Journal of the International Phonetic Association (2025), **55**, pp. 465–502
doi:10.1017/S0025100325100765

ARTICLE

Predictable vowel deletion in content words and variable vowel deletion in function words in Chichicastenango K'iche'

Elizabeth Anne Wood

The University of Texas at Austin, United States
Email: elizabethwood@utexas.edu

(Received 12 July 2024; revised 24 June 2025; accepted 4 July 2025; first published online 3 December 2025)

Abstract

Vowel deletion is frequent in the Chichicastenango dialect of K'iche' (Maya). Whereas deletion in content words is reportedly predictable based on vowel quality, syllable structure and stress, deletion in function words is much more variable. This article investigates vowel deletion in a corpus of spontaneous, monologic speech. The results show that deletion in content words is highly regular, occurring to lax vowels in unstressed, CV syllables adjacent to the stressed syllable. A difference can be observed between vowels belonging to stress domain internal morphemes and extrametrical morphemes. Deletion in extrametrical morphemes is somewhat less regular, and does not occur in word-final syllables. In function words, vowel deletion is sensitive to similar conditions to those that affect content words, but is highly variable and is influenced by the phrase-level context.

1 Introduction

Vowel deletion is common in the Mayan language family, attested in languages such as Mam, Tektitek, Ixil, Uspantek, Tz'utujil, Sakapultek, Tseltal, Yucatec Maya, Tojolab'al, Mocho', Huastec, and K'iche' (Bennett 2016; Sapón et al., 2000). In these languages, deletion is typically restricted to short vowels, and is especially common for vowels that are unstressed, non-initial, and non-final (Bennett 2016; López Ixcoy 1994; Sapón et al., 2000). Vowel deletion is particularly prevalent in the Chichicastenango dialect of K'iche', where it has a marked effect on both production and comprehension of the language. Building on the observations of Campbell (1977), López Ixcoy (1994) and England & Baird (2017), Wood (2020) describes deletion in content words as predictable based on vowel quality, syllable shape, syllable position, and stress, and unaffected by phrasal context. In function words, in contrast, deletion is practically unstudied and appears to be highly variable.

The Chichicastenango dialect is a stigmatized dialect of K'iche'. Speakers are typically aware of how people from other towns speak the language and often feel pressured to use these prescriptively correct forms when asked how their language is spoken. Therefore, a complete understanding of where vowels are deleted in Chichicastenango K'iche' requires the use of data from speakers focused on the content rather than the form of their words. This article presents the vowel deletion patterns observed in content words and function

© The Author(s), 2025. Published by Cambridge University Press on behalf of The International Phonetic Association. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

words in a corpus of Chichicastenango K'iche' spontaneous speech. The study provides support for many of the generalizations made in the previous literature, including the effects of vowel quality, onset, coda, and stress, and the observation that deletion in function words is more variable than in content words and is affected by phrasal context. However, the results also provide some novel information about the language, showing that vowels are deleted in some apparently final syllables and demonstrating differences in the deletion pattern at three levels of prosodic structure (stress domain of content words, extrametrical morphemes in content words, and function words).

The structure of the paper is as follows. Section 2 provides background on the language, its sounds, and past research on vowel deletion. Section 3 details the methods used for the corpus study. Section 4 describes the results. Section 5 discusses some of the main results of the study, including the differences observed between content words and function words and how this data informs understanding of prosodic structure in the language. Section 6 concludes the article.

2 Background

K'iche' (ISO639-3 QUC) is a Mayan language belonging to the Eastern branch of the family spoken by around one million people in the highland area of Guatemala (Campbell 2017; Instituto Nacional de Estadística 2019). The language is characterized by a relatively high degree of dialectal variation at various levels of the grammar (Sapón et al., 2000; Romero 2009, 2017).

An official K'iche' orthography was standardized by the Academia de Lenguas Mayas de Guatemala in 1987, by Decree 1046-87 (López Ixcoy 1997). This orthography indicates tense vowels with plain vowel symbols and lax vowels with diereses, and is therefore appropriate for Chichicastenango K'iche', which has this phonological contrast, as discussed further below (Wood 2020). Most other works on K'iche' use a modified orthography more suited to the vowel length contrast found in most other K'iche' dialects.

The consonant inventory of K'iche' is shown in Table 1. Where different from the IPA, the orthographic symbol is included in < >.

Stops and affricates contrast a plain series with a series referred to as 'glottalized' in the Mayanist literature (Bennett 2016). In Chichicastenango K'iche', like other Mayan languages, most of the glottalized consonants are ejective, but the bilabial is usually realized as a voiced implosive. The bilabial and uvular glottalized stops are often reduced to a glottal

Table 1. Consonant inventory of K'iche' with orthographic symbols in < >

	Bilabial	Alveolar	Postalveolar	Palatal	Velar	Uvular	Glottal
Plain stop	p	t			k	q	ʔ <'>
Glottalized stop	ɓ <b'>	t'			k'	q'	
Plain affricate		t͡s <tz>	t͡ʃ <ch>				
Glottalized affricate		t͡s' <tz'>	t͡ʃ' <ch'>				
Fricative		s	ʃ <x>			χ <j>	
Nasal	m	n					
Approximant	w	ɾ <r>		j <y>			
Lateral		l					

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

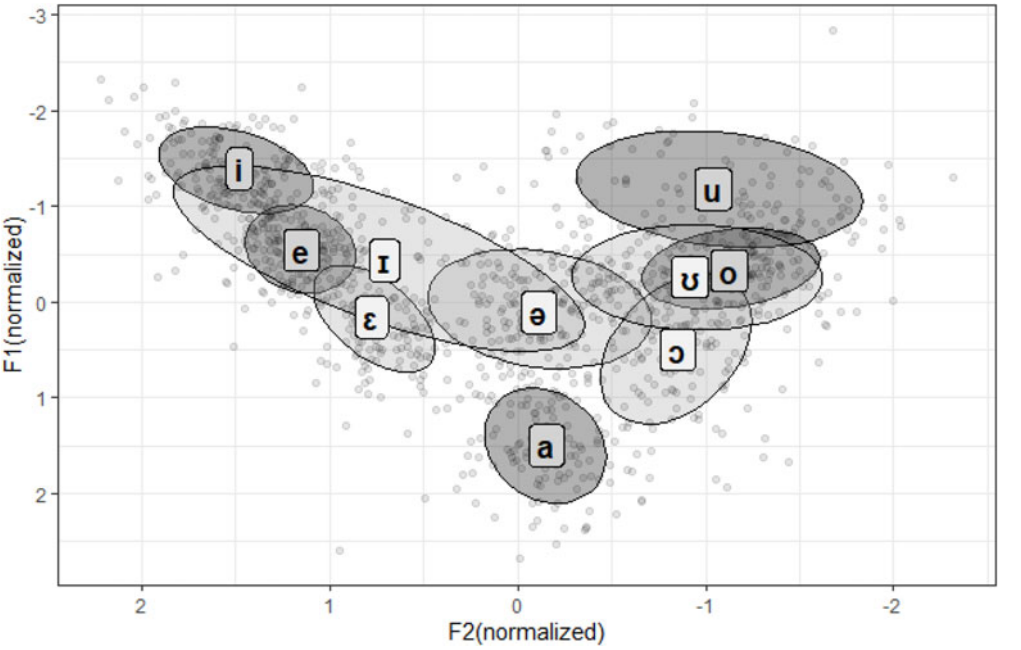


Figure 1. Realizations of the 10 vowel phonemes in Chichicastenango K'iche' in a controlled speech production task (Wood, 2025a). Tense vowels are in dark gray and lax vowels in light gray. Ellipses enclose about one standard deviation.

stop in casual speech. Whatever their realization, glottalized consonants cause glottalized phonation on adjacent vowels (Wood 2023).

Most K'iche' dialects have vowels at five places of articulation /i e u o a/, with a phonemic length contrast (Sapón et al., 2000). In Chichicastenango K'iche', however, the length contrast has been replaced by a contrast in quality, with historic long vowels appearing as 'tense' /i e u o a/ and historic short vowels as 'lax' /ɪ ɛ ʊ ɔ ə/(<ï ë ü ö ä>) (Wood, 2020). The vowels of Chichicastenango K'iche' are shown in Figure 1.

This graph shows the realizations of the tense and lax vowel phonemes of Chichicastenango K'iche' as produced by eleven speakers (three male, eight female) in a controlled speech production task (Wood, 2025a). The data was collected through a translation task from Spanish and it includes only stressed monosyllabic words. The formant values were normalized with the Lobanov method (z-scores, Lobanov 1971) to permit comparisons across speakers. The large size of the /ɪ/ ellipse is due to the fact that this phoneme is realized as [i], [e], [ɛ], or [ə], depending on the item and sometimes the speaker. This vowel appears to be merging phonetically with surrounding categories; however, it retains distinct phonological behavior and patterns as a lax vowel irrespective of its phonetic realization.

Primary stress is word-final in non-verbs in Chichicastenango K'iche', with the exclusion of two extrametrical suffixes: the superlative /-əlәχ/ and the suffix /-ɪk/ which appears on some adjectives when in phrase-final position (for more information on phrase-final suffixes in K'iche', see Henderson 2012; Royer 2021; Wood 2025b). Stress is non-final in some unadapted Spanish loanwords, which maintain the position of stress in the original Spanish and are common in everyday speech. Examples of stress in non-verbs are shown in Table 2. Here, and throughout the article, angle brackets < > enclose extrametrical morphemes. The audio files for these and all other examples cited throughout the paper are included in the

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

Table 2. Stress in non-verbs, with extrametrical morphemes in angle brackets < >

	Phonemic	Syllabification and stress	Surface	Gloss	Source
Native K'iche' words	/tikɔʔn/	ti.ˈkɔʔn	[tikɔʔn]	'crops'	history, 7:42
	/naβe/	na.ˈβe	[naβe]	'first'	3recipes, 4:38
	/səqwət͡ʃ/	səq.ˈwət͡ʃ	[səqwət͡ʃ]	'potato'	chilmol, 0:53
	/ɪʃɔqɪβ/	ɪ.ʃɔ.ˈqɪβ	[ɪʃqɪβ]	'women'	history, 5:36
	/nɪm<əlәχ>/	ˈnɪ.mə.ləχ	[nɪmləχ]	'very big'	history, 1:23
	/χɛʔl<ɪk>/	ˈχɛʔ.lɪk	[χɛʔlɪk]	'delicious'	fishing, 9:33
Adapted Spanish loans	/wakəʃ/	wa.ˈkəʃ	[wakəʃ]	'cow' (Spanish /ˈba.ka/)	caldores, 1:26
	/kintala/	kin.ta.ˈla	[kintala]	'candle' (Spanish /kan.ˈde.la/)	history, 2:26
	/kalto/	kal.ˈto	[kalto]	'broth' (Spanish /ˈkal.do/)	caldores, 0:04
Unadapted Spanish loans/code-switching	/berduras/	ber.ˈdu.ras	[berduras]	'vegetables' (Spanish /ber.ˈdu.ras/)	3recipes, 4:09
	/klase/	ˈkla.se	[klase]	'class' (Spanish /ˈkla.se/)	3recipes, 4:31
	/sabado/	ˈsa.ba.do	[sabado]	'Saturday' (Spanish /ˈsa.ba.do/)	semanasanta, 2:53

supplementary materials. Each example also appears with a code to the source recording – many of which are publicly available in the Archive of the Indigenous Languages of Latin America (K'iche' collection of Elizabeth Wood) – and timestamp. More details about the data sources can be found in the Appendix.

In verbs, stress is sensitive to vowel quality and syllable shape. Stress falls on the left-most syllable containing either a tense vowel or a coda, or if there are none, on the final syllable. The stress domain excludes most inflectional morphemes (person, aspect, and incorporated movement prefixes, and status suffixes, a set of suffixes which indicate the mood, transitivity, and phrase position of the verb). Examples of stress in verbs are shown in Table 3.

Campbell (1977) observes that all short (i.e., lax) vowels in Chichicastenango K'iche' are deleted when preceded by a consonant and followed by a stressed syllable, and provides a number of examples, all of which are disyllabic nouns. López Ixcoy (1994) notes that deletion is common but tends to affect non-final vowels, and England & Baird (2017) that it affects many unstressed vowels. Wood (2020) brings these descriptions together with original data taken from monologues and elicited materials and concludes that lax vowels are deleted in content words when in a non-final CV syllable adjacent to the stressed syllable. Tense vowels, stressed vowels, and those in onsetless, closed, or word-final syllables are never deleted. Lax vowels in unstressed, non-final CV syllables not adjacent to the stressed syllable are sometimes deleted, but less regularly than those in syllables adjacent to the stressed syllable. Very little previous research exists on the deletion of vowels in function

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

Table 3. Stress in verbs, with extrametrical morphemes in angle brackets < >

	Phonemic	Syllabification and stress	Surface	Gloss	Source
Verbs	/<ʃu>t͡ʃoməχ/	ʃu.ˈt͡ʃo.məχ	[ʃut͡ʃoməχ]	'he thought'	owl, 0:28
	/<kaw>esəχ/	ka.ˈwe.səχ	[kawesəχ]	'you remove it'	healing, 4:32
	/<kə>ɓəntəχ<ɪk>/	kə.ˈɓən.tə.χɪk	[kəɓəntχɪk]	'it is done'	history, 8:20
	/<kə>pɔq'ɔw<ɪk>/	kə.pɔ.ˈq'ɔ.wɪk	[kəpq'ɔwɪk]	'it boils'	caldores, 0:58
	/<ku>k'ʊtu/	ku.k'ʊ.ˈtu	[kuk'tu]	'he/she goes to show it'	history, 9:47

words in Chichicastenango K'iche'. Wood (2020) observes that vowels may be deleted in function words in contexts that would be prevented in content words, such as word-final or closed syllables. These deletions sometimes appear to be regular, but in other cases appear to be optional, with the same vowel appearing variably in the same word and similar phrasal contexts.

This article addresses vowel deletion in Chichicastenango K'iche' in a corpus of spontaneous speech. The primary research question is what conditions vowel deletion in the language. Specifically, the study explores (i) whether the previously identified factors of vowel quality, syllable structure, syllable position and stress are relevant, (ii) how predictable vowel deletion is according to these factors, and (iii) whether there are differences between content words and function words.

3 Methods

3.1 Data

The data used for this study comes from a corpus of spontaneous, monologic speech recorded in Chichicastenango in 2018–2019. The corpus includes a total of 2 hours and 38 minutes of audio by 11 native speakers. The corpus includes speakers of a range of ages (young adults to elders), genders (two men and nine women), and locations within the Chichicastenango area (city center and surrounding rural communities), and thus constitutes a broad sample of speakers from this dialect area. The monologues cover topics such as local history and traditions, traditional stories, personal anecdotes, instructions, and recipes. Speakers were asked in very general terms to talk about a particular topic, such as explaining how to prepare a local dish or relating something that happened to them in the past, and were allowed to speak freely for as long as they wished, which ranged from 1 to 15 minutes. These recordings were then transcribed by the author with help from native speakers. The recordings were made using a Zoom Hn4 digital recorder at a sampling rate of 44.1 kHz, using either the internal microphone or (in most cases) connected to a Shure SM10A headset microphone. Recordings for which permission was granted by the speaker are publicly available in the Archive of the Indigenous Languages of Latin America (K'iche' Collection of Elizabeth Wood) along with transcriptions and translations. The unarchived recordings are used for academic purposes with permission from the speakers.

For each morpheme, the underlying form was determined primarily based on comparison of any surface forms produced in the corpus as well as other Chichicastenango K'iche' materials collected by the author. If a vowel occurred in at least one instance of that morpheme, it was included as an underlying vowel in all instances of that morpheme. Additionally, vowels known to be present in the historical form of the language and

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

preserved in other K'iche' dialects where deletion is much less prevalent (Larsen 1988; Ajpacajá Tum 2001; Ajpacajá Tum et al. 2005; Pixabaj & Angelina, 2015) were assumed to be present underlyingly in Chichicastenango K'iche'. These two systems for identifying underlying vowels were consistent with each other.

For the purposes of this study, content words include nouns, verbs, adjectives, and adverbs. All other word types are considered to be function words, including prepositions, complementizers, determiners, pronouns, directional particles, tense–aspect–mood particles, negation markers, existential particles, and nominal modifiers. Relational nouns – a set of words in Mayan languages that are formally nominal, receiving nominal morphology such as possessive prefixes, but have functional uses (Polian 2017) – were categorized as content words. Relational nouns express meanings such as location, time, purpose, cause, and other circumstantial meanings. Temporal and locative relational nouns combine with prepositions, while others do not. Two examples of relational nouns are shown in (1)–(2).

(1) *Kärïqtäjïk lö* ***kümal.***

krɪqtɪk	lɔ	**kmal**
kə-**Ø**-rɪq-təχ-ɪk	lɔ	**k-ʊmal**
INCPL-B.3SG-find-PASSC-SS.F	DIR	A.3PL-because

'It is found by them.' (mushrooms, 2:54)

(2) *Kïnch'äw* ***puwï'*** *ju oj.*

kɪnt͡ʃˀəw	**puwɪʔ**	χu	ʔoχ
k-ɪn-t͡ʃˀəw-**Ø**	**p-u-wɪʔ**	χu	oχ
INCPL-B.1SG-talk-SS.M	PREP-A.3SG-top	DET	avocado

'I'll call out from on top of an avocado tree.' (owl, 0:35)

(1) shows the relational noun *-ümal* 'because of', which appears with the 3rd person plural possessive prefix *k-*. (17) shows the locative relational noun *-wï'* 'on top of' (literally, 'hair'), which combines with the preposition *pä* and the 3rd person singular possessive prefix *u-*.

Phonologically unadapted Spanish borrowings, whether content words (as in Table 2) or function words (e.g. *si* [si] 'if', *porque* [ˈpoɾ.ke] 'because'), were excluded from the study. It was observed during the course of the work on the language that deletion does not tend to occur in these words, which may reflect the form of the word in the original Spanish rather than the vowel deletion pattern of Chichicastenango K'iche'. Words that are composed of only a single vowel in the phonemic form, such as the plural marker *e* /e/, were also excluded, as it is not possible to know if such a word is present if its sole vowel is deleted. Instances where the word is repeated multiple times as a speech error or instances where the speaker doesn't complete a grammatical sentence and instead immediately restarts with a new sentence after the word in question were also excluded. Finally, in some cases there were gaps or uncertainties in the transcriptions, and no words from these portions were included to avoid categorization errors.

Due to time considerations and the greater number of vowels in content words as compared to function words, the content word data was taken from a subset of the corpus comprised of ten recordings produced by eight different speakers, adding up to one hour. This data includes 6967 underlying vowels (present or deleted) in 3411 words, including 891 distinct lexical items.

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

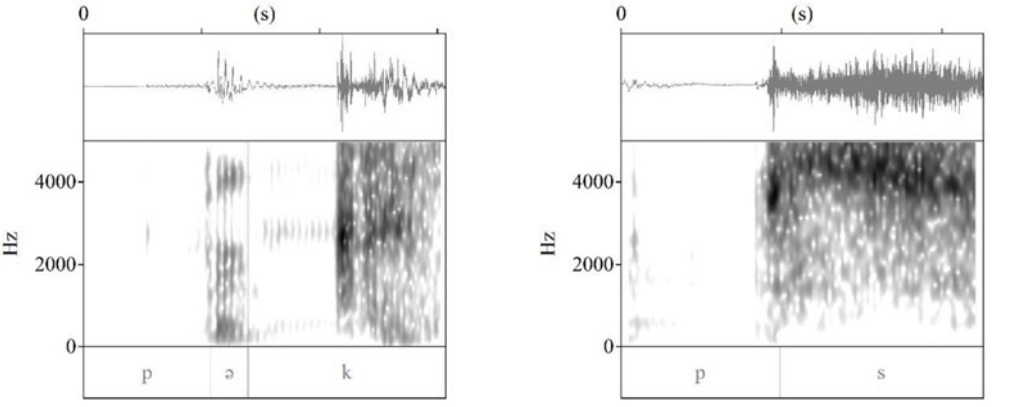


Figure 2. Present vs. deleted vowels between voiceless segments. Left: [pə k], right: [p s].

The function word data was taken from the entire corpus. Words with less than ten tokens were excluded. For words with more than 500 tokens, a randomized subset was included: every third instance of the determiner *rï* /rɪ/ (1728 total tokens, 576 included) and every other instance of the pronoun *rï'* /rɪʔ/ (1072 total tokens, 536 included). For the remaining function words, every instance of the word present in the corpus was included in the study. This led to a total of 6995 function words included in the study with 7480 underlying vowels.

The function words that were included in the study are summarized in Table 4. 'Phrase-final' refers to the form of a word found only at the end of a phrase, and 'phrase-medial' to the form of the word found elsewhere (see Henderson 2012 for more information on function words which vary by phrase position in K'iche'). 'Pre-C' refers to the form of a word found before a word beginning with a consonant and 'pre-V' to the form of the word found before a word beginning with a vowel.

3.2 Categorization

Each potential vowel was classified according to VOWEL PRESENCE (dependent variable) as well as the factors VOWEL QUALITY, CODA, ONSET, STRESS POSITION, SYLLABLE POSITION, DOMAIN, ITEM and SPEAKER (independent variables).

Vowel presence: Depending on the segmental context, the following visual metrics were used in Praat (Boersma & Weenink 2023) to determine if a vowel was present or deleted. For underlying vowels between voiceless segments, any indication of voicing (voicing bar, periodicity) was sufficient to categorize a vowel as present. Examples are shown in Figure 2.

On the left of Figure 2 is shown the preposition *pä* /pə/ and the first consonant of the following word *külew* /kʊlew/ 'their land' in the phrase *pä külew* 'on their land' (planting, 2:46). There is clear voicing after the release of the /p/ before the closure for the /k/, and the vowel in *pä* was categorized as present. On the right of the figure is shown the same preposition *pä* /pə/ followed by the first consonant of the word *sü't* /sʊʔt/ 'cloth' in the phrase *pä sü't* 'in cloths' (planting, 13:04). There is no voicing between the /p/ and the following /s/. The frication immediately follows the stop burst. In this case, the underlying vowel in *pä* was categorized as deleted.

For underlying vowels adjacent to a voiced consonant (implosive, nasal, glide or liquid), a vowel was determined to be present or deleted based on the number of segments visible in the voiced portion. The vowel was determined to be present if the voiced portion showed

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

Table 4. Function words included in the study

Category	Word	Gloss	N
Prepositions	t͡ʃɪ	general preposition	164
	pə	'in, at'	249
Complementizers	t͡ʃɪ	'that'	31
	t͡ʃɛr	'that, which'	123
	we	'if'	35
	la	polar question	29
Determiners	χu	indefinite article (pre-C)	287
	χʊn	indefinite article (pre-V)	118
	wa	proximal demonstrative	77
	le	medial demonstrative	184
	rɪ	distal demonstrative	576
Pronouns	ɪn	1SG	91
	ɔχ	1PL	150
	at	2SG	27
	arɛʔ	3SG/PL	99
	χun	'one'	87
	waʔ	proximal demonstrative	140
	laʔ	medial demonstrative	102
	leʔ	medial demonstrative	87
	rɪʔ	distal demonstrative	536
	t͡ʃɪ	'where'	29
	t͡ʃɪʔ	end of where question	28
	su	'what'	138
	tan	component of 'why'	21
	kaʔj	component of 'what' (phrase-medial)	43
	χʊmpa	'when, how much/many'	22
Directionals	ɓɪ	'here to there' (phrase-medial)	144
	ɓik	'here to there' (phrase-final)	85
	kə	'in place' (phrase-medial)	109
	kənɔq	'in place' (phrase-final)	61
	kɔq	'inwards' (phrase-final)	17

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

Table 4. Continued

Category	Word	Gloss	N
	lɔ	'there to here' (phrase-medial)	141
	lɔq	'there to here' (phrase-final)	134
	pə	'over/across' (phrase-medial)	10
Tense-aspect-mood, negation and existential	t͡ʃɪ	'already, now' (phrase-medial)	252
	t͡ʃɪk	'already, now' (phrase-final)	133
	na	'still, necessary, future'	138
	nə	negation	173
	tə	irrealis (phrase-medial)	201
	təχ	irrealis (phrase-final)	76
	k'ɔ	existential (phrase-medial)	469
	k'ɔlɪk	existential (phrase-final)	66
Discourse, syntax and information structure	are	focus	237
	ɓaʔ	'then'	14
	χeʔ	'like, as'	237
	k'ʊ	'then' (phrase-medial)	37
	tɛʔ	'then'	94
	qəs	'truly'	10
	wɪ	trace of focused prepositional phrase	95
	ʃa	'only'	111
	ʃəq	'only'	85
	ʃew	'only'	34
Nominal modifiers	t͡ʃi	'truly'	38
	ma	augmentative (pre-C)	30
	sɪ	diminutive (pre-C)	112
	sɪn	diminutive (pre-V)	28
	təq	plural/distributive	151
TOTAL			6, 995

a relatively abrupt change in formants or intensity marking a boundary between two segments. If there was only one segment in the voiced portion, the vowel was categorized as deleted. Examples are shown in Figure 3.

On the left of Figure 3 is the word *na* /na/ 'still' followed by first consonant /k/ of *komo* /komo/ 'since' in the speech fragment *köjiwï'j na komo* /kɔχiwɪʔχ na komo/ 'you wait for

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

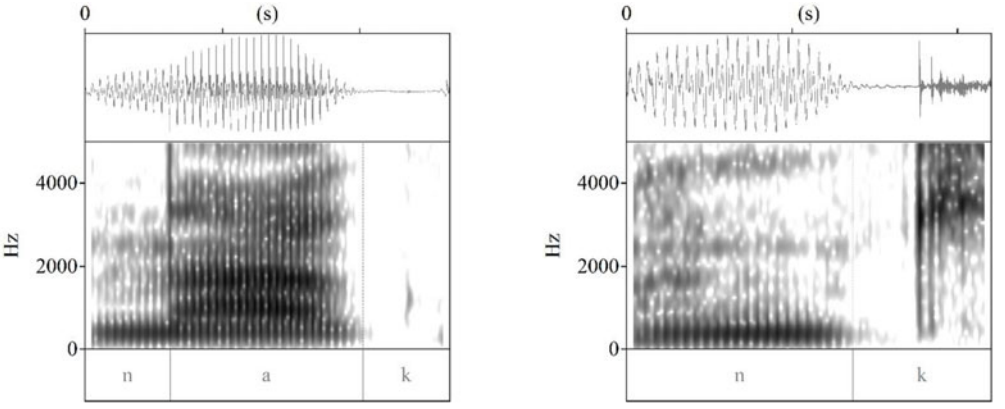


Figure 3. Present vs. deleted vowels after a voiced segment. Left: [na k], right: [n k].

us still since. . .' (marriage 1:31). There is a sharp boundary in both the formant structure and the intensity curve between the initial nasal [n] and the following vowel [a]. This vowel was categorized as present. On the right of the figure is the same word *na* followed by the first consonant [k] of *kib'* /kiɓ/ 'themselves' in the phrase *kikï'j na kib'* /kikɪʔχ na kiɓ/ 'they still wait for each other' (marriage, 4:13). Here the entire voiced portion forms one segment with no sharp boundaries between them. The whole portion is of low intensity and displays antiformants. The underlying vowel in this word was categorized as deleted in this token.

For underlying vowels adjacent to another vowel, a vowel was determined to be present if the vowel portion showed a change in the formants corresponding to the two expected vowels, an abrupt shift in intensity marking a boundary between two vowels, or evidence of glottalization in the middle of the vowel portion marking a boundary between two vowels. When only one vowel segment was present, which of the vowels was deleted was determined based on the formants. When the two underling vowels were identical in quality and there was only one vowel segment produced, the first underlying vowel was categorized as deleted and the second as present, since it is always the first vowel that is deleted when there is a change in quality.

Examples are shown in Figure 4.

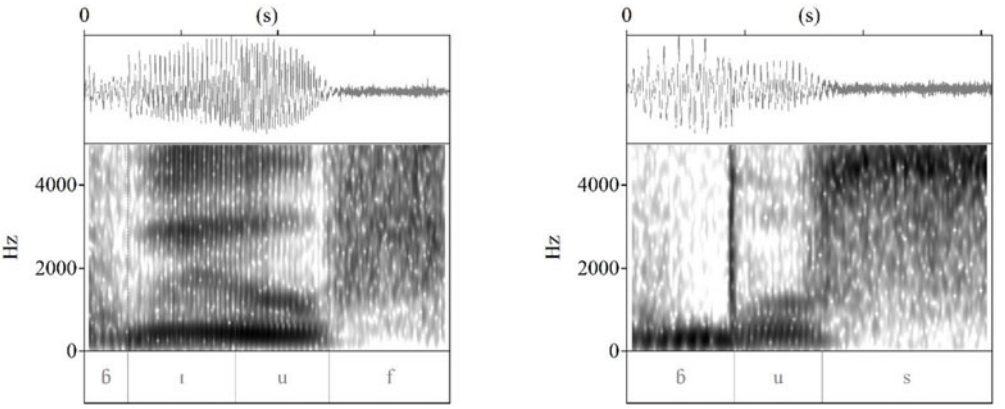


Figure 4. Present vs. deleted vowel preceding another vowel. Left: [ɓɪ uf], right: [ɓ us].

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

On the left is shown the directional particle *b'ï* /ɓɪ/ followed by the first two segments of the word *ufideo* /ufideo/ 'its noodles' in the phrase *käqäya chï b'ï ufideo* /kəqəja t͡ʃɪ ɓɪ ufideo/ 'we add its noodles' (caldores 0:44). During the vowel portion, F2 clearly shifts from a high position to a lower position (front vowel to back vowel). The vowel /ɪ/ in the directional was categorized as present in this token. On the right is shown the same directional particle followed by the first two segments of the word *usopa* /usopa/ 'its broth' in the phrase *käqäya b' ï usopa* /kəqəja ɓɪ usopa/ 'we add its broth' (caldores 0:47). Here the vowel portion forms one segment with no shift in formants or intensity. F2 is low, indicating a back vowel. The underlying vowel in the directional particle was categorized as deleted in this token.

In total there were 11499 present vowels and 2948 deleted vowels in the data.

Vowel quality: Each underlying vowel was categorized as either tense (/a e i o u/) or lax (/ə ɛ ɪ ɔ ʊ/). In most cases the phonemic identity of the vowel in a given morpheme is easily determined perceptually. However, for function words this was not always as trivial task, as the contrast between tense and lax vowel pairs can be difficult to distinguish when reduced. In this study the phonemic vowel was designated with reference to the perception of the sound over many different tokens and comparing its formant structure to those of tense and lax vowels of the same type produced by the same speaker in the surrounding context. When this was ambiguous (both tense-like and lax-like perceptions and formant comparisons were frequent in the data for a given morpheme), the designation was made through comparison to the form found in other dialects and in the historical form of the language. For example, the vowel in the complementizer *chï* /t͡ʃɪ/ varies perceptually between tense-like and lax-like in Chichicastenango, but it corresponds to a short vowel in other K'iche' dialects (Tum & Florentino, 2001; Ajpacajá Tum et al. 2005) and was categorized as lax. Meanwhile the vowel in the exclusive particle *xa* /ʃa/ 'only' also varies perceptually between tense-like and lax-like in Chichicastenango, but corresponds to a long vowel in other K'iche' dialects (Tum & Florentino, 2001; Larsen 1988) and was categorized as tense. In total, there were 8,808 lax vowels and 5,639 tense vowels in the data.

Coda and onset: Each vowel was classified according to whether its syllable has an onset and whether its syllable has a coda in the underlying form, with syllabification at the level of the morphological word (including affixes but excluding clitics). Any word-initial consonants were assumed to be the onset of the first vowel, and any word-final consonants the coda of the last vowel. A single intervocalic consonant was assumed to be the onset of the second vowel. If there were two intervocalic consonants, the first was assumed to be the coda of the first vowel and the second the onset of the second vowel. In total, the data includes 5,809 vowels in closed syllables and 8,638 vowels in open syllables. There are 13,469 vowels in syllables with onsets and 978 vowels in onsetless syllables.

Stress position: Each vowel in a content word was classified as stressed or unstressed following the pattern described in Section 2 of this article. Unstressed vowels were classified by the distance from the stressed syllable (adjacent, distant) and the direction (preceding or following).

For function words, determining whether a given syllable is stressed is more complicated. Little research exists on the prosody of function words and the interaction between word-level stress and intonation in K'iche'. Some previous works on K'iche' have argued that some function words bear stress while others are unstressed. Henderson (2012) argues that the 'phrase-final' forms of directionals and other particles are stressed and the 'phrase-medial' forms unstressed.

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

For the purposes of this study, a given word was classified as stressed if it is: (a) the phrase-final form of one of these alternating words, (b) a pronoun, or (c) a word routinely perceived as prominent (/χeʔ/, /tɛʔ/, /ɑɾe/, /k'ɔ/). In a stressed multisyllabic word, the final syllable was classified as stressed, excluding the extrametrical suffixes /-ɪk/ and /-ɔq/ – this is consistent with the stress pattern of content words as well as perceptual judgments of prominence. All other function words were classified as unstressed, including (a) the phrase-medial forms of alternating words (except for the existential /k'ɔ/, typically perceived as prominent), (b) determiners, (c) complementizers, (d) prepositions, and (e) other words usually perceived as non-prominent.

In total, there are 6,655 stressed vowels and 7792 unstressed vowels in the data.

Syllable position: Each vowel was classified as final or non-final according to the position of its syllable within the morphological word (including affixes but excluding clitics). In total there are 10,408 vowels in final syllables and 4,039 vowels in non-final syllables.

Domain: Each vowel was classified as belonging to either a function word, the stress domain of a content word, or an extrametrical morpheme in a content word. For content words, all vowels were considered to be part of the stress domain unless they belonged to verbal person, aspect, or incorporated movement prefixes, the superlative adjectival suffix, or the status suffixes *-ïk*, *-öq*, *ö*/*ü*. As discussed above, these affixes are not stressed regardless of their position or shape, including when they might be expected to be stressed, such as in word-final position for a non-verb or containing a tense vowel or closed syllable in a verb. Examples of the exclusion of these morphemes from the stress domain can be seen in Tables 2 and 3 above. In total, the data includes 5,111 vowels belonging to the stress domain of content words, 1,856 vowels in extrametrical morphemes in content words, and 7,480 vowels in function words.

Item: Each vowel was categorized according to the morpheme it belongs to, with separate items for each vowel in morphemes with multiple vowels. By this metric, there were 610 distinct items in the data.

Speaker: Each word was categorized according to the speaker who produced it. There are 11 different speakers represented in the data.

In addition to the previous factors, function words were additionally classified according to the following independent variables which address the phrasal context, which appeared to influence rates of deletion in function words, unlike content words: SYLLABLE POSITION (PHRASE), SYLLABLE POSITION (UTTERANCE), CODA (PHRASE-LEVEL), ONSET (PHRASE-LEVEL), CODA (UTTERANCE-LEVEL), ONSET (UTTERANCE-LEVEL).

Syllable position (phrase): Vowels were categorized as phrase-final if they are in the last syllable of the phonological phrase and not phrase-final if there is a following syllable within the phonological phrase. Due to a lack of previous research on phonological phrasing in K'iche', for the purposes of this study the phonological phrase was determined based on the syntactic structure, assuming that the phonological phrase should correspond roughly to the syntactic phrase as it does cross-linguistically (Bennett & Elfner 2019). Each of the major components of a sentence was considered to be a separate phrase: subject or object noun phrases, prepositional or adverbial phrases, and the verb complex (verb together with dependent particles). Any material that precedes the verb complex, such as complementizers or conjunctions, was excluded from the verb phrase and considered to form a separate phrase. Examples are shown in (3)–(4) with phrase boundaries in square brackets.

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

(3) *Të' k'ü rï' rï jün ali nä xäraj täj.*

tɛʔ	k'ʊ	ɾɪʔ	ɾ	χʊn	ali	n	ʃraχ	təχ
[tɛʔ	k'ʊ	ɾɪʔ]	[ɾɪ	χʊn	ali]	[nə	ʃə-Ø-raχ	təχ]
then	then	DEM	DET	one	girl	NEG	CPL-B.3SG-A.3SG-want	IRR

'Then the girl didn't want to.' (owl, 0:13)

(4) *Para ke r wïnäq kiköjönïk pä rï kanton.*

paɾa	ke	ɾɪ	wnəq	kikχɔnɪk	p	ɾ	kanton
[paɾa	ke]	[ɾɪ	wɪnəq]	[k-i-kɔχɔn-ɪk]	[pə	ɾɪ	kanton]
so	that	DET	person	INCPL-B.3SG-believe-SS.F	PREP	DET	community

'So that the people believe [in Christianity] in the community.' (church, 0:39)

(3) is composed of three phrases: the adverbial phrase *të' k'ü rï'* 'then', the subject noun phrases *rï jün ali* 'the girl', and the verb complex *nä xäraj täj* 'didn't want to'. (4) is composed of four phrases: the introductory material *para ke* 'so that', the subject noun phrase *rï wïnäq* 'the people', the verb *kiköjönïk* 'believe', and the prepositional phrase *pä rï kanton* 'in the community'.

In total there are 3,915 non-phrase-final tokens and 3,565 phrase-final tokens in the function word data.

Syllable position (utterance): Vowels were categorized as utterance-final if they in the last syllable before a pause, and non-utterance-final otherwise. In total there are 669 utterance-final tokens and 6,811 non-utterance-final tokens in the function word data.

Coda and onset (phrase-level and utterance-level): Words were syllabified separately at the phrase level (including all material within the phonological phrase) and the utterance level (including all material enclosed by pauses). Syllabification was determined according to the same rules used for word-level syllabification, discussed above. In total, there are 2,744 vowels in closed syllables with phrase-level syllabification, and 4,736 in open syllables. There are 495 vowels in onsetless syllables with phrase-level syllabification, and 6,985 in syllables with onsets. When syllabification is considered at the utterance level, there are 2,629 vowels in closed syllables, and 4,851 vowels in open syllables. There are 326 vowels in onsetless syllables with utterance-level syllabification, and 7,154 in syllables with onsets.

Finally, each lax vowel in a non-final, unstressed CV syllable not adjacent to a stressed syllable was categorized according to one additional factor.

Following vowel presence: Each vowel of this type was categorized according to whether the following underlying vowel was present or deleted. Of 1,476 vowels in this category, the following vowel was deleted in 661 cases and present in 815 cases.

3.3 Statistical analysis

The data was visualized in R (R Core Team 2020) with the package ggplot2 (Wickham 2016) and analyzed with mixed-effects logistic regression using the package lme4 (Bates et al. 2015). The statistical models used treatment ("dummy") coding, and the baseline condition for each variable was where deletion was most expected.

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

Table 5. Model comparison of the syllable position variable

Level of model	Degrees of freedom	AIC	BIC
Word	4	3509.682	3537.362
Phrase	4	3468.351	3496.031
Utterance	4	3475.163	3502.843

Table 6. Model comparison of the syllabification variables

Level of model	Degrees of freedom	AIC	BIC
Word	5	3502.704	3537.304
Phrase	5	3483.260	3517.859
Utterance	5	3492.601	3527.201

Deletion in function words, unlike content words, often seems to depend on the larger phrasal context. The factors of word, phrase, and utterance level are not independent of each other: all utterance-final syllables are also word-final and nearly all are phrase-final (except those that precede a phrase-internal pause). Rather than creating a complex factor with levels formed of all possible combinations of word, phrase, and utterance position – which would result in some categories with very small amounts of data – model comparison was done with the AIC (Akaike's information criterion) and BIC (Bayesian information criterion) functions in R to determine whether the word, phrase, or utterance level best predicted the presence of vowels in function words for the variables ONSET, CODA and SYLLABLE POSITION. These models predict vowel presence according to only the SYLLABLE POSITION variable or the syllabification variables (ONSET, CODA), respectively, in addition to the random effects of ITEM and SPEAKER. The results of the model comparison are shown in Tables 5 and 6.

Both AIC and BIC are lowest at the phrase level for both the syllable position variable and the syllabification variables (ONSET, CODA). Therefore, syllabification and syllable position were included at the phrase level for function words in later models (but at the word level for content words).

Next, VOWEL PRESENCE (present, deleted) was modeled in the complete data set (function words and content words). The fixed effects were VOWEL QUALITY (lax, tense), ONSET (yes, no), CODA (no, yes), SYLLABLE POSITION (non-final, final) and STRESS (unstressed, stressed). An interaction was included between each of the fixed effects and DOMAIN (stress domain internal content word, stress domain external content word, function word), with the exception of STRESS since all stress domain external vowels in content words are unstressed. These interactions were included because deletion appeared to be less regular outside of the stress domain and in function words as compared to content words. SPEAKER and ITEM were included as random effects. The equation is shown as follows.

$$\text{glmer(factor(VOWEL PRESENCE} \sim \text{VOWEL QUALITY} * \text{DOMAIN} + \text{ONSET} * \text{DOMAIN} + \text{CODA} * \text{DOMAIN} + \text{SYLLABLE POSITION} * \text{DOMAIN} + \text{STRESS} + (1|\text{SPEAKER}) + (1|\text{ITEM}), \text{family} = \text{binomial},$$

$$\text{control} = \text{glmerControl(optimizer} = \text{“bobyqa”))}$$

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

Table 7. Model comparison of the full model with and without interactions

Model	Degrees of freedom	AIC	BIC
Without interactions	10	6051.501	6127.284
With interactions	18	5747.484	5883.893

Table 8 Model comparison of the second model with and without interaction term

Model	Degress of freedom	AIC	BIC
Without interaction	6	2935.163	2972.081
With interaction	8	2657.224	2706.449

The baseline categories for each variable were those most expected to contribute to vowel deletion: lax vowel quality, onset, no coda, unstressed, non-final position, stress domain internal morpheme in content word.

Model comparison was carried out with the AIC and BIC functions to ensure that the addition of the interaction terms improved the fit of the model. The results are shown in Table 7.

Both AIC and BIC are lower for the model with interactions, indicating that this model has an improved fit.

Additionally, to explore how deletion operates for vowels that meet the qualifications for deletion, a second model was run for the subset of the data consisting of lax, unstressed vowels in non-final CV syllables. This model had fixed effects STRESS POSITION (adjacent, distant) and DOMAIN, and the interaction between them, as well as random effects of SPEAKER and ITEM. The equation is shown as follows.

$$\text{glmer}(\text{factor}(\text{VOWEL PRESENCE} \sim \text{STRESS POSITION} * \text{DOMAIN} + (1|\text{SPEAKER}) + (1|\text{ITEM}), \text{family} = \text{binomial}, \text{control} = \text{glmerControl}(\text{optimizer} = \text{“bobyqa”}))$$

The AIC and BIC values for this model compared to the same model without the interaction term is shown in Table 8. Again, both AIC and BIC are lower for the model with the interaction, indicating that the interaction improves the fit of the model.

Finally, to explore how deletion operates for vowels that meet the qualifications for deletion but are more distant from a stressed syllable, a third model was run consisting of lax, unstressed vowels in CV syllables preceding a stressed syllable by two or more syllables. This model had fixed effects FOLLOWING VOWEL PRESENCE (present, deleted) and DOMAIN (stress domain internal content word, stress domain external content word, function word), with an interaction between them, as well as random effects of SPEAKER and ITEM. The equation is shown as follows.

$$\text{glmer}(\text{factor}(\text{VOWEL PRESENCE} \sim \text{FOLLOWING VOWEL PRESENCE} * \text{DOMAIN} + (1|\text{SPEAKER}) + (1|\text{ITEM}), \text{family} = \text{binomial}, \text{control} = \text{glmerControl}(\text{optimizer} = \text{“bobyqa”}))$$

The AIC and BIC for this model compared to an identical model without the interaction term is shown in Table 9. Again, the inclusion of the interaction improves the fit of the model.

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

Table 9. Model comparison of the third model with and without interaction term

Model	Degrees of freedom	AIC	BIC
Without interaction	6	1168.089	1199.871
With interaction	8	1134.160	1176.536

4 Results

The results of the full statistical model are shown in Table 10.

There is a significant positive effect ($p < .01$) of VOWEL QUALITY, DOMAIN (function word), ONSET, CODA, and STRESS. There are also significant interactions between DOMAIN and VOWEL QUALITY, ONSET and CODA.

Vowel quality: The main effect of vowel quality shows that in content words, tense vowels are less likely to be deleted than lax vowels. The significant interaction with domain shows

Table 10. Results of the full statistical model (fixed effects)

	Estimate	Std. Error	Z-value	P-value	
Intercept	−3.274	0.377	−8.681	< .001	***
VOWEL QUALITY: tense	8.800	0.915	9.616	< .001	***
DOMAIN: extrametrical morpheme in content word	0.987	0.652	1.513	.130	
DOMAIN: function word	4.169	0.412	10.113	< .001	***
ONSET: no	23.020	3.898	5.905	< .001	***
CODA: yes	7.671	0.903	8.500	< .001	***
SYLLABLE POSITION: final	−1.060	1.184	−0.895	.371	
STRESS: stressed	6.928	0.712	9.735	< .001	***
VOWEL QUALITY: tense, DOMAIN: extrametrical morpheme in content word	−0.775	1.548	−0.501	.617	
VOWEL QUALITY: tense, DOMAIN: function word	−7.777	1.219	−6.382	< .001	***
ONSET: no, DOMAIN: extrametrical morpheme in content word	−24.420	3.909	−6.247	< .001	***
ONSET: no, DOMAIN: function word	−23.426	3.904	−6.000	< .001	***
CODA: yes, DOMAIN: extrametrical morpheme in content word	−7.185	1.071	−6.706	< .001	***
CODA: yes, DOMAIN: function word	−6.533	0.951	−6.871	< .001	***
SYLLABLE POSITION: final, DOMAIN: extrametrical morpheme in content word	23.014	15.538	1.481	.139	
SYLLABLE POSITION: final, DOMAIN: function word	2.145	1.206	1.778	.075	

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

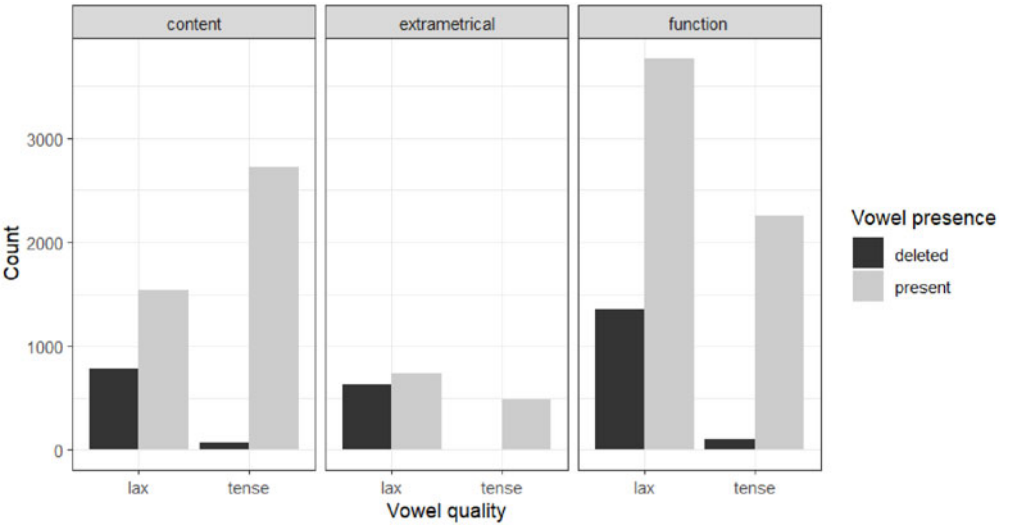


Figure 5. Deletion of vowels by vowel quality.

that the effect of vowel quality on deletion differs in content words and function words: in function words, tense vowels are more likely to be deleted than in content words. The number of present and deleted vowels in each domain is shown in Figure 5.

Examples of words with similar structures where a lax vowel is deleted and a tense vowel preserved are shown in (5). The tense vowel /a/ is preserved in (5a) while the lax vowel /ʊ/ is deleted in (5b) in a similar context: in an open, unstressed syllable preceding the stressed syllable.

(5) a. *chanim*
t͡ʃanim
t͡ʃa.ˈnim
'now' (history, 0:02)

b. *jünab'*
χnaɓ
χʊ.ˈnaɓ
'year' (history, 0:11)

The vowel /ʊ/ is known to be present underlyingly in the word *jünab'* as it is present in other contexts, e.g. *jünäb'ir* [χʊmɓir] 'a year ago' (elicited example, same speaker as 5a,b).

Onset: The main effect of onset shows that inside of the stress domain of content words, vowels in onsetless syllables are less likely to be deleted than vowels in syllables with onsets. The significant interaction with domain shows that vowels in onsetless syllables are comparatively less resistant to being deleted in extrametrical morphemes of content

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

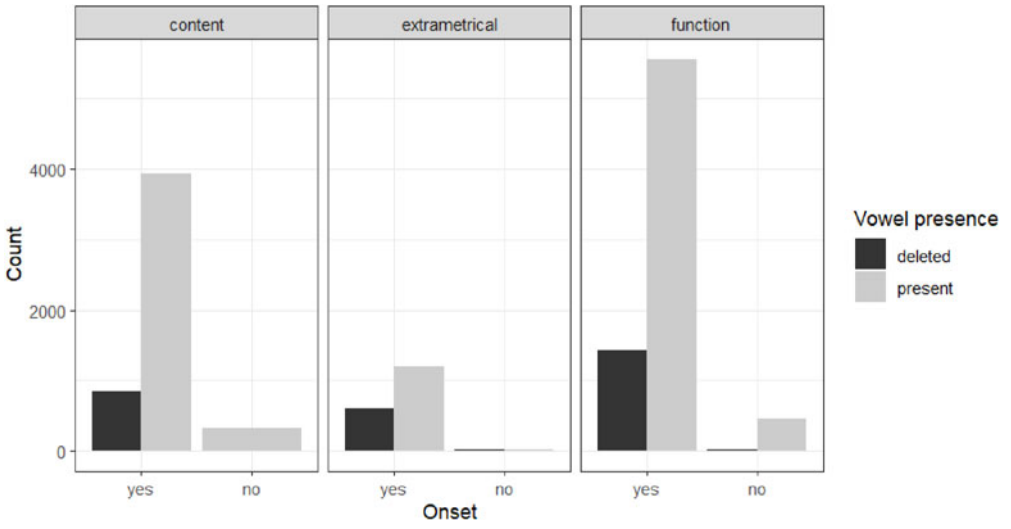


Figure 6. Deletion of vowels by onset.

words and in function words. The number of present and deleted vowels in each domain is shown in Figure 6.

Examples of words with similar structures where a vowel in an onsetless syllable is preserved while a vowel in a syllable with an onset is deleted are shown in (6). The vowel /ə/ is preserved in (6a), and realized as its tense counterpart [a] due to its word-initial position. The same vowel is deleted in (6b), where the possessive prefix provides an onset to this syllable.

(6) a. *atz'am*
ats͡'am
ə.ˈts͡'am
'salt' (3recipes, 5:53)

b. *rätz'am*
ɾts͡'am
ɾə.ˈts͡'am
'its salt' (3recipes, 3:35)

Coda: The main effect of coda shows that in stress domain internal morphemes in content words, vowels in closed syllables are less likely to be deleted than vowels in open syllables. The significant interaction terms show that vowels in closed syllables are comparatively more likely to be deleted in function words or extrametrical morphemes belonging to content words. The number of present and deleted vowels in each domain is shown in Figure 7. There are practically no deletions inside of the stress domain in content words. In extrametrical morphemes in content words and in function words, vowel deletion is more likely, though still uncommon.

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

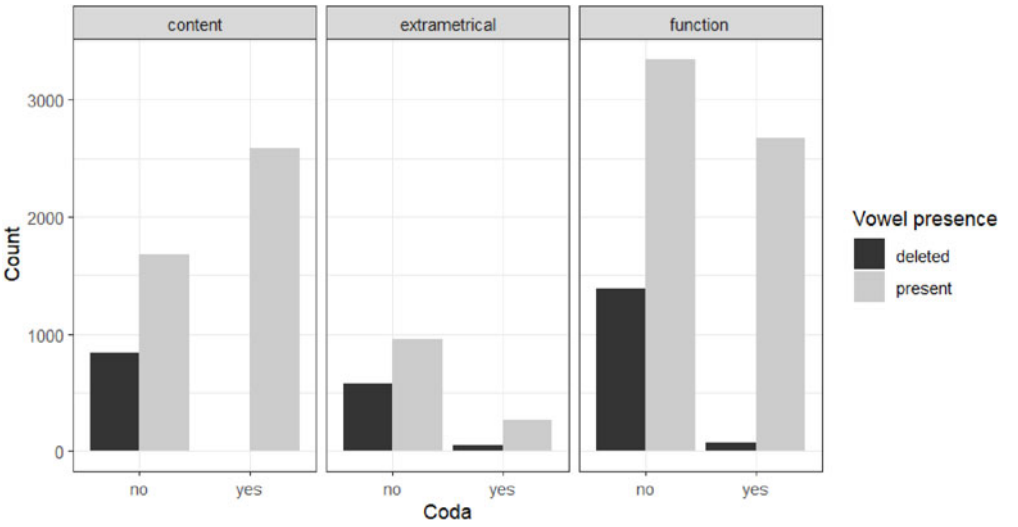


Figure 7. Deletion of vowels by coda.

The preservation of vowels in closed syllables is exemplified in (7) with two forms of the morpheme *nïm* 'big, important'. In (7a) the vowel /ɪ/ is in a closed syllable and is not deleted, whereas in (7b) it is in an open syllable, still between the same consonants, and deleted.

(7) a. *nïmq'ij*
nɪmq'iχ
nɪm.ˈq'iχ
'festival' (semanasanta, 0:32)

b. *nïma'q*
nmaʔq
nɪ.ˈmaʔq
'big' (plural) (fishing, 2:17)

Syllable position: There is no significant effect of syllable position, nor any significant interactions between syllable position and domain.

Stress: The main effect of stress shows that vowels are less likely to be deleted when stressed. The number of present and deleted vowels according to stress is shown in Figure 8. (Note that definitionally, no stressed vowels are extrametrical.)

The preservation of stressed vowels is exemplified in (8) with two forms of the verb *köj* 'wear'. In (8a), the vowel /ɔ/ in the verb root is in the stressed syllable and is not deleted. In (8b), the same vowel is in an unstressed syllable, as stress falls on the suffix /-om/. This unstressed vowel is deleted. At a segmental level, the context for these two vowels is very similar: in the penultimate syllable, which has the shape /kɔ/.

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

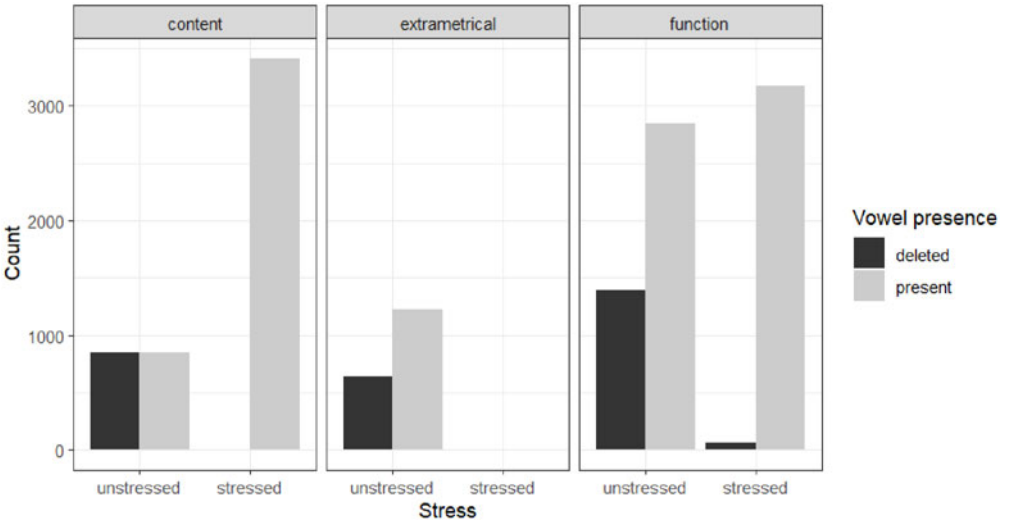


Figure 8. Deletion of vowels by stress.

(8) a. *kïköjö*

kɪkɔχɔ

kɪ.ˈkɔ.χɔ

'they wear it' (history, 10:00)

b. *köjom*

kχom

kɔ.ˈχom

'they have worn it' (history, 10:10)

Domain: The main effect of domain shows that vowels that are lax and in unstressed CV syllables are more likely to be deleted inside of the stress domain of content words than in function words. The number of deleted and preserved vowels in each domain is shown in Figure 9. Whereas these vowels are very likely to be deleted inside of the stress domain of content words, they are about equally likely to be deleted or not in extrametrical morphemes and in the stress domain of function words.

Finally, the results of the random effects in this full model are shown in Table 11. There is little variation across speakers in the overall likelihood of vowels being deleted. There is much larger variation across items, indicating that vowels in certain morphemes are more prone to being deleted than others.

The results of the statistical model of the subset of lax, unstressed vowels in non-final CV syllables are shown in Table 12.

There is a significant main effect of STRESS POSITION and both levels of DOMAIN, as well as a significant interaction between these two factors.

Stress position: For vowels belonging to the stress domain of content words, they are less likely to be deleted when two or more syllables away from a stressed syllable than when

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

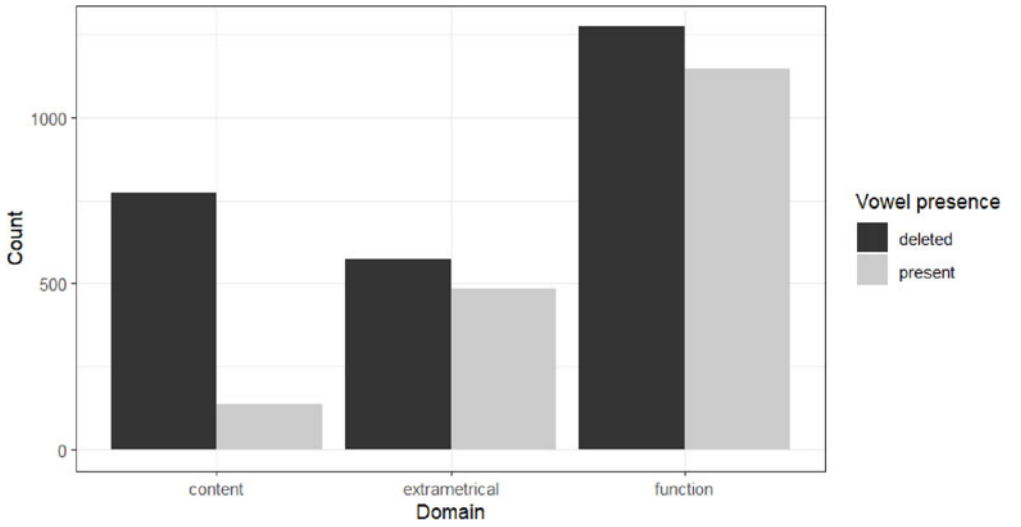


Figure 9. Deletion of unstressed lax vowels in CV syllables across domains.

Table 11. Results of full statistical model (random effects)

	Variance	Std. Deviation
ITEM	9.2767	3.0458
SPEAKER	0.0574	0.2396

Table 12. Results of the model of lax, unstressed vowels in CV syllables (fixed effects)

	Estimate	Std. Error	Z-value	P-value	
Intercept	−4.288	0.416	−10.298	< .001	***
STRESS POSITION: distant	5.562	0.684	8.129	< .001	***
DOMAIN: extrametrical syllable in content word	1.535	0.754	2.036	< .05	*
DOMAIN: function word	5.571	0.618	9.016	< .001	***
STRESS POSITION: distant, DOMAIN: extrametrical syllable in content word	−1.035	0.735	−1.408	.159	
STRESS POSITION: distant, DOMAIN: function word	−5.416	0.704	−7.698	< .001	***

adjacent to a stressed syllable. Vowels in function words are comparatively more likely to be deleted in this distant context.

Domain: For vowels adjacent to a stressed syllable, they are less likely to be deleted when in an extrametrical syllable or function word compared to inside of the stress domain of content words. Vowels more distant from a stressed syllable are comparatively more likely to be deleted in function words than inside of the stress domain of content words. The number of deleted and preserved unstressed lax vowels in non-final CV syllables according

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

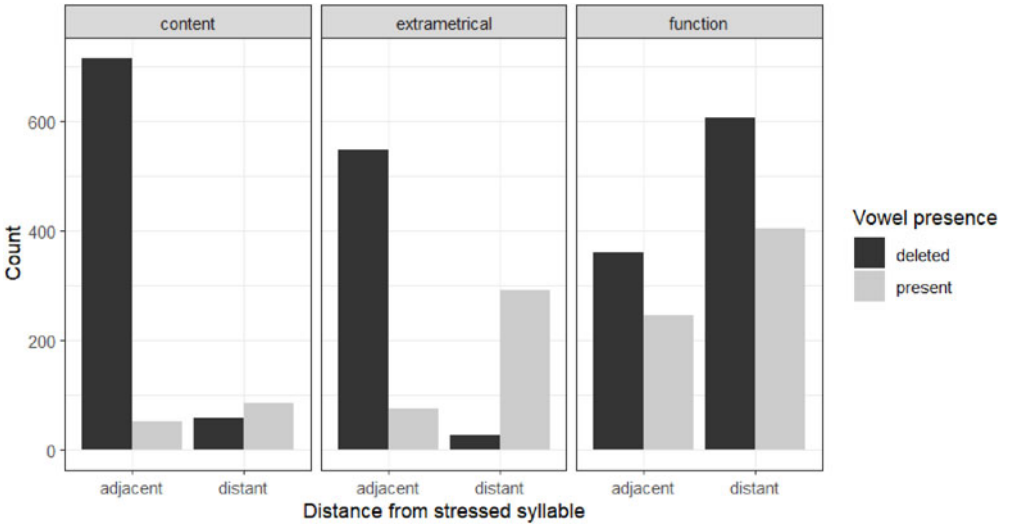


Figure 10. Deletion of unstressed lax vowels in non-final CV syllables.

to domain and position with respect to a stressed syllable is shown in Figure 10. Overall, vowels are deleted more often when adjacent to a stressed syllable than more distant from a stressed syllable. However, the contrast between positions is much smaller for vowels in function words than those in content words.

Examples of the deletion of lax vowels in non-final open syllables preceding the stressed syllable are shown in (9).

(9) a. *xïntzüküj*
ʃɪn͡tskʊχ
ʃɪn.t͡sʊ.ˈkʊχ
'I searched for it' (fishing, 6:44)

b. *ixöqïb'*
iʃqɪɓ
i.ʃɔ.ˈqɪɓ
'women' (history, 5:36)

c. *mïq'ïn*
mq'ɪn
mɪ.ˈq'ɪn
'hot' (healing, 8:50)

These vowels are known to be present underlyingly, as they appear in the same morphemes in other contexts, e.g. *utzüküxik* [ut͡sʊkʃik] 'its being searched for' (mushrooms, 6:30), *ixöq* [iʃɔq] 'woman' (healing, 2:33), and *mïq'ïna'* [mɪq'naʔ] 'hot water' (3recipes, 7:27).

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

Table 13. Results of the model of lax, unstressed vowels in CV syllables (random effects)

	Variance	Std. Deviation
ITEM	5.9223	2.4336
SPEAKER	0.0524	0.2289

Examples of the deletion of lax vowels in non-final open syllables following the stressed syllable are shown in (10). The vowel /ɛ/ is deleted in (10a) following the stressed syllable, and the vowel /ə/ in (10b).

(10) a. *käch'ab'ëxïk*

kt͡ʃ'aβʃɪk

kə.ˈt͡ʃ'a.βɛ.ʃɪk

'it is prayed for' (history, 7:44)

b. *nämäläj*

nɪmləχ

ˈnɪ.mə.ləχ

'very big' (history, 1:23)

These vowels are known to be present underlyingly because they appear in other forms of the same morphemes, e.g. *käch'äjb'ëj* [kət͡ʃ'əχβ**ɛ**χ] 'you use it to wash' (healing, 0:46) and *sïb'äläj* [sɪβ**ə**ləχ] 'very' (changes1, 3:35).

Finally, the results of the random effects in this model are shown in Table 13. Again, there is very little variation across speakers in the likelihood of vowels being deleted. There is greater variation across items, indicating that the vowels in some specific morphemes are more likely to be deleted than others.

The results of the statistical model of the subset of lax, unstressed vowels in CV syllables preceding a stressed syllable by two or more syllables are shown in Table 14.

There is a significant effect of FOLLOWING VOWEL and both levels of DOMAIN, as well as significant interactions between them.

Following vowel: The main effect of a following vowel shows that for vowels belonging to the stress domain of content words, they are much less likely to be deleted if the following vowel is deleted. Comparatively, similar vowels in extrametrical morphemes or function words are more likely to be deleted than inside of the stress domain of content words; that is, they are less dependent on the presence of the following vowel.

Domain: The main effect of domain shows that for vowels which precede a present vowel, they are less likely to be deleted if they are in extrametrical morphemes or function words than within the stress domain of content words.

The number of deleted and preserved vowels according to the presence of the following vowel and the domain are shown in Figure 11. Stress domain internal vowels in content words show the sharpest difference across context: they are almost always deleted when the following vowel is present and present when the following vowel is deleted. The same overall difference is observed for function words, but it is less predictable. For extrametrical

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

Table 14. Results of the model of lax, unstressed vowels in CV syllables distant from a stressed syllable (fixed effects)

	Estimate	Std. Error	Z-value	P-value	
Intercept	−2.724	0.800	−3.404	< .001	***
FOLLOWING VOWEL: deleted	6.156	0.859	7.168	< .001	***
DOMAIN: extrametrical morpheme in content word	4.784	1.199	3.992	< .001	***
DOMAIN: function	3.082	0.965	3.194	< .01	**
FOLLOWING VOWEL: deleted, DOMAIN: extrametrical morpheme in content word	−5.404	1.048	−5.156	< .001	***
FOLLOWING VOWEL: deleted, DOMAIN: function word	−3.570	0.876	−4.074	< .001	***

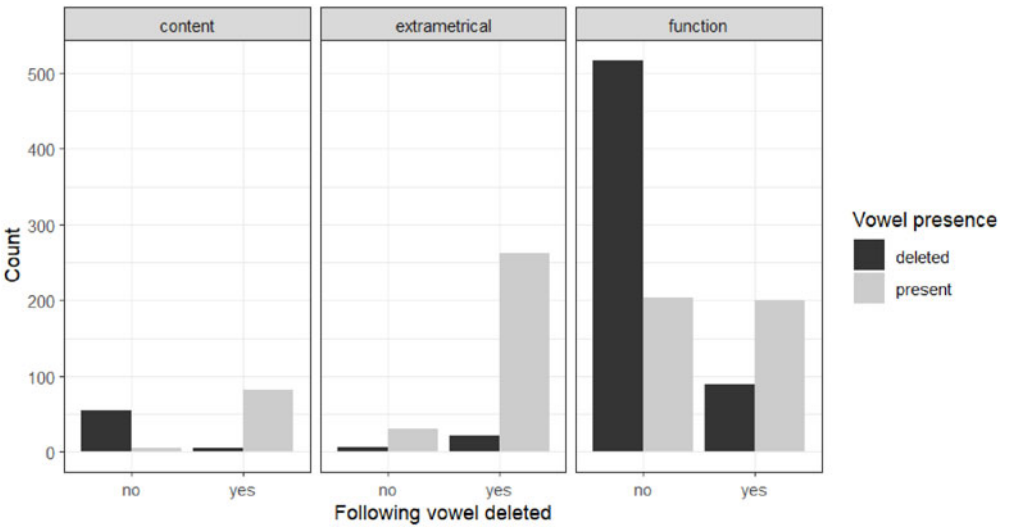


Figure 11. Deletion of lax vowels in non-final CV syllables distant from a stressed syllable.

vowels in content words, they are usually present in both contexts, but proportionally more often deleted if the following vowel is present.

Examples of how deletion affects lax vowels in non-final open syllables more distant from the stressed syllable are shown in (11). In (11a), the vowel of the initial syllable is deleted, preceding a syllable with a present vowel. In (11b), the vowel of the initial syllable is present, and it precedes a syllable with a deleted vowel.

(11) a. *söqʺölbʺäl*

sq'ɔlβəl

sɔ.q'ɔl.ˈβəl

'field' (LXE, 9:15)

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

b. *tïnämït*

tɪnmɪt

tɪ.nə.ˈmɪt

'town' (history, 1:24)

Finally, the results of the random effects in this model are shown in Table 15. As in previous models, there is little variation across speakers. There is greater variation across items, indicating that the vowels in certain morphemes are more likely to be deleted than those in others.

Table 15. Results of the model of lax, unstressed vowels in CV syllables distant from a stressed syllable (random effects)

	Variance	Std. Deviation
ITEM	3.1266	1.7682
SPEAKER	0.1413	0.3759

5 Discussion

As noted above, building on prior work by Campbell (1977), López Ixcoy (1994) and England & Baird (2017), Wood (2020) concluded that lax vowels are deleted in content words when in a non-final CV syllable adjacent to the stressed syllable. Tense vowels, stressed vowels, and those in onsetless, closed, or word-final syllables are not deleted. Lax vowels in unstressed, non-final CV syllables not adjacent to the stressed syllable are sometimes deleted, but less regularly than those in syllables adjacent to the stressed syllable. Very little previous research existed on the deletion of vowels in function words in Chichicastenango K'iche', but Wood (2020) observed that vowels may be deleted in function words in contexts that would be prevented in content words, such as word-final. These deletions sometimes appear to be regular, but in other cases appear to be optional, with the same vowel appearing variably in the same word and similar phrasal contexts.

Most of these observations are well substantiated in the corpus data. Vowels are significantly less likely to be deleted when tense, in closed syllables, in onsetless syllables, or stressed. Deletion is very frequent when the vowel is lax and in an unstressed CV syllable adjacent to a stressed syllable. When in a similar context but more distant from a stressed syllable, deletion is less frequent. In function words, vowel deletion follows similar restrictions to content words but is much more variable. However, the corpus study results find no significant effect of syllable position on vowel deletion, nor an interaction with domain. The corpus data also clarifies the differences in the vowel deletion pattern across domains (within the stress domain of content words, extrametrical morphemes, and function words), including differences in the predictability of the pattern.

The following sections discuss some of the major results of the study. Section 5.1 discusses the effect of syllable position, which was identified as a factor in previous literature but was not significant in the statistical model. Section 5.2 discusses differences in the deletion pattern between different domains. Finally, Section 5.3 addresses the limitations of the study and possibilities for future research.

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

5.1 Deletion in word-final syllables

Previous descriptions of vowel deletion in Chichicastenango K'iche' had indicated that deletion is limited to non-final syllables (Campbell 1977; López Ixcoy 1994; Wood 2020). However, this factor was not significant in the statistical model, either as a fixed effect or as an interaction with domain.

An exploration of these final syllables sheds some light on why this factor appeared in previous descriptions but was not significant in the statistical model. The statistical model included as the baseline categories those most expected to contribute to vowel deletion: lax vowel quality, unstressed syllable, syllable with an onset, syllable without a coda. The baseline for domain was vowel belonging to the stress domain of content words. But lax vowels only ever occur in word-final position in extrametrical morphemes or in function words. Any final syllables of content words that belong to the stress domain of the word must always have a tense vowel, a coda, or both. Therefore, vowel deletion would be fully prevented in word-final syllables belonging to the stress domain of content words through the effects of the restriction against deleting tense vowels and vowels in closed syllables.

Vowels belonging to extrametrical morphemes may, however, be lax in word-final position without the need of a following consonant. Specifically, the status suffix -ɔ/-ʊ, which appears on root transitive verbs when at the end of an intonational phrase, creates verbs with a final unstressed CV syllable with a lax vowel. These vowels are rarely or never deleted, as they practically always appear when expected; that is, at the end of an intonational phrase (Wood 2025b). An example is shown in (15). Extrametrical morphemes are indicated in angle brackets < >.

(15) a. *kawïlö*
kawɪlɔ
<ka>.ˈwɪ.l<ɔ>
'you see it' (3recipes, 5:10)

Here, the lax vowel /ɔ/ is not deleted despite occurring in an unstressed CV syllable that follows the stressed syllable. However, it is not clear whether this is an effect of syllable position per se, as assumed in Wood (2020), or rather an effect of this specific suffix, which as it consists of only a single vowel would result in a complete loss of the morpheme if the vowel were deleted.

Function words may also end in a lax vowel, including when they are the final syllable of a phonological phrase. Of the 2,420 lax vowels in unstressed CV syllables in function words, 802 are phrase-final. Of these, 306 are deleted (38.2%). The remaining 1,618 are not phrase-final. Of these, 967 are deleted (59.8%). Therefore vowel deletion is more frequent in a function word that is not at the end of a phrase, indicating that there is some effect of syllable position though it was not evident in the statistical results.

It is also interesting to consider the small number of cases where vowel deletion does occur in final syllables, counter to expectations. As noted above, vowel deletion would be expected to be avoided entirely in final syllables inside of the stress domain of content words due to the effects of vowel quality and syllable shape. However, of 3,151 stress domain internal vowels in word-final syllables, five are deleted. These cases might seem negligible due to the tiny number of occurrences. However, four of these five tokens are similar: they end in a CVC syllable with a glide onset and a lax vowel, and they follow the stressed syllable. Interestingly, there are no tokens in the data that end in this type of syllable where the vowel is not deleted. An example is shown in (12).

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

(12) *xik'iyär*

ʃikˈijɾ

ʃi.ˈkˈi.jəɾ

'they grew'(history, 3:23)

This vowel is known to be present underlyingly as it appears in other forms of the same morpheme /-əɾ/, e.g. *säqärïk* [sqərɪk] 'it dawned' (fishing, 0:03).

In order to understand why vowel deletion occurs in this position, but not in similar contexts where the first consonant is not a glide, it is necessary to understand the structure of verbs in K'iche'. As is common across the Mayan language family, K'iche' has a set of verbal suffixes referred to as 'status suffixes' which indicate the mood and transitivity of the verb stem (Polian 2017). In K'iche', these suffixes also provide information about the phrase position of the verb, as they are absent or change form when the verb is not at the end of an intonational phrase (Henderson 2012; Royer 2021; Wood 2025b). Status suffixes appear on intransitive and root transitive verbs (a transitive root with no further derivational morphology), but not on derived transitive verbs (derived from a root that is not a transitive verb). When present, status suffixes are always the last morpheme in the word and are extrametrical. An example of the status suffix alternation is shown in (13).

(13) a. *Ya kächäq'ïj rï qäti'*

ja	**kətʃ͡q ' ɪχ**	ɾɪ	qtiʔ
ˈja	kə.tʃ͡ə.ˈqˈɪχ	ɾɪ	qə.ˈtiʔ
already	it.cooks	DET	our.meat

'Our meat is already cooking.' (3recipes, 3:31)

b. *Después ya kächäq'ïjïk*

despues	ja	**kətʃ͡q ' ɪχɪk**
des.ˈpues	ˈja	kə.tʃ͡ə.ˈqˈɪ.χɪk
after	already	it.cooks

'After that it is already cooking.' (3recipes, 8:10)

In (13a), the verb *kächäq'ïj* 'it cooks' is followed by the noun phrase *rï qäti'* 'our meat' within the intonational phrase, and it has the (null) phrase-medial status suffix. In (13b), the same verb occurs at the end of the intonational phrase and appears with the phrase-final suffix *-ïk*.

The alternation between phrase-medial and phrase-final forms is relevant to understanding the vowel deletion pattern because in Chichicastenango K'iche', like in the related languages Q'anjob'al (Mateo Toledo 2017) and the San Mateo Ixtatán variety of Chuj (Coon 2019; Royer 2021), phrase-final status suffixes are additionally used in phrase-medial contexts in order to prevent disallowed word-final consonant clusters. The only surface clusters which occur in word-final position in Chichicastenango K'iche' consist of either a glide or a glottal stop followed by another consonant. If a verb would otherwise end in a disallowed consonant cluster, the relevant phrase-final status suffix appears even in phrase-medial position. Notably, these potentially word-final consonant clusters which are

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

avoided through the appearance of the phrase-final status suffix always result from vowel deletion. An example is shown in (14).

(14) *kataq'änïk b 'ï chupam*

kataq'nık ɓɪ t͡ʃpam

ka.ˈta.q'ə.nık ɓɪ t͡ʃu.ˈpam

'you go up inside of it' (fishing, 3:05)

Here the verb *kataq'änïk* [kataq'nık] 'you go up' appears with the phrase-final suffix *-ïk* despite its phrase-medial position. The vowel /ə/ in the verb is deleted, which would result in a word- final [q'n] cluster if the suffix were absent.

The question that is raised by forms like *kataq'nïk* 'you go up' is why vowel deletion should occur in these verbs to begin with. If these vowels were not deleted, there would be no need for the phrase-final status suffix to appear in medial contexts. Analytical choices about the underlying structure of verbs, and whether or not the status suffixes are present in the underlying form of intonational phrase-medial verbs, result in different conclusions about the cause of vowel deletion in these cases. The typical assumption in the Mayanist literature is that phrase-final suffixes are removed in phrase-medial contexts rather than inserted in phrase-final contexts (e.g., Larsen 1988; Coon 2016; Pixabaj & Telma, 2017; del Prado & Curiel, 2017). This is certainly the case from a historical perspective, as status suffixes would have been present in all phrasal environments in Proto-Mayan. However, it is not clear whether it is the best analysis for the individual modern languages in which the appearance of the suffix is conditioned by phrasal context. Some authors describe the suffixes as being added in phrase-final position, instead (e.g., Baird 2014; Pye, Pfeiler & Mateo Pedro 2017).

If it is assumed that the phrase-final forms of the status suffixes are present underlyingly in all cases, no matter the prosodic context of the verb, then the fact that vowel deletion occurs in verbs like *xik'iyr* and *kataq'nïk* is not surprising. The presence of the status suffix means that the preceding syllable is not closed, as the last consonant of the verb stem becomes the onset of the status suffix syllable rather than the coda of the preceding syllable. Therefore, vowel deletion in these cases affects lax vowels in unstressed CV syllables adjacent to the stressed syllable, exactly where it occurs in other contexts. Under such an analysis, the status suffix would then be unpronounced in medial contexts unless required for phonotactic reasons.

If it is instead assumed that the phrase-final forms of the status suffixes are not present underlyingly, at least in intonational phrase-medial context, this means that there is a second context in which vowel deletion occurs. This second deletion rule would affect lax vowels in unstressed CVC syllables which follow the stressed syllable. This rule applies if it results in either a permissible word-final consonant cluster, as in (12), or a non-permissible cluster which can be avoided by adding in the relevant status suffix for that type of verb, as in (14). If the application of the rule would result in a non-permissible cluster which cannot be avoided in this way – as is the case in derived transitive verbs which have no status suffixes in any context – then deletion is prevented. An example is shown in (15).

(15) *käqesäj*

kəqesəχ

kə.ˈqe.səχ

'we remove it' (caldores, 1:04)

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

The verb *esäj* 'remove' does not have a status suffix that appears in phrase-final contexts. Therefore, deletion of the vowel /ə/ in the final syllable would result in a word-final cluster [sχ] which cannot be prevented by inserting a status suffix, and this vowel is not deleted.

Neither of these possibilities were initially suggested by the corpus study because syllable position was determined according to the morphemes that appear in the surface form of the word. Phrase-final status suffixes were counted in the syllabification if and only if they were present in the surface form, irrespective of the phrase position of the verb.

In sum, there is no conclusive evidence that syllable position affects vowel deletion in content words. Extrametrical vowels are not deleted when expected in word-final syllables, but this may be due to the specific morphemes involved rather than their position in the word per se. The likelihood of vowel deletion in function words does appear to be influenced by phrase position. Finally, the minority of cases where vowel deletion does occur unexpectedly in word- final (closed) syllables in content words reveals the possible existence of a second vowel deletion rule: lax vowels in CVC syllables following the stressed syllable. Whether or not deletion occurs in such an environment depends on assumptions about the underlying structure of verbal status suffixes, which would be a fruitful avenue for future research.

5.2 *Differences across domains*

The statistical results show that there are important differences in how deletion operates between stress domain internal morphemes of content words, extrametrical morphemes, and function words. The likelihood of deletion is most predictable for stress domain internal morphemes in content words, where the difference between contexts where deletion is expected and not expected is very sharp. Whereas vowels are practically never deleted when tense, stressed, in closed syllables, or in onsetless syllables, they are almost always deleted when lax, unstressed, and in CV syllables adjacent to a stressed syllable. Extrametrical vowels in content words follow similar patterns, but less starkly. For example, deletion of vowels in closed syllables, though not the norm, is more common in extrametrical morphemes than those belonging to the stress domain. An example is shown in (16), with extrametrical morphemes in angle brackets < >.

(16) *ïncha*
n.ˈt͡ʃa
<ɪn>.ˈt͡ʃa
'I say' (fishing, 1:27)

The vowel is deleted in the first syllable, which has a coda /n/. This vowel belongs to the extrametrical 1st person singular personal prefix *ïn-*.

Similarly, the preservation of lax vowels in CV syllables that are adjacent to a stressed syllable, though not the dominant pattern, is more likely for extrametrical vowels than those belonging to the stress domain. An example is shown in (17).

(17) *köjok*
kɔχok
<kɔ.ˈχ>ok
'we begin' (lit. 'we enter') (3recipes, 8:15)

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

Here the lax vowel /ə/ is not deleted in the syllable /kɔ/, which is extrametrical and precedes the stressed syllable.

Extrametrical lax vowels in CV syllables more distant from a stressed syllable are also less consistently preserved when the following vowel is deleted, or deleted when the following vowel is preserved. That is, they are less predictable based on the presence or absence of the following vowel. Examples are shown in (18).

(18) a *käqäch'äja*

kəqət͡ʃˈχa

<kə.qə>.t͡ʃˈə.ˈχa

'we go to wash it' (3recipes 3:08)

b. *käkïb'änö*

kkɓənɔ

<kə.kɪ>.ˈɓə.n<ɔ>

'they do it' (history 1:54)

In (18a) the lax vowel /ə/ is not deleted in the first syllable, despite the fact that it precedes a syllable with a present vowel. In (18b), conversely, the vowel in the first syllable is deleted, despite the fact that the following, unstressed vowel is also deleted.

Compared to extrametrical morphemes in content words, deletion in function words is even more variable. Although vowels are unlikely to be deleted when tense, stressed, in closed syllables, or in onsetless syllables, deletion is more likely in this context than for vowels in content words. Even more notably, although vowels in function words are most likely to be deleted when lax and in unstressed, CV syllables adjacent to a stressed syllable, the rate of deletion in this context is little more than half. Examples of the same function words with and without the vowel in very similar contexts are shown in (19)–(21).

(19) a. *Katïj rï **sïn** u'al sï salbiasanta.*

katɪχ	ɾ	**sɪn**	uʔal	sɪ	salbiasanta
ka.ˈtɪχ	ɾɪ	**sɪn**	u.ˈʔal	sɪ	sal.bia.ˈsan.ta
you.drink	the	DIM	broth	DIM	sage

'You drink the infusion of sage.' (healing, 1:47)

b. *Kutïj b'ik rï **sïn** u'al limon.*

kutɪχ	ɓik	ɾ	**sn**	uʔal	limon
ku.ˈtɪχ	ˈɓik	ɾɪ	**sɪn**	u.ˈʔal	li.ˈmon
he/she.drinks	DIR	the	DIM	broth	lime

'He/she drinks the lime juice.' (healing, 12:34)

(20) a. *Näj i **k'ö wï** känöq.*

nəχ	i	**kˈ**	**wɪ**	kənɔq
ˈnəχ	i	ˈ**kˈɔ**	**wɪ**	ˈkənɔq

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

far	PL	EXIST	PPTRACE	DIR

'They are far away.' (mushrooms, 5:21)

b. *Chpam ïn* ***k'ö wï*** *känöq.*

t͡ʃpam	ɪn	**k'ɔ**	**w**	ˈkə.nɔq
t͡ʃu.ˈpam	ˈɪn	ˈ**k'ɔ**	**wɪ**	ˈkə.nɔq
inside	1SG	EXIST	PPTRACE	DIR.F

'I was inside.' (earthquake, 1:14)

(21) a. *Xok* ***chï*** *b'ï rï nümama.*

ʃok	**t͡ʃɪ**	ɓɪ	ɾɪ	nmama
ˈʃok	**t͡ʃɪ**	ɓɪ	ɾɪ	nʊ.ma.ˈma
she.entered	again	DIR	DET	my.mother

'My mother entered again.' (earthquake, 1:10)

b. *Kel* ***chï*** *b'ï rï ju chïköp*

kel	**t͡ʃ**	ɓɪ	ɾ	χu	t͡ʃkɔp
ˈkel	**t͡ʃɪ**	ɓɪ	ɾɪ	χu	t͡ʃɪ.ˈkɔp
it.goes.outagain	DIR	DET	one	animal	

'The animal goes out again.' (kot, 1:12)

In (19a), the vowel in the diminutive *sïn* /sɪn/ is preserved, whereas it is deleted in (19b). These words occur in very similar contexts: following the determiner *rï*, whose vowel is deleted, and preceding the noun *u'al* 'broth, juice'. In (20a), the vowel in the existential *k'ö* /k'ɔ/ is deleted and the vowel in the following particle *wï* /wɪ/ is present. In (20b), the vowel in the existential is present and that in the following particle is deleted. In both cases these words are followed by the same directional particle *känöq*. In (21a), the vowel in the particle *chï* /t͡ʃɪ/ 'already, again' is preserved, whereas it is deleted in (21b). In both cases, it follows a stressed syllable and precedes the directional particle *b'ï* /ɓɪ/ 'here to there'.

In addition to the greater variability of vowel deletion in function words, there is a second important difference compared to content words. In content words, phrasal context is irrelevant. For instance, the extrametrical superlative suffix /-ələχ/ appears 33 times in the data. The first vowel, which follows the stressed syllable, is typically deleted. The second vowel, however, is never deleted, no matter the following word. Examples are shown in (22).

(22) a. *nïmäläj tïnämït*

nɪmləχ	tɪnmɪt
ˈnɪ.m<ə.ləχ>	tɪ.nə.ˈmɪt

'very big town' (history, 1:23)

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

b. *nïmäläj ulew*

nɪmləχ	ulew
ˈnɪ.m<ə.ləχ>	ʊ.ˈlew
very.big	land

'very big land' (history, 8:41)

In (22a), the suffix precedes the word *tïnämït* 'town', which begins with a consonant. Therefore, resyllabification at the phrase level would not affect the shape of the final syllable of the word *nïmäläj*. In (22b), however, the following word *ulew* 'land' begins with a vowel. If resyllabification occurred at the phrase level, the final consonant would become the onset of the second word, leaving an open syllable at the end of the first word. This vowel is not deleted, however – neither in this instance nor in any of the other instances in the data where the following word begins with a vowel. The phrasal context does not affect where vowels are deleted in content words.

Similarly, the addition of an affix as compared to a function word to a given content word has a markedly different effect on deletion. Prefixes permit the deletion of an otherwise word-initial vowel. Function words never have this effect – even when they could provide an onset to the following word-initial vowel – and the word-initial vowel is realized as tense in this context regardless of its underlying quality. This contrast is shown in (23).

(23) a. *külew*

klew

kʊ.ˈlew

'their land' (history, 0:58)

b.

täq	*ulew*
təq	ulew
təq	ʊ.ˈlew

'the lands' (fishing, 4:27)

The intermediate lax /ʊ/ in *külew* 'their land' in (23a), which is in a CV syllable preceding a stressed syllable, is deleted. The underlyingly lax vowel /ʊ/ at the beginning of the word *ulew* 'land' in (23b), when preceded by the proclitic *täq*, occurs in the same type of segmental context: preceded by a single consonant and followed by a stressed syllable. However, it is realized as tense and is not deleted.

In sum, the results of the study suggest the existence of three different levels of prosodic structure in the language. The stress domain includes content word roots, derivational affixes, and some inflectional affixes. At this level, vowel deletion is highly predictable based on vowel quality, syllable structure, and stress. The morphological word incorporates the remaining, extrametrical inflectional affixes. At this level, vowel deletion remains fairly predictable based on the same factors, though less so than inside of the stress domain. Finally, the phrase level incorporates function words. Vowel deletion is highly variable in this context, but is influenced by the same types of factors that constrain deletion at lower levels: vowel quality, syllable shape, and stress.

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

5.3 Limitations and future research

Studies on vowel deletion and other forms of speech reduction have identified a number of factors not considered in this study that are relevant cross-linguistically. Many studies have shown that reduction is more common for high-frequency words than low frequency words, which correlates with greater predictability and lower cognitive effort (e.g., Ernestus & Warner 2011; Clopper & Turnbull 2018, and references therein). The differences observed between content words and function words in the present study may be partially explained by word frequency effects. Deletion occurs in unexpected environments (such as a closed syllable or tense vowel) more commonly in function words than in content words, which may be due to their greater predictability and higher frequency. However, vowel preservation also occurs in unexpected environments more commonly in function words than in content words. This cannot be explained as a result of preserving information where it is less predictable. The effects of word frequency would be of great interest for future research. Due to the understudied status of the language, frequency estimates for content words are currently unknown.

Similarly, speech tempo is a common factor in vowel deletion cross-linguistically, with deletion occurring more at faster speech rates (e.g., Lindblom 1990). This study did not consider speech rate as a factor. The results do not suggest a likely effect of speech rate for content words, at least within the stress domain, as the contexts where deletion occurs or does not occur are highly predictable based on vowel quality, syllable structure, and stress alone. The author's experience is that deletion also occurs in the same words in contexts which induce phonetic lengthening, such as focus or utterance-final position. Speech rate may have an effect on deletion in more variable contexts, however, such as in extrametrical morphemes in content words and in function words. Similarly, speech style is known to affect vowel deletion across languages, with deletion being more common in casual speech than clear or formal speech (Ernestus & Warner 2011). Speech style could not be included as a factor for the corpus study, since all of the data belongs to the same speech style: spontaneous, fairly casual, but with the speaker aware that they were being recorded by someone interested in how their language is spoken. However, anecdotal evidence suggests that deletion in content words occurs predictably in the same contexts in other speech styles, such as in very clear speech produced to correct the pronunciation of the researcher, who is not a native speaker of the language, or in controlled speech production experiment data. Therefore, it is not the case that deletion is restricted only to casual speech. For function words, however, it may have a greater effect. Future research on the possible effects of speech rate and style could test these predictions.

Finally, the present study considered deletion as binary: vowels were categorized as either present or deleted. Even minor evidence of the presence of the vowel, such as voicing between otherwise adjacent voiceless consonants, was considered sufficient to code a vowel as present. These present vowels therefore cover a range from hyperarticulated vowels to extreme reductions. It is possible that deletion might not be a categorical phenomenon but rather a realization of gradient vowel reduction. Bennett, Henderson & Harvey (2023) argue that in the closely related language Uspanteko, the apparent deletion of vowels is actually better understood as a result of high levels of gestural overlap. Although the vowel is articulated, it is acoustically masked by the flanking consonants and therefore inaudible. They provide pilot electroglottographic (EGG) data that shows that at least in some cases there was covert voicing detected in the EGG signal that was not detectable in the acoustic signal. They further argue that this gestural overlap is not phonetic in nature, but phonologically controlled: it is not conditioned by speech rate or style, it is a language-specific pattern, and it shows phonological and morphological conditioning. Bennett et al. (2023) do not discuss word class specifically, but all of the provided examples are of content words.

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

There are many similarities between the vowel deletion patterns of Uspanteko and Chichicastenango K'iche', at least for content words. In both languages, deletion targets unstressed vowels preceding the stressed syllable (when the word has final stress) or following the stressed syllable (when the word has non-final stress). In both languages, deletion produces highly marked consonant clusters which are rare or unattested underlyingly in the language, especially morpheme-internally, and which violate common cross-linguistic principles like sonority sequencing. In both languages, deletion occurs not only in fast, casual speech but also in careful elicited speech and in contexts that induce phonetic lengthening, like utterance-final position. The languages differ primarily in the specific phonological and morphological conditions that deletion follows. For instance, deletion is avoided in prefixes and between identical adjacent consonants in Uspanteko, but it is not restricted in either of these contexts in Chichicastenango K'iche'. An example of deletion in a prefix can be seen in (16) above; examples of deletions between identical consonants include *anïnäq* /anɪnəq/ [annəq] 'quickly' (3recipes, 3:06), *qäqän* /qəqən/ [qqən] 'our legs' (changes1, 1:56), and *kkaj* /kəkaχ/ [kkaχ] 'they want' (fishing, 5:41).

The languages also appear to differ, however, in the predictability of the deletion pattern. Bennett et al. (2023) note that deletion is more common in post-tonic position than pre-tonic position in Uspanteko, but in both contexts it is variable. The same speakers sometimes produce the same words with and without the vowels in quick succession. In Chichicastenango K'iche', in contrast, deletion within the stress domain of content words is highly predictable, occurring over 90% of the time when expected.

In sum, there are many similarities between the apparent patterns of vowel deletion in Chichicastenango K'iche' and the closely related language Uspanteko, where "deletion" has been argued to result from a gradient reduction process of gesture overlap. However, the Chichicastenango K'iche' pattern is much more predictable, at least for content words. Further research is needed to assess the possibility of understanding the Chichicastenango K'iche' pattern as a phonologically controlled gestural coordination.

6 Conclusion

The results of the corpus study provide empirical support for most of the previous generalizations about vowel deletion in content words, within the stress domain: deletion is prevented for tense vowels, those in closed syllables, those in onsetless syllables, and those that are stressed. Even though very reduced realizations are typical in casual, spontaneous speech, these restrictions are practically absolute in the data. Deletion occurs very regularly for lax vowels in unstressed CV syllables adjacent to the stressed syllable. Additionally, depending on how status suffixes are analyzed, there may be a second context where deletion occurs: lax vowels in unstressed post-tonic closed word-final syllables, if they have a glide onset or if an appropriate phrase-final status suffix can be recruited to prevent the subsequent consonant cluster from being word-final. Finally, for vowels of a type that could in principle be deleted but which are not adjacent to the stressed syllable, deletion typically occurs when the following vowel is present and does not occur when the following vowel is deleted; that is, deletion is avoided in adjacent syllables.

Outside of the stress domain, deletion follows similar restrictions. However, it is less predictable in this context than within the stress domain. Vowels in extrametrical syllables are more likely to be deleted when expected to be present, or present when expected to be deleted, than vowels within the stress domain.

For function words, in turn, vowel deletion is highly unpredictable. Although deletion is uncommon for tense vowels, those in closed syllables, those in onsetless syllables, and those that are stressed, it occurs in this context more frequently than in content words.

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

Furthermore, vowels of the type most likely to be deleted – lax, in unstressed CV syllables – are deleted in only a little over half of all instances in the data compared to over 90% of the time within the stress domain of content words.

These results suggest the existence of three different levels of prosodic structure in Chichicastenango K'iche'. At each of these levels, similar phonological rules apply, however the regularity of the application differs. The existence of similar phonological rules that apply differently at different levels of prosodic structure is attested cross-linguistically. For example, the assimilation of English nasals is obligatory within words but optional across word boundaries (Kiparsky 1985). Similarly, palatalization in English is gradient and variable at the post-lexical level but categorical at the level of stem derivation (Zsiga 2000). These types of patterns have been argued to result from rule scattering; that is, stabilization of a previously gradient phonetic process into a categorical rule without the previous process disappearing (Robinson 1976; Bermúdez-Otero & Trousdale 2012; Iosad 2016). Over time, the level at which a phonological pattern applies may change in the language, and the result may be the existence of multiple similar processes at different levels of structure where one is categorical and the other gradient.

The greater likelihood of deletion in unexpected contexts in function words makes sense from an articulatory perspective, as function words are more often reduced than content words, and a high degree of variability is characteristic of reduction (Warner 2019). Variability in vowel deletion in casual speech is found in patterns such as schwa deletion in English (Oshika et al. 1975) and French (Dell 1981; Bürki et al. 2011). The fact that vowels in these contexts are *less* likely to be deleted when expected, however, cannot be easily explained as resulting from reduction, and rather appears to be due to a phonological effect of prosodic structure.

Acknowledgements *Sib'alaj maltyox* to all of the K'iche' speakers who shared their language with me and made this work possible. Thank you to Scott Myers, Brandon Baird, Danny Law, Megan Crowhurst, Ryan Bennett and Joshua Lieberstein for helpful conversations on many aspects of this work. Thanks also to two anonymous reviewers and the editor for their comments, which substantially improved the manuscript. This material is based upon work supported by the National Science Foundation 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.

Competing interests The author declares none.

Supplementary material To view supplementary material for this article (including audio files to accompany the language examples), please visit https://doi.org/10.1017/S0025100325100765

Appendix A. Abbreviations

The following abbreviations are used in example glosses throughout the text.

1	first person	INCPL	incompletive aspect
2	second person	IRR	irrealis
3	third person	M	phrase-medial
A	set A (ergative and possessive)	NEG	negation
B	set B (absolutive)	PASSC	completive passive
CPL	completive aspect	PL	plural

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

DEM	demonstrative	PPTRACE	trace of fronted prepositional phrase
DET	determiner		
DIM	diminutive	PREP	preposition
DIR	directional	SG	singular
EXIST	existential	SS	status suffix (verbal category suffix)
F	phrase-final		

Appendix B. Data sources

Unless otherwise noted, all examples cited throughout the text come from the collection of spontaneous, monologic speech used for the corpus study and are identified with a code to the recording and timestamp. Audio files of each of these examples are included in the supplementary materials. Information about each of the recordings that make up the corpus is shown in the following table. The archive PID is included for all recordings archived in the Archive of the Indigenous Languages of Latin America (AILLA, the K'iche' Collection of Elizabeth Wood, https://ailla.utexas.org/collections/563/).

Table B1. Information about the recordings that comprise the corpus

Speaker	Code	Duration	Topic	PIA
JT	3recipes	5:36	Recipes	4620
MJL	cooking	4:22	Traditions	4621
	chilmol	1:42	Recipes	4622
	church	9:08	History	4624
	changes1	3:43	History	4623
	changes2	3:34	History	4623
MCX	history	11:06	History	4626
	kot	3:15	Traditional story	4628
	planting	13:55	Traditions	4625
	mushrooms	10:59	Traditions	4627
	fishing	10:43	Personal life	4629
MXM	marriage	10:56	Traditions	4631
	earthquake	2:37	History	4630
MACM	healing	14:44	Traditions	4632
RSI	caldores	2:20	Recipes	4633
	owl	2:09	Traditional story	4634
TJL	semanasanta	2:58	Traditions	4638

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

Table B1. Continued

Speaker	Code	Duration	Topic	PIA
	talentos	6:20	Personal life	4637
	atolblanco	2:58	Recipes	4636
LRI	dailyactivities	3:00	Personal life	Unarchived
	magicegg	2:15	Traditional story	Unarchived
LXE	lxe	6:00	Traditions and traditional stories	Unarchived
MRT	mr	22:10	History and traditional stories	Unarchived
TRI	tri	2:46	Traditional story	Unarchived

References

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

Ajpacajá Tum, Pedro Florentino, M. I. Chox Tum, F. L. Tepax Raxulew, & D. A. Guarchaj Ajtzalam. 2005. *Diccionario K'iche'*. 2nd edition. Guatemala: Proyecto Lingüístico Francisco Marroquín.

Baird, Brandon Orrin. 2014. *An acoustic analysis of contrastive focus marking in Spanish–K'ichee' (Mayan) bilingual intonation*. PhD thesis. The University of Texas at Austin.

Bates, Douglas, Martin Mächler, Ben Bolker, & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67.1, 1–48. https://doi.org/10.18637/jss.v067.i01.

Bennett, Ryan. 2016. Mayan phonology. *Language and Linguistics Compass* 10(10), 469–514.

Bennett, Ryan, & Emily Elfner. 2019. The syntax–prosody interface. *Annual Review of Linguistics* 5(1), 151–171.

Bennett, Ryan, Robert Henderson, & Meg Harvey. 2023. Vowel deletion as grammatically controlled gestural overlap in Uspanteko. *Language* 99.3, 399–456.

Bermúdez-Otero, Ricardo & Graeme Trousdale. 2012. Cycles and continua: On unidirectionality and gradualness in language change. In: *The Oxford handbook on the history of English*, Article 55, pp. 691–720. Ed. by Terttu Nevalainen & Elizabeth Closs Traugott. Oxford: Oxford University Press.

Boersma, Paul & David Weenink. 2023. Praat: Doing phonetics by computer. [Computer program]. Version 6.3.14. http://www.praat.org/.

Bürki, Audrey, Cécile Fougeron, Cedric Gendrot, & Ulrich H Frauenfelder. 2011. Phonetic reduction versus phonological deletion of French schwa: Some methodological issues. *Journal of Phonetics* 39.3, 279–288.

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

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

Can Pixabaj, Telma Angelina. 2015. *Complement and purpose clauses in K'iche'*. PhD thesis. The University of Texas at Austin.

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

Clopper, Cynthia G. & Rory Turnbull. 2018. Exploring variation in phonetic reduction: Linguistic, social, and cognitive factors. In: *Rethinking reduction: Interdisciplinary perspectives on conditions, mechanisms, and domains for phonetic variation*, Chap. 2, pp.25–72. Ed. by Francesco Cangemi, Meghan Clayards, Oliver Niebuhr, Barbara Schuppler, and Margaret Zellers. Berlin, Boston: De Gruyter Mouton.

Coon, Jessica. 2016. Mayan morphosyntax. *Language and Linguistics* Compass 10(10), 515–550.

Coon, Jessica. 2019. Building verbs in Chuj: Consequences for the nature of roots. *Journal of Linguistics* 55(1), 35–81.

Dell, François. 1981. On the learnability of optional phonological rules. *Linguistic Inquiry* 12.1, 31–37.

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

Ernestus, Mirjam & Natasha Warner. 2011. An introduction to reduced pronunciation variants. *Journal of Phonetics* 39, 253–260.

Henderson, Robert. 2012. Morphological alternations at the intonational phrase edge: The case of K'ichee'. *Natural Language & Linguistic Theory* 30.3, 741–787.

Instituto Nacional de Estadística, Guatemala. 2019. XII Censo Nacional de Población y VII de Vivienda. https://www.ine.gob.gt/

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press

Iosad, Pavel. 2016. Rule scattering and vowel length in Northern Romance. *Papers in Historical Phonology* 1, 218–237.
Kiparsky, Paul. 1985. Some consequences of lexical phonology. *Phonology* 2, 85–138.
Larsen, Thomas Walter. 1988. *Manifestations of ergativity in Quiché grammar*. PhD thesis, University of California, Berkeley.
Lindblom, Björn. 1990. Explaining phonetic variation: A sketch of the H&H theory. *Speech production and speech modelling*, pp. 403–439. Dordrecht: Springer Netherlands.
Lobanov, Boris. 1971. Classification of Russian vowels spoken by different speakers. *The Journal of the Acoustical Society of America* 49(2B), 606–608. https://doi.org/10.1121/1.1912396
López Ixcoy, Saqijix Candelaria Dominga. 1994. *Las vocales en K'ichee'*. Guatemala: Oxlajuuj Keej Maya' Ajtz'iib'.
López Ixcoy, Candelaria Dominga. 1997. *Gramática K'ichee'*. Guatemala: Fundación Cholsamaj.
Mateo Toledo, Eladio. 2017. Q'anjob'al. In: *The Mayan languages*, Chap. 20, pp. 533–569. Ed. by Judith Aissen, Nora C. England, and Roberto Zavala Maldonado. 1st edition. London: Routledge Ltd.
Oshika, Beatrice, Victor Zue, Rollin Weeks, Helene Neu, & Joseph Aurbach. 1975. The role of phonological rules in speech understanding research. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 23(1), 104–112.
Par Sapón, María Beatriz & Telma Angelina Can Pixabaj. 2000. *Ujunamaxiik ri K'ichee' Ch'ab'al: Variación Dialectal en K'ichee'*. Guatemala: Oxlajuuj Keej Maya' Ajtz'iib', Cholsamaj.
Polian, Gilles. 2017. Morphology. In: *The Mayan languages*, Chap. 8, pp. 201–225. Ed. by Judith Aissen, Nora C. England, and Roberto Zavala Maldonado. 1st edition. London: Routledge Ltd.
Pye, Clifton, Barbara Pfeiler, & Pedro Mateo Pedro. 2017. Mayan language acquisition. In: *The Mayan languages*, Chap. 2, pp. 19–42. Ed. by Judith Aissen, Nora C. England, and Roberto Zavala Maldonado. 1st edition. London: Routledge Ltd.
Ramírez del Prado, Alejandro Curiel. 2017. Tojolabal. In: *The Mayan languages*, Chap. 21, pp. 570–609. Ed. by Judith Aissen, Nora C. England, and Roberto Zavala Maldonado. 1st edition. London: Routledge Ltd.
R Core Team. 2020. R: A language and environment for statistical computing. [Computer program]. Version 4.0.3. Vienna, Austria. https://www.R-project.org/
Robinson, Orrin Warner. 1976. A 'scattered' rule in Swiss German. *Language* 52(1), 148–162.
Romero, Sergio. 2009. Phonological markedness, regional identity, and sex in Mayan: The fricativization of intervocalic /l/ in K'iche'. *Variation in Indigenous Minority Languages* 25, 281–297.
Romero, Sergio. 2017. The labyrinth of diversity: The sociolinguistics of Mayan languages. In: *The Mayan languages*, Chap. 15, pp. 379–400. Ed. by Judith Aissen, Nora C. England, and Roberto Zavala Maldonado. 1st edition. London: Routledge Ltd.
Royer, Justin. 2021. Prosody as syntactic evidence: The view from Mayan. *Natural Language & Linguistic Theory* 40(1), 1–46.
Warner, Natasha. 2019. Reduced speech: All is variability. *Wiley Interdisciplinary Reviews: Cognitive Science* 10(4), e1496.
Wickham, Hadley. 2016. ggplot2: Elegant graphics for data analysis. New York: Springer-Verlag. https://ggplot2.tidyverse.org
Wood, Elizabeth. 2019. The K'iche' Collection of Elizabeth Wood. The Archive of the Indigenous Languages of Latin America (AILLA), ailla.utexas.org. PID collection: 563. https://www.ailla.utexas.org/collections/563/
Wood, Elizabeth Anne. 2020. *Vowels and stress in Chichicastenango K'iche'*. MA thesis. University of Texas at Austin.
Wood, Elizabeth. 2023. Vowel-initial words and glottalization: A corpus study of Chichicastenango K'iche'. In: *Proceedings of the 20th International Congress of Phonetic Sciences*, pp. 1766–1770.
Wood, Eilzabeth Anne. 2025a. An acoustic study of tense, lax, and glottalized vowels in Chichicastenango K'iche' (Maya). *Journal of the International Phonetic Association*, 1–37. doi:10.1017/S0025100325100698
Wood, Elizabeth Anne. 2025b. Phrase-final suffixes at the syntax–prosody interface: A corpus study of Chichicastenango K'iche'. *International Journal of American Linguistics* 91(4), 543–580. doi:10.1086/737188
Zsiga, Elizabeth C. 2000. Phonetic alignment constraints: Consonant overlap and palatalization in English and Russian. *Journal of Phonetics* 28, 69–102.

Cite this article: Wood Elizabeth Anne (2025). Predictable vowel deletion in content words and variable vowel deletion in function words in Chichicastenango K'iche'. *Journal of the International Phonetic Association* **55**, 465–502. https://doi.org/10.1017/S0025100325100765

https://doi.org/10.1017/S0025100325100765 Published online by Cambridge University Press