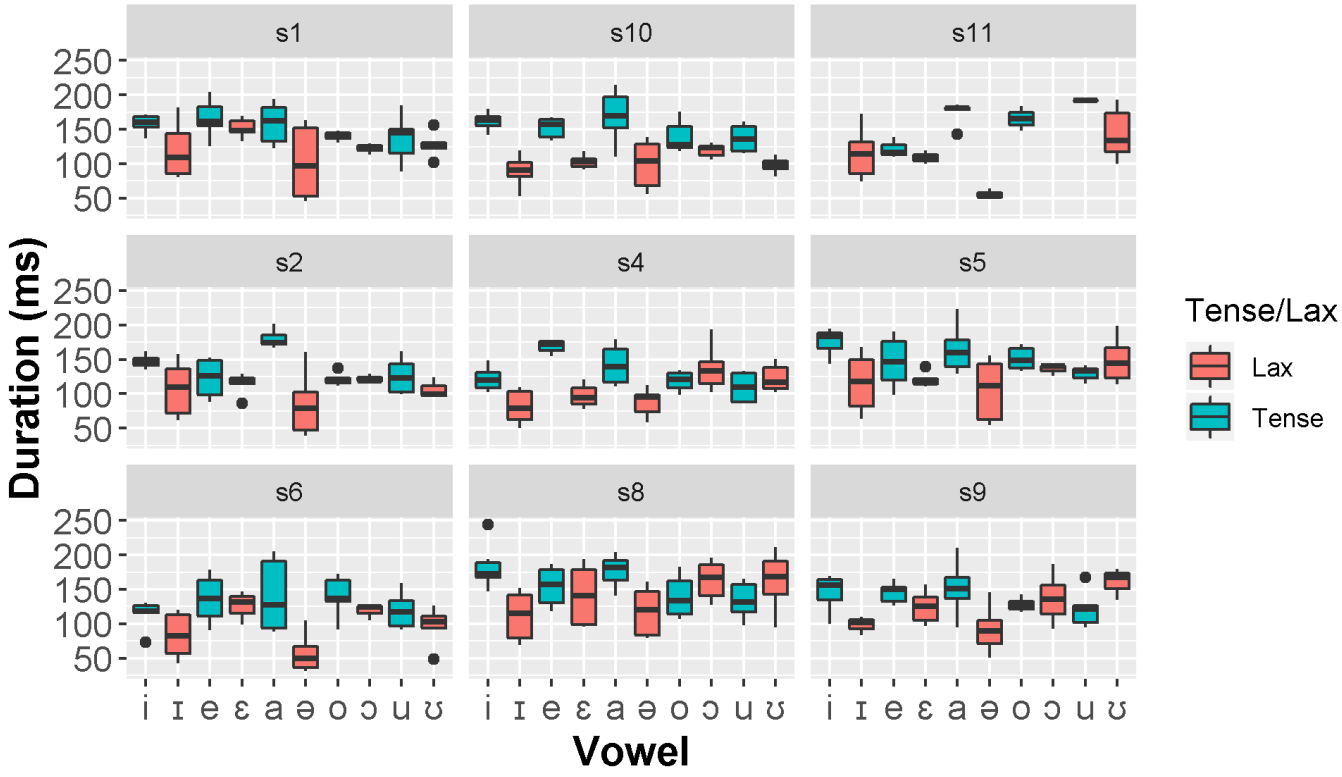

Figure 2.5: Duration results by vowel for each individual participant

In this figure it can be seen that there are some differences between speakers in this data, but the general pattern described above does hold for all or most. All speakers show a difference in the /i ɪ/ and /a ə/ pairs, where the tense vowel is longer. The majority show a difference in the /e ɛ/ pair, where the tense vowel is longer, although the degree varies by participant. There is more variability in the back vowels, where some speakers show no average difference between tense and lax vowels, some have tense vowels that are on average longer, and some have lax vowels that are on average longer.

Finally, durations of each vowel for participant 3 (the participant with much longer vowel durations than all others) are shown in Figure 2.6.

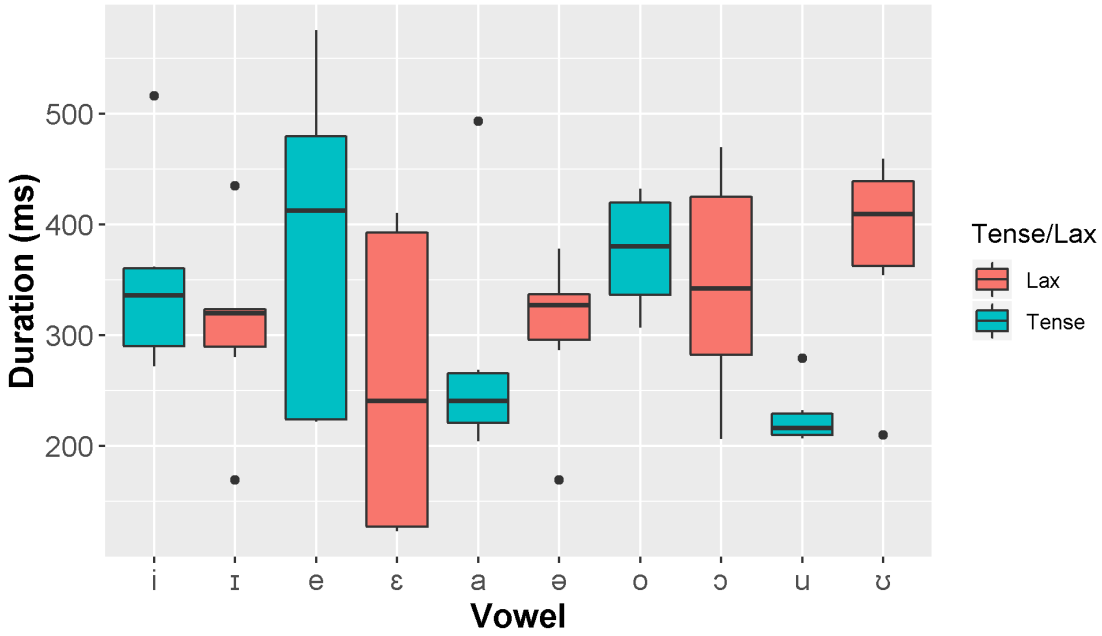

Figure 2.6: Duration results by vowel for subject 3

This speaker does not show the same pattern that is found in other speakers regarding which vowels are longer and which shorter, and for some vowels the range of durations is very large (up to several hundred ms). I hypothesize that any duration differences between vowels in this person's speech may have been lost in this data due to the unnaturally slow pronunciation.

Next, investigating these duration differences by item shows a slightly more complicated picture, where the short average durations of the vowels /ə/ and /ɪ/ are pulled down by only one of the items in each case. Consider Figure 2.7 which shows the durations of tokens of the words with /a/ and /ə/. Here it can be seen that the average duration of tokens of the word k'əʃ (*k'äx*) is not much shorter than the average duration of tokens of words with /a/. However, the average duration of tokens of the word səq (*säq*) is very short, which causes the large average difference between the two sets. These differences can furthermore be seen fairly systematically in each individual speaker, as shown in Figure 2.8.

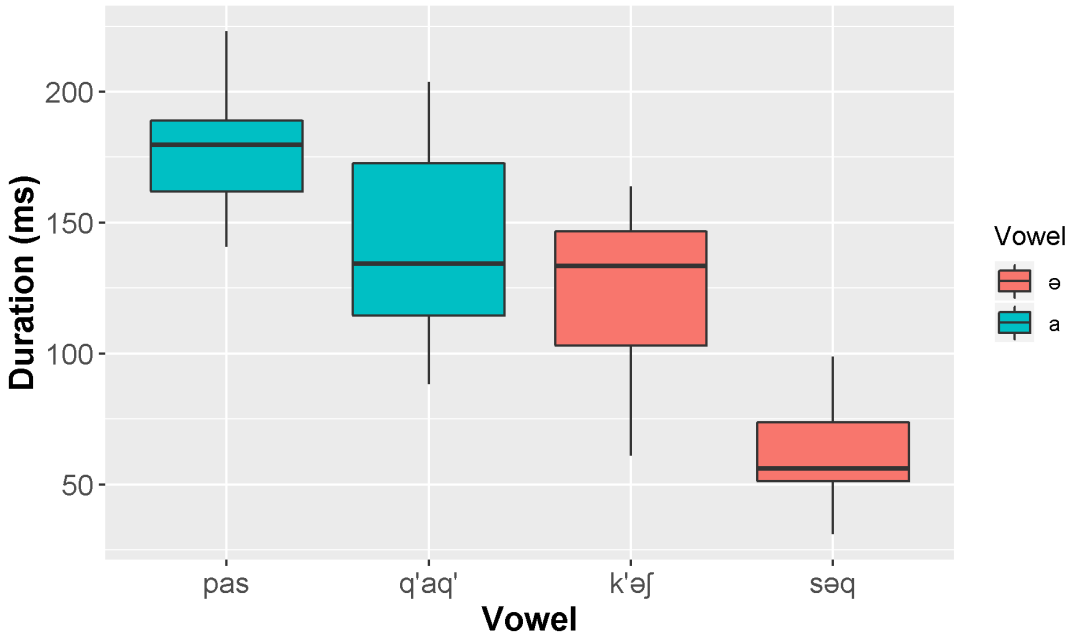

Figure 2.7: Duration by word for vowels /a/ and /ə/

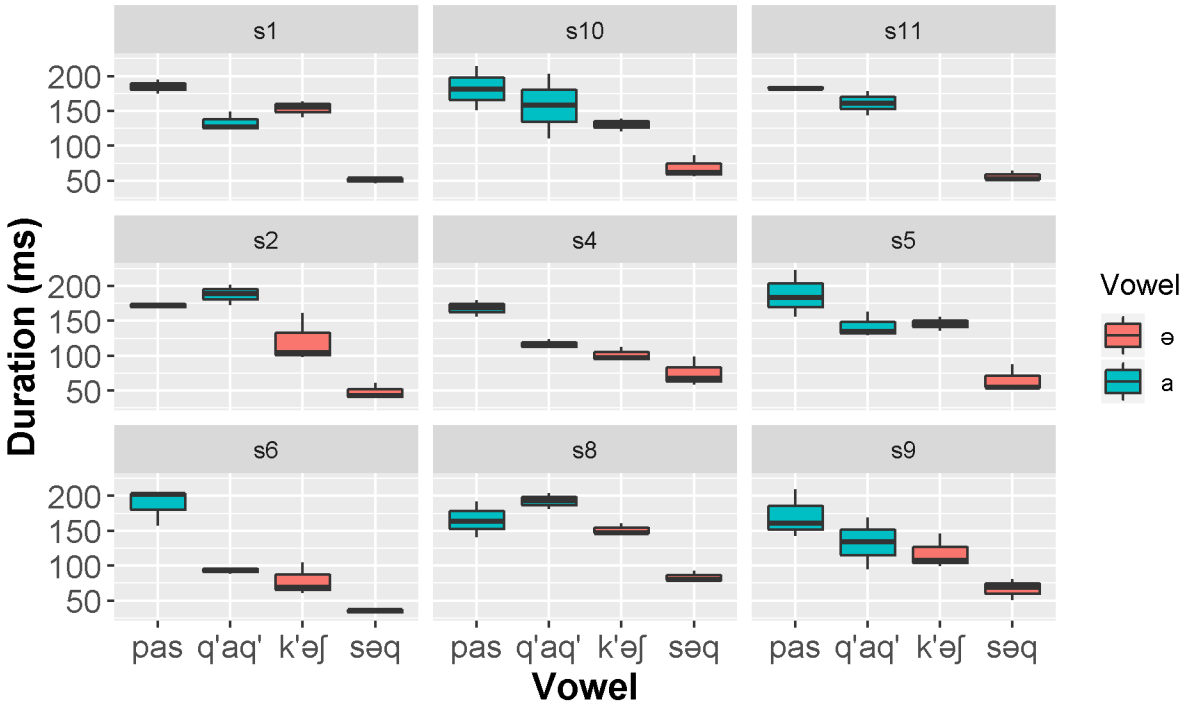

Figure 2.8: Duration by word for vowels /a/ and /ə/ for each individual speaker

The same can be seen of the vowels /i/ and /ɪ/, as shown in Figure 2.9. Here, the average duration of the vowel /ɪ/ is significantly pulled down by the word t͡ʃɪqʼ (*chïqʼ*), while the difference between pɪʃ (*pïx*), kʼiʃ (*kʼix*) and qʼiχ (*qʼij*) is much smaller. This pattern also holds fairly well across speakers, as can be seen in Figure 2.10.

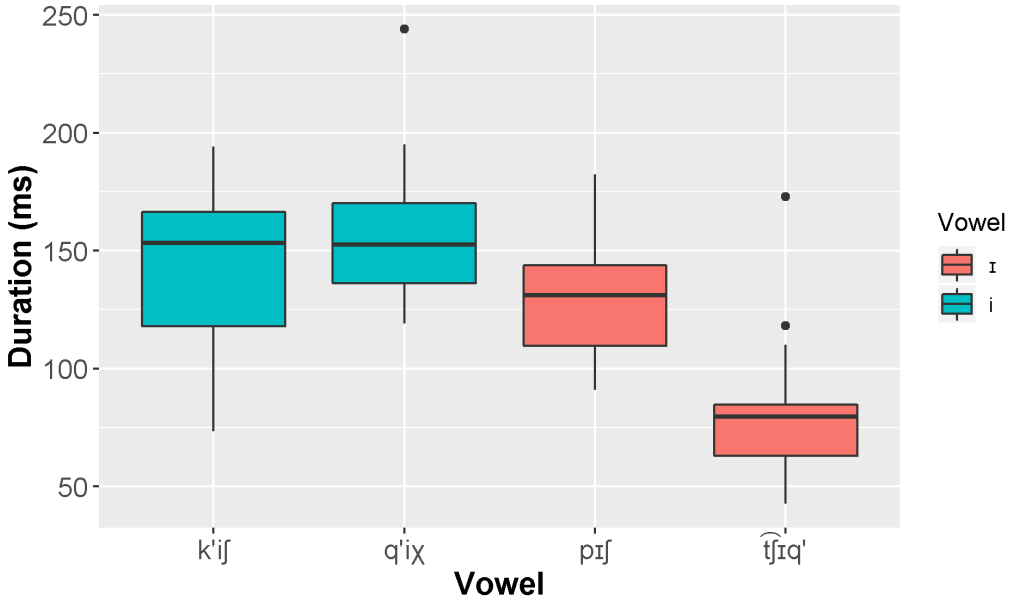

Figure 2.9: Duration by word for vowels /i/ and /ɪ/

It is worth noting that the word t͡ʃɪqʼ (*chïqʼ*) was the only experiment item that was not found in the dictionary of a related dialect. Given that a word spelled *chäq* was later found in another dictionary (Christenson n.d.) with this same meaning, it is likely that the correct phonemic representation of this word would be t͡ʃəqʼ and the vowel was miscategorized in this experiment. If this is the case, then it seems that truly it is tokens of /ə/ that are very short, while other lax vowels are not.

The average duration of each item of the mid front vowels can be seen in Figure 2.11.

This graph shows that, once again, there is a duration difference between items, in this case primarily between the two items of tense /e/. The average duration of ʃeʃ (*xex*) is much longer than that of t͡ʃʼek (*chʼek*), which is very similar to the average durations of the items of lax /ɛ/. This pattern is additionally seen in several individual speakers, although not in all; this can be seen in Figure 2.12.

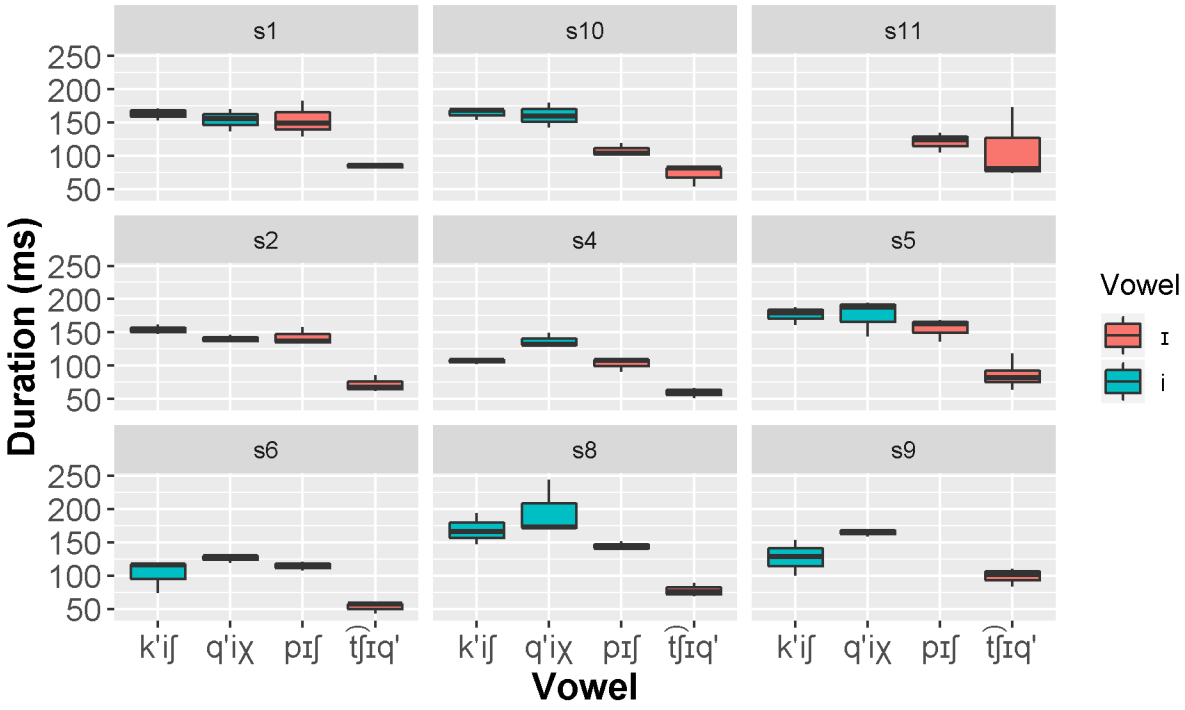

Figure 2.10: Duration by word for vowels /i/ and /ɪ/ for each individual speaker

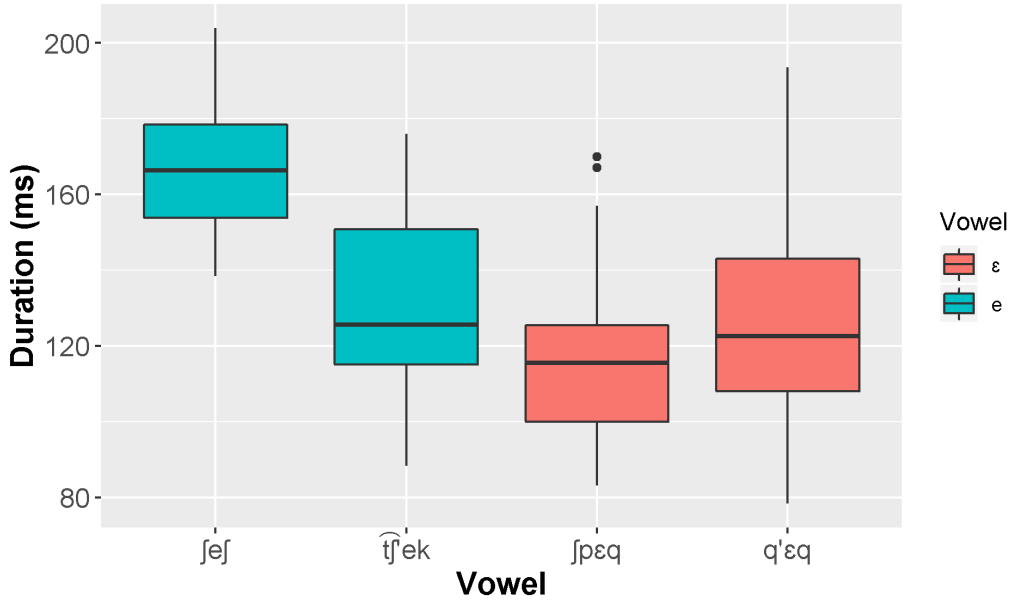

Figure 2.11: Duration by word for vowels /e/ and /ɛ/

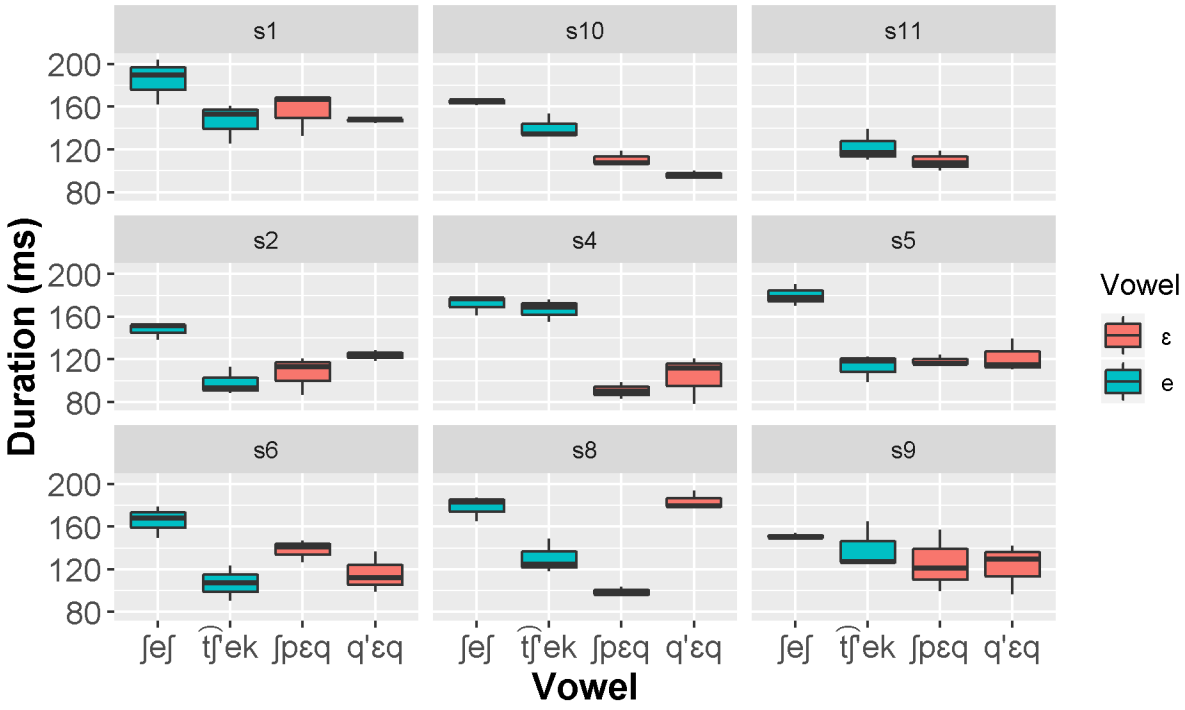

Figure 2.12: Duration by word for vowels /e/ and /ɛ/ for each individual speaker

The average duration of each item of the mid back vowels can be seen in Figure 2.13.

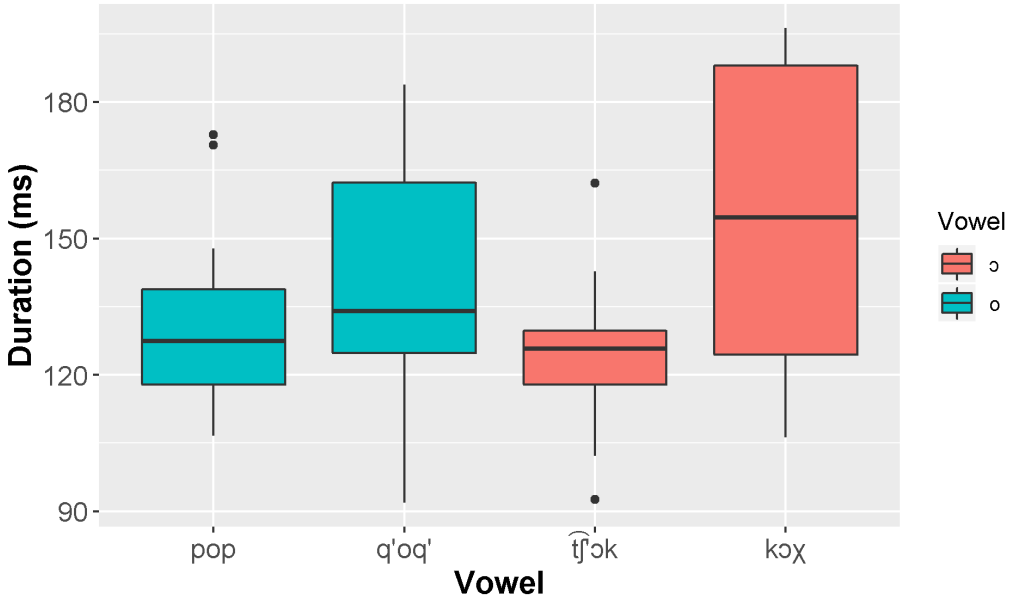

Figure 2.13: Duration by word for vowels /o/ and /ɔ/

This graph shows a different pattern from the previous graphs. Although there is once again some amount of difference by item within each vowel category, the four averages are all relatively close to each other and there is considerable overlap. None of the four items stands out as particularly distinct from the rest. Furthermore, this pattern is not found robustly in individual subjects, as shown in Figure 2.14.

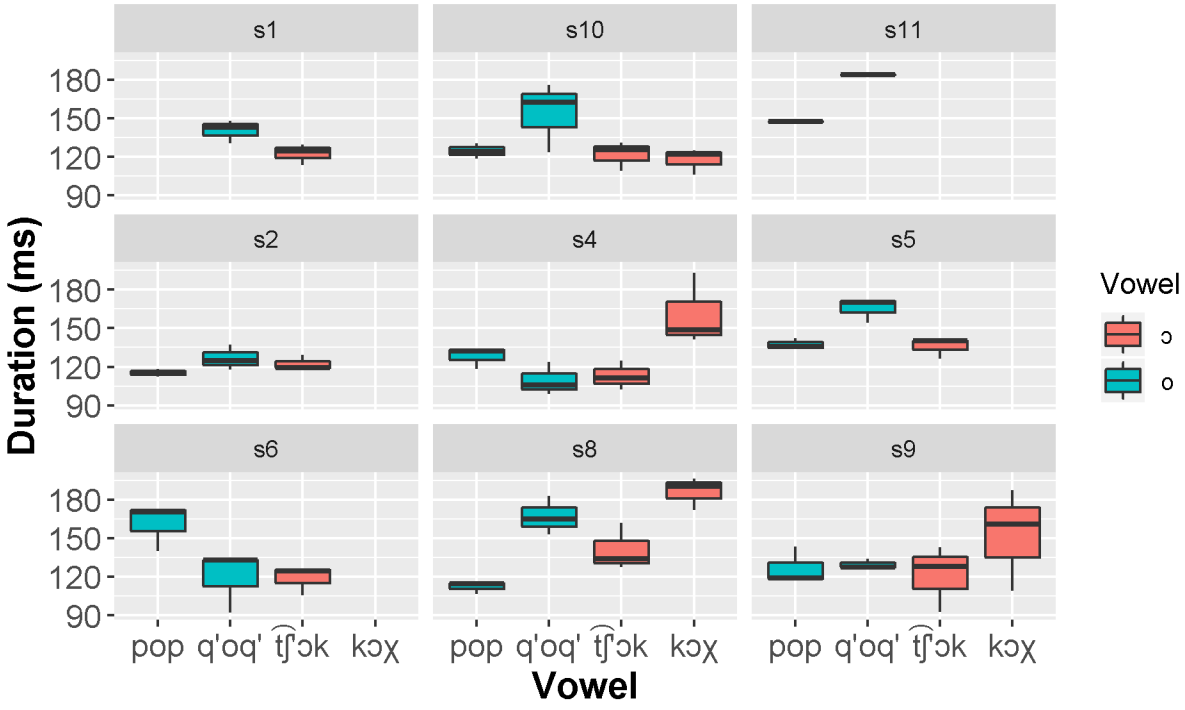

Figure 2.14: Duration by word for vowels /o/ and /ɔ/ for each individual speaker

Patterns in this graph are somewhat obscured by the lack of tokens of the word kɔχ (*köj*) for several speakers. However, it is still apparent that the same pattern is not found for all speakers. While some have very similar average durations for all items, others have large differences in duration between items of the same vowel category, and which items have a longer duration in each case is not the same across speakers.

A similar pattern to this is found for the high back vowels. The average duration of each item of the high back vowels can be seen in Figure 2.15.

Similarly to what was observed for the mid back vowels, here there are some differences between items in each vowel category, but the four words do not fall naturally into

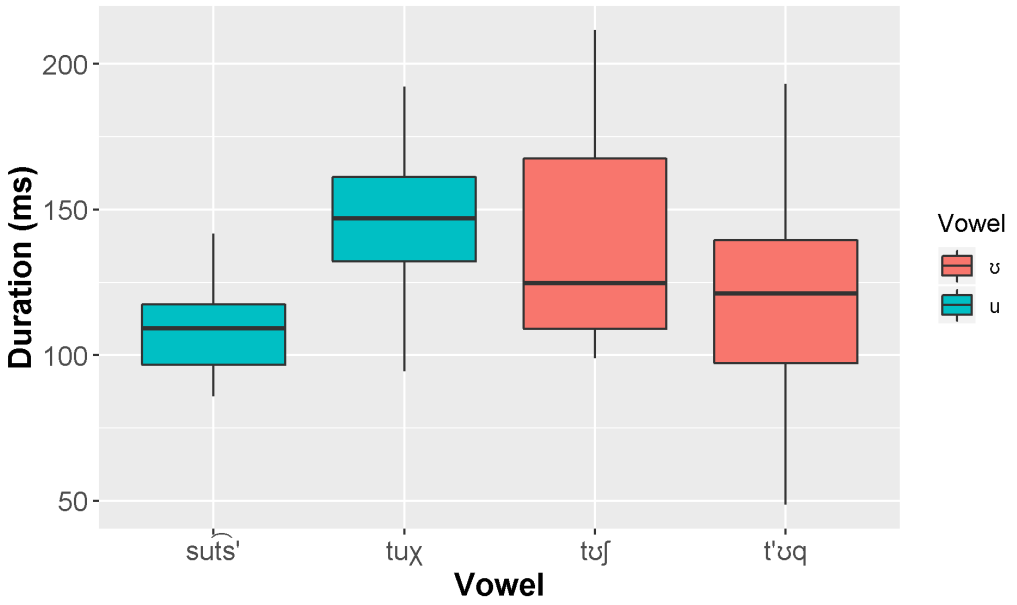

Figure 2.15: Duration by word for vowels /u/ and /ʊ/

subgroupings. Rather, the overall average durations of each item are quite similar, and none of the items stands out as particularly distinct. Again, similarly to the mid back vowels, the same pattern is not robustly observed among individual speakers, as shown in Figure 2.16. While some speakers have similar average durations for all items, others show one item which is notably longer or shorter, but it is different in each case.

2.3.2 Discussion of Duration Results

This preliminary data suggests that the question of whether there is a duration difference between tense and lax vowels in Chichicastenango K'iche' does not have a simple answer. There may be a duration difference between tense and lax vowels at some places of articulation (in particular for front and central vowels), but these are not found at all places of articulation. There is not a clear effect of vowel height on duration as expected from cross-linguistically observed intrinsic duration differences, as higher vowels are not consistently shorter on average than lower vowels when all else is equal (e.g., there is no observed duration difference between high /u/ and mid /o/). Furthermore, there are notable differences by lexical item within the same vowel category. Finally, the pattern observed in back

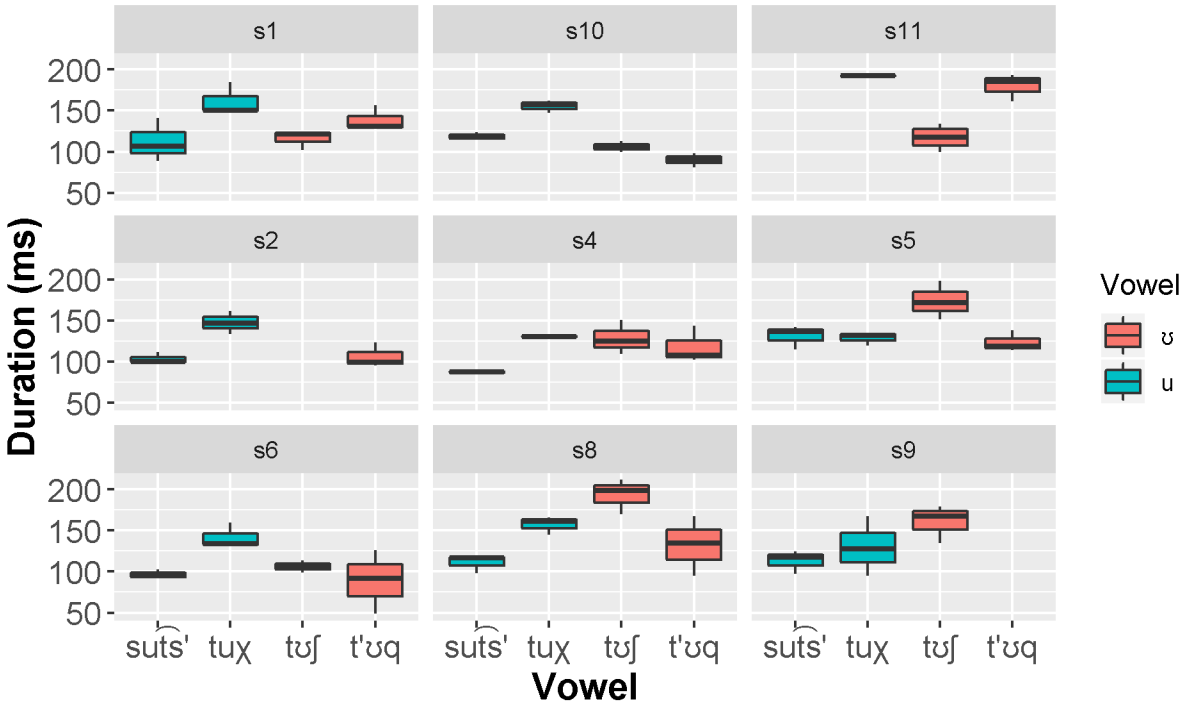

Figure 2.16: Duration by word for vowels /u/ and /ʊ/ for each individual speaker

vowels is different from that of front and central vowels, where there are notable individual differences for back vowels but the same approximate pattern is observed across speakers for front and central vowels.

From this data, it would appear that duration is not an important cue for distinguishing tense and lax back vowels, although it may play a role in distinguishing front and low vowels. However, further research is needed on this question. Due to the variation observed by lexical item and speaker, a follow-up study should include more participants and a larger number of items for each vowel category. This would help tease out the effects of lexical item or consonant context from those related to vowel category.

2.3.3 Formant Results

Figure 2.17 shows the z-scores for the first and second formant values for each speaker in the dataset, with tense vowels in blue and lax vowels in red.

In this graph it can be seen that lax vowels at all places of articulation are on average

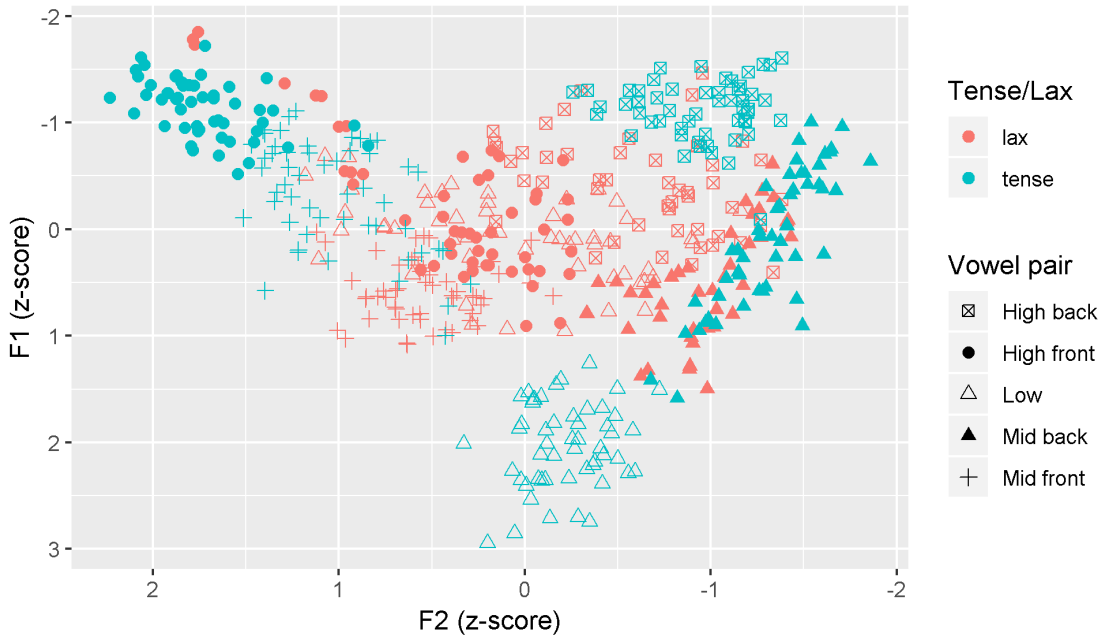

Figure 2.17: Z-scores of tense and lax vowels

closer to the center of the vowel space and tense vowels are more peripheral, with some exceptional tokens (in particular of lax high front vowel /ɪ/). Furthermore, this general pattern holds for each individual speaker, as shown in Figure 2.18.

Figure 2.19 shows the vowels classified by height.

Tense high vowels /i u/ have lower F1 values than tense mid vowels /e o/, which have lower F1 values than tense low vowel /a/. High lax vowel /ʊ/ for the most part has lower F1 values than those of mid /ɛ ɔ/, but this is not true of high /ɪ/, whose F1 values range from as low as those of tense high vowel /i/ to as high as those of mid /ɛ ɔ/. Tokens of lax low vowel /ə/ are mostly found to have F1 values between those of high /ʊ/ and mid /ɛ ɔ/. Across the dataset, most tense high vowels have lower F1 values than all mid vowels. However, many lax high vowels have higher F1 values than some mid vowels. Similarly, most tense mid vowels have lower F1 values than all low vowels, but many lax mid vowels have higher F1 values than lax low vowels (although they have lower F1 values than all tense low vowels).

Figure 2.20 shows the vowels classified by frontness.

Figure 2.18: Tense and lax vowels for each individual speaker, showing centralization of lax vowels

Figure 2.19: Z-scores of vowels by height

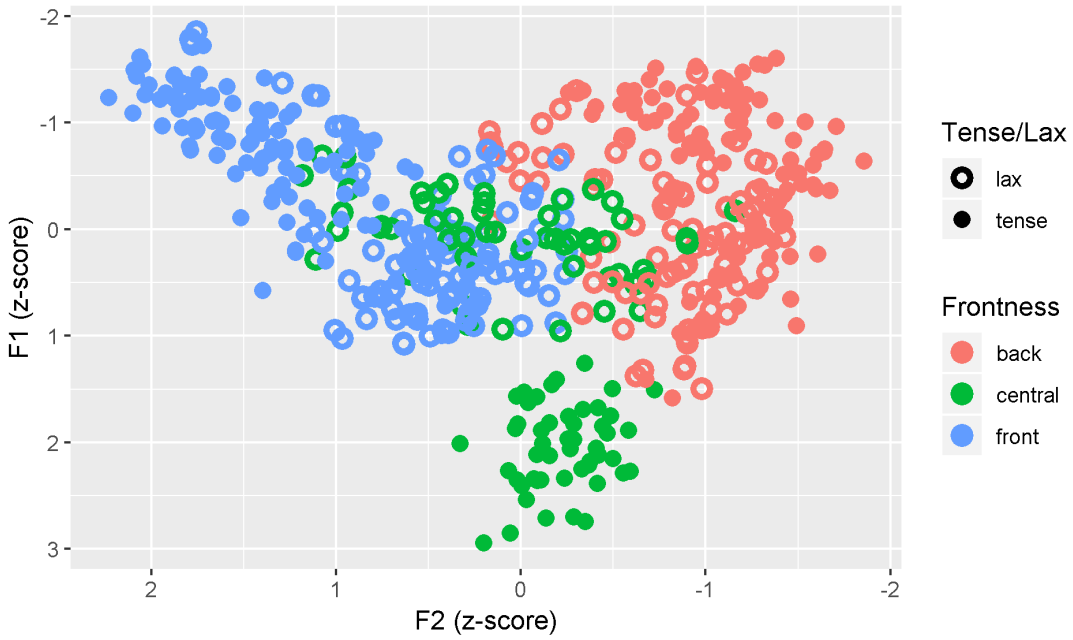

Figure 2.20: Z-scores of vowels by frontness

Tense back vowels /u o/ have lower F2 values than tense central vowel /a/, which has lower F2 values than tense front vowels /i e/. For the most part, back lax vowels /ʊ ɔ/ have lower F2 values than those of front lax vowels /ɪ ɛ/, but tokens of central /ə/ range in F2 from as low as that of most lax back vowels to as high as that of most lax front vowels. Across the dataset, most front vowels, both tense and lax, have higher F2 values than central vowels, which have higher F2 values than both tense and lax back vowels; exceptions are mostly due to the large F2 range of lax central vowel /ə/.

Next, I will show results for the vowel pair at each place of articulation individually.

2.3.3.1 High Back Vowels

The high back vowels have been previously labeled as tense /u/ and lax /ʊ/ (written in the orthography as <u> and <ü>, respectively). The average F1 and F2 values of these vowels for male and female speakers are shown in Table 2.6.

Tokens of these vowels from the experiment data are shown in Figure 2.21, separated by male and female speakers.

Vowel	Item	F1 (Hz) Female	F1 (Hz) Male	F2 (Hz) Female	F2 (Hz) Male
	sut͡s’	469	377	1340	1254
u	tuχ	457	392	1282	1114
	total	463	384	1310	1184
	tʊʃ	542	460	1607	1464
ʊ	t’ʊq	575	492	1348	1118
	total	557	478	1493	1267

Table 2.6: Average F1 and F2 values of high back vowels by item and gender

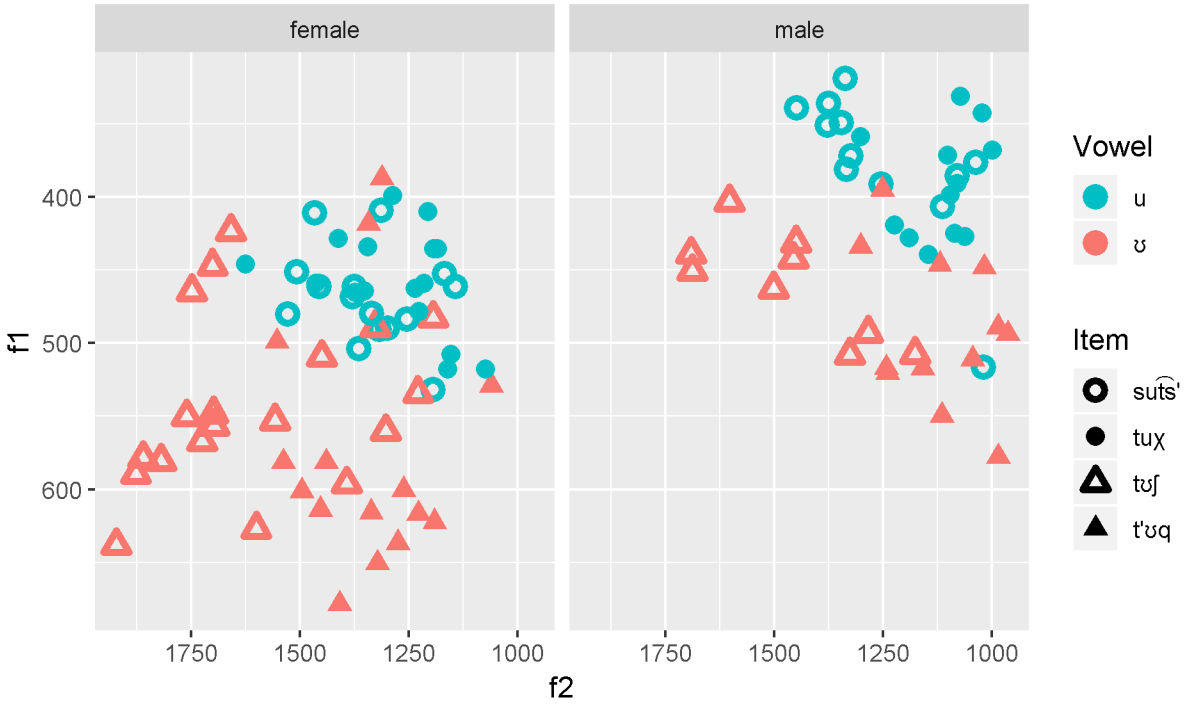

Figure 2.21: High back vowels by gender

From the table and these graphs, a difference can be seen between tense /u/ and lax /ʊ/ in both the first and second formants. Tokens of /u/ have on average lower F1 and F2 values than tokens of /ʊ/ (they are higher and backer). That is, the lax vowels have formant values closer to the center of the vowel space in both dimensions. Among the female speakers, some tokens of /ʊ/ have low F1 and F2 values that are not higher than those of

/u/, but the pattern does hold for the majority of the data.

Additionally, a difference can be observed between the two items for each vowel. Most tokens of tʊʃ (*tüx*) have higher F2 and slightly lower F1 values than tokens of t'ʊq (*t'üq*), and most tokens of sut͡s' (*sutz'*) have higher F2 values than tokens of tuχ (*tuj*). The vowels which precede alveolar and postalveolar consonants in this case have higher F2 values (they are fronter) than the vowels preceding uvular consonants. Between the items of /ʊ/, the vowels preceding the postalveolar also have lower F1 values (they are higher) than the vowels preceding the uvular, but this F1 difference is not apparent for the two items with /u/.

The pattern seen when averaging over all speakers is found individually in several speakers, although not all. This is shown in Figure 2.22.

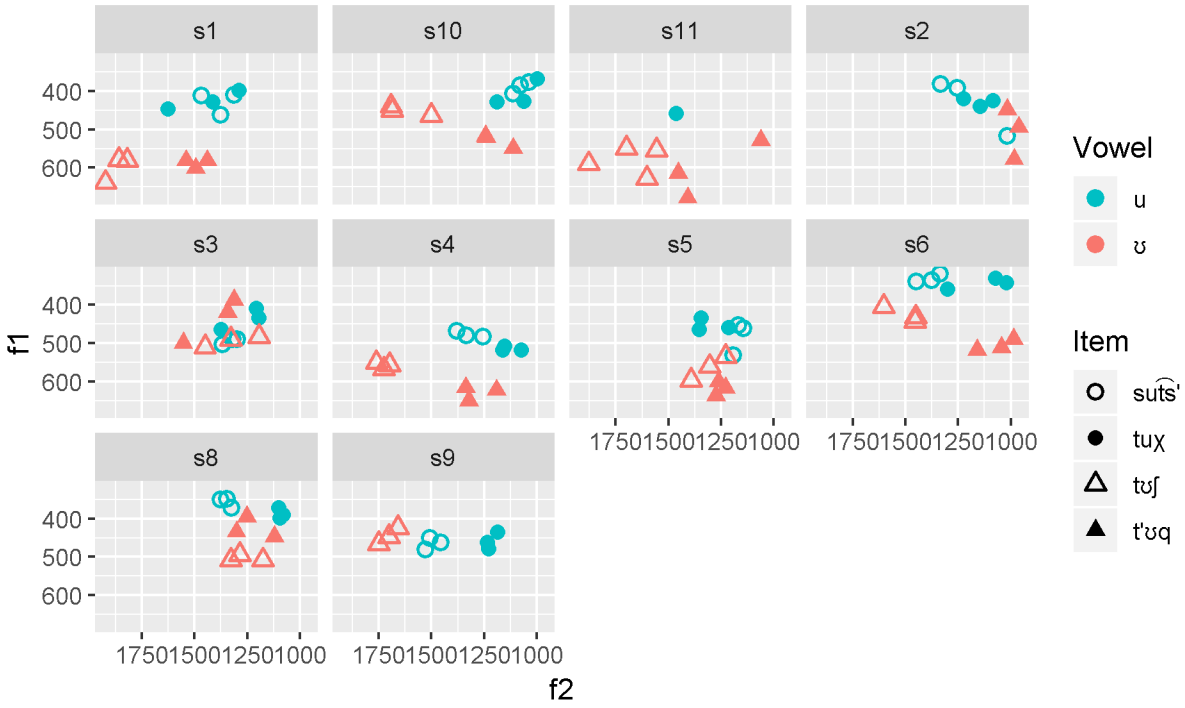

Figure 2.22: High back vowels by individual speaker

As shown in this figure, about half of the speakers individually show a pattern where tense /u/ is higher and backer than lax /ʊ/ (e.g. speakers 1, 10 and 4). However, others have different patterns. For speakers 2, 5 and 8, the only difference between the tense and

lax vowels seems to be height, while for speaker 9 it is only frontness. Finally, for speaker 3 no real difference is noticeable between the two vowels. With regards to lexical differences within each vowel, there are also differences. While differences between the formant values of the the two items of the vowel /ʊ/ are found in all speakers for which there is data, the same is not true of the two items of the vowel /u/, for which differences are visible in some speakers (e.g. 4, 8 and 9) but not others (e.g. 1, 2 and 10).

2.3.3.2 Mid Back Vowels

The mid back vowels have been previously labeled as tense /o/ and lax /ɔ/ (written in the orthography as <o> and <ö>, respectively). The average F1 and F2 values of these vowels for male and female speakers are shown in Table 2.7.

		F1 (Hz)		F2 (Hz)	
Vowel	Item	Female	Male	Female	Male
o	pop	537	487	990	919
	q'oq'	733	530	1274	998
	total	645	509	1147	958
ɔ	t͡ʃ'ɔk	715	614	1439	1156
	kɔχ	645	523	1215	994
	total	689	577	1355	1102

Table 2.7: Average F1 and F2 values of mid back vowels by item and gender

Tokens of these vowels from the experiment data are shown in Figure 2.23, separated by male and female speakers.

In these graphs, tense /o/ and lax /ɔ/ differ in both F1 and F2. Tokens of /o/ have on average lower F1 and F2 values than do tokens of /ɔ/ (they are backer and higher). That is, the lax vowels have formant values closer to the center of the vowel space in both dimensions.

Additionally, a difference can be observed between the two items for each vowel. Tokens of pop (*pop*) have on average higher F1 and F2 values (they are higher and backer) than tokens of q'oq' (*q'oq'*). Tokens of t͡ʃ'ɔk (*ch'ök*) have on average lower F1 and F2 values than tokens of kɔχ (*köj*). In this case, vowels preceding uvulars are higher and backer for

Figure 2.23: Mid back vowels by gender

tokens of lax vowel /ɔ/, but lower and fronter for tokens of tense vowel /o/, as compared to the other item in the data.

The pattern seen when averaging over all speakers is found individually in several speakers, although not all. This is shown in Figure 2.24.

As this graph shows, the difference between tense /o/ and lax /ɔ/ is one of both height and frontness for most speakers (e.g. speakers 2, 6 and 9). However, for a few speakers it more closely resembles a plain frontness difference (e.g. speakers 1 and 4). Lexical differences between items with the vowel /o/ are found for almost all speakers; those with less obvious patterns or not at all may be speakers 10, 6 and 8. Differences between the two items with the vowel /ɔ/ are less easily seen due to the frequent absence of tokens of kɔχ 'lion', a rather infrequent word which many speakers did not know. However, among those speakers with tokens of this word, differences between the two lexical items are present for some (e.g. 4, 8 and 9) but not others (e.g. 3).

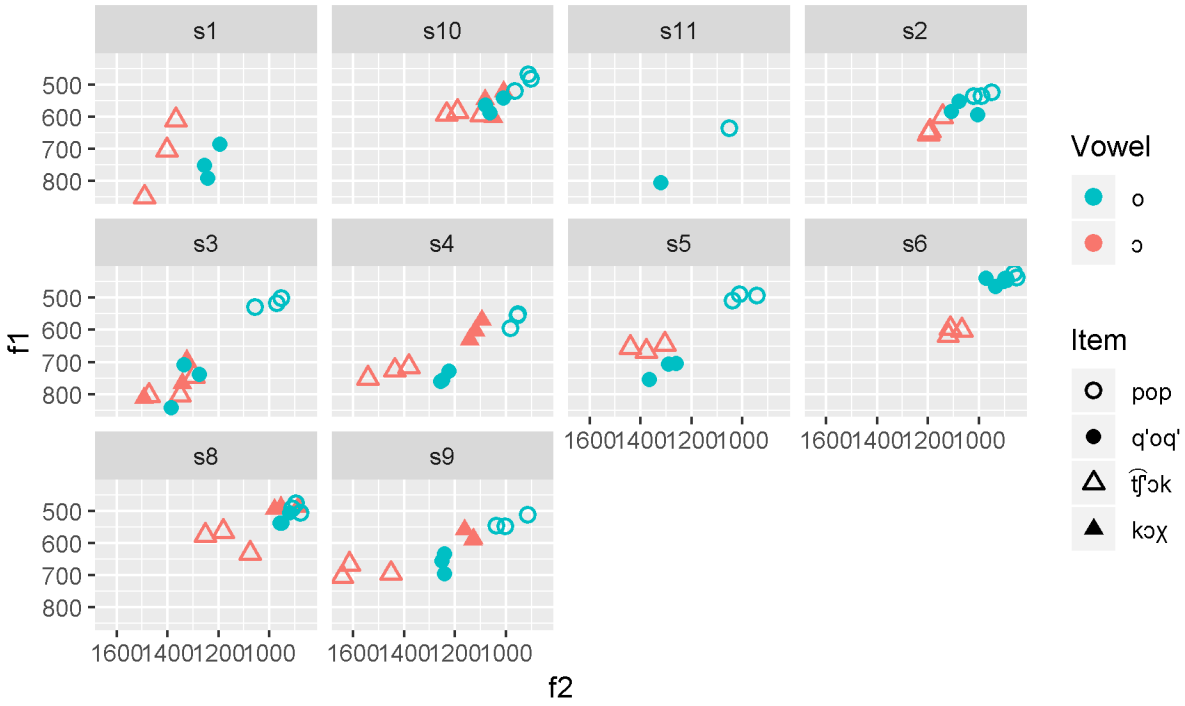

Figure 2.24: Mid back vowels by individual speaker

2.3.3.3 Low Vowels

The low vowels have been previously labeled as tense /a/ and lax /ə/ (written in the orthography as <a> and <ä>, respectively). The average F1 and F2 values of these vowels for male and female speakers are shown in Table 2.8.

		F1 (Hz)		F2 (Hz)	
Vowel	Item	Female	Male	Female	Male
	pas	864	723	1676	1434
a	q'aq'	914	803	1680	1495
	total	888	761	1678	1463
	k'əʃ	606	515	1983	1928
ə	səq	635	551	1511	1405
	total	622	533	1726	1666

Table 2.8: Average F1 and F2 values of low vowels by item and gender

Tokens of these vowels from the experiment data are shown in Figure 2.25, separated

by male and female speakers.

Figure 2.25: Low vowels by gender

As these graphs show, there are clear height differences between the two vowels. Tokens of tense /a/ have much higher F1 values than those of lax /ə/. Among male speakers, there is also a large difference in average F2 values for the two vowels, driven by an item split. Tokens of səq (*säq*) have F2 values extremely close to those of both items of the vowel /a/, but tokens of k'əʃ (*k'äx*) have much higher F2 values. This pattern is not found among female speakers, where tokens of both items of the vowel /a/ have F2 values that are on average intermediate between those of the two items with /ə/. In both male and female speakers there is a large difference in F2 between the two items of the vowel /ə/: tokens of the word səq, where the vowel precedes a uvular consonant, have much lower F2 values than those of the word k'əʃ, where the vowel precedes a postalveolar consonant. There is also a small difference in the average F1 values of these two items, where tokens of səq tend to have slightly higher F1 values than those of k'əʃ. Between the two items of the vowel /a/ there is no clear difference over all, although among the male speakers many of the tokens

of pas (*pas*), where the vowel precedes an alveolar consonant, can be seen to have slightly lower F1 and F2 values (they are higher and backer) than tokens of q'aq' (*q'aq'*), where the vowel precedes a uvular.

The pattern seen when averaging over all speakers is found very robustly in individual speakers. This is shown in Figure 2.26.

Figure 2.26: Low vowels by individual speaker

Every speaker has a clear height difference between tense /a/ and lax /ə/. Furthermore, differences between the two lexical items are found for all speakers for the /ə/ category, with k'əʃ fronter than səq. For the /a/ category, there are rarely differences between the two items, although a small difference may be found for a few speakers (e.g. 6 and 9).

2.3.3.4 Mid Front Vowels

The mid front vowels have been previously labeled as tense /e/ and lax /ɛ/ (written in the orthography as <e> and <ë>, respectively). The average F1 and F2 values of these vowels for male and female speakers are shown in Table 2.9.

Vowel	Item	F1 (Hz) Female	F1 (Hz) Male	F2 (Hz) Female	F2 (Hz) Male
e	ʃeʃ	527	429	2351	2063
	t͡ʃ'ek	662	500	2120	2054
	total	598	464	2228	2058
ɛ	ʃpɛq	700	568	1976	1739
	q'ɛq	705	596	2109	1900
	total	702	582	2036	1820

Table 2.9: Average F1 and F2 values of mid front vowels by item and gender

Tokens of these vowels from the experiment data are shown in Figure 2.27, separated by male and female speakers.

Figure 2.27: Mid front vowels by gender

In these graphs the two vowels differ in both F1 and F2 for most speakers. On average, tokens of tense /e/ have lower F1 and higher F2 values (they are higher and fronter) than those of lax /ɛ/.

Additionally, some differences may be observed between items for each vowel. Tokens of item ʃeʃ (*xex*), where the vowel precedes a postalveolar, have notably lower F1 values on average than those of t͡ʃ'ek (*ch'ek*), where the vowel precedes a velar. They also have higher F2 values among female speakers, although the same is not apparent among male speakers. Tokens of item ʃpɛq (*xpëq*) additionally have on average slightly lower F2 values than those of q'ɛq (*q'ëq*); in both words, the vowel precedes a uvular stop.

It is also worth noting that for four of the six female speakers, and one of the four male speakers, the pronunciation of the vowel in the word t͡ʃ'ek is located within the cloud of tokens of the vowel /ɛ/. It is possible that this word has two pronunciation variants, one with the vowel /e/ and one with the vowel /ɛ/, both of which are accepted.

The pattern seen when averaging over all speakers is found individually in nearly all speakers. This is shown in Figure 2.28.

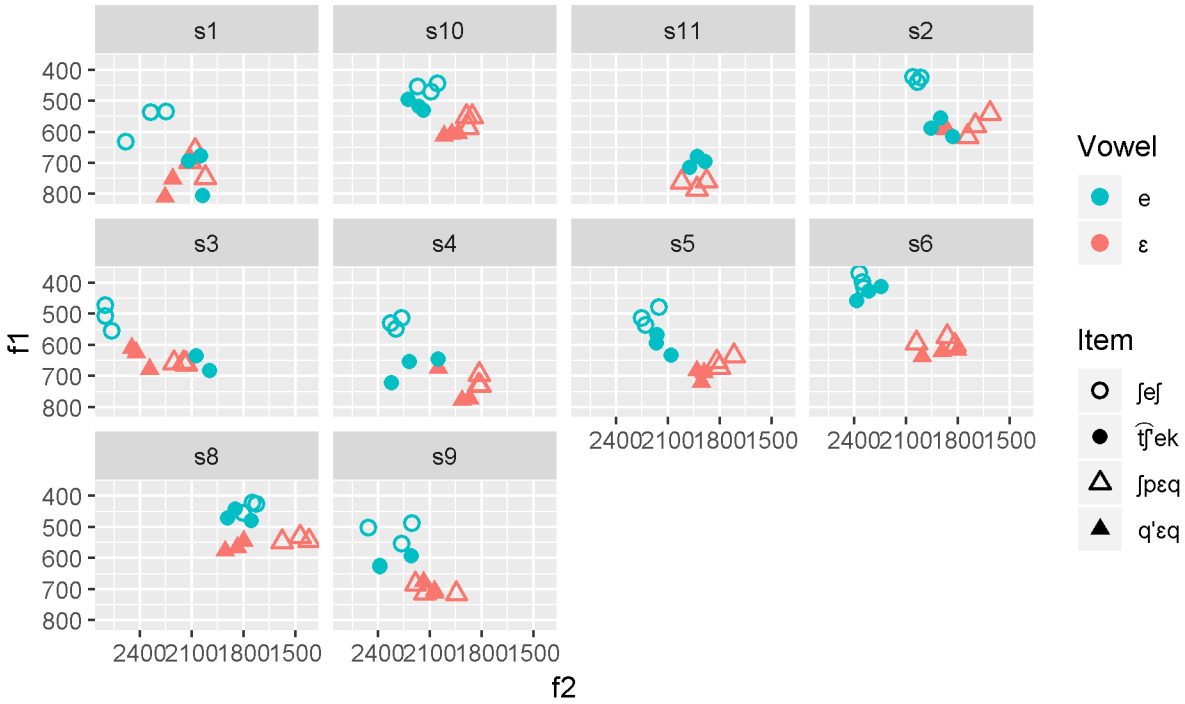

Figure 2.28: Mid front vowels by individual speaker

As this figure shows, both F1 and F2 differences can be seen between tense /e/ and lax /ɛ/ for nearly all speakers, with the possible exception of speaker 11 for whom there is

very little data. Differences between the two lexical items for vowel /e/ are found for several speakers, where ʃeʃ is has lower F1 and higher F2 values than t͡ʃ'ek (e.g. speakers 1, 2, 3 and 9). However, others have very little difference (e.g. speakers 6 and 8). For the vowel /ɛ/ lexical differences are smaller, but visible in some speakers (e.g. 4, 5 and 8) where ʃpɛq has lower F2 values than q'ɛq.

2.3.3.5 High Front Vowels

The high front vowels have been previously labeled as tense /i/ and lax /ɪ/ (written in the orthography as <i> and <ï>, respectively). The average F1 and F2 values of these vowels for male and female speakers are shown in Table 2.10.

		F1 (Hz)		F2 (Hz)	
Vowel	Item	Female	Male	Female	Male
i	k'iʃ	446	381	2609	2425
	q'iχ	470	388	2577	2332
	total	458	384	2593	2379
ɪ	pɪʃ	597	386	1834	2163
	t͡ʃɪq'	639	532	1848	1633
	total	620	459	1842	1898

Table 2.10: Average F1 and F2 values of high front vowels by item and gender

Tokens of these vowels from the experiment data are shown in Figure 2.29, separated by male and female speakers.

As these graphs show, tense /i/ and lax /ɪ/ differ in the data in both height and frontness. Tokens of the vowel /i/ have on average lower F1 and higher F2 values than those of /ɪ/, that is, they are higher and fronter. Additionally, differences may be seen by item in each vowel category. Tokens of k'iʃ (*k'ix*), where the vowel precedes a postalveolar consonant, may be seen to have slightly lower F1 and higher F2 values on average than tokens of q'iχ (*q'ij*), where the vowel precedes a uuvlar. Finally, for the two items of the vowel /ɪ/, a different pattern is observed for male and female speakers. For female speakers, no difference between the average formant values of pɪʃ (*pïx*) and t͡ʃɪq' (*chïq'*) is readily

Figure 2.29: High front vowels by gender

apparent. However, for male speakers, there is a large difference between the two, where tokens of pɪʃ can be seen to have much lower F1 and higher F2 values than those of t͡ʃɪq'. In fact, tokens of pɪʃ are very close in formant values to those of the vowel /i/.

It is also worth noting that the formant values found for tokens of the vowel /ɪ/ in this data, with the exception of those of pɪʃ for male speakers, place this vowel at a much more central position than would be expected. Auditorily, these vowels are difficult to distinguish from /ə/, while male speakers' vowels in pɪʃ are indistinguishable from /i/.

To further investigate this, it is helpful to visualize the three vowels /i ɪ ə/ together. Consider Figure 2.30, which shows the vowels of all tokens of these three vowels in the experiment data, divided by gender.

These graphs show that for female speakers, the formant values of /ɪ/ in the experiment data are very close to those of /ə/, while for male speakers there is a split by item where tokens of pɪʃ have formant values very close to those of /i/ but tokens of t͡ʃɪq' have formant values very close to those of /ə/.

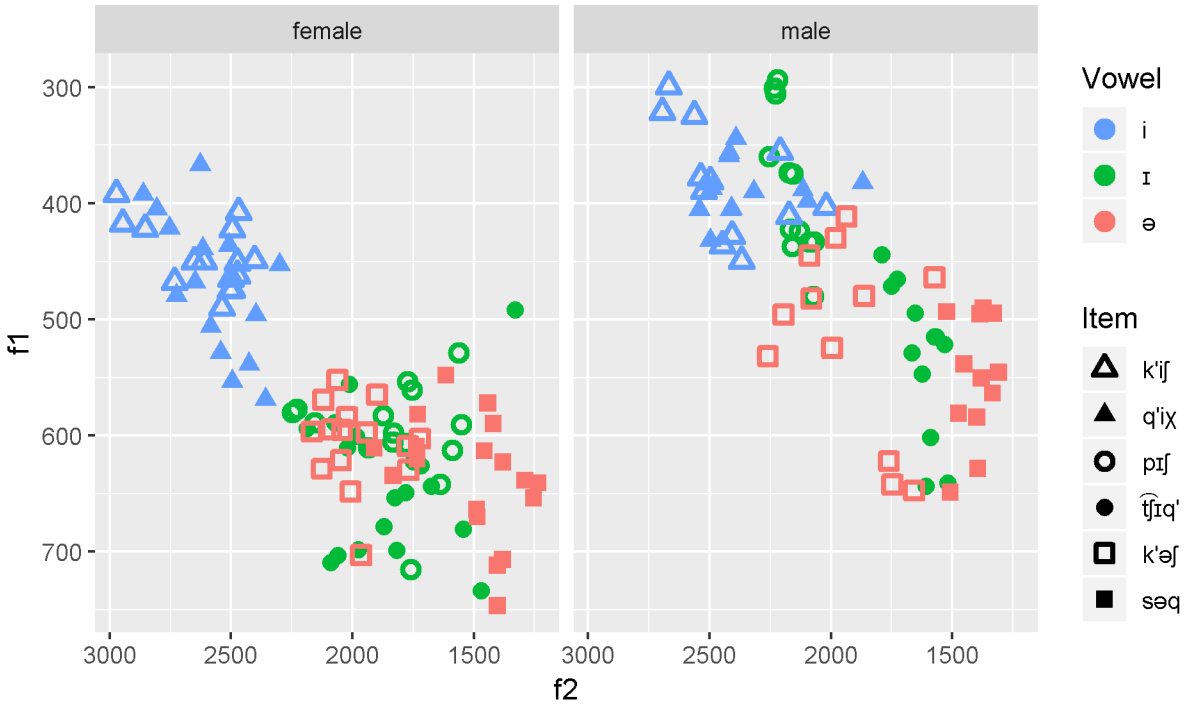

Figure 2.30: Overlap of phonemes /ɪ ə i/ in experiment data, by gender

The pattern seen when averaging over all speakers is found very robustly in individual speakers. This is shown in Figure 2.31.

This figure shows two distinct patterns: that of speakers 1, 3, 4, 5 and 9, where tense /i/ and lax /ɪ/ are distinct and differ in both F1 and F2, and that of speakers 2, 6, 8 and 10, where there is overlap or extreme similarity between the formant values of some tokens of the two vowels (specifically, tokens of pɪʃ overlap with those of the two /i/ words while those of t͡ʃɪq' are lower and backer). These groups line up exactly with the split between male and female speakers. The slight lexical differences noted above between items of the vowel /i/ are found individually for some speakers (e.g. 4 and 6) but not visibly for others (e.g. 1, 3, 8 and 10).

2.3.4 Discussion of Formant Results

It was expected that the experiment results would show ten contrastive vowel qualities, with hypotheses addressing vowel height, backness and tenseness. Differences were not

Figure 2.31: High front vowels by individual speaker

expected by lexical item, speaker or gender nor between the type of contrast between tense and lax for each vowel pair. Although some of these hypotheses were consistent with the results, in many ways the results are different from what was expected. Various aspects of the results are discussed as follows.

2.3.4.1 Tense-Lax Contrast as Centralization

The formant results show lax vowels which have F1 and F2 values closer to the center of the vowel space than those of their tense counterparts at all places of articulation, with the possible exception of the high front vowels where a merger may be occurring. Therefore, the difference between the two sets of vowels referred to as "tense" and "lax" appears to be one of centralization. This is what was expected from previous impressionistic descriptions of this vowel inventory and the IPA labels chosen for each sound (e.g. López Ixcoy 1994; Par Sapón and Can Pixabaj 2000). This is also consistent with the historical origin of these vowels, in coming from an inventory of long and short vowels at the peripheral places of articulation

(Campbell 1977). Vowel reduction of the shorter vowel phonemes in each pair due to their shorter duration could easily lead to a quality difference describable as centralization.

The terms "tense" and "lax" are cover terms which refer to a wide range of phonetic contrasts in different languages, including tongue root position (e.g. Fulop et al. 1998), vowel height (e.g. Dalton 2011), duration (e.g. D. C. Bennett 1968) and voice quality (e.g. Di Paolo and Faber 1990). In Kaqchikel, a language closely related to K'iche', these terms refer to a centralization contrast very similar to the one found in Chichicastenango K'iche' (R. Bennett 2016b). However, these terms are perhaps most often used to signify a relationship where the lax vowel is lower than its tense counterpart (has higher F1 value). Because of this, they may be misleading if applied to Chichicastenango vowels. Although the lax variant is lower than the tense variant for the high and mid vowels in this data (in addition to backness differences in each case), this is not true for the low vowels. The "lax" equivalent of the vowel /a/ is not lower than it, but higher (the vowel /ə/). Therefore, more precise terms to distinguish these two sets of vowels would be "peripheral" (for tense) and "centralized" (for lax). In the following chapters of this work, these terms will be used in place of tense and lax, with the exception of cases when explicit reference is made to previous works which use the terms tense and lax.

2.3.4.2 Height and Backness

With regard to vowel height, it was hypothesized that high vowels would have lower F1 values than mid vowels, which would have lower F1 values than the tense low vowel /a/. F1 values of the lax low vowel /ə/ were expected to be close to those of the mid vowels. Results show that within the tense and lax groups, this is mostly accurate, with the exception of the behavior of lax high vowel /ɪ/, many tokens of which have quite high F1 values similar to those of the mid vowels and /ə/. However, comparing height for tense and lax vowels together, it can be seen that F1 values of the lax high vowel /ʊ/ are not always lower than those of the tense mid vowels. Therefore, tenseness has a significant effect on height contrasts, and it is not the case that all high vowels have lower F1 values than all mid vowels or that all mid vowels have lower F1 values than all low vowels. Classification of lax vowels as high, mid or low is to be interpreted with respect to other lax vowels and

by association with the height of the tense vowel that they alternate with; it is in this sense that the vowel /ə/ is classified as low, not because of a phonetic realization with high F1 values but because it behaves as the lax variant of the low vowel /a/.

With regard to vowel backness, it was hypothesized that front vowels would have higher F2 values than central vowels, which would have higher F2 values than back vowels. Results show that this is for the most part very accurate. F2 values of most front vowels, both tense and lax, are higher than those of most tense central vowels, which are higher than those of most back vowels, both tense and lax. The exception is the lax central vowel /ə/, tokens of which surface with a wide range of F2 values, from as high as many front vowels to as low as some back vowels. Similarly to vowel height, tenseness has an effect on backness, in that lax front vowels have lower F2 values than their tense counterparts while lax back vowels have higher F2 values than their tense counterparts.

In summary, although across the board height and backness results were roughly as expected, some unexpected results were found (in particular for the lax high front vowel /ɪ/). The effect of tenseness on both height and backness was also found to be strong.

2.3.4.3 Effects of Consonant Context or Lexical Item

For nearly all vowel categories in the experiment, differences in formant values were found between the two items belonging to that category. These were often fairly consistent across speakers, although not in all cases.

One notable pattern observed across the data is that frequently vowels of the same category preceding the uvular consonants /q q' χ/ surface with higher F1 and lower F2 values than vowels preceding other consonants. For example, tokens of tuχ have on average higher F1 and lower F2 values than those of su$\widehat{\text{ts}}$', and the same is true of t'ʊq as compared to tʊʃ, səq as compared to k'əʃ and q'iχ as compared to k'iʃ. However, this pattern does not hold for all sets of items for the same vowel category. The vowels in kɔχ, preceding uvular /χ/, have on average lower F1 values than those in t$\widehat{\text{ʃ}}$'ɔk. Similarly, the vowels in q'oq' and q'aq', each preceding a uvular ejective, have on average higher F2 values than the other item in their category, pop and pas, respectively. This is shown in the following two tables, where

uvular consonants are in bold. Table 2.11 shows relative F1 differences by item for each vowel phoneme. The first column shows the item for each vowel category that was usually realized with relatively lower F1 values across speakers as compared to the second column which shows the item that was realized with relatively higher F1 values.

Vowel	Item with relatively lower F1 (vowel realized higher)	Item with relatively higher F1 (vowel realized lower)
/u/	su$\widehat{\text{ts}}$’	tu**χ**
/ʊ/	tʊʃ	t’ʊ**q’**
/o/	pop	**q’**o**q’**
/ɔ/	kɔ**χ**	$\widehat{\text{tʃ}}$’ɔk
/a/	pas	**q’**a**q’**
/ə/	k’əʃ	sə**q**
/e/	ʃeʃ	$\widehat{\text{tʃ}}$’ek
/i/	k’iʃ	**q’**i**χ**
/ɪ/	pɪʃ	$\widehat{\text{tʃ}}$ɪ**q’**

Table 2.11: Relative F1 differences between items of the same vowel category

Table 2.12 shows relative F2 differences by item for each vowel phoneme. The second column shows the item for each vowel category that was usually realized with relatively lower F2 values across speakers as compared to the second column which shows the item that was realized with relatively higher F2 values.

Vowel	Item with relatively lower F2 (vowel realized backer)	Item with relatively higher F2 (vowel realized fronter)
/u/	tu**χ**	su$\widehat{\text{ts}}$’
/ʊ/	t’ʊ**q’**	tʊʃ
/o/	pop	**q’**o**q’**
/ɔ/	kɔ**χ**	$\widehat{\text{tʃ}}$’ɔk
/a/	pas	**q’**a**q’**
/ə/	sə**q**	k’əʃ
/e/	$\widehat{\text{tʃ}}$’ek	ʃeʃ
/ɛ/	ʃpɛ**q**	**q’**ɛ**q**
/i/	**q’**i**χ**	k’iʃ

Table 2.12: Relative F2 differences between items of the same vowel category

The general pattern of vowels adjacent to uvular consonants surfacing with higher F1 and lower F2 values has been reported cross-linguistically for many different languages and language families. Secondary uvularization in Arabic is reported to be characterized by lowered F2 and slightly raised F1 (Zawaydeh 1997). In Western Aleut, vowels adjacent to uvulars have higher F1 and lower F2 relative to those adjacent to velars, and this effect is stronger in the consonant transition than in the middle of the vowel (Gordon et al. 2002). High vowels in the related language Inuktitut are also lowered adjacent to uvulars (Schultz-Lorentzen 1945, in Gordon et al. 2002). In Montana Salish, F2 is lowered in vowels adjacent to rounded uvulars as compared to adjacent to rounded velars (Gordon et al. 2002). Quechua high vowels (and low vowels to a lesser extent) are lowered following uvular consonants, and lowering is consistent throughout the duration of the vowel, not just during the consonant transitions. High front vowels following uvulars have lower F2 values, but high back vowels have higher F2 values; that is, there is centralization of the high vowels in this context (Holliday and Martin 2018). In some varieties of Qiang, acoustic correlates of uvularization include raising of F1 and lowering of F2 in vowels (Evans et al. 2016). In Jul|'hoan, uvularization elevates F1 (Miller 2007, in Evans et al. 2016). The vowels /i iː/ are found to lower following most uvulars in Nuu-cha-nulth, although the vowels /u uː/ are unaffected in this same environement (Wilson 2007).

The lowering and backing effect of uvular consonants on adjacent vowels, found in many languages, is suggestive of coarticulation. However, although similar effects may be found across languages, the particulars of degree and exact effect may differ in each case. For example, raising F1 is found for some but not all vowels in Nuu-cha-nulth, and in Quechua the effect on F2 seems to be not lowering in all cases but centralization. Therefore, coarticulation in itself cannot be a sufficient explanation for this process in any given language. Furthermore, in the Chichicastenango K'iche' data, vowels adjacent to uvulars are not found with higher F1 and lower F2 values in all cases; rather, there are differences between some items with the same vowel which are exactly contrary to this pattern, such as the realization of most tokens of the vowel in /pop/ with F2 values that are lower than those in most tokens of /q'oq'/. It is possible that the effects observed in the Chichicastenango K'iche' data of adjacent uvular consonants are in part due to coarticulatory processes, but in ways that

are more complicated than are visible in this data set due to the lack of sufficient distinct items (for example, there may be an interaction between the effects of the preceding and following consonants on the vowel quality). Alternatively, there may be lexical effects that do not stem directly from consonant context. Further research, involving a larger number of speakers and items for each vowel, would help shed light on this question.

Finally, one other pattern observed in the Chichicastenango K'iche' data is suggestive of a coarticulatory explanation. Tokens of the vowel /ə/ show a much greater differentiation by item than that seen for most other vowels. That is, tokens of the two words with this vowel phoneme are very separate in the F2 dimension, where tokens of səq have much lower F2 values than those of k'əʃ. Assuming that the difference in F2 in these vowels is due to effects of the surrounding consonants, the larger effect for this vowel category as compared to the others may be due to the relatively short duration of this vowel as seen in the duration results previously detailed. A shorter total duration for each vowel permits a much greater influence on it from surrounding segments than would be expected for longer vowels.

2.3.4.4 Differences between Male and Female Speakers and the Status of /ɪ/

No differences were expected between the speech of male and female speakers, with the exception of overall higher formant values for female speakers in keeping with smaller average vocal tracts. However, there are multiple ways in which the data of female and male participants differs. Most of these differences are small and possibly insignificant, such as larger differences between the formant values of the vowels in the words t͡ʃ'ɔk and kɔχ for male speakers than for female speakers. However, there is one area in which the difference between the speech of male and female speakers is quite large: production of words with the vowel /ɪ/. Female speakers produced all tokens of /ɪ/ in both pɪʃ and t͡ʃɪq' with formant values very close to those of /ə/, while male speakers produced tokens of pɪʃ with formant values very close to those of /i/ but tokens of t͡ʃɪq' with formant values very close to those of /ə/. This is shown again in Figure 2.32: all tokens of /ɪ/ (green circles) cluster together for female speakers, while for male speakers there is a large difference between the two items (green open circles and green filled circles).

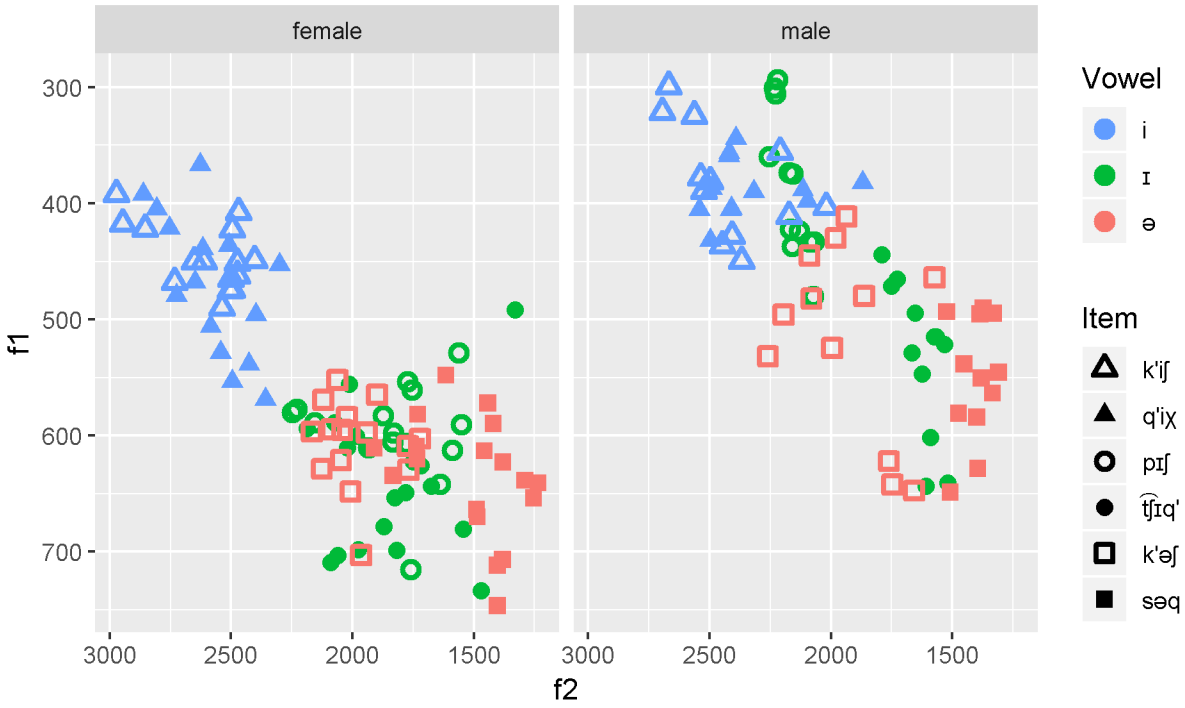

Figure 2.32: Vowels /i ɪ ə/ for male and female speakers

This figure also shows that the tokens of the vowel /ɪ/ do not occupy an easily definable area in the vowel space which is distinct from that of other vowels. All tokens of this vowel are very close in formant values to tokens of either /i/ or /ə/, depending on the lexical item and gender of the speaker. Although finding overlapping formant values for distinct vowel phonemes is common (e.g. the mapping of English vowels shown by Peterson and Barney (1952)), this overlap is particularly extensive, and coincides with a much larger range of realizations for this vowel than for the others in the data. This suggests that the phoneme /ɪ/, the expected modern reflex of the historical short vowel /i/, may no longer exist as a distinct category in modern Chichicastenango K'iche', or may be merging to some degree with surrounding sounds. This is similar to what has been described for Sololá Kaqchikel, a closely related language spoken in a town neighboring Chichicastenango, where vowels /ɪ/ and /i/ appear to be merging (R. Bennett 2016b).

Although the experiment data suggests that phoneme /ɪ/ may be merging with surrounding sounds, this is on the basis of only two lexical items. Furthermore, it is quite likely

that the vowel in one of these two lexical items, transcribed as /t͡ʃɪq'/, was miscategorized, and should rather be /ə/ (explaining why formant values of this vowel are quite close to those of /ə/ for all speakers). Therefore, it is also possible that /ɪ/ is a distinct category in Chichicastenango K'iche', but the quality difference is more subtle that it is possible to see from the data in this experiment. Because of this, I considered some additional data from a small corpus of texts and elicitation materials. Stressed and/or final vowels were segmented by the same rules described for this experiment and then categorized on the basis of cross-dialectal correspondence (long to tense, short to lax), and formant values were obtained with the same script used in the experiment. The following graphs show additional data from three of the experiment participants. Figure 2.33 shows formant results of instances of /i ɪ e ə/ in about two minutes of a free narration by participant 2, a young male speaker (data source code [txt;kot]). Figure 2.34 shows formant results of instances of /i ɪ e ə/ in about two minutes of a free narration by participant 4, a young female speaker (data source code [txt;owl]).

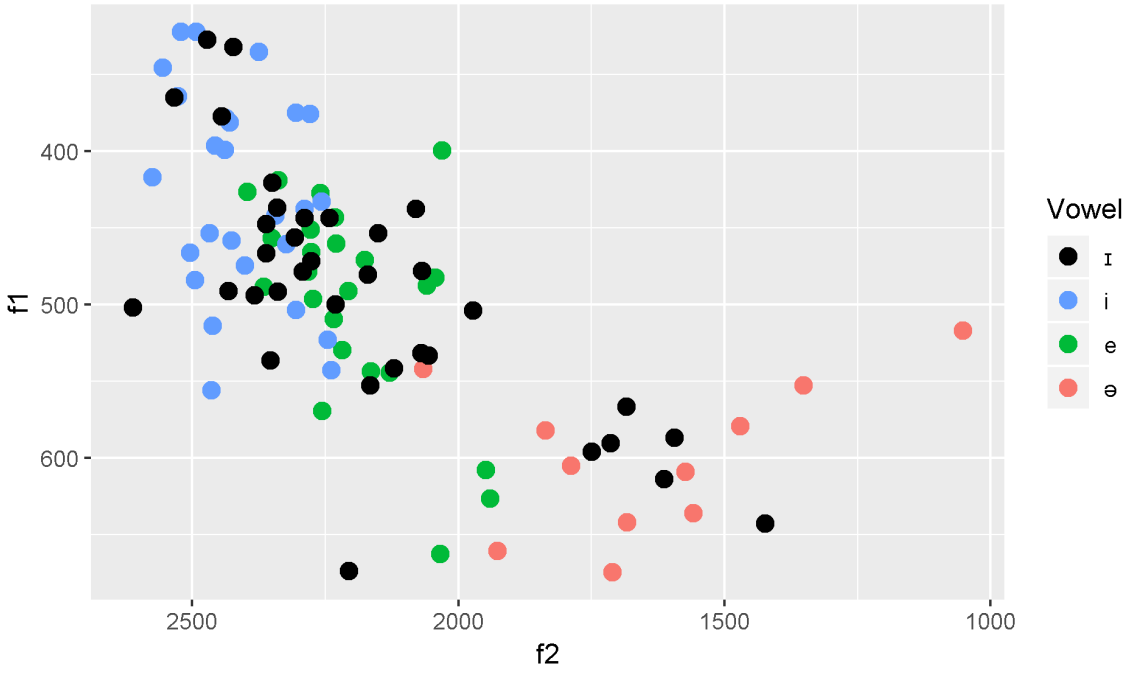

Figure 2.33: Phonetic realization of phonemes /i ɪ e ə/ in text data from one male speaker

Figure 2.35 shows formant results of instances of the words ɓe (*b'e*) 'road' and ɓɪʔ

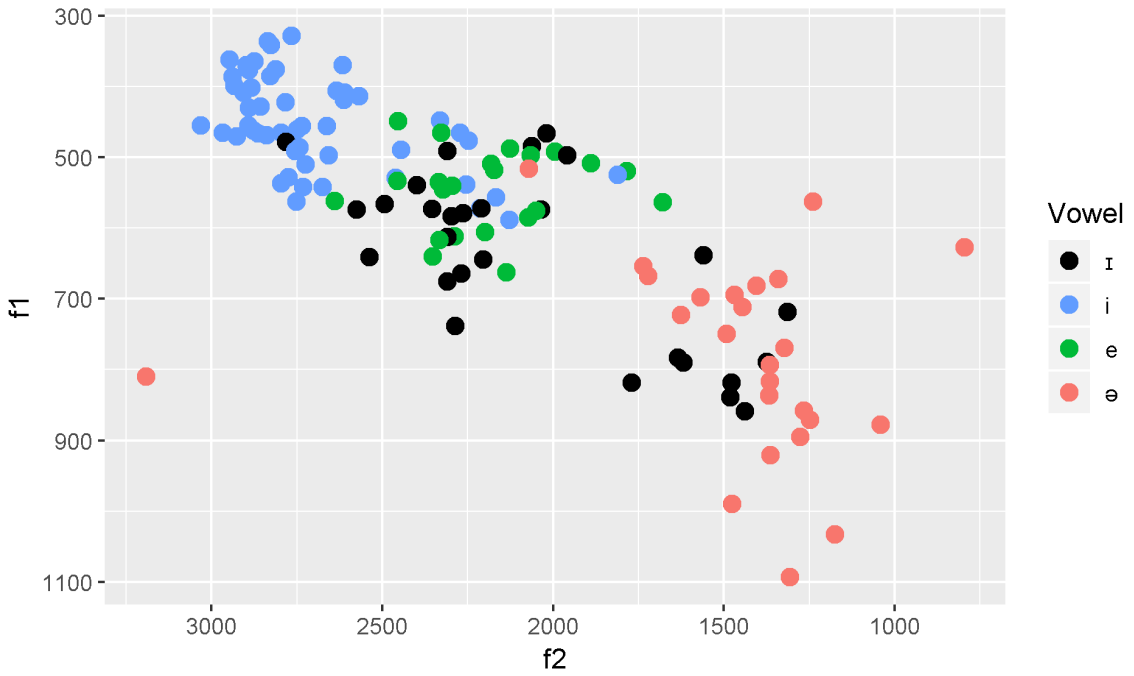

Figure 2.34: Phonetic realization of phonemes /i ɪ e ə/ in text data from one female speaker

(*bʔiʔ*) 'name' from elicitation data of participant 11, an older female speaker (data source code [elic;MJL4]). These words were produced in isolation, in some cases with possessive prefixes, in response to my asking what the difference between the two words is.

The formant values of the vowels in the textual data shown in the first two figures indicate that the phonetic realization of phoneme /ɪ/ (represented as black circles) covers a wide range, overlapping with tokens of /i/, /e/ and /ə/ for both the male and female speaker. Results from the elicitation data show that the realization of the vowels in the words ɓe and ɓɪʔ are very similar to each other. Therefore, the supposed phoneme /ɪ/ is behaving in unexpected ways in additional data, and the effect observed in the experiment described in this chapter is not likely to be due simply to the two lexical items used. Rather, it seems quite possible that this vowel category may have merged or be merging with surrounding vowels. However, much more data is ultimately necessary in order to address this question systematically, including a much larger sample of items for each vowel category and a greater number of speakers. This would be a very productive avenue for future research.

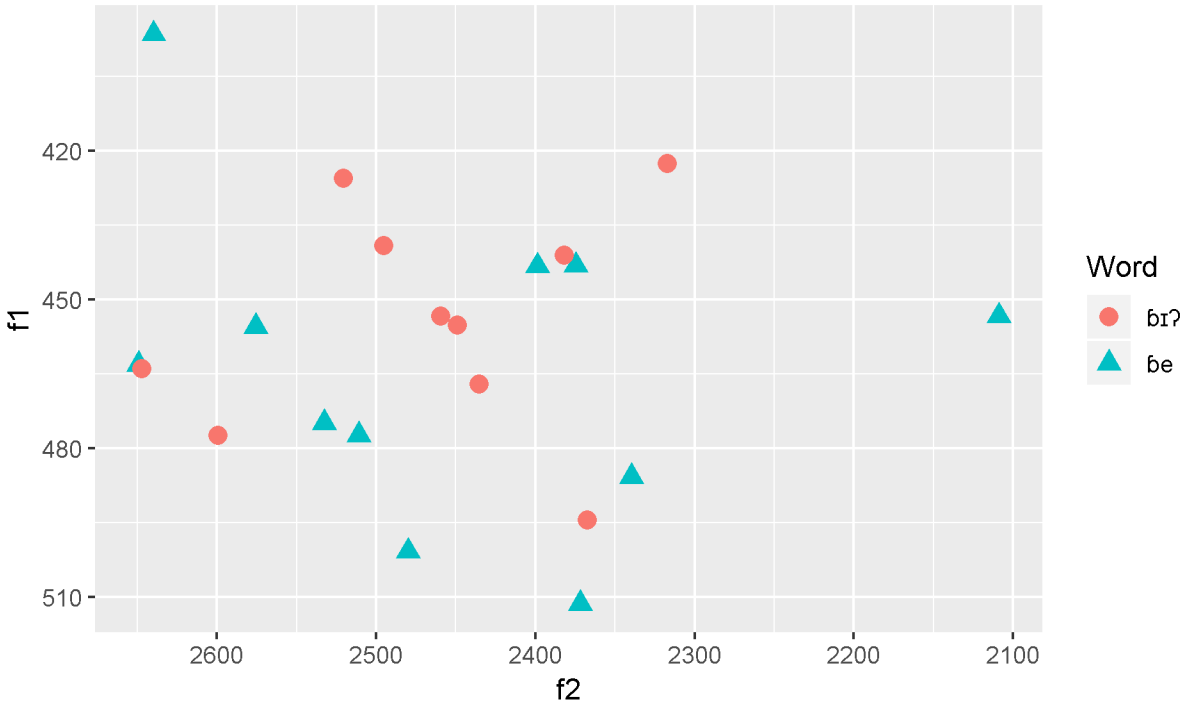

Figure 2.35: Tokens of /βe/ and /βɪʔ/ in elicited data from one female speaker

2.3.4.5 Individual Variation

Formant results reveal that although some patterns observed in the data can be found consistently in each individual speaker, in other areas there is a lot of variability across speakers.

One area that displays individual variability is the dimensions across which tense and lax vowels differ at each place of articulation. Averages over all speakers show that there is a difference in both F1 and F2 (height and frontness) between tense and lax high and mid vowels. Lax high and mid front vowels are both lower and backer on average than their tense counterparts, and lax high and mid back vowels are both lower and fronter on average than their tense counterparts. However, although this pattern is found for some speakers for each place of articulation, it is not found consistently for all. For example, some speakers show a difference between tense /u/ and lax /ʊ/ which depends only on F1 values, with very similar F2 values for both categories, while at least one has a pattern that depends only on F2 values, with very similar F1 values for both categories.

Another locus of individual variation is in the degree to which differences in formant values are found between lexical items of the same vowel category. Across the board, differences are found between the two lexical items for almost all of the vowel phonemes. However, these differences are not consistently found in each individual speaker and may vary in degree across speakers. For example, for several individual speakers there is a discernible difference between the formant values of the vowel /o/ in the words pop and q'oq', where tokens of the vowel in pop have lower F1 and F2 values than those of q'oq'. However, other speakers produced vowels in these two words which do not differ discernibly in formant values.

This experiment has insufficient data to be able to handle this variability effectively. Future research with a larger number of participants and lexical items would help overcome this problem in order to reach more firm conclusions about the vowel categories of Chichicastenango K'iche'.

2.4 Conclusion

This experiment had intended to provide experimental support for the impressionistically described vowel inventory of Chichicastenango K'iche', consisting of five tense vowels and five lax vowels with no length contrast. Results do provide some support for this picture, however it is mixed with a revelation of a great deal of complexity and variability present in the system.

Duration results indicate that average duration difference are to be found between tense and lax sets for some vowel pairs, but not for others. Furthermore, these differences are not consistent across items for the same vowels. This suggests that although duration may not be an important cue for distinguishing between vowels across the board, it may still be involved in distinguishing particular vowel pairs or items, especially the front and central vowels.

Formant results suggest that the main difference between tense and lax vowels in Chichicastenango K'iche' is, as expected from previous descriptions, one of centralization: lax vowels across the board have formant values that are closer to those of schwa in both height and frontness dimensions than those of their tense counterparts. However, results also

show a large degree of variability in formant values across lexical items for each vowel category, possibly attributable to consonant context in some cases but likely not in others, and variability across individual speakers. Finally, results reveal a difference in the production of the phoneme /ɪ/ as close to /i/ or /ə/ depending on the lexical item and gender of the speaker. The absence of a clear /ɪ/ category points towards a possible loss or near merger of this phoneme category with surrounding categories that warrants further investigation.

Due to the large amount of variability in the data in this experiment, firm conclusions cannot be reached. However, this experiment can serve as a pilot study for a larger experiment involving a greater number of participants and many more items for each vowel category. More data may be able to shed some light on the variability observed in this study, finding sources for this variation across vowel pairs, lexical items, individuals, and genders.

Chapter 3

Stress Pattern and Acoustic Cues to Stress

3.1 Introduction

This chapter discusses stress in K'iche'. The first section presents the final stress pattern found in the majority of K'iche' dialects, as well as reviewing the literature on phonetic cues to stress in K'iche' and related languages. The second section presents a preliminary analysis of the stress pattern of Chichicastenango K'iche'.

3.2 Stress in Other K'iche' Dialects

3.2.1 Location of Stress

Stress in most dialects of K'iche' is fixed on the final syllable (López Ixcoy 1997; Baird 2014). In at least some dialects three affixes fall outside of the domain of stress, resulting in stress falling on the penultimate syllable of some adjectives and verbs. These are the suffix -a (*-a*) which appears on some attributive adjectives, the suffix -alaχ (*-alaj*) which appears on superlative adjectives and on the modifier siɓalaχ (*sib'alaj*) 'very', and the suffix -a (*-a*) which appears on phrase-medial intransitive verbs in the dependent mood (imperatives and incorporated movement constructions). Examples of these suffixes in Nahualá K'iche' are shown in Table 3.1.

Orthography	IPA	Gloss
nima kaminaq	[ˈni.m]a [ka.mi.ˈnaq]	'large carcass'
jun nimalaj sib'	χun [ˈni.m]a.laχ [ˈsiɓ]	'a very big (cloud of) smoke'
kinuxlana jub'e'q	[ki.nuʃ.ˈla.n]a [χu.ˈbeʔq]	'I rest a little'

Table 3.1: Suffixes outside of the stress domain in Nahualá K'iche' [txt;3stories]

When none of these suffixes are present in the word, stress can be seen to fall on the final syllable for all stress-bearing words. Examples are shown in Table 3.2.

Orthography	IPA	Gloss
wachajiil	wa.t͡ʃa.ˈχiːl	'my husband'
saqmo'l	saq.ˈmoʔl	'egg'
asaroon	a.sa.ˈɾoːn	'hoe'
qarajawaxik	qa.ɾa.χa.wa.ˈʃik	'our needs'
kinchoomaj	kin.t͡ʃoː.ˈmaχ	'I think'
xupiso	ʃu.pi.ˈso	'she tied it'
kinkoosik	kin.koː.ˈsik	'I am tired'
xb'ee le ixoq pa k'ayb'al	ˈʃɓeː le i.ˈʃoq pa k'aj.ˈɓal	'the woman went to the market'

Table 3.2: Final stress in Nahualá K'iche' [txt;3stories]

3.2.2 Phonetic Cues to Stress

In an acoustic study of three K'iche' dialects (Nahualá, Zunil and Cantel), Baird (2014) finds that for all three dialects there is a significant rise in F0 in the stressed syllable. However, tonic vowels are only found to be significantly longer than atonic vowels in the Cantel dialect, which has lost the historic phonemic length contrast. In the Nahualá and Zunil dialects, which maintain a phonemic length contrast, no significant duration differences were found between short tonic and atonic vowels. For the closely related language Kaqchikel, which like Chichicastenango K'iche' is described as having ten phonemic vowel qualities with no length contrast, Berinstein (1979) reports based on production and perception data that both duration and pitch are cues to stress.

3.3 Stress in Chichicastenango K'iche'

3.3.1 Location of Stress

Previous descriptions of Chichicastenango K'iche' have assumed that stress is final in this dialect just as it is in other dialects. Recently it has been noted that this is not always the case (e.g. Can Pixabaj 2019), however no analysis of the stress pattern is available. This constitutes the first description that I know of for this stress pattern. This analysis is based on elicitation and text data from 8 speakers with forms covering a wide range of syllable shapes, morphological contexts, word classes, etc. In these forms I marked the location of stress impressionistically. In some cases this was very clear, and judgements from

various listeners all agreed with mine. However, in others it was very unclear, and various listeners produced different judgments, leaving much room for error. Therefore, this should be taken as a preliminary analysis which is subject to change and motivates hypotheses for future study. Based on the impressionistic marking of stress in the corpus, the following generalizations emerge:

1. In non-verbs (nouns, adjectives, adverbs), stress falls on the final syllable, irrespective of syllable weight

2. In verbs, stress is assigned according to a hierarchy of syllable weights

The following sections detail the stress pattern for verbs and non-verbs. Since stress must be assigned prior to vowel deletion, as discussed further in Chapter 4, the following examples show deleted vowels that are assumed to be present underlyingly in parentheses in the second tier of the transcription.

3.3.1.1 Stress in Non-verbs

Stress in non-verbs appears to be fixed on the final syllable in the data in my corpus. The suffix -ləχ (*-läj*) which appears on superlative adjectives is outside of the stress domain, as in other dialects. Unlike other dialects, there is no linking suffix -a (*-a*) on attributive adjectives in this dialect, and therefore stress falls on the final syllable of these words as well.

Some examples of -ləχ are shown in Table 3.3. Here, and throughout this chapter, the stress domain is marked in square brackets in words in which there is some material which falls outside it.

Orthography	IPA	Gloss
tzätzläj	/[ˈt͡sət͡s].ləχ/	‘very thick’
nïmläj	/[ˈnɪm].ləχ/	‘very big’

Table 3.3: Stress with the -ləχ suffix in Chichicastenango K’iche’ [txt;3recipes, txt;kot]

Some examples of attributive adjectives with no linking suffix are shown in Table 3.4.

Orthography	IPA	Gloss
ju nïm oj	/χu ˈnɪm ˈʔoχ/	‘a big avocado tree’
chom pan	/ˈt͡ʃom ˈpan/	‘large (loaf of) bread’

Table 3.4: Absence of linking suffix on attribute adjectives in Chichicastenango K’iche’ [txt;owl, elic;MXM2]

Examples of nouns and adverbs of various syllable weights and counts are shown in Table 3.5; stress falls on the final syllable in all.

Orthography	IPA	Gloss	Data source
mnäq	/m(ɪ).ˈnəq/	‘person’	txt;kot
chanim	/t͡ʃa.ˈnim/	‘now’	txt;kot
päjb’äl	/pəχ.ˈɓəl/	‘scale’	elic;MXM2
meb’a’	/me.ˈɓaʔ/	‘poor’	elic;SAGB1
ukk’el	/u.k(ɪ).ˈk’el/	‘his/her blood’	txt;church
k’che’laj	/k’(ə).t͡ʃe.ˈlaχ/	‘forest’	elic;SAGB1
uk’äslmal	/u.k’əs.l(ɛ).ˈmal/	‘his/her life’	txt;church
arko’	/aɾ.ˈkoʔ/	‘bridge’	elic;MCX3
nab’e	/na.ˈɓe/	‘first’	txt;church
wakäx	/wa.ˈkəʃ/	‘cow’	txt;caldores
xïb’il	/ʃɪ.ˈɓɪl/	‘big’	elic;MCX3

Table 3.5: Stress in non-verbs in Chichicastenango K’iche’

3.3.1.2 Stress in Verbs

Domain of Stress

The stress domain of verbs excludes person and aspect prefixes/proclitics as well as the status suffixes -ɪk (*-ïk*), -ɔ (*-ö*), -ʊ (*-ü*) and -ɔq (*-öq*). These are never stressed, despite the fact that several of them form heavy syllables which, as will be discussed below, usually attract stress in verbs. One consequence of the status suffixes being excluded from the stress domain is that stress always falls on the same syllable of the verb with and without these suffixes. Since these suffixes appear obligatorily on phrase-final verbs but are usually absent from phrase-medial verbs, this means that the same vowel is stressed in a given verb both phrase-finally and phrase-medially. This can be seen in the following examples:

(4) a. *kawïl rï pö't*

ka.w[ˈɪl] ɾɪ ˈpɔʔt

k-∅-aw-ɪl ɾɪ pɔʔt

INCPL-B.3S-A.2S-see DET BLOUSE

'You see the blouses.' [txt;kot]

b. *kawïlö*

ka.w[ˈɪ.l]ɔ

k-∅-aw-ɪl-ɔ

INCPL-B.3S-A.2S-see-SS

'You see it.' [txt;3recipes]

(5) a. *kïnchkün farmacia*

kɪn.[t͡ʃ(ə).ˈkʊn] faɾ.ˈma.sja

k-ɪn-t͡ʃək-ʊ-n farmasja

INCPL-B.1S-work-TV-ANT pharmacy

'I work in pharmacy.' [txt;talentos]

b. *kïnchkünïk*

kɪn.[t͡ʃ(ə).ˈkʊ.n]ɪk

k-ɪn-t͡ʃək-ʊ-n-ɪk

INCPL-B.1S-work-TV-ANT-SS

'I work.' [txt;talentos]

Further examples of the exclusion of person markers and status suffixes from the stress domain are given as follows.

(6) a. *xkämïk*

ˈʃ[kə.m]ɪk

ʃ-∅-kəm-ɪk

CPL-B.3S-die-SS

'He/she/it died.' [txt;church]

b. *xb'sönïk*
x[b(ɪ).ˈsɔ.n]ik
x-∅-b'ɪs-ɔ-n-ik
CPL-B.3S-sad-TV-ANT-SS
'He was sad.' [txt;owl]

c. *kqrükü*
kq[ˈrʊk].ʊ
k-∅-q-rʊk-ʊ
INCPL-B.3S-A.1P-stir.liquid-SS
'We stir it (the liquid).' [txt;3recipes]

d. *katpöq'öq*
kat.[ˈpɔ.q']ɔq
k-at-pɔq'-ɔq
INCPL-B.2S-bloom-SS
'Bloom!' [elic;JT2]

e. *katwäröq*
kat.[ˈwə.ɾ]ɔq
k-at-wəɾ-oq
INCPL-B.2S-sleep-SS
'Sleep!' [elic;JT2]

It is worth noting that status suffixes are not excluded from the stress domain in other dialects of K'iche', where they may be stressed. In Nahualá K'iche', stress is final, and status suffixes are always stressed since they are always the last syllable of the word that they are on. An example is given in 7.

(7) a. *karilo*
[ka.ɾi.ˈlo]
ka-∅-ɾ-il-o
INCPL-B.3S-A.3S-see-SS
'He sees it.' (Nah) [txt;3stories]

Stress Rule

Within the stress domain, according to the available data, it appears that primary stress in verbs in Chichicastenango K'iche' is assigned according to a syllable weight hierarchy as follows. Stress falls:

1. **On the left-most peripheral vowel in a closed syllable**[1]**, if available** (as in 8);

(8) a. *kqmub'a'*
kq[mu.ˈɓaʔ]
k-∅-q-mu-ɓaʔ
INCPL-B.3S-A.1P-soak-CAUS.POS
'We soak it.' [txt;3recipes]

b. *xinkchpa'*
ʃin.k[t͡ʃ(ə).ˈpaʔ]
ʃ-in-k-t͡ʃp-aʔ
CPL-B.1S-A.3P-catch-SS
'They caught me.' [elic;MCX3]

[1]Syllabification throughout this document is based on the following rules: a) every syllable has one vowel; b) if there is only one consonant between two vowels it is counted as the onset of the second syllable; c) if there are two consonants between vowels the first is counted as the coda of the first syllable and the second as the onset of the second syllable; d) if there are three consonants between vowels the first is counted as the coda of the first syllable and the second and third as onsets of the second syllable; e) all consonants preceded by no vowels at the beginning of a word are counted as onsets of the first syllable; f) all consonants followed by no vowels at the end of a word are counted as codas of the last syllable

c. *utz'ib'am*
u.[t͡s'i.ˈɓam]
u-t͡s'iɓ-am
A.3S-write-PERF
'He/she has written'. [txt;church]

d. *kqna'tsäj*
kq[ˈnaʔ.tsəχ]
k-∅-q-naʔt-sə-χ
INCPL-B.3S-A.1P-remember-CAUS-ACT
'We remember him.' [txt;church]

2. **If not, on the left-most peripheral vowel in an open syllable, if available** (as in 9);

(9) a. *ktäkmayïk*
k[təkˈma.j]ɪk
k-∅-tək-maj-ɪk
INCPL-B.3S-limp-AFF-SS
'He/she is limping.' [elic;MXC3]

b. *knwesäj*
k(ɪ)n.w[ˈe.səχ]
k-∅-nw-e-sə-χ
INCPL-B.3S-A.1S-go.out-CAUS-ACT
'I take it out.' [txt;owl]

c. *karumäj awib'*

ka.[ˈɾu.məχ]	[a.ˈwiɓ]
k-∅-a-ɾum-ə-χ	aw-iɓ
INCPL-B.3S-A.2S-run-TV-ACT	A.2S-self

'You run.' [elic;MCX3]

d. *kq'jomnïk*
k[q'(ɔ).ˈχo.m(ə).n]ɪk
k-∅-q'ɔχom-ə-n-ɪk
INCPL-B.3S-marimba-TV-ANT-SS
'He/she plays marimba.' [elic;SAGB1]

3. **If not, on the left-most centralized vowel in a closed syllable** (as in 10).

(10) a. *kuknäj*
ku.[k(ʊ).ˈnəχ]
k-∅-u-kʊn-ə-χ
INCPL-B.3S-A.3S-cure-TV-ACT
'It cures (it).' [txt;healing]

b. *kupätnïj*
ku.[pə.t(ə).ˈnɪχ]
k-∅-u-pətən-ɪ-χ
INCPL-B.3S-A.3S-function-TV-ACT
'It functions.' [txt;healing]

c. *k-'äntäk*
k[ˈʔən.tək]
k-∅-ʔən-tək
INCPL-B.3S-do-PASS
'It is done.' [txt;3recipes]

d. *kqpqöwsäj*
kq[p(ɔ).ˈqɔw.səχ]
k-∅-q-pɔqɔw-sə-χ
INCPL-B.3S-A.1P-boil-CAUS-ACT
'We boil it.' [txt;3recipes]

e. *katkli'*
kat.[k(ʊ).ˈlɪʔ]
k-at-kl-ɪʔ
INCPL-B.2S-marry-POS.V
'You get married.' [txt;owl]

f. *xkämsxïk*
x[ˈkəm.s(ə).x]ɪk
x-∅-kəm-sə-χ-ɪk
INCPL-B.3S-die-CAUS-PASS-SS
'He was killed.' [txt;church]

4. **If not, on the final syllable** (as in 11)

(11) a. *xb'söniik*
x[b(ɪ).ˈsɔ.n]ik
x-∅-b'ɪs-ɔ-n-ik
CPL-B.3S-sad-TV-ANT-SS
'He was sad.' [txt;owl]

b. *kïnchküniik*
kɪm.[t͡ʃ(ə).ˈkʊ.n]ɪk
k-ɪm-t͡ʃək-ʊ-n-ɪk
INCPL-B.1S-work-TV-ANT-SS
'I work.' [txt;talentos]

That is, stress appears to falls on the left-most of the heaviest syllable type available, or in the case of no heavy syllables default to the right-most (final) syllable.

3.3.2 Phonetic Cues to Stress

Acoustic studies of other K'iche' dialects and related languages have shown different phonetic cues to stress in different dialects. Although pitch has been shown to be associated

with tonic syllables in several dialects, both those with phonemic vowel length contrasts and those without, duration has only been found to be associated with stress in dialects of K'iche' and other Mayan languages which lack phonemic vowel length (Berinstein 1979; Baird 2014). As detailed in Chapter 2, Chichicastenango K'iche' is described as having no phonemic length contrast (López Ixcoy 1994; López Ixcoy 1997; Par Sapón and Can Pixabaj 2000). Rather, historically long vowels are realized as peripheral and historically short as centralized, and there does not seem to be any consistent duration difference between these two sets. Therefore, it might be expected that tonic syllables in Chichicastenango K'iche' would be associated with longer duration. A quick glance at the data does seem to support this. However, no conclusions can be drawn without a controlled study which takes into account all of the possible reasons that might affect vowel duration other than stress, including vowel height, centralization, surrounding sounds, syllable position, etc. This would be an interesting direction for future research.

3.3.3 Typological Note

The stress pattern described here, if accurate, would make Chichicastenango K'iche' typologically unusual in some ways.

Quality-sensitive stress patterns which make reference to vowel centralization, avoiding stressing central vowels, have been described for a number of languages across the world in a wide range of language families (Goedemans and Hulst 2013). Examples include Aljutor (Chukotko-Kamchatkan, Kenstowicz 1997), Archi (Nakh-Daghestanian, Chumakina et al. 2016), Indonesian (Austronesian, McCarthy and Cohn 1998), French (Indo-European, Collins and Mees 2013), Mari (Uralic, Kenstowicz 1997) and Sakao (Austronesian, Guy 1974). However, these languages usually categorize only one or two vowels in their inventory as central for the purposes of stress. This is most commonly the vowel /ə/, although /ɨ ʌ/ and possibly others are also attested. In contrast to the small number of vowels which avoid stress, some of these languages have large inventories of vowels which do not. For example, Parisian French avoids stressing /ə/ alone as compared to its non-centralized vowels /i y e ø ɛ œ a u o ɔ ɑ œ̃ ɛ̃ ɔ̃ ɑ̃/ (Collins and Mees 2013), and Sakao avoids stressing /ɨ/ as compared to /i e ɛ a u o ɔ ɒ y ø œ/ (Guy 1974). Chichicastenango K'iche', in contrast, has a balanced

system with the same number of peripheral and centralized vowels. Furthermore, vowels at the same places of articulation as those treated as centralized in Chichicastenango K'iche' are commonly treated as peripheral in other languages (e.g. /ɛ/ and /ɔ/ are stressable peripheral vowels in both Sakao and Parisian French).

In some languages with quality-sensitive stress it has been observed that vowels treated as central (most commonly /ə/) have shorter duration than vowels treated as peripheral, and this can explain why these vowels avoid being stressed. One example is Mari (Kenstowicz 1997). In Chichicastenango K'iche', however, there does not appear to be a consistent duration difference between these two sets of vowels based on the results of the acoustic study detailed in the previous chapter. For example, stress falls on peripheral /o/ over centralized /ɔ/ despite the fact that the average duration of these two vowels would appear to be the same. This can be seen in the following example, repeated from above:

(12) *kq'jomnïk*
k[q'(ɔ).ˈχo.m(ə).n]ɪk
k-∅-q'ɔχom-ə-n-ɪk
INCPL-B.3S-marimba-TV-ANT-SS
'He/she plays marimba.' [elic;SAGB1]

This leaves the question of what motivates stress falling on peripheral vowels over centralized vowels.

Finally, the stress pattern described here for Chichicastenango K'iche' would be unusual in its separate treatment of verbs and non-verbs, as non-verbs have fixed final stress and verbs have quality- and quantity-sensitive stress.

3.4 Conclusion

Previous descriptions of stress in K'iche' have focused on other dialects; no descriptions are available for Chichicastenango K'iche'. In other dialects, stress is described as fixed on the final syllable, and depending on the specific dialect the most relevant phonetic cues to stress are identified as being pitch and/or duration. Chichicastenango K'iche' differs from

these other dialects in the distribution of stress. According to this preliminary analysis, stress assignment appears to differ between verbs and non-verbs: it is final for non-verbs in all cases, but depends on syllable weight for verbs (where both vowel quality and syllable closure are relevant factors). If accurate, this would mean that Chichicastenango K'iche' has a typologically rather unusual stress pattern.

This analysis, though apparently consistent with the available data, is very preliminary and relies on impressionistic marking of stress. There are many possible problems with this method, including L1 bias on the part of the listener, which make it impossible to consider these firm conclusions. Rather, this first attempt at a description of the stress pattern of Chichicastenango K'iche' will serve as a motivation for hypotheses for further research on this topic.

Chapter 4

Vowel Deletion

4.1 Introduction

Campbell (1977) describes vowel deletion in Chichicastenango K'iche' by saying that all short vowels are deleted when preceded by a consonant and followed by a stressed syllable, presumably assuming that stress is final as is common in K'iche' dialects. He provides examples in support, shown in Table 4.1.[1]

Surface form	Source	Gloss
ɸt͡ʃaːx	w-**i**t͡ʃaːx	'my vegetable'
ɾɓaːx	ɾ-**a**ɓaːx	'his rock'
ɸxiːx	w-**a**xiːx	'my cane'
t͡ʃkop	t͡ʃ**i**kop	'animal, bird'
t͡sʼnuːn	t͡sʼ**u**nun	'hummingbird'
kmat͡s	k**u**mat͡s	'snake'
ɸnon	w**o**non	'wasp'
snik	s**a**nik	'ant'

Table 4.1: Chichicastenango vowel deletion according to Campbell (1977) (modified to match modern IPA; deleted vowels indicated in bold)

However, these words are all disyllabic nouns, with the structure CVCVC. Even if this description is accurate for CVCVC nouns, the question remains what occurs in longer words, other word classes, or those with different syllable structure. Other mentions of vowel deletion in Chichicastenango K'iche' have not attempted complete descriptions, but note various tendencies such as that it affects many non-final vowels (López Ixcoy 1994) and unstressed vowels (England and Baird 2017).

[1]Note that Campbell writes long and short vowels rather than tense/peripheral and lax/centralized; deletion of short vowels is equivalent to deletion of lax/centralized vowels.

Considering the larger picture of vowel deletion across the Mayan family, we find that it is not an uncommon process in Mayan languages: various types of deletion patterns have been reported for at least Mam, Tektitek, Ixil, Uspantek, Tz'utujil, Sakapultek, Tseltal, Yucatec, Tojolab'al, Mocho' and Huastec, as well as many different dialects of K'iche' (R. Bennett 2016b; Par Sapón and Can Pixabaj 2000). Despite the prevalence of deletion patterns across the family, it is not yet well understood; however, there are some frequently observed patterns. Vowel deletion is typically restricted to short, and not long, vowels, and to those in unstressed syllables, in particular adjacent to stressed syllables (R. Bennett 2016a). Additionally, other restrictions may apply, including limiting deletion to affixes, prohibiting deletions from verbal prefixes, or enforcing certain phonotactic rules. Word-initial vowels may also be exempt from deletion, although not always (R. Bennett 2016b). In K'iche' dialects, it is also reported that vowel deletion tends to occur especially in non-final and non-initial syllables, which are almost always unstressed (López Ixcoy 1994), and in particular in intermediate syllables in longer words (Par Sapón and Can Pixabaj 2000). The loss of initial vowels is registered in only one dialect, that of Cunén (Par Sapón and Can Pixabaj 2000).

Previous descriptions of vowel deletion in Chichicastenango K'iche' and other dialects have presented a range of factors that may be relevant to deletion. However, a comprehensive description which addresses deletion in different word classes and which takes into account the changes in the stress pattern of this dialect is not yet available. This chapter will attempt to shed more light on this issue, by showing that there appear to be different restrictions on vowel deletion in content and function words and by attempting to distinguish between diachronic vowel loss and synchronic deletion. Since there are a wide range of factors which could be relevant to deletion, this chapter will focus on a few that appear to be relevant in the data, including vowel quality (peripheral vs. centralized), stress, syllable shape (CV, VC, CVC) and syllable position (non-final, final). This chapter addresses vowel deletion in content words based on data from a database of multisyllabic forms (primarily verbs and nouns) collected from a corpus of texts and elicitation materials. Deletion in function words is considered from observing their behavior primarily in the text corpus, with some reference to elicitation materials.

Finally, a note on the presentation of examples in this chapter. It appears that stress assignment must occur prior to vowel deletion, since deletion is sensitive to stress. Evidence for this comes from word pairs where the only difference between two vowels in terms of context and position appears to be stress, and the unstressed vowel is deleted while the stressed one is not. Consider the following two verbs:

(13) a. *xïnkchäpö*
ʃɪn.k[ˈt͡ʃə.p]ɔ
ʃ-ɪn-k-t͡ʃəp-ɔ
CPL-B.1S-A.3P-hold-SS
'They caught me.'

b. *xïnkchpa'*
ʃɪn.k[t͡ʃ(ə).ˈpaʔ]
ʃ-ɪn-k-t͡ʃ(ə)p-aʔ
CPL-B.1S-A.3P-hold-SS
'They caught me.'

In 13a, the schwa vowel in the root is stressed, and that vowel is not deleted. In 13b, this vowel is unstressed, since stress falls on the suffix which has a peripheral vowel. The vowel in this case is deleted. Given that stress is believed to occur prior to vowel deletion, underlying vowels were shown in parentheses in examples in the previous chapter. This same convention is used in this chapter, where parentheses indicate a posited underlying vowel which is absent in the form in question.

4.2 Types of Vowel Deletion and Loss

Prior to discussing the factors which affect vowel deletion, it is important to define the process that is being described. It is quite unclear from previous descriptions of vowel deletion in Chichicastenango K'iche' (e.g. Campbell 1977) and other dialects whether the process referred to is diachronic or synchronic. In this chapter, I use the term diachronic vowel loss to refer to a historical process whereby vowels that previously existed are no

longer present, with the historical vowel posited to be no longer part of the underlying representation. By synchronic vowel deletion I refer to a phonological process whereby vowels which exist in the underlying representation are not pronounced in some contexts, leading to alternations between forms with the vowel and forms without it. Previous descriptions of vowel deletion in Chichicastenango K'iche' provide no alternations or sets of words with the vowel present in one and absent in the other, and almost all examples are of disyllabic CVCVC nouns where the first vowel is deleted. Therefore, it is not possible to evaluate this question from previously provided data. However, a closer look at contexts of vowel deletion in Chichicastenango from new data indicates that at least three distinct processes which could be referred to as deletion may be posited to occur (or have ocurred) in Chichicastenango K'iche': diachronic vowel loss, synchronic vowel deletion, and vowel optionality.

4.2.1 Diachronic Vowel Loss

Diachronic vowel loss can be observed in many function words and grammatical morphemes which appear in the data exclusively without the historic vowel. Examples include the negation markers *n* (*nə) and *t* (*tə), the preposition *p* (*pə) or the person prefixes *k-* (*kɪ-), *q-* (*qə-) and *n* (*nʊ).[2] Some of these can be seen in the following examples, where historic vowels are not written in the second and third tiers as they are posited to be fully absent from the underlying representation. Example 14 shows the three historically CV particles *nə, *tə and *t͡ʃɪ, all of which appear exclusively without the vowels.

(14) *n k'o t ch uchak*

n	k'o	t	t͡ʃ	u.ˈt͡ʃak
n	k'o	t	t͡ʃ	u-t͡ʃak
NEG	EXIST	NEG	anymore	A.3S-work

'He doesn't have work anymore.' [elic;MACM2]

The examples in 15 show the first person singular possessive prefix, historically *nʊ-. This prefix appears without the vowel no matter what word it attaches to, including when

[2]Historic reconstruction based on the form that is found in the Nahualá dialect, which does not delete these vowels, and with reference to the long-peripheral, short-centralized correspondence.

followed by consonant clusters created by other deletions as in 15a and 15b and with Spanish loanwords beginning in consonant clusters as in 15d. The nasal also assimilates to the place of the following consonant as can be seen in 15b and 15c.

(15) a. *nrkil*
nɾ(ɪ).ˈkil
n-ɾɪkil
A.1S-food
'my food' [elic;TJL1]

b. *mbq'il*
nɓ(ə).ˈqil
n-ɓəq-il
A.1S-bone-IN.POS
'my body' [elic;MCX8]

c. *mb'i'*
ˈnɓɪʔ
n-ɓɪʔ
A.1S-name
'my name' [txt;talentos]

d. *ngrupo*
ˈngɾu.po
n-grupo
A.1S-group
'my group' (Spanish loan) [txt;talentos]

Since these forms show no alternations with vowel-bearing forms, no matter the context they appear in, there is no evidence that the historic vowels are present in any form in the modern language. The same could be said for a variety of historically CVCVC nouns which seem to never appear with any suffixes, and therefore for which the deleted vowel is likely unrecoverable. An example might be /knəq'/ (*kɪnəq') 'black bean'. However, further

research on this question is necessary; it may be that some grammatical context can be found in which nouns like this do recover their vowels.

4.2.2 Synchronic Vowel Deletion

Synchronic deletion can be observed in many nouns and verbs which show alternations between vowels that are present and absent, depending on the phonological context. For instance, the vowel /ə/ is present in the causative suffix /-sə/ in active forms, but deleted in the corresponding passives, as in the following example:

(16) a. *kukämsäj*
ku.[ˈkəm.səχ]
k-∅-u-kəm-sə-χ
INCPL-B.3S-A.3S-die-CAUS-ACT
'It kills him.' [txt;healing]

b. *xkämsxïk*
ʃ[ˈkəm.s(ə).ʃ]ɪk
ʃ-∅-kəm-sə-ʃ-ɪk
CPL-B.3S-die-CAUS-PASS-SS
'He was killed.' [txt;church]

Synchronic vowel deletion is expected to be regular and depend on context, affecting all vowels of the relevant type which are in the correct environment for deletion.

4.2.3 Optional Vowels

In addition to these regular patterns of synchronic and diachronic vowel deletion, optional vowels can be observed in certain function words and grammatical morphemes. Some examples are the demonstrative /ɾ/∼/ɾɪ/, the directional particle /ɓ/∼/ɓɪ/, the particle /k/∼/kʊ/ and the person markers /n/∼/ɪn/ and /ʃ/∼/ɪʃ/. In each of these, the presence of the vowel does not seem to depend on the environment in any obvious way, and rather may be due to free variation or personal preference (although it is also possible that these

vowels do follow a more regular pattern that is just more complicated than I have been able to determine). For example, in 17a, the vowel in the determiner ɾɪ is present, whereas in 17b it is absent, although the context is quite similar.

(17) a. *kawïl rï pö't*

ka.[ˈwɪl]	ɾɪ	ˈpɔʔt
k-∅-aw-ɪl	ɾɪ	pɔʔt
INCPL-B.3S-A.2S-see	DET	blouse

'You see the blouses.' [txt;kot]

b. *kel lö r pö't*

k[ˈel]	lɔ	r(ɪ)	ˈpɔʔt
k-∅-el	lɔ	rɪ	pɔʔt
INCPL-B.3S-go.out	DIR	DET	blouse

'The blouses are made' (lit: 'The blouses go out.') [txt;kot]

Similarly, the vowel in the set A and set B first person singular marker /ɪm(w)/ is often deleted, but not always, as can be seen in the following examples. The vowel is absent in 18a and 18b, which have syllabic nasals, but present in 18c and 18d.

(18) a. *knk'lï' b'ï*

k(ɪ)n.[k'(ʊ).ˈlɪʔ]	ɓɪ
k-ɪm-k'ʊl-ɪʔ	ɓɪ
INCPL-B.1S-marry-POS.V	DIR

'I marry.' [txt;owl]

b. *ütz xnwïlö*

ʊt͡s	ʃ(ɪ)n.w[ˈɪ.l]ɔ
ʊt͡s	ʃ-∅-ɪnw-ɪl-ɔ
good	CPL-B.3S-A.1S-see-SS

'I liked it.' (lit: 'I saw it well.') [txt;talentos]

c. *kïnwesäj uwäch*

kɪn.w[ˈe.səχ]	u.ˈwət͡ʃ
k-∅-ɪnw-e-sə-χ	u-wət͡ʃ
INCPL-B.3S-A.1S-go.out-CAUS-ACT	A.3S-face

'I imitate it.' (lit: 'I take out its face.') [txt;owl]

d. *su kïn'änö*

su	kɪn.[ˈʔə.n]ɔ
su	k-∅-ɪn-ʔən-ɔ
what	INCPL-B.3S-A.1S-do-SS

'What do I do?' [txt;owl]

This chapter does not attempt to address all of these types of deletion. Rather, it will focus on synchronic vowel deletion, using as data those forms for which alternations are found between the presence and absence of the vowel depending on the context, and vowels never appear to be optional in any given form. Diachronic vowel loss and the question of optional vowels are left for future research.

4.3 Synchronic Vowel Deletion

This section describes some factors which appear to influence vowel deletion, including vowel quality, stress, syllable structure and syllable position. Additionally, vowel deletion appears to apply differently in content and function words, that is, following different restrictions. Therefore, these are discussed separately in the next two sections.

4.3.1 Deletion in Content Words

Deletion in content words (thus far, observed in verbs and nouns) appears to affect unstressed centralized vowels when they are found in non-final CV syllables. That is, vowels are not deleted when they are initial, in final syllables, in closed syllables, or stressed. Periph-

eral vowels appear to resist deletion. Additionally, vowels are often deleted when adjacent to a stressed syllable. Each of these restrictions is explored as follows.

4.3.1.1 Effect of Vowel Quality

In content words, deleted vowels are centralized (corresponding to short in other dialects), and the difference between centralized and peripheral vowels often seems to be the only difference between two words which can explain the deletion pattern. For example, consider the following examples in which peripheral vowels in the first (unstressed) syllable of disyllabic CVCVC words are not deleted.

(19) a. *chanim*
t͡ʃa.ˈnim
t͡ʃanim
now
'now' [txt;kot]

b. *wakäx*
wa.ˈkəʃ
wakəʃ
cow
'cow' [txt;caldores]

Words of the same shape and syllable structure that have centralized vowels in the same position can be seen to have that vowel deleted, as in the following examples:

(20) a. *jnäb'*
j(ʊ).ˈnəɓ
jʊnəɓ
year
'year' [elic;MCX5]

b. *jümb'ir*
jʊ.n(ə).ˈɓiɾ
jʊnəɓ-iɾ
year-PST
'last year' [elic;MCX5]

(21) a. *qlew*
q(ʊ).ˈlew
q-ʊlew
A.1P-land
'our land' [txt;church]

b. *ülew*
ʊ.ˈlew
ʊlew
land
'land' [txt;church]

The vowels in the first syllable of each of these words are all unstressed, non-initial, and in non-final open syllables. The centralized vowels which exist in other forms of the word (as shown in 20b and 21b) are deleted in 20a and 21a in this position, while the peripheral vowels in the same position in 19a and 19b are not.

K'iche' underwent a historical process of vowel neutralization in which non-final long vowels regularly became short. This means that the only long vowels in non-final positions in many dialects of K'iche' are those which result from other sources. One source of these long vowels in non-final syllables is a later historical change through which short vowels were lengthened upon the loss of a following glottal fricative (Larsen 1988). Another source is stressed syllables in Spanish loanwords (Larsen 1988). This historical process of shortening of long vowels in non-final syllables appears to have affected Chichicastenango K'iche' in the same way as other dialects. Those peripheral vowels which do appear in non-final syllables, as in examples 19a and 19b, seem to originate from these exceptions: the /a/ in t͡ʃanim from

a *Vh sequence (*tya+ʔahn.iim, Kaufman and Justeson 2002, p. 1289) and the /a/ in wakəʃ from a stressed vowel in Spanish ˈba.ka. These vowels surface as non-final long vowels in other dialects of K'iche' (Ajpacajá Tum et al. 2005). Other vowels which are peripheral when in the final syllable can be deleted (and therefore should be analyzed as centralized) when no longer in the final syllable, as can be seen upon the addition of suffixes. The following examples show this occurring with the -Vɓ pluralizing suffix.

(22) a. *äk'al*
ə.ˈk'al
ək'al
child
'child' [elic;TJL2]

b. *äk'lab'*
ə.k'(ə).ˈlaɓ
ək'al-aɓ
child-PL
'children' [elic;TJL2]

(23) a. *ajchak*
aχ.ˈt͡ʃak
aχ-t͡ʃak
AG-work
'worker' [elic;TJL2]

b. *ajchkib'*
aχ.t͡ʃ(ə).ˈkiɓ
aχ-t͡ʃak-iɓ
AG-work-PL
'workers' [elic;TJL2]

(24) a. *ilq'om*
il.ˈq'om
ilq'-om
steal-AG
'thief' [elic;TJL2]

b. *ilq'mab'*
il.q'(ɔ).ˈmaɓ
ilq'-om-aɓ
steal-AG-PL
'thieves' [elic;TJL2]

(25) a. *q'äb'rel*
q'ə.ɓ(ə).ˈrel
q'əɓər-el
drunk-AG
'drunkard' [elic;TJL2]

b. *q'bärlab'*
q'(ə).ɓə.ɾ(ɛ).ˈlaɓ
q'əɓər-el-aɓ
drunk-AG-PL
'drunkards' [elic;TJL2]

(26) a. *b'nel q'jom*
ɓ(ə).ˈnel q'(ɔ).ˈχom
ɓən-el q'ɔχom
do-AG marimba
'marimba player' [elic;TJL2]

b. *ajq'öjmab'*
aχ.q'ɔ.χ(ɔ).ˈmaɓ
aχ-q'ɔχom-aɓ
AG-marimba-PL
'marimba players' [elic;TJL2]

In each of these examples, vowels which are peripheral in the final syllable are deleted when a suffix is added, indicating that they had become centralized.

4.3.1.2 Effect of (Primary) Stress

The analysis of stress in Chichicastenango K'iche' presented in the previous chapter assumes that stress assignment is prior to deletion, in order to account for cases where vowels in similar positions are stressed and not deleted in one form and deleted (with stress on a different syllable) in another. The example from the introduction of this chapter is repeated here:

(27) a. *xïnkchäpö*
ʃɪn.kˈt͡ʃə.p]ɔ
ʃ-ɪn-k-t͡ʃəp-ɔ
CPL-B.1S-A.3P-hold-SS
'They caught me.' [elic;MCX3]

b. *xïnkchpa'*
ʃɪn.k[t͡ʃ(ə).ˈpaʔ]
ʃ-ɪn-k-t͡ʃ(ə)p-aʔ
CPL-B.1S-A.3P-hold-SS
'They caught me.' [elic;MCX3]

These two verb forms are from the same root /t͡ʃəp/ 'hold'. In 27a, the root vowel is not deleted, while in 27b it is. These vowels are both centralized and in non-final CV syllables. However, the vowel in 27a is stressed, while the vowel in 27b is not (see the previous chapter

for the posited stress rule which accounts for this based on the vowel quality difference in the two suffixes).

The effect of stress can also explain the difference between examples like the following:

(28) a. *kuköjö*
ku.[ˈkɔ.χ]ɔ
k-∅-u-kɔχ-ɔ
INCPL-B.3S-A.3S-wear-SS
'She wears it.' [elic;MXM2]

b. *ukjom*
u.[k(ɔ).ˈχom]
∅-u-kɔχ-om
B.3S-A.3S-wear-PERF
'She has worn it.' [elic;MXM2]

In 28a, the root vowel is not deleted, while in 28b it is. These vowels are both centralized and in non-final CV syllables. However, the vowel in 28a is stressed, while the vowel in 28b is not.

4.3.1.3 Effect of Syllable Closure

In content words, all deleted vowels appear to be in open syllables.[3] Vowels in closed syllables never seem to be deleted in my data. For example, compare the following:

[3]As noted in the previous chapter, syllabification throughout this document is based on the following rules: a) every syllable has one vowel; b) if there is only one consonant between two vowels it is counted as the onset of the second syllable; c) if there are two consonants between vowels the first is counted as the coda of the first syllable and the second as the onset of the second syllable; d) if there are three consonants between vowels the first is counted as the coda of the first syllable and the second and third as onsets of the second syllable; e) all consonants preceded by no vowels at the beginning of a word are counted as onsets of the first syllable; f) all consonants followed by no vowels at the end of a word are counted as codas of the last syllable

(29) a. *kümb'äl*
kʊn.ɓəl
kʊn-ɓəl
heal-INSTR
'medicine' [txt;healing]

b. *kuknäj*
ku.[k(ʊ).ˈnəχ]
k-∅-u-kʊn-ə-χ
INCPL-B.3S-A.3S-heal-TV-ACT
'It heals it.' [txt;healing]

These two words both come from the root kʊn 'heal'. In 29b the vowel is deleted, while in 29a it is not. These are both centralized vowels in unstressed non-final syllables. The vowel in 29a is in a closed syllable and is not deleted, while the vowel in 29b is in an open syllable and is deleted. Further examples of unstressed centralized vowels in non-final CVC syllables that are not deleted are shown as follows:

(30) a. *säqwäch*
səq.ˈwət͡ʃ
səq-wət͡ʃ
white-face
'potato' [txt;healing]

b. *päjb'äl*
pəχ.ˈɓəl
pəχ-ɓəl
weigh-INSTR
'scale' [elic;MXM2]

4.3.1.4 Effect of Syllable Onset

In content words, word-initial vowels never appear to be deleted; that is, all deleted vowels are in syllables with onsets.[4] For example, consider the following:

(31) a. *ätz'am*
ə.ˈt͡s'am
ət͡s'am
salt
'salt' [txt;3recipes]

b. *rtz'am*
ɾ(ə).ˈt͡s'am
ɾ-ət͡s'am
A.3S-salt
'its salt' [txt;3recipes]

The initial vowel in *ätz'am* in 31a is centralized, unstressed, non-final, and in an open syllable, but it is initial and is not deleted. When a prefix is added as in 31b, this vowel is deleted.

4.3.1.5 Effect of Adjacency to a Stressed Syllable

It has been previously stated that deletion occurs in pre-tonic syllables (Campbell 1977), presumably assuming that stress is final as in other dialects. As noted in the previous chapter, stress is not always final in verbs in Chichicastenango K'iche'. However, the position of a vowel with respect to the stressed syllable does seem to be a relevant factor to consider. This can be seen in longer words where there are multiple centralized vowels in unstressed CV syllables.

[4] K'iche' does not allow adjacent vowels within words, so non-initial syllables always have onsets.

(32) a. *ïxöq*
ɪ.ˈʃɔq
ɪʃɔq
woman
‘woman’ [elic;MXM2]

b. *wïxqil*
wɪ.ʃ(ɔ).ˈqɪl
w-ɪʃɔq-ɪl
A.1S-woman-IN.POS
‘my wife’ [elic;TJL2]

Both of the vowels in the noun /ɪʃɔq/ in 32a are expected to become suited to deletion in 32b, when the possessive morphology is added, since these are both centralized vowels in unstressed, non-final CV syllables. However, only the second vowel is deleted. This is the one which immediately precedes the stressed syllable. Therefore, adjacency to the stressed syllable could explain this. However, there are also longer forms in which multiple vowels are deleted. Consider the following:

(33) *kq’jomnïk*
k[q’(ɔ).ˈjo.m(ə).n]ɪk
k-∅-q’ɔjom-ən-ɪk
INCPL-B.3S-marimba-ANT-SS
‘He/she plays marimba.’ [elic;SAGB1]

In 33, two vowels are deleted. Both are centralized vowels in unstressed non-final CV syllables, one preceding the stressed syllable and the other following it. It is possible that post-tonic vowels are also subject to deletion, or it is possible that there is an effect of secondary stress. Since there is no description of secondary stress in this dialect, and addressing this would require a specific study and larger database, this question is left to future research.

4.3.1.6 Effect of Syllable Position

In content words, the vowel in the final syllable is never deleted in the data. Consider the following examples:

(34) a. *kawïlö*
ka.w[ˈɪ.l]ɔ
k-∅-aw-ɪl-ɔ
INCPL-B.3S-A.2S-see-SS
'You see it.' [txt;3recipes]

b. *kqrükü*
kq[ˈrʊ.k]ʊ
k-∅-q-rʊk-ʊ
INCPL-B.3S-A.1P-stir.liquid-SS
'We stir it (the liquid).' [txt;3recipes]

Here, the final vowels -ɔ and -ʊ are not deleted, despite being centralized vowels in unstressed CV syllables adjacent to the stressed syllable.

4.3.1.7 Additional Examples

Some more examples of deletion in which multiple factors are relevant at once are shown here:

(35) a. *kïk'*
ˈkɪk'
kɪk'
blood
'blood' [txt;3recipes]

b. *ukk'el*
u.k(ɪ).ˈk'el
u-k(ɪ)k'-el
A.3S-blood-IN.POS
'his blood' [txt;church]

In example 35, the centralized vowel /ɪ/ is present in the bare noun in 35a, but absent from the root when the suffix is added in 35b. The deleted vowel is unstressed and in a non-final open syllable, while the non-deleted vowel is stressed and in a final closed syllable.

(36) a. *ïxöq*
ɪ.ˈʃɔq
ɪʃɔq
woman
'woman' [elic;MXM2]

b. *ïxqib'*
ɪ.ʃ(ɔ).ˈqiɓ
ɪʃɔq-iɓ
woman-PL
'women' [txt;kot]

Similarly, in example 36, the centralized vowel /ɔ/ is present in the bare noun in 36a but absent when the plural marking is added in 36b. The deleted vowel is unstressed and in a non-final open syllable adjacent to the stressed syllable, while the non-deleted vowel is stressed and in a final closed syllable.

4.3.2 Deletion in Function Words

Unlike deletion in content words, deletion in function words does not appear to be a consistent process; that is, not all vowels are deleted in the same types of contexts. The only

word (yet identified) for which there seems to be a consistent pattern of when the vowel is present and when it is absent is the existential particle /k'o/ (*k'o), which appears without the vowel when followed by the locative focus particle /wi/ but with the vowel elsewhere; this vowel which is deleted is a peripheral vowel. An example is shown in 37. While the vowel in k'o is deleted in 37a preceding wi, it is present in 37b.

(37)

a. *chu k' wi baño chi'*

t͡ʃu	k'(o)	wi	ˈba.ɲo	t͡ʃiʔ
t͡ʃu	k'o	wi	baɲo	t͡ʃiʔ
where	EXIST	FOC.LOC	bathroom	here

'Where is the bathroom here?' [elic;MACM1]

b. *k'o täj*

k'o	təχ
k'o	təχ
EXIST	NEG

'There is none.' [elic;MACM1]

It is unclear whether this is a lexicalized deletion, or depends on some other factor such as stress, especially since there is no available description of stress in function words nor sentence level stress.

Other function words never appear to delete vowels in the data. Many of these appear to have peripheral vowels, such as /ʃa/ (*ʃa) 'only', the question marker /la/ (*la), or the person marker /at/ (*at). However, others with centralized vowels, which would be expected to be more susceptible to deletion, also never seem to delete their vowels; an example is /tɛ/ (*tɛ) 'then'. Deletion of vowels in CVC particles seems to be entirely absent; some examples are the pluralizer /təq/ (*təq), the phrase-final negator /təχ/ (*təχ), the directional particle /lɔq/ (*lɔq) and the discourse particle /ɾɪʔ/ (*ɾɪʔ). However, deletion or vowel optionality does appear in CVC syllables in longer words, for example in verbal person prefixes.

Finally, as noted above, many other function words either have optional vowels, whose presence or absence does not appear to depend on the phonological context in any way yet determined, or vowels which are fully lost and never appear in any context.

In conclusion, there is much that is unclear about deletion in function words. In order to address this question, a more systematic study and more data are necessary.

4.4 Conclusion

Previous descriptions of vowel deletion in Chichicastenango were able to identity some of the environment restrictions on deletion, but left many open questions. Additionally, the distinction between diachronic and synchronic deletion was not addressed This chapter presented further data on deletion in Chichicastenango K'iche', distinguishing between diachronic vowel loss, synchronic deletion, and optional vowels. Focusing on synchronic deletion, it was noted that different patterns can be found for content and function words. Deletion in content words appears to be restricted to unstressed centralized vowels in non-final CV syllables. Deletion in function words, however, may affect vowels in any syllable shape, final syllables are not immune, and the effects of stress are unclear. Furthermore, there seem to be many examples of optional vowels in function words.

Although this description makes some progress in expanding our knowledge of vowel deletion in Chichicastenango K'iche', there are many questions that are still unanswered. In addition to further research on the factors relevant to deletion considered in this chapter, many others could be considered, such as possible effects of speech rate, register, speaker preference or possible phonotactic constraints on pronounceability. Additionally, due to the preliminary nature of the stress pattern rules detailed in Chapter 4 and the effect of stress on deletion, further investigation is needed in this area. Therefore, there is much further work that could be done on the topic of vowel deletion in Chichicastenango K'iche'.

Chapter 5

Conclusion

5.1 Discussion of Overall Results

This paper addresses the questions of vowel quality, duration and deletion in Chichicastenango K'iche', a phonologically innovative dialect which differs considerably from other dialects of the language. Additionally, a preliminary description of the stress pattern of this dialect is presented. Each of these topics would benefit greatly from additional research, as results are inconclusive. This paper builds on previous work on these topics, expanding upon what has previously been done, and pointing out areas where there are open questions.

After the introduction to the topics and previous literature in Chapter 1, Chapter 2 addresses the phonetic realization of the 10 reported vowel phonemes, identified as tense /a e i o u/ and lax /ə ɛ ɪ ɔ ʊ/ in previous literature. Through a phonetic experiment, quality (F1 and F2) and duration of each vowel were measured. Duration results suggest that duration is not an important cue for distinguishing these two sets of vowels, since average duration differences can only be found for some vowel pairs, and these are mostly upheld by a few particularly short items. Formant results indicate that the primary difference between the two sets of vowels is one of centralization: the vowels called tense have more peripheral qualities than those called lax, which are more centralized. Additionally, the phonetic status of supposed phoneme /ɪ/, corresponding cross-linguistically to short /i/, was inconclusive. Realization of words with this phoneme was found to be close to [i] or [ə] in the experiment data, depending on the item and gender of the speaker. An additional brief study of textual and elicitation data found realizations of this phoneme close to [e] as well. This suggests that this phoneme category may have merged or be in the process of merging with surrounding vowel categories. Finally, a large degree of variability was observed in formants for all vowels in the experiment data, comparing across speaker, item and gender.

Chapter 3 addresses the stress pattern of Chichicastenango K'iche', providing a preliminary description based on impressionistic judgments of syllable prominence. This constitutes the first known description of this stress pattern. Unlike most other dialects of K'iche', which place stress on the final syllable of the word for all stress-bearing word classes, the Chichicastenango dialect appears to behave differently for verbs and non-verbs. Non-verbs, like all words in other dialects, have final stress, no matter the weight of the syllables in the word. Verbs, however, appear to place stress according to a hierarchy of syllable weight, identifying the heaviest available syllable in the verb stem. Stress is placed on the left-most peripheral vowel in a closed syllable, if present, and if not on the left-most peripheral vowel in an open syllable, if present, and if not on the left-most centralized vowel in a closed syllable, if present. In the absence of any heavy syllables, stress appears to default to the final syllable. Additionally, a brief glance at phonetic cues to stress suggested that syllables identified as stressed had what seemed to be longer vowels than unstressed syllables.

Chapter 4 addresses patterns of vowel deletion and loss in Chichicastenango K'iche'. Three distinct types of deletion were identified in the data: diachronic vowel loss, synchronic vowel deletion, and optional vowels. Of these, synchronic vowel deletion was studied in more detail, and in particular focusing on content words. Deletion in this context was found to be restricted to unstressed centralized vowels in non-final CV syllables, especially adjacent to stressed syllables. In content words, these restrictions did not appear to hold, at least not as strictly.

5.2 Future Directions

This paper presents results from analyses of vowel quality, duration, stress and deletion in Chichicastenango K'iche' which considerably expand upon our understanding of this vowel system. However, there are many unanswered questions remaining. Future research is needed to further knowledge on this topic.

Experimental results revealed a large degree of variability in the data, when comparing across speakers, genders, items and vowel pairs. This variability was such that it was difficult to generalize to the approximate place of articulation of each vowel. A more thorough

production experiment is needed, with more subjects and more words per category, to more fully address this question.

Additionally, the production experiment revealed an unexpected split in the realization of the supposed phoneme /ɪ/, which surfaced as close to [i] for male speakers for one item, and [ə] for female speakers for one item and both male and female for the second. No tokens of /ɪ/ were close to what might be an expected realization of [ɪ]. Further investigation through elicitation and textual data indicated that realization of this phoneme is the same as or very close to that of the three phonemes /i e ə/, depending on the word and sometimes the gender of the speaker. It is unclear whether this is a near merger, where speakers rely on cues to distinguish these sounds that are more subtle than the present experiment could discover, or a complete phonetic merger of these categories. Further research is needed to address this question.

This paper presents the first analysis of the stress pattern of this dialect. However, these stress rules should be tested on more constructions, to verify that they truly hold for all words. Additionally, stressed syllables were identified based on auditory impression of prominence, not using any strict acoustic measure. Therefore, my biases as a English and Spanish speaker may have influenced the results. Further research on the phonetic cues to stress in this dialect, as well as a careful study of vowel duration in different syllables, would be very useful.

Deletion results presented in this paper were able to show distinct environments in which deletion is observed for content and function words. However, there are many additional factors which could be investigated which could influence deletion, including register, speech rate, speaker variability and other phonotactic constraints. Deletion rules in longer words in which more than one vowel is a potential candidate for deletion are also unclear from the present data, as is the effect of adjacency to a stressed syllable and the question of peripheral vs. centralized vowels. Future research should investigate these questions in a wider range of morphological forms and with larger corpora.

Finally, there are additional questions about Chichicastenango K'iche' that have been mentioned in the literature but which were not addressed in this paper. One large question

relates to the reported neutralization of the realization of the peripheral/centralized contrast found outside of final (or possibly stressed) syllables. The deletion pattern discussed in Chapter 5 suggests that a historic neutralization of peripheral and centralized vowels outside of the final syllable did occur, but later changes resulted in the re-introduction of peripheral vowels in non-final syllables and therefore a contrast remains. However, there also appear to be some peripheral-sounding vowels in non-final syllables which correspond to short vowels in other dialects. Therefore, there is likely more to be said on this topic Due to the comparatively small number of non-deleted unstressed vowels, this question would rely on elicitation of specific types of words or a much larger corpus to be answered.

In conclusion, this paper provides new information on a variety of topics related to vowels in Chichicastenango K'iche'. However, there are many questions remaining which would be productive avenues for future research.

Appendices

Appendix A

Abbreviations

The following abbreviations are used in example glosses in this document:

1: first person

2: second person

3: third person

A: set A (ergative and possessive)

ACT: active

AFF: affect

AG: agentive

ANT: antipassive

B: set B (absolutive)

CAUS: causative

CAUS.POS: causative of a positional

CPL: completive aspect

DET: determiner

DIR: directional

EXIST: existential

FOC.LOC: focused locative

INCPL: incompletive aspect

IN.POS: inalienable possession

INSTR: instrumental

NEG: negation

P: plural

PASS: passive

PERF: perfect

PL: pluralizer

POS.V: positional verb

S: singular

SS: status suffix

TV: thematic vowel

The following abbreviations are used for different K'iche' dialects:

Chi: Chichicastenango dialect

Nah: Nahualá dialect

Appendix B

Data Sources

All original data cited in this document comes from Elizabeth Wood's fieldwork recordings in the towns of Chichicastenango and Nahualá. Data comes from either recordings of texts, most of which are available in the Archive of Indigenous Languages of Latin America (The K'iche' Collection of Elizabeth Wood), or recordings of unarchived elicitation sessions. Metadata for each of these soundfiles, along with the code used to reference it throughout the text, is listed here.

Code	Resource
elic;JT2	Elicitation session with Juana Tol (Speaker) and Elizabeth Wood (Researcher). Chichicastenango, July 11th, 2019. Unarchived.
elic;MACM1	Elicitation session with Mildred Alediz Clarivel Mejía (Speaker) and Elizabeth Wood (Researcher). Chichicastenango, June 13th, 2019. Unarchived.
elic;MACM2	Elicitation session with Mildred Alediz Clarivel Mejía (Speaker) and Elizabeth Wood (Researcher). Chichicastenango, June 19th, 2019. Unarchived.
elic;MCX3	Elicitation session with Manuel Chicoj Xirum (Speaker) and Elizabeth Wood (Researcher). Chichicastenango, June 29th, 2019. Unarchived.
elic;MCX5	Elicitation session with Manuel Chicoj Xirum (Speaker) and Elizabeth Wood (Researcher). Chichicastenango, July 12th, 2019. Unarchived.
elic;MCX8	Elicitation session with Manuel Chicoj Xirum (Speaker) and Elizabeth Wood (Researcher). Chichicastenango, July 12th, 2019. Unarchived.
elic;MXM2	Elicitation session with María Xirum Mejía (Speaker) and Elizabeth Wood (Researcher). Chichicastenango, July 2nd, 2019. Unarchived.
elic;SAGB1	Elicitation session with Sonia Angélica González Bocel (Speaker) and Elizabeth Wood (Researcher). Chichicastenango, June 17th, 2019. Unarchived.
elic;TJL1	Elicitation session with Teresa Jerónimo Leom (Speaker) and Elizabeth Wood (Researcher). Chichicastenango, June 17th, 2019. Unarchived.
elic;TJL2	Elicitation session with Teresa Jerónimo Leom (Speaker) and Elizabeth Wood (Researcher). Chichicastenango, June 24th, 2019. Unarchived.

Table B.1: Data sources referenced in this document: elicitation sessions

Code	Resource
txt;3recipes	Tol, Juana (Speaker) and Elizabeth Wood (Researcher). 2019. Text: Cooking (3 recipes). The K'iche' Collection of Elizabeth Wood. The Archive of the Indigenous Languages of Latin America, ailla.utexas.org. Access: public. PID ailla:271541.
txt;3stories	Recording of Manuel de Jesús Tahay Gómez (Speaker), who tells three stories/jokes, and Elizabeth Wood (Researcher). Nahualá, July 19th, 2018. Unarchived.
txt;caldores	Sen Ixtuc, Rafaela (Speaker) and Elizabeth Wood (researcher). 2019. Text: Beef soup recipe. The K'iche' Collection of Elizabeth Wood. The Archive of the Indigenous Languages of Latin America, ailla.utexas.org. Access: public. PID ailla:271534.
txt;church	Jerónimo Leom, Magdalena (Speaker) and Elizabeth Wood (Researcher). 2019. Text: History of Chij Tinamit Church of Christ. The K'iche' Collection of Elizabeth Wood. The Archive of the Indigenous Languages of Latin America, ailla.utexas.org. Access: public. PID ailla:271549.
txt;healing	Mejía, Mildred Alediz Clarivel (Speaker) and Elizabeth Wood (Researcher). 2019. Text: Medicinal plants. The K'iche' Collection of Elizabeth Wood. The Archive of the Indigenous Languages of Latin America, ailla.utexas.org. Access: public. PID ailla:271557.
txt;kot	Chicoj Xirum, Manuel (Speaker) and Elizabeth Wood (Researcher). 2019. Text: Legend of the double-headed eagle. The K'iche' Collection of Elizabeth Wood. The Archive of the Indigenous Languages of Latin America, ailla.utexas.org. Access: public. PID ailla:271553.
txt;owl	Sen Ixtuc, Rafaela (Speaker) and Elizabeth Wood (Researcher). 2019. Text: A story about a man who pretends to be an owl. The K'iche' Collection of Elizabeth Wood. The Archive of the Indigenous Languages of Latin America, ailla.utexas.org. Access: public. PID ailla:271526.
txt;planting	Chicoj Xirum, Manuel (Speaker) and Elizabeth Wood (Researcher). 2019. Text: Planting. The K'iche' Collection of Elizabeth Wood. The Archive of the Indigenous Languages of Latin America, ailla.utexas.org. Access: public. PID ailla:271561.
txt;talentos	Jerónimo Leom, Teresa (Speaker) and Elizabeth Wood (Researcher). 2019. Text: Health Talents International. The K'iche' Collection of Elizabeth Wood. The Archive of the Indigenous Languages of Latin America, ailla.utexas.org. Access: public. PID ailla:271545.

Table B.2: Data sources referenced in this document: texts

References

Ajpacajá Tum, Pedro Florentino. 2001. *K'ichee' Choltziij*. Guatemala: Editorial Cholsamaj.

Ajpacajá Tum, Pedro Florentino, Manuel Isidro Chox Tum, Francisco Lucas Tepaz Raxulew, and Diego Adrian Guarchaj Ajtzalam. 2005. *Diccionario K'iche'*. Antigua Guatemala: Proyecto Lingüístico Francisco Marroquín, Editorial Cholsamaj.

Baird, Brandon O. 2010. Un análisis acústico de las vocales de K'ichee'. In: *Conference on indigenous languages of Latin America (CILLA) IV, Austin, Texas: The Center for Indigenous Languages of Latin America (CILLA) at the University of Texas at Austin*.

Baird, Brandon O. 2014. Acoustic correlates of stress in K'ichee': a preliminary analysis. In: Proceedings of Formal Approaches to Mayan Linguistics (FAMLi) II.

Baird, Brandon O. 2018. Dialectal evolution of the vowel systems of K'ichee'. In: *Production and Perception Mechanisms of Sound Change*. Ed. by Daniel Recasens and Fernando Sánchez-Miret. Lincom GmbH.

Bennett, David C. 1968. Spectral form and duration as cues in the recognition of English and German vowels. *Language and Speech* 11.2, pp. 65–85.

Bennett, Ryan. 2016a. La tensión vocálica en el kaqchikel de Sololá, Guatemala: un estudio preliminar. *Las Actas del Seminario Phonologica*.

Bennett, Ryan. 2016b. Mayan Phonology. *Language and Linguistics Compass*, pp. 469–514.

Berinstein, Ava E. 1979. A cross-linguistic study on the perception and production of stress. MA thesis. UCLA, Linguistics.

Boersma, Paul and David Weenik. 2019. *Praat: doing phonetics by computer*. [Computer program]. Version 6.0.49, retrieved from http://www.praat.org/.

Campbell, Lyle. 1977. *Quichean linguistic prehistory*. Berkeley: University of California Press.

Campbell, Lyle. 2017. Mayan history and comparison. In: *The Mayan Languages*. Ed. by Judith Aissen, Nora C. England, and Roberto Zavala Maldonado. 1st. London: Routledge Ltd. Chap. 3, pp. 43–61.

Can Pixabaj, Telma A. 2017. K'iche'. In: *The Mayan Languages*. Ed. by Judith Aissen, Nora C. England, and Roberto Zavala Maldonado. 1st. London: Routledge Ltd. Chap. 18, pp. 461–499.

Can Pixabaj, Telma A. 2019. *Correlatos fono-sintácticos en las oraciones de complemento en k'iche'*. Paper presented at Simposio Maya' Chii', Antigua Guatemala, Guatemala.

Christenson, Allen J. N.d. K'iche'-English Dictionary and Guide to Pronunciation of the K'iche'-Maya Alphabet. Unpublished manuscript, Bringham Young University.

Chumakina, Marina, Oliver Bond, and Greville G Corbett. 2016. Essentials of Archi grammar.

Cojtí Ren, Iyaxel Ixkan. 2019. The Emergence of the Ancient Kaqchikel Polity: A case of Ethnogenesis in the Guatemalan Highlands. PhD thesis. Vanderbilt University.

Collins, Beverley and Inger M Mees. 2013. *Practical phonetics and phonology: A resource book for students*. Routledge.

Dalton, Will. 2011. An ultrasound imaging study of the tense-lax distinction in Canadian French vowels. *Canadian Acoustics* 39.3, pp. 168–169.

Di Paolo, Marianna and Alice Faber. 1990. Phonation differences and the phonetic content of the tense-lax contrast in Utah English. *Language Variation and Change* 2.2, pp. 155–204.

England, Nora C. and Brandon O. Baird. 2017. Phonology and phonetics. In: *The Mayan Languages*. Ed. by Judith Aissen, Nora C. England, and Roberto Zavala Maldonado. 1st. London: Routledge Ltd. Chap. 7, pp. 175–200.

Evans, Jonathan P, Jackson T-S Sun, Chenhao Chiu, and Michelle Liou. 2016. Uvular approximation as an articulatory vowel feature. *Journal of the International Phonetic Association* 46.1, pp. 1–31.

Fulop, Sean A., Ethelbert Kari, and Peter Ladefoged. 1998. An acoustic study of the tongue root contrast in Degema vowels. *Phonetica* 55.1-2, pp. 80–98.

Goedemans, Rob and Harry van der Hulst. 2013. Weight Factors in Weight-Sensitive Stress Systems. In: *The World Atlas of Language Structures Online*. Ed. by Matthew S. Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology. https://wals.info/chapter/16.

Gordon, Matthew, Paul Barthmaier, and Kathy Sands. 2002. A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association* 32.2, pp. 141–174.

Guy, J.B. 1974. *A grammar of the northern dialect of Sakao*. Canberra: Department of Linguistics, Research School of Pacific Studies, The Australian National University.

Holliday, Nicole and Sean Martin. 2018. Vowel categories and allophonic lowering among Bolivian Quechua–Spanish bilinguals. *Journal of the International Phonetic Association* 48.2, pp. 199–222.

Kaufman, Terrence and John S. Justeson. 2002. *A preliminary Mayan etymological dictionary*. Available online at http://www.famsi.org/reports/01051/. Foundation for the Advancement of Mesoamerican Studies.

Kenstowicz, Michael. 1997. Quality-sensitive stress. *Rivista di linguistica* 9, pp. 157–188.

Larsen, Thomas Walter. 1988. Manifestations of Ergativity in Quiché Grammar. PhD thesis. University of California, Berkeley.

Lindblom, Björn. 1968. Vowel duration and a model of lip mandible coordination. *Speech Transmission Laboratory Quarterly Progress Status Report* 4, pp. 1–29.

López Ixcoy, Candelaria Dominga. 1997. *Gramática K'ichee'*. Guatemala: Fundación Cholsamaj.

López Ixcoy, Saqijix Candelaria Dominga. 1994. *Las vocales en K'ichee'*. Guatemala: OKMA.

McCarthy, John J and Abigail Cohn. 1998. Alignment and parallelism in Indonesian phonology. Working Papers of the Cornell Phonetics Laboratory 12. 6.

Par Sapón, María Beatriz and Telma Angelina Can Pixabaj. 2000. *Ujunamaxiik ri K'ichee' Ch'ab'al: Variación Dialectal en K'ichee'*. Guatemala: OKMA, Cholsamaj.

Peterson, Gordon E and Harold L Barney. 1952. Control methods used in a study of the vowels. *The Journal of the acoustical society of America* 24.2, pp. 175–184.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/.

Richards, Michael. 2003. *Atlas lingüístico de Guatemala*. Editorial Serviprensa Guatemala.

Romero, Sergio. 2016. Bill Gates speaks K'ichee'! The corporatization of linguistic revitalization in Guatemala. *Language & Communication* 47, pp. 154–166.

Romero, Sergio. 2017. The labyrinth of diversity: The sociolinguistics of Mayan languages. In: *The Mayan Languages*. Ed. by Judith Aissen, Nora C. England, and Roberto Zavala Maldonado. 1st. London: Routledge Ltd. Chap. 15, pp. 379–400.

Velleman, Leah Bridges. 2014. Focus and movement in a variety of K'ichee'. PhD thesis. The University of Texas at Austin.

Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN: 978-3-319-24277-4. https://ggplot2.tidyverse.org.

Wilson, Ian. 2007. The Effects of Post-Velar Consonants on Vowels in Nuu-chah-nulth: Auditory, Acoustic, and Articulatory Evidence. *Canadian Journal of Linguistics* 52.1-2, pp. 243–70.

Wood, Elizabeth. *The K'iche' Collection of Elizabeth Wood*. The Archive of Indigenous Languages of Latin America, ailla.utexas.org. Access: public. PID ailla: 271517.

Zawaydeh, Bushra Adnan. 1997. An acoustic analysis of uvularization spread in Ammani-Jordanian Arabic. *Studies in the Linguistic Sciences* 27.1.