An Acoustic Phonetic Description of Ecuadorian Siona Monophthongs
Damonte, Oleksandra

Citation
Damonte, O. (2023). *An Acoustic Phonetic Description of Ecuadorian Siona Monophthongs*.

Version:	Not Applicable (or Unknown)
License:	License to inclusion and publication of a Bachelor or Master Thesis, 2023
Downloaded from:	https://hdl.handle.net/1887/3636617

Note: To cite this publication please use the final published version (if applicable).

An Acoustic Phonetic Description of Ecuadorian Siona Monophthongs

by

Oleksandra Damonte-Matveienko

s3225437

Submitted for the partial satisfaction of the requirements for the degree of

MASTER OF ARTS IN LINGUISTICS

specialization of

LANGUAGE DOCUMENTATION AND DESCRIPTION

at

LEIDEN UNIVERSITY

Written under the supervision of:
Dr. Martine Bruil
Dr. Bert Botma

Second reader:
Dr. Yiya Chen

June 2023

Abstract

In this thesis, I provide the first acoustic description of Ecuadorian Siona phonemic vowels, six oral vowels and six corresponding nasal vowels. Two phonetic dimensions – vowel height, measured through the first formant frequency (F_1), and vowel backness, measured through the second formant frequency (F_2) – are taken as the descriptors of vowel quality. These dimensions are used to illustrate the target vowels in their acoustic space. Vowel quantity, which refers to vowel duration, is also measured. For each target vowel, the mean frequencies of F_1 and F_2, as well as the durational means, are presented. In addition, the effects of different phonological environment on the realization of target vowels are investigated. Ultimately, I construct the acoustic vowel space for oral and nasal vowels, and I compare the acoustic properties of the two types of vowels. The results for oral vowels demonstrate six distinct qualities, similar to the ones described by Bruil (2014), with the back vowels appearing consistently lower in the vowel space. Nasal vowel space shows more variability and a general shrinking effect of vocalic contrasts. The findings also demonstrate a nasalization effect whereby all nasal vowels are on average lower than their oral counterparts. Finally, the analysis revealed that the vowels /i, ĩ/ are phonetically long, at least in the context that they appeared.

Table of Contents

Acknowledgements

My deepest gratitude goes towards my supervisors, Bert Botma and Martine Bruil, for the invaluable knowledge, patience, and support that they provided throughout my thesis journey. I am very lucky to have had not just one but two supervisors, each being a highly experienced and skilled professional in different fields of Linguistics and each providing a unique contribution to my thesis. I also express my gratitude to Yiya Chen for accepting the role of the second reader.

My appreciation extends to Leiden University and LUCL for giving me the financial opportunity to continue with my studies in the time of war and hardship faced by my people. If it were not for this great gesture of kindness, I would have not been able to write this thesis and obtain my degree in the foreseeable future.

I further express my humble appreciation to the Siona people of Ecuador who had shared and continue to share their language and their knowledge with the scientific community. I would also like to thank a Siona speaker and language consultant, Ligia Criollo, who worked with Martine Bruil and without whom this paper would not be possible.

Finally, I am very grateful to my family and friends for always believing in me and supporting me in difficult times. I thank my parents for leading by example, my partner for being my rock, and my mentor and friend, Marijn, for everything.

List of Glossing Abbreviations

2	second person
3	third person
ANIM	animate
ASS	assertive
CAUS	causative
CLS	classifier
CONTAIN	container
DEM	demonstrative
EXIS	(wide-scope) existential
F	feminine
FUT	future
IMPF	imperfective
INF	infinitive
INT	interrogative
LONG	long (classifier)
M	masculine
NEG	negation
N.ASS	non-assertive
OTH	other
PL	plural
PRS	present
PRX	proximate
PST	past
REP	reportative
RIVER	river (classifier)
S	singular
THIN	thin (classifier)
TRS	transtivizer

1. Introduction

Ecuadorian Siona (or Siona), an endangered Western Tukanoan language, spoken in the lowlands of eastern Ecuador, has been lacking a phonetic description of vowels. The first formal description of Ecuadorian Siona phonology was done by Bruil (2014) as a part of a larger documentation and description project of the language. Siona vowel inventory consists of six oral vowels /i, ɨ, u, e, o, a/ and six corresponding nasal vowels /ĩ, ɨ̃, ũ, ẽ, õ, ã/, as described by Bruil (2014: 87). The aim of this thesis is to complement the phonological description in Bruil (2014) with an acoustically based phonetic description in order to provide a more objective and precise description of the vowels of the language.

Typical for Tukanoan family, Ecuadorian Siona vowel system is based on six vocalic contrasts, namely /i, ɨ, u, e, o, a/; the same vocalic categories are found in sound systems of other related languages (cf. Silva, 2012; Johnson & Levinsohn, 1990; Stenzel, 2013). To my knowledge, no study focusing on acoustic properties of vowels has yet been published for Tukanoan languages. However, acoustic vowel space appears in a couple of descriptive grammars. In a description of Kubeo (Eastern Tukanoan), Chacon (2012: 20) constructs vowel space based on the speech of one male speaker, although mean formant values, which the chart is based on, are not provided, making it more difficult to compare to other languages. In a description of Desano (Eastern Tukanoan), Silva (2012: 34) provides both, mean formant values and the acoustic vowel space, based on the speech of two male speakers. One of the contributions of the current study is to better our understanding of the acoustic properties of the six-vowel system found in Tukanoan languages.

Siona vowel system constitutes an interesting case for the Tukanoan family since most of its languages do not possess inherently nasal vowels. The closely related Colombian Siona has been reported as a language with phonemic nasal vowels (Wheeler, 1987a, 1987b). For the rest of Tukanoan languages, the phonological reports are either contradictory or describe an exclusively contextual nasalization of vowels. Sekoya (Western Tukanoan) is first reported to only possess contextually nasalized vowels (Johnson & Levinsohn, 1990), and in a more recent publication (Vallejos, 2013), it is reanalyzed as possessing inherently nasal vowels. Similarly, Cook & Criswell (1993) provide two alternative analyses for Koreguaje (Western Tukanoan) – one in favor and the other one against phonemic nasal vowels. In Eastern Tukanoan languages, nasalization is not phonemic and is usually reported as an auto- or supra-segmental feature that is independent from the segmental level (see e.g. Chacon (2012) for Kubeo; Stenzel (2013) for Kotiria; Gómez-Imbert (2004) for Barasana and Tatuyo; Kaye (1971) for Desano).

In addition to contrastive nasal vowels, oral vowels can be nasalized in Siona as a result of nasal harmony, the process that is also found in other Tukanoan languages. However, the degree of nasalization between inherently nasal vowels and contextually nasalized vowels can differ significantly. For example, during articulation of phonemically nasal and contextually nasalized vowels in French, the difference in the movement of the soft palate (responsible for nasalization) between the two is drastic, so that the contrast between nasal and oral vowels is maintained (Cohn, 1990). This thesis therefore focuses on phonemic nasal vowels found in oral contexts, excluding the vowels that are nasalized from nasal harmony.

The aim of this thesis is to provide the first acoustic description of the twelve monophthong vowels of Ecuadorian Siona and to show them in an acoustic vowel space. The acoustic analysis of this study entails determining vowel quality (high/low and front/back) and vowel quantity (duration). Vowel height is measured through the first formant and vowel backness through the second formant, while duration merely stands for the temporal vowel length. Oral vowels and nasal vowels are measured separately and then compared. The effects of the phonological environment on vowel quality are also investigated. In the end, Ecuadorian Siona vowel system is represented in its acoustic vowel space, which provides more precise phonetic representations of Siona vowels. The findings support the idea that our understanding of phonological structures is augmented by phonetic studies.

Chapter 2 of this paper is dedicated to Ecuadorian Siona, namely language background (2.1), overview of transcription and orthography used in language examples (2.2), and a phonological sketch covering relevant phonological aspects (2.3). Chapter 3 lays out a general background of concepts pertinent to the study, namely vowel space (3.1), nasality (3.2), and stylistic and articulatory effects on the realization of vowels (3.3). Chapter 4 presents the methodology, viz. data description (4.1), and the filtering and measurements that were applied (4.2). Chapter 5 presents the results for oral vowels (5.1) and for nasal vowels (5.2), and then, compares the two types of vowels (5.3). Chapter 6 provides a conclusion and an in-depth discussion of the general findings (6.1), the case of the lowering of back vowels (6.2), and the occurrence of long vowels (6.3).

2. Ecuadorian Siona

2.1 Language background

Ecuadorian Siona[1] is a Western Tukanoan language that is spoken in the rural areas of the Sucumbíos province of Eastern Ecuador (see figure 1). The language variety is spoken by an estimate of 250 people (Mejeant, 2001) and is critically endangered. Most of the speakers live in six villages of the Sucumbíos province: Puerto Bolívar, Tarabëaya, Sototsiaya, Orahüeaya, Aboqüehuira and Bi'ãna. The Siona people call themselves *bãĩ*, meaning 'people', and their language – *bãĩkoka*, meaning 'the language of the people' (Bruil, 2014: 4-5).

The Siona people are native to Ecuador and Colombia and their language is closely related to Sekoya (native to Ecuador and Peru) (Chacon, 2014). Bruil (2014) proposes a split of the Siona language into Ecuadorian Siona and Colombian Siona based on their lexical, phonological and morphosyntactic differences. In some ways, Ecuadorian Siona is, in fact, closer to Sekoya than to Colombian Siona. For instance, Ecuadorian Siona and Sekoya both lost word-internal voiced velar stop, unlike the Colombian variety. Hence, Bruil (2014) proposes that the three varieties constitute a tripartite dialectal continuum, called Siona–Sekoya (or Baicoca–Siecoca), with Ecuadorian Siona in the middle.

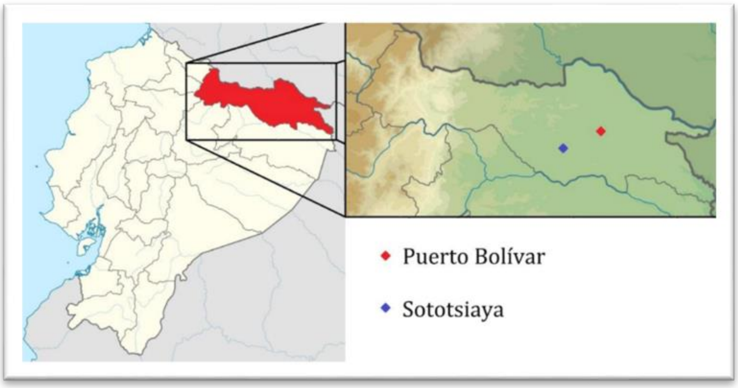


Figure 1. Province of Sucumbíos in red (left side); fieldwork sites of the language documentation project (right side) (Bruil, 2014: 5).

The first attempt at formal language documentation appeared in 1955 after Maria and Orville Johnson, two missionaries from the Summer Institute of Linguistics, arrived in Sucumbíos (Bruil, 2014: 4-6). Based on the Spanish orthography, the Johnsons developed a writing system for Siona, which is used by the native speakers to this day. Over time, the influence of Spanish on Siona reached critical mass. In Tarabëaya, for example, new generations no longer speak Siona. In Sototsiaya, however, the situation looks more optimistic: children acquire Siona as their first language, and Spanish when they go to school. External factors, such as tourism and missionary activity, as well as internal factors, such as migration and intermarriage, contributed to the spread of Spanish as a lingua franca and as a language of higher prestige, while giving Siona a status of a less important language. Fortunately, this did not go unnoticed, and people started to develop revitalization policies in different Siona communities (ibid.).

Earlier research on Siona has focused on the Colombian variety (Wheeler, 1967, 1970, 1987a, 1987b, 2000; Wheeler & Wheeler, 1975). More recently, a documentation project on Ecuadorian Siona was carried out by Bruil (2012), who collected a total of 124 minutes of audio recordings in the fieldwork sites of Puerto Bolívar and Sototsiaya (see figure 1) between June 2010 and September 2011, as well as in September of 2012. The annotated recordings are available in an online repository, the ELAR or Endangered Languages Archive (Bruil, 2012). Since then, Ecuadorian Siona has received more attention in the linguistic community; several studies have been published (Bruil, 2014; Bruil, 2015; Bruil, 2018; Bruil & Stewart, 2022), and one is forthcoming (Veer et al., in press). Most notably, Bruil (2014) presents a synchronic and diachronic description of the clause-typing system in Ecuadorian

[1] The ISO 639-3 code for Siona is [snn]; no distinction is made between the Ecuadorian and the Colombian varieties. Throughout the paper, 'Siona' is used to refer to Ecuadorian Siona unless the distinction is made between the Colombian and Ecuadorian varieties.

Siona with a sketch of the relevant aspects of the grammar for the historical analysis. To date, no studies have been carried out on the vowels of Siona. This thesis is the first to address this topic.

2.2 Transcription and orthography

Throughout the paper, all language examples are presented according to the conventions of phonetic transcription and adapted Siona orthography. Every language example can have up to five lines in the following order (from top): surface representation in IPA, underlying representation in IPA, morphologically parsed orthographic representation, followed by corresponding glosses, and English interpretation. Surface forms are presented in square brackets and underlying forms in slashes. The orthography used is not the original Siona orthography but the adapted version (Bruil, 2014: 129-132). In-text language examples are also shown in adapted orthography. Accordingly, /p̰, t̰, k̰, s̰/ are represented as <b, d, g, z> respectively, /ʔ/ is represented as <’>, /t͡ʃ/ as <ch>, and long vowels as <VV>. For the rest of the sounds, the adapted orthography coincides with the underlying representations in IPA. Generally, allophonic variation is not expressed in orthography, such as the palatal nasal [ɲ] is underlyingly /j/ and is thus represented as <j>.

2.3 Phonology

The following brief overview of Siona phonology is largely based on the description provided by Bruil (2014) and partly on a study by Bruil & Stewart (2022), which focuses on nasal harmony but also provides an updated version of Siona sound inventory. Some of the examples have been slightly adapted from the original source according to IPA, and phonological rules and orthographic conventions of the language. For example, underlying representations using IPA were added where it is relevant, and long vowels were changed from [VV] to [Vː]. The following sections present a brief overview of Siona consonants (section 2.3.1), vowels (section 2.3.2), prosody and syllable structure (section 2.3.3), nasal harmony (section 2.3.4), and laryngealization (section 2.3.5).

2.3.1 Consonants

The consonantal phoneme inventory of Siona, presented in table 1, is unusual due the presence of laryngealized phonemes, especially the laryngealized sibilant /s̰/, which is cross-linguistically very rare (Ladefoged & Maddieson, 1996: 178).

Table 1. The consonantal phoneme inventory (Bruil & Stewart, 2022).

		Bilabial	Denti-alveolar	Palatal	Velar		Glottal	
					Plain	Round	Plain	Round
Nasal		m	n					
Plosive	Plain	p	t		k	k^w		
	Laryng.	p̰	t̰		k̰	k̰^{w}	ʔ	
Affricate				t͡ʃ				
Fricative	Plain		s				h	h^w
	Laryng.		s̰					
Approximant		w		j				

More often, languages contrast voiced and voiceless obstruents, but in Siona there is a contrast between a series of plain /p, t, k, kʷ, s/ and laryngealized /p̰, t̰, k̰, k̰ʷ, s̰/ obstruents, illustrated in (1) for /p/ vs. /p̰/, and in (2) for /t/ vs. /t̰/.

(1) a. [paɨ]
/paɨɨ/
paɨ-ɨ
scare.off-2/3S.M.PST.N.ASS
'Did he scare (it) off?'
(Bruil, 2014: 88)

b. [p̰aɨ]
/p̰aɨ/
ba-ɨ
have-2/3S.M.PST.N.ASS
'Did he have (it)?'

(2) a. [te.ʔo]
/teʔo/
te'-o
one-CLS:F
'one woman'
(Bruil, 2014: 88)

b. [t̰e.ʔo]
/t̰eʔoo/
de'o-o
be.good-3S.F.PST.ASS
'She was good.'

Word-initially, the laryngealized consonants are realized as [p̰, t̰, k̰, k̰ʷ, s̰]; phonetically, this laryngealization spreads to the following vowel (more on laryngealization in Siona, see section 2.3.5). In other positions, the picture is not so uniform. Generally, word-initial [p̰, t̰] appear in complementary distribution with word-internal [β, ɾ] respectively. Bruil (2014: 92–95) postulates that /p̰, t̰/ undergo lenition when they appear intervocalically, more specifically in V(ʔ)_V position, as shown in (3).

(3) a. [p̰ã.βi.je]
/p̰ã.p̰i.je/
'to touch'
(Bruil, 2014: 93)

b. [t̰oʔ.ɾo.wɨ]
/t̰oʔ.t̰o.wɨ/
'basket'

Lenition of /p̰, t̰/ also tends to happen in rapid speech (example 4a) or at the beginning of the second word in a compound (example 4b). The realization of the laryngealized /t̰/ is further complicated for a set of suffixes, such as /t̰aʔ/, /t̰owɨ/ and /t̰aʔka/, where it can be realized as [t̰], [d̰] or [d] (Bruil, 2014: 94).

(4) a. [p̰eo.ɾoʔ.ɾo.wɨ]
/p̰eo.t̰oʔ.t̰o.wɨ/
beo-do'do-wɨ
NEG.EXIS-basket-CLS:CONTAIN
'containing nothing'
(Bruil, 2014: 94)

b. [ɨ.ha.βãɨ̃]
/ɨ.ha.p̰ãɨ̃/
ɨha-bãɨ̃
foreign-people
'non-Sionas'

The (labialized) velar stops /k̰, k̰ʷ/ are always realized as laryngealized [k̰, k̰ʷ], and intervocalically, they only occur at the beginning of the second member of a compound.[2] The sibilant /s̰/ is realized as laryngealized [s̰] in word-initial positions unless it starts the second part of a compound; then it is realized as plain [s], which leads to neutralization between /s/ and /s̰/, illustrated in (5). Additionally,

[2] The word-internal voiced velar stop [g] was lost in Ecuadorian Siona (Bruil, 2014: 11). Note that <g> appears in examples only as an orthographic representation of [k̰].

the sibilants /s, s̰/ can undergo occasional affrication and be realized as [t͡s, t͡s̰] (Bruil, 2014: 99–100), illustrated in (6).

(5) [soh.to.sia.ja]
/soh.to.s̰ia.ja/
sohto-zia-ja
clay-river-CLS:RIVER
‘Clay River’
(Bruil, 2014: 100)

(6) a. [sai.je] ~ [t͡sai.je] ‘To go’
b. [s̰ia.je] ~ [t͡s̰ia.ja] ‘River’
(Bruil, 2014: 100)

In addition to the surface forms discussed above, Siona also has [ɲ, w̃, h̃], which are allophones of /j, w, h/ in nasal context and emerge as a result of nasal harmony – the process considered in 2.3.4. When it comes to phonemic nasal consonants, there seems to be a split between the Western Tukanoan languages (e.g., Ecuadorian Siona (Bruil, 2014), Colombian Siona (Wheeler, 1987a, 1987b), Sekoya (Vallejos, 2013), and Koreguaje (Cook & Criswell, 1993)), which are reported to possess /m, n/, and the Eastern Tukanoan languages (e.g., Desano (Silva, 2012), Kubeo (Chacon, 2012), Kotiria (Stenzel, 2013), Barasana and Tatuyo (Gómez-Imbert, 2004)), which do not have underlyingly nasal segments at all.

2.3.2 Vowels

The phonemic vowel inventory of Siona is presented in table 2; it consists of twelve vowels: six oral vowels and six corresponding nasal vowels (vowels can also be nasalized through nasal harmony, discussed in 2.3.4). The vowels are organized in terms of phonological features based on Bruil & Stewart (2022; adapted from Bruil, 2014). The distinctive features that Siona vowels possess are [±nasal], [±front], and [±round], where only [–front] vowels are specified for [±round]. For the purpose of this study, these features are only used as a baseline for the hypothesized target vowel, such as /i/ is expected to be articulated as high and front.

Table 2. The vowel phoneme inventory (Bruil & Stewart, 2022).

	+front		–front			
			–round		+round	
	–nasal	+nasal	–nasal	+nasal	–nasal	+nasal
High	i	ĩ	ɨ	ɨ̃	u	ũ
Mid	e	ẽ			o	õ
Low			a	ã		

The six vowel qualities are commonly found in other Tukanoan languages but nasality as a distinctive feature of segments is not. Phonemic nasal vowels are sometimes included in sound inventories of Western Tukanoan languages (see e.g. Wheeler, 1987a; Vallejos, 2013) but are usually absent in the inventories of Eastern Tukanoan languages (see e.g. Stenzel, 2013; Chacon, 2012; Silva, 2012), the latter being a much larger group. This makes Siona vowel inventory twice the size of the vowel inventories of many Tukanoan languages. An example of an oral and a nasal vowel contrast in Siona is shown in (7) where /a/ in /kahka/ contrasts with /ã/ in /kãh/.

(7) a. [kah.kaɨ.ɲã]
kahka-ɨ-jã
enter-2/3S.M.PST.N.ASS-REP
'... (mas.) entered, they say.'
(Bruil, 2012: 20120914elicr008)[3]

b. [kãh̃.kɨ.ɲã]
kãh-kɨ-jã
sleep-2/3S.M.PST.N.ASS-REP
'… (mas.) slept, they say.'

Furthermore, Siona vowels can undergo vowel coalescence, vowel assimilation, vowel harmony, vowel deletion, reduction, and dissimilation (Bruil, 2014: 115–123). Various vowel combinations, forming diphthongs and long vowels, appear in the language and can be explained by some of these processes. For example, the occurrence of the diphthong [ɨo] is caused by vowel dissimilation: root-final /u/ dissimilates in the sequence /uo/ when the causative suffix *-o* is added.

Bruil (2014: 115) describes three types of vowel coalescence. The first type occurs when a suffix, made up of a single vowel, is added to a morpheme that ends in the same vowel, resulting in the fusion of the two vowels into one, illustrated in (8a). The second type of vowel coalescence happens when /i/ is attached to a monomoraic root that ends in /e, ẽ/, which results in the long vowels [eː] or [ẽː], as illustrated in (8b). The third type happens when the high vowel /ɨ/ is fused with one of the other high vowels /i, u/, resulting in /i, u/ respectively, as shown in (8c).

(8) a. [p̰õ.nɨ̃]
bõnɨ-ɨ
turn.around-2/3S.M.PST.N.ASS
'Did you (M) / he turn around?'
(Bruil, 2014: 115)

b. [weː.ko]
we-i-ko
lie.in.hammock-IMPF-3S.F.ASS
'She is lying in a hammock.'
(Bruil, 2014: 116)

c. [pũʔ.pu]
pũ'pu-ɨ
smoke-2/3S.M.PST.N.ASS
'Did you (M) / he smoke?'
(Bruil, 2014: 117)

Partial vowel assimilation frequently happens in diphthongs, namely /ɨ/ assimilates in backness and roundedness to the preceding vowels /e, o/, resulting in [ei] and [ou] respectively. As a monophthong, the vowel /ɨ/ is also harmonized to the vowel in the preceding syllable when they are separated by a glottal stop, illustrated in (9a) for partial harmony and in (9b) for full vowel harmony.

(9) a. [te.ʔi]
te'-ɨ
one-CLS:ANIM.M
'one man'
(Bruil, 2014: 118)

b. [kɨa.si.ʔi]
kɨa-si-ʔɨ
tell-FUT-OTH.ASS
'I am going to tell.'

In disyllabic roots, the final vowel is deleted (example 10a) or reduced (example 10b) when a derivational suffix *-a* 'transitive' or *-o* 'causative' is added. In (10a) the final vowel of the root is deleted and in (10b) the root-final /ku/ is reduced to [k^w] when the suffix *-a* is added.[4]

[3] This example is taken from the archived collection which is why the bundle name is given instead of the page number.
[4] The vowel reduction/deletion process is motivated by a bimoraic stem constraint (Bruil, 2014: 121), which is discussed in section 2.3.3 below.

(10)	a.	[jeʔ.ja.je]	b.	[õh.kʷa.je][5]
		je'je-a-je		ũhku-a-je
		learn-TRS-INF		drink-TRS-INF
		'to teach'		'to give someone something to drink'
		(Bruil, 2014: 119)		(Bruil, 2014: 120)

Vowel coalescence, as described for the example in (8b), is not the only source for the occurrence of long monophthongs in Siona. They also appear in stems that are derived when the causative suffix *-a* is added to monomoraic roots that end in identical vowel, as shown in (11). In addition, long vowels appear in verb roots, which Bruil (2014: 85) considers to have underlyingly bimoraic structure, such as the verb *kaa* 'say' illustrated in (12). [6]

(11)	[saː.sio]	(12)	[kaː.kɨʔ.nẽ]
	sa-a-si-o		kaa-kɨ-'ne
	go-CAUS-FUT-3S.F.ASS		say-2/3S.M.PRS.N.ASS-INT
	'She will take.'		'Do(es) you (m) he say?'
	(Bruil, 2014: 100)		(Bruil, 2014: 84)

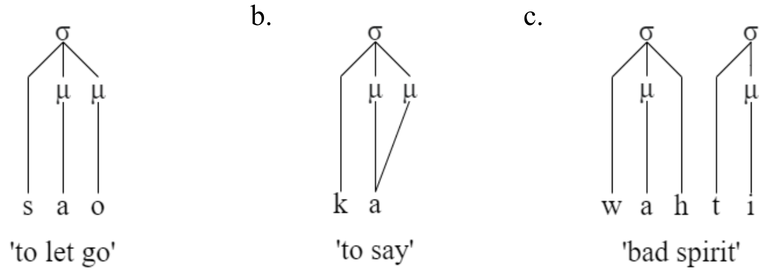
There is some occasional lowering of nearly all non-low vowels. As can be seen from the example in (10b) above, the initial vowel /ũ/ is realized as [õ], but it is unclear whether the lowering of /ũ/ is a productive process in Siona (Bruil, 2014: 120). Bruil & Stewart (2022: 8) mention that some vowels are often lowered, namely the mid vowels /e, ẽ, o, õ/ can be pronounced as the low-mid [ɛ, ɛ̃, ɔ, ɔ̃] and the high central vowels /ɨ, ɨ̃/ as the mid central [ɘ, ɘ̃], although it is unclear when the lowering occurs or what causes it.

2.3.3 Prosody and syllable structure

The basic syllable template of Siona is (C)(V)V(H), where H stands for a glottal stop /ʔ/ or a glottal fricative /h/, meaning that no other consonant can occupy coda position. Long vowels and diphthongs are generally restricted to open syllables, which means that syllable *CVVH is deemed ill-formed. Furthermore, Siona words do not end with a closed syllable except for borrowings, such as a Spanish loanword [moh.tor] 'motor'. Bruil (2014) analyzes syllables with two (identical) vowels as bimoraic (illustrated in 13a and 13b) and the coda H as non-moraic (illustrated in 13c). Similar to other Tukanoan languages, such as Barasana (Gómez-Imbert, 1997) and Kotiria (Stenzel, 2013), Ecuadorian Siona is analyzed as a language with a bimoraic stem constraint, which restricts stems to a minimum and a maximum of two moras (Bruil, 2014: 83–86).[7]

[5] Note that in this examples /ũ/ is lowered to [õ]; the lowering of vowels is discussed below.
[6] Whether the long vowels in this case should be considered underlyingly long has potential implications for the phoneme inventory, but for this study, it is sufficient to recognize that they appear at the surface level.
[7] There is however a set of trimoraic animal names that appear as an exception to the bimoraic constraint, such as *tahkado* 'hawk' and wãjũmi 'anaconda' (Bruil, 2014: 148); Bruil proposes that the suffixes *-do* and *-mi* may be frozen noun classifiers.

(13) a. b. c.



(adapted from Bruil, 2014: 84–85)

A stem can consist of a monomoraic root followed by at most one derivational suffix, such as the one in (13a) where *sao* consists of a monomoraic root *sa* 'go' and a derivational suffix *-o* 'causative'. Alternatively, a stem can consist of a bimoraic root, such as the ones in (13b) and (13c) where *kaa* and *wahti* represent bimoraic roots. Stems are for the most part same as roots, except for some verbs that have a monomoraic root, such as *sa* above. These verbal stems are derived using the causative suffixes *-a* and *-o* (for more on these verbs refer to Bruil, 2014, § 5.4.2).

Lastly, Siona is described as having neither stress nor tone, although an in-depth analysis of these issues is still needed (Bruil, 2014: 86). Among Tukanoan languages, some are described as tone languages, e.g., Kotiria (Stenzel, 2013), some as stress languages, e.g., Colombian Siona (Wheeler & Wheeler, 1975), and some as both, e.g., Kubeo (Chacon, 2012). There is currently no evidence indicating that Ecuadorian Siona relies on either stress or tone. Bruil (2014: 86) observed an intonational pattern in Siona whereby the pitch is said to rise on the last syllable of a word. In a more recent, preliminary study on intonation, Stewart and Bruil (2018) observed that Siona does not possess a "rich inventory of pitch contours" with almost all utterances falling into a standard "plateau pattern". For this thesis, I will assume that Siona does not display a reduced vowel quality in unstressed syllables (no such reduction has been observed in the sources that I consulted).

2.3.4 Nasal harmony

In addition to phonemic nasals, sounds in Siona can be nasalized as a result of nasal harmony, a process that is also found in other Tukanoan languages, such as Sekoya (Johnson & Levinsohn,1990), Tatuyo and Barasana (Gómez-Imbert, 2004), Tucano (Noske, 1995), Desano (Kaye, 1971), and Kotiria (Stenzel, 2013). The variables involved during nasalization, that is the direction, domain and the behavior of the segments, vary within the language family, especially when comparing the Western Tukanoan branch to the Eastern one (for an overview, see e.g. Botma, 2004, §3.2, or Bruil & Stewart, 2022).

In Siona, the nasality feature is said to spread from underlyingly nasal segments, specifically /ĩ, ɨ̃, ũ, ẽ, õ, ã/ and /m, n/. While nasal phonemes trigger nasal harmony, all oral obstruents block the nasality from spreading. In addition, newly grammaticalized suffixes also act as blockers; one such example is the plural suffix *-wa'i* (Bruil & Stewart, 2022: 23). The rest of the segments – the oral vowels, the approximants /w, j/, and the glottal consonants /ʔ, h, hʷ/ – are targeted by nasal harmony. According to Bruil & Stewart (2022), the glottal stop can fulfill an additional role of a transparent segment; when /ʔ/ is realized as laryngealization, it acts as a target, but when /ʔ/ is realized as a full stop,[8] the segment becomes transparent to the nasal harmony. To illustrate the different roles that segments and suffixes fulfill, consider examples in (14).

[8] More on realization of the glottal stop, see section 2.3.5 below.

(14) a. [ãĩ.h̃ĩ̵]
/ãi.hɨ/
'while we are eating'
(Brui & Stewart, 2022: 19)

b. [ãh̃.kɨ]
/ãh.kɨ/
'did you (M) / he eat?'
(Bruil, 2014: 128)

c. [mẽ.h̃ã.w̃ĩ̵]
/me.ha.wɨ/
'beach'
(Brui & Stewart, 2022: 17)

d. [pũʔ.pu.je]
/pũʔ.pu.je/
'to smoke'
(Bruil, 2014: 90)

e. [ĩõ.wa.ʔi]
/ĩo.wa.ʔi/
ĩ-o-wa'i
DEM.PRX-CLS:ANIM.F-PL
'they'
(Brui, 2014: 126)

In (14a), nasality spreads from the vowel /ã/ to the end of the prosodic word, while in (14b), nasality spreads from the same underlyingly nasal vowel until it is blocked by the velar stop. In (14c), all of the segments are targeted by the nasality spreading from the initial /m/. In (14d), the vowel /ũ/ fails to trigger nasal harmony because the bilabial stop /p/ blocks the nasality from spreading in both directions, while the glottal stop in this case acts as a segment transparent to nasality. Finally, (14e) illustrates that the newly grammaticalized suffix *-wa'i* does not undergo nasalization.

In Siona, the nasality spreads in a bidirectional manner, that is both progressively (towards the end of the word), and regressively (towards the beginning of the word). Progressive nasal harmony applies to the whole prosodic word, while regressive is restricted only to the syllable where the nasal segment occurs. Examples below demonstrate this difference: in (15b), the vowel /ã/ appears in the second syllable and triggers nasal harmony only in that syllable, whereas in (15d), the vowel /ã/ appears in the first syllable and triggers nasal harmony in both the first and the second syllable. In addition, example in (14c) above illustrates a trisyllabic word where nasality spreads from the first syllable to the end of the prosodic word.

(15) a. [ui.jo]
ui-jo
spear-CLS:LONG.THIN
'spear'

b. [ui.ɲõã]
ui-jo-ã
spear-CLS:LONG.THIN-PL
'spears'

c. [kaː.jɨ]
kaa-jɨ
say-OTH.PRS.ASS
'I/you/we/you (PL), they see.'
(Bruil, 2014: 125-126)

d. [ɲãː.ɲĩ̵]
jãã-jɨ
see-OTH.PRS.ASS
'I/you/we/you (PL), they see.'

Furthermore, examples in (15) demonstrate a complementary distribution between an oral sonorant and its nasal counterpart. In particular, the palatal approximant /j/ is realized as [j] in oral context (examples 15a and 15c) and as [ɲ] in nasal context (examples 15b and 15d). Nasality feature can then create additional nasalized allomorphs. For example, the suffix *-jo* has an oral allomorph [jo] in an oral context (example 15a) and a nasalized allomorph [ɲõ] when nasal harmony occurs (example 15b). In addition to inherently nasal segments, Siona also exhibits inherently nasal suffixes, such as the plural suffix *-ã* (example 15b).

2.3.5 Laryngealization

As already mentioned, the consonants /p, t, k, kʷ, s/ are realized as laryngealized [p̰, t̰, k̰, k̰ʷ, s̰] in word-initial positions and some of them also in compounds. Importantly, the creaky voice from these consonants tends to spread to the beginning of the following vowel, making the vowels partially creaky, exemplified in (16).

(16) [p̰ḭa] 'pepper'
[p̰ḛo.je] 'to not be/have'
(Bruil, 2014: 93)

The glottal stop is another segment that can manifest itself as laryngealization. Bruil (2014: 96) states that the glottal stop is frequently articulated without the full occlusion; instead, it is "realized as a creaky voice on the vocalic stream". Creaky phonation usually happens in intervocalic positions, such as in (17), while a complete closure before consonants, such as in (18). The absence of a full closure in glottal stops, especially in intervocalic positions, is also common cross-linguistically (Ladefoged & Maddieson, 1996: 75).

(17) [ma.ʔa] 'path'
[jɨ.ʔɨ] 'I'
(Bruil, 2014: 96)

(18) [kʷaʔ.ko.je] 'to cook'
[jaʔ.hi.je] 'to ripen'
(Bruil, 2014: 96)

Bruil (2014: 98) observes that the glottal stop, as a non-contrastive segment, occasionally appears in word-initial positions, such as in (19), but in fast speech it tends to dissapear. Considering that [ʔ] in this position appears before a vowel and not a consonant and is likely to be preceded by a vowel in connected speech, it is reasonable to assume that here also, [ʔ] has a tendency to lack a full occlusion.

(19) [ʔiha] ~ [iha] 'foreign'
(Bruil, 2014: 98)

3. Phonetic Properties of Vowels

3.1 Vowel space

Vowels are distinguished from one another through their quality, as well as quantity, which is an umbrella term for the position of the tongue, the lips, and the lower jaw. Changes in these articulatory organs can result in audible differences, producing different vowels. Conventionally, vowel quality is described in terms of two dimensions that create vowel space: high–low (a.k.a. close–open) and front–back. Originally, these labels were used to describe the tongue body position (relative to the roof of the mouth): in [i] the tongue is high and front, while in [ɑ], it is low and back. However, Ladefoged (2006) clarifies that while these labels were originally created as articulatory descriptions of the tongue position, they are not absolute descriptions thereof. Alternatively, the terms ‘close’ and ‘open’ refer to the position of the lower jaw: the jaw is close (to the roof of the mouth) in the vowel [i], whereas in [ɑ] the jaw is open. None of these labels, however, represent precise articulatory descriptions, which becomes more evident with the help of imaging techniques, such as X-rays and MRI scans. When phoneticians originally described vowel space based on these two dimensions, they thought they were describing the tongue body position, but they were, in fact, describing auditory quality of vowels (Ladefoged, 1967; Ladefoged, 2006).

Ladefoged (2006) describes vowel sounds as a continuum, as one vowel can glide into another. Despite the fluid nature of vowels, every language creates distinct vowels that contrast with each other (i.e., vowel phonemes), generating different word meanings. Vowels can be distinguished from one another based on their quality and/or quantity.[9] In Hungarian, for example, vowels have seven distinct qualities, which also contrast in length, creating minimal pairs such as /vis/ ‘carry’ and /viːʒ/ ‘water’. Some short vs. long vowel pairs in Hungarian also have a somewhat different quality, namely /ɛ/ contrasts with /eː/ and /ɑ/ contrasts with /aː/ (Szende, 1994).

Cross-linguistically, five vowel systems that we find in Spanish and Swahili, and six vowel systems found in Tukanoan languages, are more common than systems with more or fewer vowels (sample of 3,020 languages from Moran & McCloy, 2019). Jalapa Mazatec, too, has a five-vowel system, illustrated in an auditory/acoustic vowel space in figure 2 (righthand side). Jalapa Mazatec vowels are dispersed in their acoustic space, creating audibly very distinct vocalic categories. While cross-linguistic parallels are drawn between vowel systems, the precise vowel space is still language-specific, and the labels high–low and front–back rather describe how vowels sound relative to one another within a language. Hence, vowel space represents *relative* vowel quality (Ladefoged & Maddieson, 1996; Ladefoged, 2006).

[9] Of course, tone can also create meaningful contrasts, but it is not discussed in this paper.

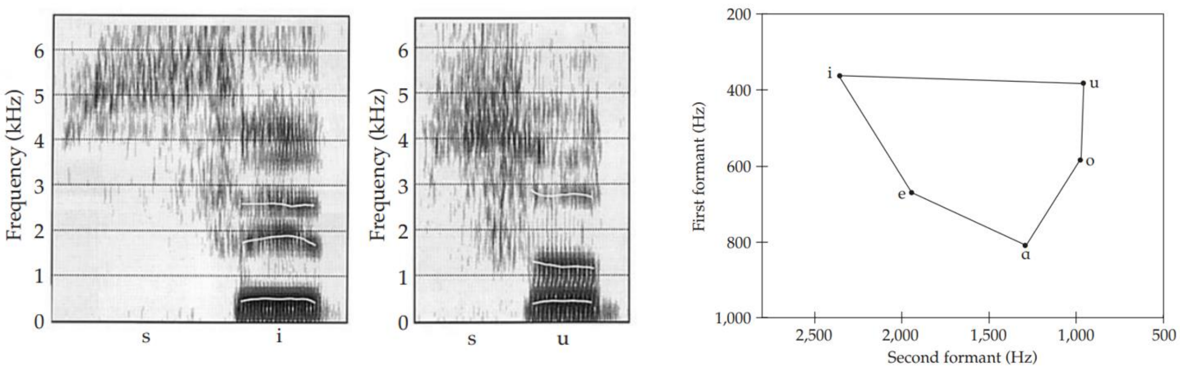


Figure 2. An example of spectrograms and vowel space in Jalapa Mazatec. On the left: Spectrograms of the Jalapa Mazatec words, [si] 'dirty' and [su] 'lukewarm' (the first three formants of the vowels are highlighted by the white lines). On the right: vowel space in Jalapa Mazatec (Johnson, 2012: 143-144).

The particular shape of the vowel space in figure 2 (righthand side) is of course not arbitrary; its dimensions are based on the phonetic properties of vowels. Specifically, vowel height is correlated with the frequency of the first formant (F_1), whereas vowel backness is correlated with the frequency of the second formant (F_2). Frequency is the rate with which a sound wave is repeated, normally expressed in hertz (Hz) – the cycles of the waveform per second (cf. Johnson, 2012). The vocal tract has differing resonant frequencies, called formants (F_1, F_2, F_3, etc.). Spectrograms can be understood as a temporal slice of a sound, representing frequency (y-axis) over time (x-axis) and intensity (relative darkness). Spectrograms of two Jalapa Mazatec words are shown in figure 2 (lefthand side); note that the vowels have similar F_1, but F_2 is considerably higher in [i] than in [u], which is reflected in the vowel chart on the right.

With the development of spectrographs, Joos (1948) was one of the first to notice the correlation between acoustic analysis, specifically between the spectrographic analysis, and vowel properties. Consequently, Peterson & Barney (1952) demonstrate that perceptual vowel categorization is correlated with its first and second formants. Lindau (1978) examines both articulatory and acoustic correlates of vowel features and concludes that the two features High (or height) and Back (or backness) are indeed the most prevalent contrastive features cross-linguistically, and that they receive a special place in vowel description.

Based on Ladefoged (1975), Lindau (1978) supports the idea that the Back feature is better correlated with the difference in the frequencies of F_1 and F_2 rather than just F_2. The difference between F_1 and F_2 is said to reflect the auditory perception more accurately and results in a vowel chart with a greater resemblance to the positions of cardinal vowels[10] in the chart. This presents an alternative way of describing the vowel system in a language that would reflect on the vowel contrast perceived by the listener. The current study, however, makes use of pure F_1 and F_2 measurements to describe vowel space, as it is a well-grounded method, the results of which are comparable cross-linguistically.

Since the 1970s, many more phonetic papers were published describing acoustic vowel space in various languages and dialects, as well as making cross-linguistic comparisons. For example, Klein et al. (1970) describe acoustic vowel space in Dutch, while Lindau et al. (1976) compare acoustic spaces in two Kwa languages, Yoruba and Ewe. Deterding (1997) compares vowel space based on connected speech of BBC broadcasters to the vowel space from Deterding (1990), which was based on citation speech of British English speakers. Studies on vowel space readily appear among larger, national

[10] Cardinal vowels, first developed by Daniel Jones, are standardized vowel categories located at the most extreme or peripheral points on the vowel chart, often used as a reference when describing vowels in a language (see e.g. Ladefoged, 2006: 211–218).

languages, such as French (e.g., Carignan, 2014), Finnish (e.g., Iivonen, 1995), and especially English dialects (e.g., Deterding, 1997; Ferragne & Pellegrino, 2010; Pillai et al., 2010). At the same time, vowel space has been described for many smaller, indigenous languages, such as Chickasaw (Gordon et al., 2000), Plains Cree (Muehlbauer, 2012), and Nungon (Sarvasy et al., 2020); no such studies have so far been published for Tukanoan languages.

Admittedly, formant-based vowel space is a reliable tool in determining more precise vowel realizations, which has aided researchers in fine-tuning vowel systems. For example, Tuttle et al. (2011) provide a more precise description of the Upper Tanana vowel system via this method, establishing the presence of the mid central unrounded vowels (/ɘ/ and /ɘː/), which is in contrast to previous studies reporting on a very different quality of this vowel. In a phonetic study of Suzhou Chinese vowels, Feng (2009) describes the so-called "fricative vowels" that have lower F_2 than their oral counterparts and demonstrates three distinguishable vowel heights, contrary to the five-way distinction previously described.

In addition to height and backness, there are other properties that can add to the description of vowel quality, such as rhotacization, rounding, tongue root advancement, and nasalization (cf. Ladefoged, 2006). One of these (secondary) features is of particular interest in Ecuadorian Siona, namely nasalization; vowel nasalization is the topic of the next section.

3.2 Nasality

3.2.1 On the typology of nasality

Cross-linguistically, phonemic nasal consonants are very common (96% of the world's languages have [m] and 78% have [n]), while phonemic nasal vowels are relatively uncommon (18% and 17% of the world's language have [ĩ] and [ã] respectively) (sample of 3,183 languages from Moran & McCloy, 2019). Aside from height and backness, nasality is the most common phonological feature, appearing in at least one in five languages (Maddieson, 1984). Languages with contrastive nasal vowels have either an equal number of nasal and oral vowels or a smaller number of nasal than oral vowels in their sound systems (Maddieson, 2007). Among languages of Central and South America (sample of 96 languages from Maddieson, 2007), nearly half of them (45 languages) have phonemic nasal vowels, among which an equal number of oral and nasal vowels is more common (32 out of the 45).

Phonetic nasal coarticulation on vowels neighboring nasal consonants is extremely common if not universal, although the acoustic patterns of nasalization vary across languages (cf. Hajek, 1997; Cohn, 1990). In English, for example, anticipatory nasalization has more variability and is more "pervasive" (i.e., the entire segment may be nasalized) than in French and Sundanese (Cohn, 1990: 306). Cohn (1990) measured oral and nasal airflow in English, French and Sundanese, demonstrating that inherently nasal vowels are nasal throughout or for the most part of the vowel duration, while contextually nasalized oral vowels have a lot more variability. For French vowels, in particular, the difference in the velum movement between phonemically nasal and contextually nasalized vowels was drastic, so that the contrast between nasal and oral vowels is maintained.

Furthermore, vowels have a degree of nasality, which means they can be more or less nasal. Some languages make a phonetic three-way surface distinction between oral, "lightly nasalized", and "heavily nasalized" vowels (Ladefoged & Maddieson: 1996: 298). For example, Ladefoged (1971) points out this three-way contrast for Palantla Chinantec, demonstrating a greater rate of nasal airflow in fully nasalized vowels than in partly nasalized vowels. Usually, this contrast appears in a language that has contextually nasalized vowels in addition to phonologically nasalized vowels, the latter being more nasal.

When nasality spreads beyond adjacent segments in a systematic manner, we are dealing with nasal harmony, although the variables involved differ from language to language (see e.g. Walker, 2011). For example, in Eastern Tukanoan languages nasal harmony is more common and uniform than in Western Tukanoan languages, to which Siona belongs (section 2.3.4 summarizes nasal harmony in Siona). In the Eastern Tukanoan branch, nasal harmony has been generally described as a suprasegmental feature, operating at the morpheme level; this is illustrated in (20) where an oral and a nasal morpheme (as opposed to segments) create a minimal pair in Desano. Western Tukanoan languages show a lot more variation in the domain: nasal harmony can operate at the level of the segment, syllable, morpheme, and word (cf. Botma, 2004, §3.2; Walker, 2011; Bruil & Stewart, 2022).

(20) *Desano*
[w̃ãĩ] ‘name’ [wai] ‘fish’
(Kaye 1971: 37)

Languages also differ with regards to which segments are harmonized and which ones are neutral to the harmony. Harmonic segments can act as triggers (initiators of nasal harmony) or targets (subjects to harmony); neutral segments can be opaque (blockers) or transparent (invisible) to nasal harmony. Another varying factor has to do with the direction: nasality can spread leftward (regressive), rightward (progressive), or bidirectionally. Sometimes stress is an important variable in nasal harmony.

(21) *Guaraní*
[ʔĩɲãˌkãɾ̃ãˈku]
/ijaˌkãraˈku/
‘is hot-headed’
(Gregores & Suárez, 1967: 69)

For example, in Guaraní, stressed nasal vowels trigger nasal harmony while stressed oral vowels block it; voiceless obstruents are transparent to the harmony and do not block it as they do in Siona (Walker, 2011). Nasal harmony in Guaraní is illustrated in (21) above.

3.2.2 Nasalization in vowels

This section explores the effects of nasalization in vowels: how nasal vowels differ from oral vowels acoustically, perceptually, and articulatory, what happens in the nasal vowel space, such as how nasalization affects vowel height and backness, whether duration plays a role, and how can acoustic effects be accounted for by nasal articulation.

Nasal vowel articulation is achieved through the lowering of the soft palate, allowing the air to flow through the nose – the opposite of oral vowel articulation, whereby the velum is raised. During articulation of nasal vowels, both the oral cavity and the nasal cavity are open, the nasal tract becoming a side-branch resonator of the oral tract, which leads to the development of additional, nasal formants (N_1, N_2, etc.) and nasal anti-formants (NZ_1, NZ_2, etc.) (Harrington, 2010). The additional formants of nasal vowels have some acoustic consequences for the oral formants at lower frequencies. On one hand, F_1, N_1, and NZ_1 can merge and become indistinguishable, which makes the bandwidth of the first formant broader. On the other hand, nasal formants can be visibly distinct from the oral formants on a spectrogram, usually appearing between F_1 and F_2 in high vowels (ibid.).[11]

[11] The concept of oral formants (F_1, F_2, etc.) was introduced in 3.1, and F_1 and F_2 are taken as the main measurements of vowel quality; oral formants are also referred to as ‘formants’.

The vowel contrasts tend to be neutralized in nasalized vowels, such as when the distinction between *pin* and *pen* is lost in some American English dialects (Johnson, 2012: 201). In a study on vowel perception, Wright (1986) describes a general shrinking effect of the vowel space for nasalized vowels. Wright (1986) recorded the production of American English vowels as a set of oral and nasalized vowels with the only difference in the velum position, using optical palatography to ensure the correct position of the tongue body. Listeners were presented with these oral-nasal pairs to judge their similarity. The results reveal that the listeners' perceptual vowel space was consistently different along the height dimension: low nasal vowels were perceived higher, while high and mid nasal vowels – lower, than their oral pairs. Wright postulated that the perception of nasal vowel height might be affected by the relative position of the first nasal formant to the first oral formant.

Furthermore, the coupling of the nasal aperture and oral aperture creates an opening of the velopharyngeal port (VP), referred to as VP coupling, and has some acoustic consequences for formant frequencies, F_1 in particular (cf. Carignan, 2014). Feng & Castelli (1996), and Serrurier & Badin (2008) demonstrate that the coupling of nasal and oral cavities results in centralization of French vowels along the height dimension: F_1 is increased for high vowels and decreased for low vowels. In another study on French vowels, Carignan (2014) concludes that the effect of VP coupling on the formant frequencies is enhanced by various "oral articulatory configurations". Carignan (2014) also concludes that while generalizations can be made for all (twelve) speakers, there are idiosyncratic differences with regards to the "oral articulatory strategies" used to produce oral-nasal contrasts. Generally, the results demonstrate a decrease in F_1 from the oral [a] to the nasal [ɑ̃], and an increase in F_1 from [ɛ] to [ɛ̃], which is in line with the shrinking effect along the height dimension. However, the nasalization effect from the vowel [o] to the vowel [ɔ̃] varied across speakers: a decrease for F_1 was observed for eight speakers, an increase of F_1 – for two speakers, and no change in F_1 – for another two speakers.

A centralizing effect along the front-back dimension, which can be attributed to VP coupling, has also been demonstrated, although in a lot less consistent fashion. French non-back nasal vowels were reported to have a decrease in F_2 in studies using simulation models (Feng & Castelli, 1996; Serrurier & Badin, 2008) and acoustic analysis of experimental data (Carignan, 2014). A centralization effect was also reported for English non-low front nasal vowels in the perceptual study by Wright (1986). When it comes to the nasalization effects on back vowels, however, there is no uniform pattern. Wright (1986) reports an increase in F_2 for [õ], and a slight decrease in F_2 for [ũ] and [ʊ̃] in the perception of English vowels. Similarly, French speakers (for the most part) produced [ɔ̃] as more back, i.e., with a lower F_2, than [o] (Carignan, 2014). Contrary to these findings, Feng & Castelli (1996) find more fronting, or increased F_2, for non-low back vowels.

It becomes clear that the effects of nasalization in vowels are more pronounced and consistent in vowel height than in vowel backness. In a cross-linguistic study on the effects of nasalization on vowel height, Beddor (1983) examined 75 languages, taking into account vowel height, vowel backness, and vowel context (i.e., whether a vowel is contextually nasalized or underlyingly nasal). She concludes that low nasal vowels tend to raise, whereas high nasal vowels tend to lower. For mid nasal vowels, the split is made between contextual and non-contextual (i.e. phonemic): non-contextual vowels tend to lower, while contextual are further divided into front and back. Back (mid, contextual) nasal vowels tend to be raised, while front (mid, contextual) tend be lowered ("unless the mid *back* vowel is raised in that language, in which case *both* front and back vowels raise") (Beddor, 1993: 189). These typological patterns established by Beddor (1983) are summarized in table 3 below. It is important to keep in mind that these are tendencies that languages exhibit and idiosyncratic cases that do not fit these patterns do exist.

Table 3. Vowel nasalization patterns of the 75 languages from Beddor (1983), as summarized by Beddor (1993: 189).

Vowel Height	Vowel context	Vowel Backness	Nasalization effect
High nasal vowels	NA	NA	Lower
Low nasal vowels	NA	NA	Raise
Mid nasal vowels	Non-contextual	NA	Lower
	Contextual	Back	Raise
		Front	Lower

Lastly, vowel duration can act as a cue for oral-nasal contrasts. Experimental studies on perception of nasal vowels by French and American English speakers report that the participants consistently identified shorter vowels as oral and the corresponding longer vowels as nasal (Beddor, 1993). The findings suggest that perceived vowel nasality is augmented by increased duration. This durational effect is also reported to be vowel-independent and language-independent. In French, nasal vowels are indeed longer, so it is not surprising that French speakers perceived longer vowels as nasal. In English, however, contextually nasalized vowels are shorter than their non-nasalized pairs, suggesting that the durationally-enhanced perception of nasality is language-independent (ibid.).

3.3 Stylistic and articulatory effects on vowel realization

Aside from nasalization, the precise vowel pronunciation depends on a number of factors, and it is crucial for the sake of this study to understand and account for these effects as much as possible. Therefore, in this section, I explore the potential effects of speech style, data type (such as words or sentences), phonological environment, and phonation type on realization of vowels; the effects of vowel length on vowel quality are also discussed. This section does not present an exhaustive list of factors that influence how vowels are realized; for example, reduced vowel quality in unstressed syllables is not discussed here since it is assumed that stress does not play a role in Siona (see section 2.3.3), and between-speaker variation is also not mentioned since this study includes only one speaker.

Among studies on vowel space, data can represent clear or citation speech (e.g., Klein et al., 1970; Choi, 1991; Iivonen, 1995; Yusuf et al., 2021), and less frequently, connected, naturalistic speech (e.g., Deterding, 1997; Muehlbauer, 2012). Clear speech is typically characterized by over-articulation of sounds, often in a form of syllables, citation speech represents prompted speech, such as utterances read in isolation, and naturalistic speech strives to represent how people talk in real life. Over-articulated speech can have an overshoot effect: vowels tend to have longer duration and less reduced quality when compared to more "normal" or citation speech (Moon & Lindblom, 1994). Connected speech is generally associated with more articulatory simplification (such as when /r/ is vocalized to [ɐ]) and sound coarticulation, which minimizes the transitions between adjacent sounds (Farnetani & Recasens, 2010).

Natural speech has a lot of stylistic effects and variation; speakers of different social backgrounds adapt their speech according to the circumstances (cf. Labov, 1985). According to the H&H theory, coined by Lindblom (1990), speakers may either hyper- or hypoarticulate depending on the perceptual needs of the listeners. Hypoarticulation refers to more relaxed speech, which requires less energy and is associated with more simplification and reduction of sound qualities, whereas hyperarticulation refers to clearer, often slower speech with rather exaggerated articulation; the two represent opposite endpoints of a continuum. For instance, in noisy environments, speakers are likely to hyperarticulate, but if the conditions are optimal for transmission of speech, they default to low energy hypoarticulation. In Siona, for example, there is an occasional articulation of [ʔ] in word-initial position if there is no initial consonant, as in [ʔiha] 'foreign'; however, in fast speech, the glottal stop tends to disappear, as in [iha] 'foreign' (Bruil, 2014: 98). As such, [ʔiha] could be considered a case of hyperarticulation, while [iha] – hypoarticulation.

The corpus data available for Siona presents a choice between citation speech (elicited word pairs and utterances) and connected, naturalistic speech (dialogues and narratives). While citation speech may not accurately represent how the language is spoken, it allows for some control of stylistic and coarticulatory effects on the realization of vowels. Natural-like speech is more likely to be hypoarticulated, such as produced more rapidly and potentially with some reduction in vowel quality. Even though elicited speech is not immune to such changes in articulation, it has less variability compared to natural speech.

Furthermore, citation speech can be in a form of word lists (e.g., Peterson & Barney, 1952; Iivonen, 1995; Gordon et al., 2000; le Roux & le Roux, 2008) or in a form of words inserted into carrier sentences (e.g., Harshman et al., 1977; Pillai et al., 2010; Carignan, 2014; Yusuf et al., 2021). For example, in a study on tongue shapes in English vowels, Harshman et al. (1977) use the /hVd/ word shape, inserted into a carrier sentence 'Say *h*(vowel)*d* again'. In an acoustic and articulatory investigation of Northern Metropolitan French vowels (Carignan, 2014), the speakers were recorded in a laboratory setting saying the carrier sentence *Il retape X parfois* 'He retypes X sometimes', where X is the inserted lexical item with the target vowels. Although not explicitly stated in these studies, the reason to opt for carrier sentences might be due to potential prosodic effects. Reading long repetitive wordlists may produce unwanted prosodic changes. Carrier sentences are also repetitive, but they allow to control for prosodic position of the target word within a sentence, making the intonational patterns more consistent across the data.

The word shape /hVd/ mentioned in the previous paragraph is a deliberate choice and has been commonly used in studies on English (e.g., Peterson & Barney, 1952; Ferragne & Pellegrino, 2010); similarly, /hVt/ was used in at least one study on Dutch (Klein et al., 1970). The choice for this word shape lies in the minimal effects of the preceding consonant on the vocalic nucleus. In particular, the voiceless glottal fricative adapts to the following vowel so much that the vowel is said to reach its target formant frequencies better than in other environments. In an experimental study on Finnish vowels, Iivonen (1995) compared the formant frequencies in syllables /tVt/ and /hVh/[12]. The target vowels were considerably more peripheral in /hVh/ than in /tVt/. Specifically, more close and more front positions were achieved for the vowels /e/, /ø/, /i/, /y/ in /hVh/ than in /tVt/ context. For the rest of the vowels, namely /u/, /o/, /a/, /æ/, the /tVt/ context yielded points nearer the upper left corner, while the /hVh/ context generated more peripheral points in the vowel chart. This can be attributed to the position of the tongue during the consonantal articulation.

While [h] presents an ideal environment due to its minimal effects on the vowel, approximants, nasals and rhotics are known to have strong coarticulatory effects. Vowels adjacent to nasal consonants are likely to be (partially) nasalized. Semivowels, such as [w] and [j], have particularly severe coarticulation with the following vowel due to their proximity to vocalic sounds. For example, phonetic and auditory differences between syllable [wa] and diphthong [ua] can be hardly discernible. Overall, postalveolar consonants tend to exhibit more tongue-body coarticulation than alveolar consonants (Farnetani & Recasens, 2010). With regards to the manner, fricatives tend to have the least degree of coarticulation, followed by stops, followed by liquids (ibid.). Furthermore, the degree of C-to-V coarticulation differs depending on the vowel. Recasens (1991) reports that dental, alveolar, and palatal contexts have larger coarticulatory effects on back vowels and schwa than on front vowels, explained

[12] In Finnish, syllable /hVh/ appears only as interjections and syllable-final /h/ is possible in other words.

by tongue dorsum raising and fronting. Greater variation in vowel articulation was also correlated with greater variation in F_2, and [ə] was reported to have the highest variability (ibid.).

Vowel duration is another factor that can influence vowel quality. Lindblom (1963) was the first to demonstrate that reduced vowels are associated with decreased duration, and vice versa, i.e., increased vowel duration is correlated with the acoustic target for which the vowel formants aim. In a study on the Plain Cree vowel system, Muehlbauer (2012) measured the first two formants of the hypothesized short and long vowels in a frequent segment sequence /kVt/, concluding that long vowels consistently occupy more extreme or peripheral positions in the vowel space than their short counterparts. In a description of phonetic structures of Chickasaw, Gordon et al. (2000) examine vowel formants in three vocalic groups: phonemically short, phonemically long, and rhythmically lengthened vowels.[13] The general pattern for Chickasaw vowels is in line with other research: long vowels (phonemic or not) tend to occupy more peripheral positions, while short vowels occupy relatively centralized positions in the vowel space. Iivonen (1995), Tuttle et al. (2011), and Iivonen & Harnud (2005) show similar results with long vowels occupying more peripheral positions. Contrary to these findings, Kharlamov & Oberly (2021) report that the effects of phonemic vowel length on F_1 and F_2 were statistically not significant in Southern Ute, despite the significant differences in vowel duration.

Finally, different phonation types, i.e., creaky, breathy, and modal (or regular) voice, may also have different effects on vowel quality. The effects of creaky voice are of particular interest here, since in Siona, vowels following laryngealized consonants are realized as (partially) creaky (see section 2.3.5). During creaky voicing, the vocal cords are more tense or compressed together, and the larynx is more constricted than during modal voicing (cf. Ladefoged, 2001; Johnson, 2012). The muscular tension characteristic for creaky voice often makes the vocal cords vibrate more slowly, resulting in a lower fundamental frequency. This difference is visually apparent on a spectrogram: the vertical bands representing pulses of the vocal cords are further apart in creaky voice than in modal voice. In terms of quantifying phonation types, the relation between the second and first harmonics[14] has proven to be a reliable measurement (cf. Johnson, 2012), but what about the formant frequencies?

Examining the three phonation types in Jalapa Mazatec, Kirk et al. (1993) report that the formant frequencies generally have the same values across phonation types, but F_1 may appear at a slightly higher frequency in the creaky vowel than in the modal vowel. Analyzing the waveforms, they observe that the distance between pulses is slightly larger and the amplitude of the wave denoting F_1 is greater in creaky vowels than in the modal ones. Furthermore, creaky vowels are consistently longer than their modal counterparts (at least in Jalapa Mazatec), and, as already stated, longer vowels reach their target better. Lastly, formants of creaky vowels show more energy (especially at higher frequencies) than formants of modal vowels (Kirk et al., 1993; Ladefoged, 2001).

[13] In Chickasaw, the duration of rhythmically lengthened vowels falls between phonemically short and phonemically long vowels.

[14] Harmonics are produced by the movements and shapes of the vocal cords, whereas formants by the different configurations of the whole vocal tract.

4. Methodology

4.1 Data description

The recordings used for the analysis were collected by Martine Bruil during her fieldwork in 2012 as a part of a larger documentation project and can be accessed in an online repository, ELAR (Bruil, 2012). The analyzed speech was produced by a native Siona speaker and language consultant Ligia Criollo (female, 52) during elicitation sessions in September 2012. The elicitation sessions consisted of a reading task, which was carried out in the kitchen of the consultant's house (Puerto Bolívar cite, see figure 1, § 1), and were recorded on a Marantz PMD620 device. The analyzed recordings therefore represent citation speech, which was deemed preferable to the more natural, spontaneous speech due to potential effects on the realization of sounds.

The target vowels represent six phonemic oral vowels /i, ɨ, u, e, o, a/ and six phonemic nasal vowels /ĩ, ɨ̃, ũ, ẽ, õ, ã/. Cases where realization of vowels is affected by the (morpho-)phonological processes described in Siona (see section 2.3) were minimized in the data. Hence, contextually nasalized vowels that surface as a result of nasal harmony were not a part of the data; likewise, long vowels were avoided as much as possible. However, it was not always possible to predict the variation in vowel realization, such as when non-low vowels are occasionally lowered.

The analyzed data largely consist of declarative clauses in the form of carrier sentences, less frequently word pairs or other elicited sentences, and two instances of elicited questions. The carrier sentences show some uniformity in that they reoccur across different vowels in the data. For example, a carrier sentence *Go'je mo'se* ____ *si'awa'i* 'Everyone ____ yesterday', where a verb is inserted, appears several times across different vowels. However, since fieldwork data involves different types of elicited sentences, it is next to impossible to entirely control for this variable. In most cases, the target vowels occur in the penultimate word of the utterance and in the first syllable of a foot, but in many cases, they do not (it is still unclear if and how the prosodic position affects vowel realization in Siona). Appendix A contains a full list of the analyzed recordings, i.e., all the utterances and word pairs with the target vowels, as well as their location within the archive. Table 4 summarizes the data: the target vowels, the phonological contexts in which they occur, and the number of unique occurrences.

Table 4. A summary of analyzed data (V = target vowel; N = number of unique occurrences; dot represents syllable boundary; # represents word boundary and acoustically, a period of silence).

V	Context	N	Ṽ	Context	N
/i/	[sih]	31	/ĩ/	[#ĩh̃]~[ʔĩh̃]	20
/ɨ/	[kɨ.β]	10	/ɨ̃/	[kɨ̃h̃], [tɨ̃h̃]	9 (8, 1)
/u/	[tuh]	14	/ũ/	[#ũh̃]~[ʔũh̃], [pũʔ]	18 (9, 9)
/e/	[te.k], [te.h]	12 (9, 3)	/ẽ/	[sẽh̃], [sẽʔ]	6 (4, 2)
/o/	[toh], [koh], [soh], [#oh]~[ʔoh]	14 (2, 3, 2, 7)	/õ/	[sõh̃]	4
/a/	[kah], [sah]	20 (10, 10)	/ã/	[tãh̃], [kãh̃]	33 (23, 10)

The phonological context and position of the target vowels has a fair amount of variation, which does not allow for a total control for any potential coarticulatory and prosodic effects. Initially, however, a narrower set of criteria was established for selecting the phonological context of the vowel. Considering Siona phonology, such as segment distribution and frequencies, prosodic position, and syllable types, it seemed advantageous to opt for the closed syllable of the type /CVh/, where C is one of /k, t, s/. Such syllables are relatively frequent in the data, contain only short vowels and appear in the same prosodic position (i.e., first syllable of a foot). However, this narrow selection produced many gaps in the Siona vowel inventory. The selection was expanded to the sequence $C_1V(.)C_2$, where C_1 is any plain voiceless oral obstruent and C_2 is any obstruent. The vowel /ĩ/ still did not appear in this context, but very frequently occurred in the word *ĩhjo* [ĩh̃ɲõ] 'here'. Thus, I also included word-initial syllables of the type /Vh/. However, this created an additional problem since such syllables are optionally articulated with a glottal stop [ʔVh].

Admittedly, the glottal stop is far from an ideal environment as it tends to lack full occlusion and is often articulated as creaky voice on the surrounding vowels, particularly in intervocalic position (see section 2.3.5). Indeed, during the analysis it was observed that a true glottal stop is rarely realized; more commonly, there is creaky or stiff voice that is superimposed on the vocalic stream to a differing degree. However, since [ʔ] did not occupy an intervocalic position, the creak was often short and generally not very pronounced. In syllables where [ʔ] appears after the target vowels, i.e., /pũʔ/ and /sẽʔ/, it is followed by a plain stop and does not seem to make the preceding vowel creaky. In syllables where it optionally occurs before the target vowel (e.g., [oh] ~ [ʔoh]), the glottal stop is present in most cases, often realized as a creak spreading to the start of the following vowel. Luckily, in most cases the vowels were not creaky throughout and it was possible to separate the vowel from the glottal sound. Only one token where the vowel was creaky throughout was discarded (see Appendix A for discarded tokens). Considering the lack of alternative syllables for some vowels, in particular /ĩ/, the syllable [Vh] ~ [ʔVh] was kept for the analysis. Altogether, the consonants banned from the target syllables are nasals, approximants, rhotics, and laryngealized obstruents.

4.2 Measurements and filtering

Praat (Boersma & Weenink, 2021) was used to trim the recordings down to the sentences (sometimes words) containing the target syllables, to annotate the sound files, and to mark the boundaries of the target vowels. The demarcation of the vowel boundaries was largely based on three factors: the formant transitions, vertical striations (representing voicing or vocal folds vibration), and relative intensity (see figures 3 and 4 below). Every boundary was marked on a zero crossing of the waveform. A Praat script (Appendix B) was written to find the frequencies of F_1 and F_2 at the vowel target, measured as the mean of frequencies within 10 ms exactly at the durational midpoint of the vowel. The script returns a list of unique vowel occurrences, their formant values, durations, and file names (for precise location).

Closer inspection of the spectrograms revealed that some data contained the so-called 'ghost formants'[15] and nasal formants[16], which interfered with the measurements and skewed the results. All such instances were noted down during the slicing of the recordings. Ghost formants occurred sporadically and were especially prevalent in syllables with higher levels of noise, such as [sih] and [sẽh̃]. There were also some nasal formants, which were visibly distinct from oral formants. For the purpose of this study, the identified nasal formants and ghost formants were filtered in Praat. All filtering was done in a similar manner, namely by applying a formula to completely block the frequencies in the specified band containing interfering formants and allowing for the proper (oral) formant frequencies to pass, called band-stop filtering (for an overview of different types of frequency filtering, see Weenink, 2022: 135–136).

The vast majority of vowel tokens in the syllable [sih] contained ghost formants, which is why a second Praat script (Appendix C) was created specifically for this syllable. This second script first

---

[15] Praat can sometimes pick up on extra formants that aren't a part of the vowel acoustic properties, which I refer to as 'ghost formant'; the term was pointed out to me by one of my supervisors, Bert Botma.

[16] The concept of nasal formants is introduced in section 3.2.2 together with nasal anti-formants; however, here I refer to them as simply "nasal formants" since it is unclear which ones were the interfering formants in the data, the nasal formants, or the nasal anti-formants, and such distinction is not important here.

filters the sound files, and then retrieves the vowel measurements. The rest of the formant filtering was done manually by applying a specific formula that would be suitable for a given syllable (see Appendix D for the full list of manually filtered data and the formulas that were used). An example of filtering the ghost formants is illustrated for syllable [sih] in figure 3, where the unfiltered vowel target[17] has formant frequencies of 340 Hz (F_1) and 2011 Hz (F_2), and the filtered one – 341 Hz (F_1) and 2530 Hz (F_2). Figure 4 shows an example of filtering an entire nasal formant band in syllable [ʔĩh], where the unfiltered vowel target has 335 Hz (F_1) and 2077 Hz (F_2), and the filtered one – 409 Hz (F_1) and 2769 Hz (F_2).

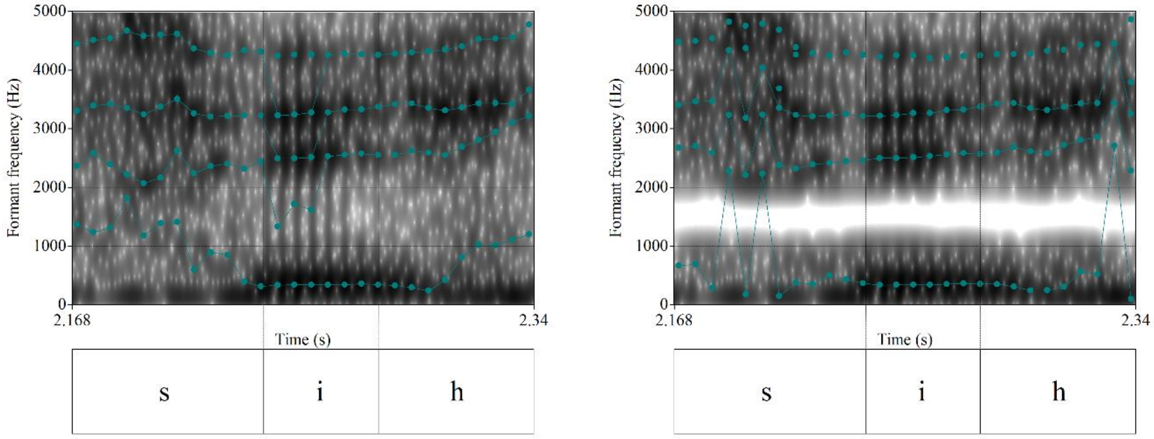


Figure 3. Band-stop filtering of ghost formants in syllable [sih] (unfiltered spectrogram on the left and filtered spectrogram on the right).

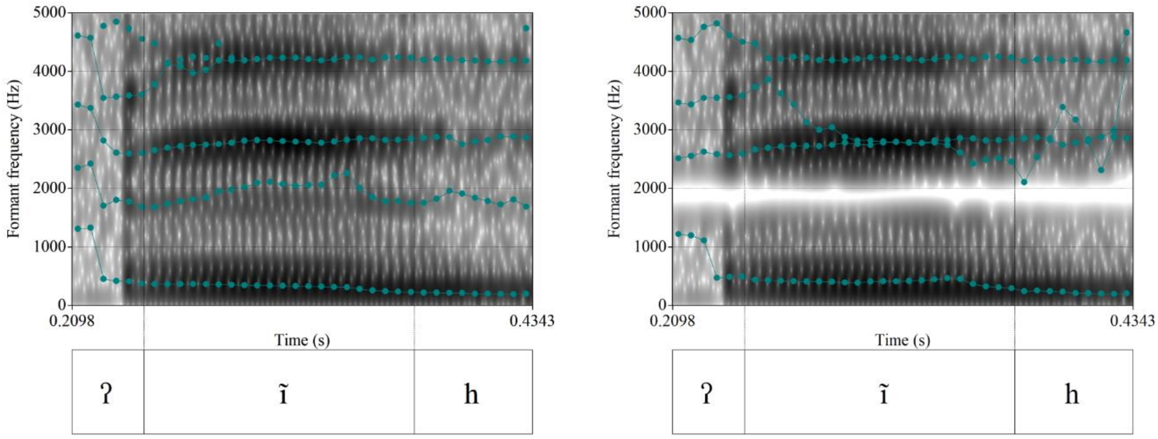


Figure 4. Band-stop filtering of the nasal formant band in syllable [ʔĩh] (unfiltered spectrogram on the left and filtered spectrogram on the right).

Admittedly, this is probably not the best way of filtering interfering formants, especially the nasal formants, which often overlap with the oral formants, making them harder to discern and impossible to filter using a simple band filter. However, this was deemed to be the most efficient and straight-forward method for this study.

[17] To reiterate, vowel target is measured as the average of formant frequencies within 10 ms at the durational midpoint of the vowel; not to be confused with target vowel, which is a vocalic phone that I aim to describe.

5. Results

Data analysis and visualization were conducted via R (R Core Team, 2021; RStudio Team, 2021), using the formant frequencies and durations that were obtained from the two Praat scripts (Appendix B and C). First, the results of oral vowels are presented (in 5.1), then, the results of nasal vowels (in 5.2), and finally, a comparison of oral and nasal vowels is presented (in 5.3). In the oral vowel section, I first summarize the results for different phonological environments, showing how context affected oral vowel quality (in 5.1.1), and then, I present the values of formant frequencies and durations for all oral vowels and visualize them in vowel space (in 5.1.2). Similarly, the nasal vowel section first summarizes the formant results in different phonological environments (in 5.2.2), and then presents the formant and durational results for all nasal vowels, subsequently depicting them in vowel space (in 5.2.3). Lastly, the results for oral vowels are compared to the results for nasal vowels (in 5.3).

5.1 Oral vowels

5.1.1 Effects of phonological environment

A preliminary analysis was conducted to determine whether the formants of the target vowels are significantly affected by different consonantal environments. For oral vowels, the phonological context is different in /a, o, e/. Table 5 summarizes the results of formant frequencies and statistical tests. For the vowels /a/ and /e/, two-sample *t*-tests were applied, since there are two different contexts for each vowel. For the vowel /o/, a one-way ANOVA was applied since there are four different contexts in which this vowel occurs. For both the two-sample *t*-test and ANOVA the null hypothesis (H_0) states that the means of the different groups (here, phonological contexts) are the same, or do not differ significantly.

Table 5. Formant frequencies of oral vowels in different contexts (N = number of unique tokens; SD = standard deviation from the mean; * indicates rejection of H_0).

Vowel	Context	N	F_1			F_2		
			Mean	SD	*p*-value	Mean	SD	*p*-value
/a/	[kah]	10	756	42.5	0.748	1777	107.4	0.8419
	[sah]	10	762	36.5		1785	71.0	
/o/	[toh]	2	625	63.7	0.119	1342	80.4	6.51e-05*
	[koh]	3	614	29.6		1034	19.7	
	[soh]	2	650	1.7		1268	16.6	
	[oh]~[ʔoh]	7	586	26.8		1047	58.0	
/e/	[tek]	9	515	24.6	0.913	2380	87.2	0.7316
	[teh]	3	517	21.8		2362	69.9	

For the vowel /a/, the *p*-values are greater than the significance level ($p > 0.05$) for both F_1 and F_2, which means that H_0 is not rejected. Thus, there is no sufficient evidence to say that the mean frequency values of both F_1 and F_2 between [kah] and [sah] are significantly different. For the vowel /o/, the *p*-value is greater than 0.05 in F_1, and thus H_0 is not rejected; F_1 does not appear to differ significantly between the different contexts of /o/. The results of F_2 for the vowel /o/, however, yield *p*-value below the significance level, thus rejecting H_0. The contexts in which the vowel /o/ appears seem to affect F_2 of the target vowel. For the vowel /e/, the p-values are greater than 0.05 for both F_1 and F_2, and thus H_0 is not rejected; the following consonantal environment does not seem to have any effect on the formants of the target /e/.

A closer look at the mean values in /o/ shows that the frequencies of F_2 are similar after the alveolar segments [t, s], which are different from the other two environments. At the same time, the velar stop [k] and the context alternating between no onset and the glottal stop [ʔ] yield very similar values of F_2 in /o/. As a result, the four contexts were further divided into two groups: [toh]/[soh] and [koh]/[oh]~[ʔoh]. Next, two-sample *t*-tests were run for each group, the results of which are summarized in table 6.

Table 6. Formant frequencies of the vowel /o/ in two groups determined based on F_2 values (N = number of unique tokens; SD = standard deviation from the mean).

Vowel	Group	Context	N	F_1			F_2		
				Mean	SD	*p*-value	Mean	SD	*p*-value
/o/	1	[toh]	2	625	63.7	0.6771	1342	80.4	0.4108
		[soh]	2	650	1.7		1268	16.6	
	2	[koh]	3	614	29.6	0.2307	1034	19.7	0.6048
		[oh]~[ʔoh]	7	586	26.8		1047	58.0	

All the *p*-values exceed the significance level for the two groups in table 6. This means that there is no statistical evidence to suggest that neither [toh] and [soh] have different formant means, nor that [koh] and [oh]~[ʔoh] have different formant means. The results suggest that the quality of /o/ in group 1 is different from that in group 2, specifically along the F_2 parameter. Figure 5 illustrates this difference. On the left (figure 5A), the ellipse depicts a wide variation in F_2 values, for which the *p*-value is below the significance level. On the other hand, the split between two groups on the right side (figure 5B) illustrates a much narrower variation within each group. Group 2 ([koh]/[oh]~[ʔoh]) shows a great deal of uniformity unlike group 1 ([toh]/[soh]) possibly due to the total number of unique tokens for each group (see table 6).

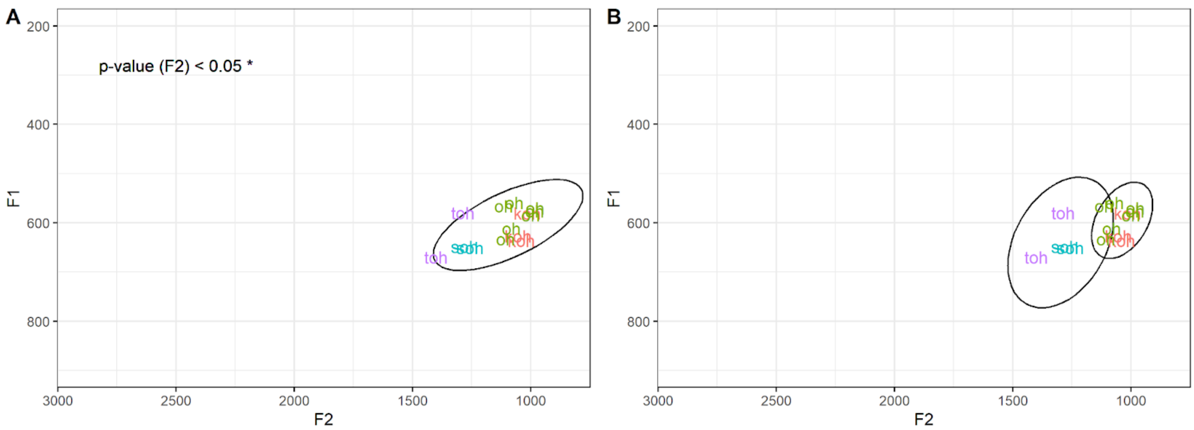


Figure 5. The effect of phonological environment on the vowel /o/ in vowel space. **A:** the ellipse (based on a bivariate *t*-distribution) for the vowel /o/ in all contexts. **B:** the ellipses (based on a bivariate *t*-distribution) for the vowel /o/ split into two groups conditioned by F_2 values.

5.1.2 Oral vowel space and duration

The results for oral vowels are summarized in table 7. The first column shows phonemic vowels as they were described by Bruil (2014) and presented in section 2.3.2. The second column shows their phonetic realization, based on the results. While the second column gives phonetically more accurate representations, it is important to keep in mind that these vowels appear in specific context, namely after [k, t, s] or less often after a period of silence or after [ʔ], and before [h] in most cases but also before [k, β] (for the full description of data refer to section 4.1).

Table 7. Formant frequencies and durations of Siona oral vowels (SD = standard deviation).

Vowel		F_1 (Hz)		F_2 (Hz)		Duration (ms.)	
		Mean	SD	Mean	SD	Mean	SD
/i/	[i]	330	20.6	2601	63.0	53	10.5
/ɨ/	[ɨː]	345	15.9	1917	80.8	121	28.4
/u/	[ʊ]	439	30.3	1138	71.4	72	11.4
/e/	[e]	516	23.0	2376	80.6	76	14.2
/o/	[ɔ]~[ɔ̝]	606	36.9	1118	133.0	60	12.6
/a/	[a]	759	38.7	1781	88.7	82	13.3

The durational results show that vowel length generally increases from high vowels to low vowels. However, the duration of high central /ɨ/ is exceptionally long: it is more than twice as long as the duration of /i/ and /ĩ/, and about two times as long as that of /o/ (see table 7 above and table 9 below). The results suggest that /ɨ/ is phonetically long and is better transcribed as [ɨː] in the analyzed data. Because [ɨː] turned out to be long, the formant values may have been skewed towards being more peripheral.

The particular choice of phonetic realizations becomes evident when the vowel occurrences are plotted on the vowel space according to the frequencies of F_1 (reversed y-axis) and F_2 (reversed x-axis) in figure 6 below. Figure 6(A) shows all occurrences of oral vowels in the data, and figure 6(C) – the vowel space constructed based on the mean frequency values for each vowel. The same is show in figure 6(B) and 6(D) but the vowel /o/ is split according to the results from 5.1.1 above. The ellipses (figure 6, A & B) are generated according to the bivariate *t*-distribution.

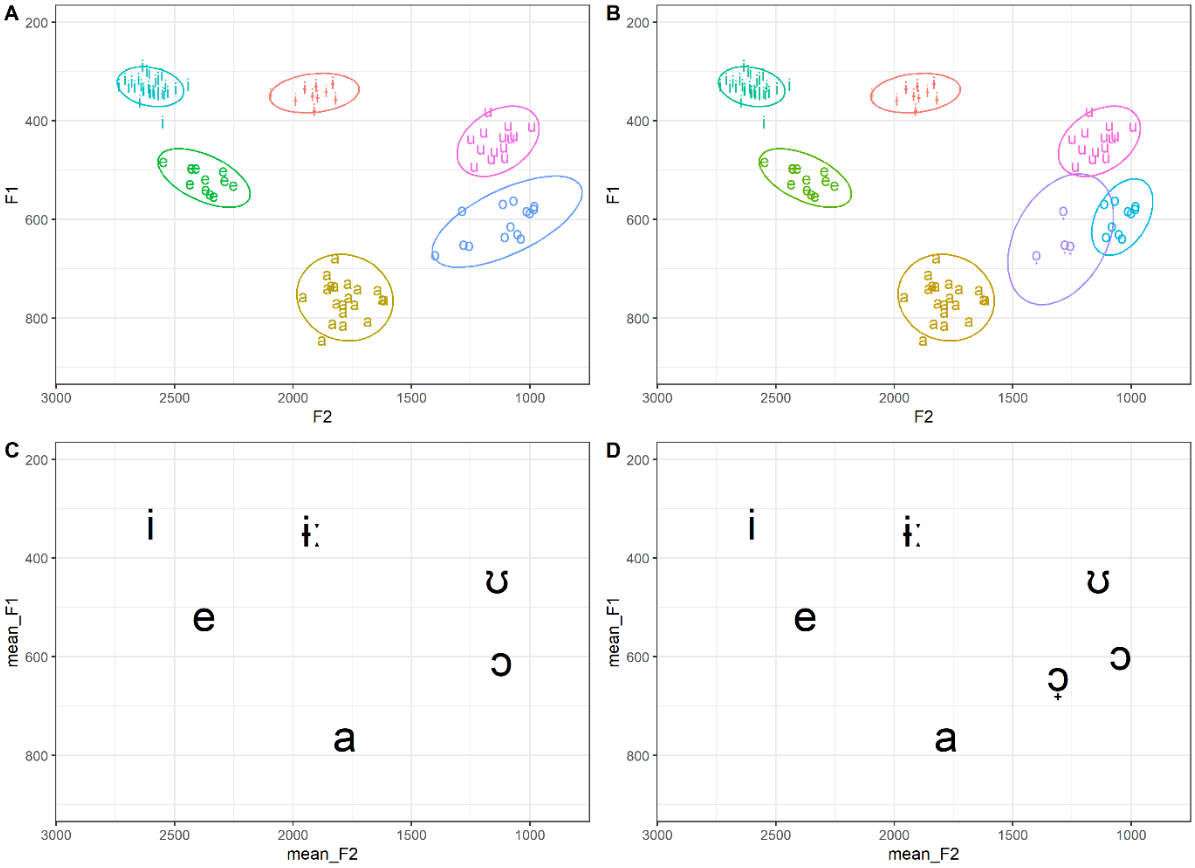


Figure 6. Oral vowel space in Siona. **A:** unique vowel tokens plotted over F_1xF_2 (ellipses based on bivariate *t*-distribution). **B:** unique vowel tokens plotted over F_1xF_2 with a split /o/ (ellipses based on bivariate *t*-distribution). **C:** Vowel space based on F_1 and F_2 means from A above. **D:** Vowel space based on F_1 and F_2 means from B above.

The rounded back vowels /u, o/ appear to be lower in vowel space and are phonetically better transcribed as [ʊ, ɔ] respectively. The variation of F_2 in [ɔ] is visible from the relatively large radius of the ellipse in figure 6(A), as well as from the relatively high standard deviation in table 7 above. This variation can be attributed to the different consonantal environment: before the alveolar consonants [t, s] it appears as the more advanced [ɔ̟].

5.2 Nasal vowels

5.2.1 Effects of phonological environment

The results of the preliminary analysis on the effects of phonological context on nasal vowels are summarized in table 8. Among the target nasal vowels, /ɨ̃, ũ, ẽ, ã/ appear in different contexts, as shown in table 8. For the central high /ɨ̃/, statistical tests are not applicable because there is only one instance where it appears in the [tɨ̃h̃] context. One can only compare the raw formant values without any statistical certainty: there seems to be a considerable difference in F_1 (582 vs. 742 Hz) and possibly in F_2 (1737 vs. 1609 Hz) between [kɨ̃h̃] and [tɨ̃h̃]. For the rest of these vowels, a two-sample *t*-test was applied, given that every vowel has two different contexts. The results are presented in table 8.

Table 8. Formant frequencies of nasal vowels in different contexts (N = number of unique tokens; SD = standard deviation from the mean; * indicates rejection of H_0).

Vowel	Context	N	F_1			F_2		
			Mean	SD	*p*-value	Mean	SD	*p*-value
/ɨ̃/	[kɨ̃h̃]	8	582	83.1	NA	1737	172.1	NA
	[tɨ̃h̃]	1	742	NA		1609	NA	
/ũ/	[pũʔ]	9	722	59.1	1.561e-08*	2068	127.5	0.03718*
	[ũh̃]~[ʔũh̃]	9	394	71		1922	144.2	
/ẽ/	[sẽh̃]	4	742	17.8	0.05961	2151	23.6	0.1343
	[sẽʔ]	2	716	6		2127	5.7	
/ã/	[tãh̃]	23	878	56.4	0.5185	1719	91.7	0.1708
	[kãh̃]	10	864	54.2		1768	88.4	

For the vowels /ẽ/ and /ã/, the *p*-values are greater than the significance level (> 0.5), which means that there is no statistical evidence that would indicate that the mean formant values between the different contexts are different. On the other hand, the two contexts where the vowel /u/ appears, yield *p*-values below the significance level for both F_1 and F_2. The null hypothesis, which states that the means of the two different groups are equal, is rejected. This means that both F_1 and F_2 of the target vowel in [pũʔ] are different from those in [ũh̃]~[ʔũh̃]. This difference is illustraetd in figure 7, where A (lefthand side) presents all /ũ/ tokens and B (righthand side) shows a split between the two contexts, which is primarly visible along the F_1 dimension.

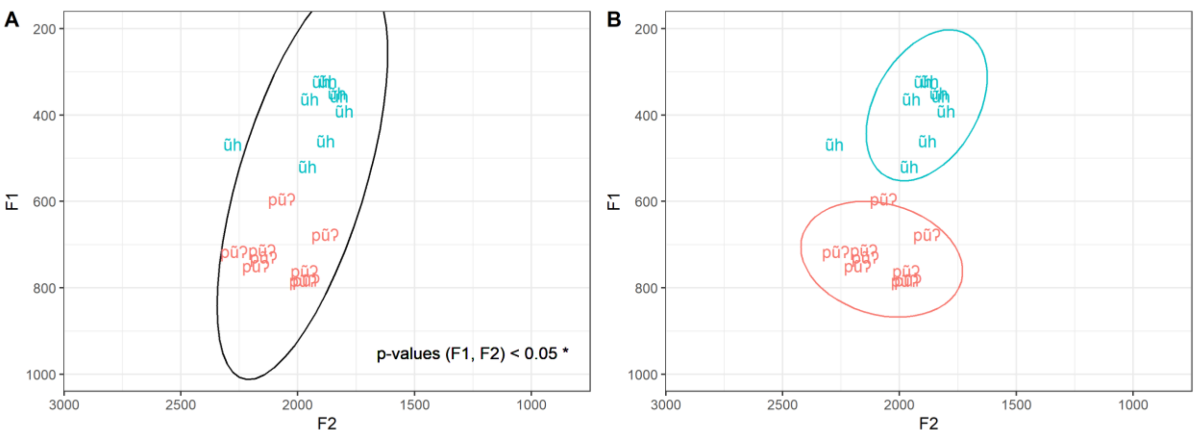


Figure 7. The effect of phonological environment on the vowel /ũ/ in vowel space. **A:** the ellipse (based on a bivariate *t*-distribution) for the vowel /ũ/ in all contexts. **B:** the ellipses (based on a bivariate *t*-distribution) for the vowel /ũ/ split into two groups conditioned by different contexts.

5.3.2 Nasal vowel space and duration

The results of nasal vowel formant frequencies and durations are presented in table 9. The target vowel phonemes are given in the first column, and their corresponding phonetic realizations, which were based on the results, in the second column. The phonetic realizations are more precise representations of the vowel quality and quantity, at least for the context in which they appeared in the data. The nasal vowels appeared before [p, t, k, s] or an empty onset (phonetically, a period of silence), occasionally filled in with [ʔ]. The consonants following the nasal vowels are [h̃, ʔ], where [h̃] is nasalized due to the preceding nasal vowel (for a detailed data description, see section 4.1).

Table 9. Formant frequencies and durations of Siona nasal vowels (SD = standard deviation).

Vowel		F_1 (Hz)		F_2 (Hz)		Duration (ms.)	
		Mean	SD	Mean	SD	Mean	SD
/ĩ/	[ĩː]	414	30.6	2794	52.3	94	19.8
/ɨ̃/	[ə̃]	600	94.3	1722	166.5	50	4.9
/ũ/	[ʉ̃]~[œ̃]	558	180.3	1995	151.9	67	25.0
/ẽ/	[ɛ̃]	733	19.4	2143	22.3	77	17.1
/õ/	[ɔ̃]	760	158.4	1107	182.2	83	12.1
/ã/	[ã]	874	55.2	1734	92.1	80	15.6

The durational results suggest that the undelying /ĩ/ is phonetically long, and is thus, more accuartely transcribed as [ĩː] (see table 9). Aside from the vowel /ĩ/, the vowel length generally increases from high vowels to low vowels. On average, the duration of [ĩː] (94 ms) is nearly twice the duration of [ə̃] (50 ms) and is considerably longer than the oral [i] (53 ms). The vowel [ĩː] has on average the longest duration of all nasal vowels, which is expected to be the shortest if it were indeed short.

Next, the nasal vowel tokens are plotted in the acoustic vowel space, shown in figure 8 below. All nasal vowel tokens are plotted in figure 8(A), and the corresponding vowel categories based on the means in 8(C). Similarly, figure 8(B) illustrates all nasal vowels tokens but with a split /ũ/, as well as added ellipses, and the corresponding vowel categories, including the split /ũ/, in figure 8(D). There is a considerable amount of variation in the realization of vowels /ũ, ɨ̃, õ/. This variation is apparent from the formants' relatively high SD values (table 9) and is especially visible from the large radius and overlaps of the different vowel categories in figure 8.

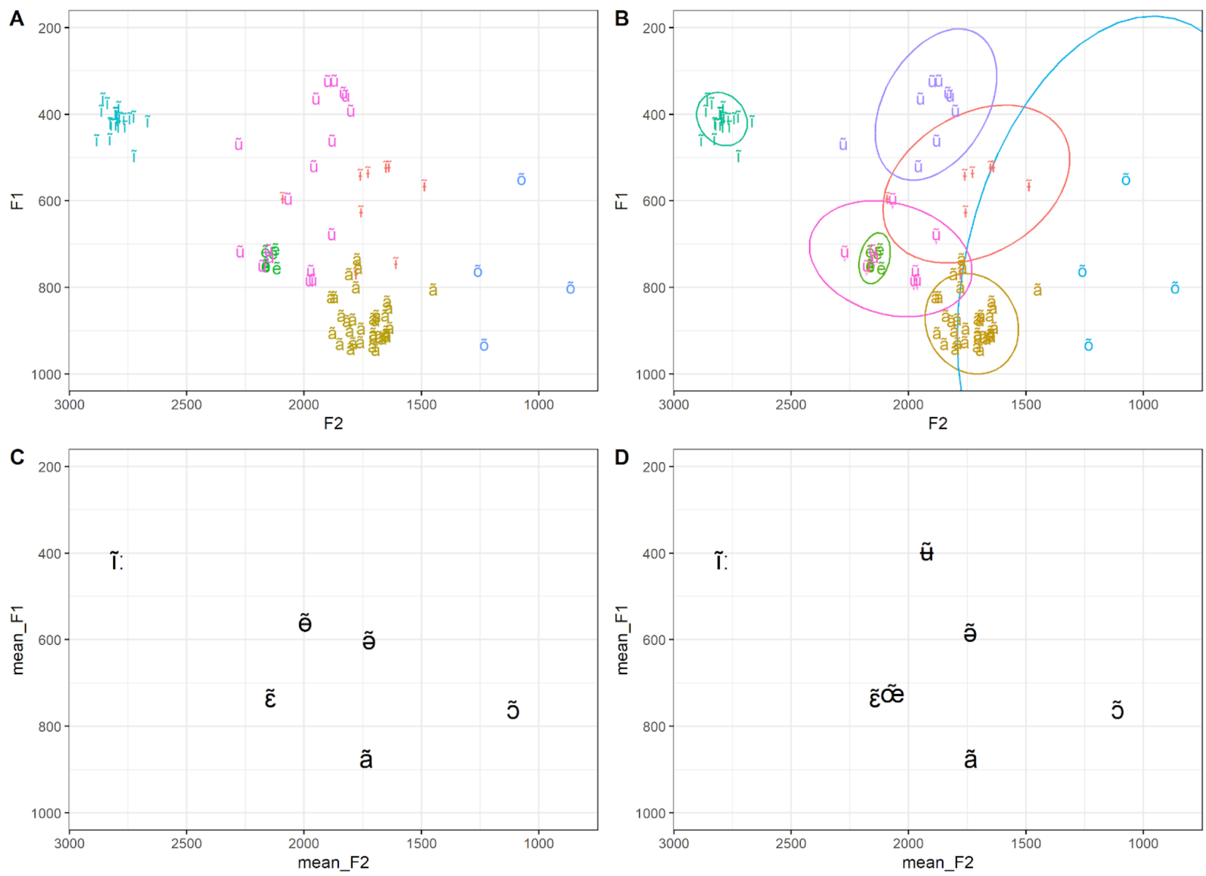


Figure 8. Nasal vowel space in Siona. **A:** unique vowel tokens plotted over F_1xF_2. **B:** unique vowel tokens plotted over F_1xF_2 with a split /ũ/ (ellipses based on bivariate *t*-distribution). **C:** Nasal vowel space based on F_1 and F_2 means from A above. **D:** Nasal vowel space based on F_1 and F_2 means from B above.

The choices for the phonetic representation of nasal vowels lie in the relative vowel quality and quantity, rather than purely absolute values. The vowels /ɨ̃, ẽ, õ/ all appear considerably lower along the height dimension, which is why phonetically, they are better transcribed as [ə̃, ɛ̃, ɔ̃] respectively. The vowel /ũ/ shows the greatest amount of variation, which is contextually determined, as was demonstrated in the previous section. In all contexts, /ũ/ is considerably more front, very broadly representing what would be the vowel [ɵ̃]. A more narrow and precise representation of /ũ/ is a split between the high central [ʉ̃] (in the /ũh/ context) and the low-mid front [œ̃] (in the /pũʔ/ context). This split is visually apparent in figure 8(B) and 8(D); in fact, [ɛ̃] (underlyingly /ẽ/) and [œ̃] (underlyingly /ũ/) have nearly identical qualities, except for their roundedness. In comparison, F_1 means equal 733 Hz for [ɛ̃] and 722 Hz for [œ̃], and F_2 means are 2143 Hz for [ɛ̃] and 2068 Hz for [œ̃].

5.3 Comparison of oral and nasal vowels

5.3.1 General comparison

Nasal vowels show a great deal of variation, especially when they are compared to oral vowels. The oral vowels show clearly distinct vocalic categories in the acoustic space (see figure 6 above), whereas the nasal vowels show considerable overlaps between different vocalic categories (see figure 8 above). Table 10 summarized the results for oral and nasal vowels together and figure 9 illustrates how oral vowel space differs from nasal vowel space. On average, all nasal vowels have higher F_1 values and thus, appear lower in the acoustic space than their oral counterparts (note that when /ũ/ is split into two, this may no longer be the case for one of the realizations).

Table 10. Comparison of formant frequencies and durations between oral and nasal vowels.

Vowel		F_1 (Hz)		F_2 (Hz)		Duration (ms.)	
Oral	Nasal	Oral	Nasal	Oral	Nasal	Oral	Nasal
/i/ [i]	/ĩ/ [ĩː]	330	414	2601	2794	53	94
/ɨ/ [ɨː]	/ɨ̃/ [ə̃]	345	600	1917	1722	121	50
/u/ [ʊ]	/ũ/ [ʉ̃]~[œ̃]	439	558	1138	1995	72	67
/e/ [e]	/ẽ/ [ɛ̃]	516	733	2376	2143	76	77
/o/ [ɔ]~[ɔ̥]	/õ/ [ɔ̃]	606	760	1118	1107	60	83
/a/ [a]	/ã/ [ã]	759	874	1781	1734	82	80

The differences for the front–back quality are not as uniform as for the high–low quality. The back nasal /ũ/ appears a lot more fronted (F_2 = 1995 Hz) than its oral counterpart /u/ (F_2 = 1138 Hz), occupying a rather central position in the acoustic space as either [ʉ̃] or [œ̃]. The back nasal /õ/, realized as [ɔ̃], however, does not seem to differ from its oral counterpart /o/, realized as [ɔ], along the F_2 dimension. Among non-back vowels, /ẽ, ɨ̃/ or [ɛ̃, ə̃] have higher F_2 values and thus, appear more back in the vowel space than /e, ɨ/ or [e, ɨː], while /a/ and /ã/ show practically no difference in F_2. Lastly, the high non-back /ĩ/ or [ĩː] has higher F_2 and thus, occupies a more front position than /i/ or [i].

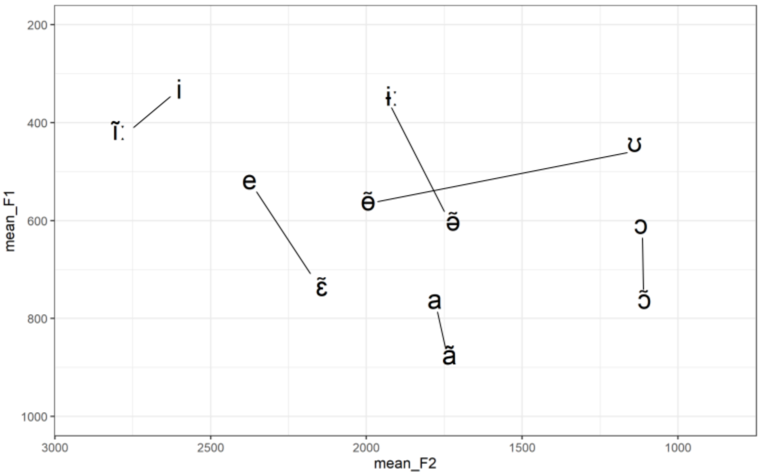


Figure 9. Comparison of oral and nasal vowels in the acoustic space (the lines connect oral and nasal counterparts).

Durational results in table 10 indicate that the vowel /ɨ/ turned out to be not only phonetically long [ɨː] but also the longest vowel in the data, and considerably longer than [ĩː]. Setting aside the surface long vowels [ɨː, ĩː], there seems to be no considerable durational difference between oral and nasal vowels except in the case of the mid back /o/ and /õ/, where the nasal one appears longer than its oral counterpart.

5.3.1 The case of /a/ versus /ã/

Here, I present the results of statistical comparison between [a] and [ã], which are the only vowels with a decent size sample, summarized in table 11. To control for the effects of phonological environment, the first group only includes the syllable /kVh/, whereas the second group includes all tokens where the two target vowels appeared. The durational results show practically no difference between [a] and [ã], as can be gleaned from the (nearly) identical means of vowel duration in table 11. To test the differences in formant values, a two-sample *t*-test was applied to each group (the null hypothesis states that the two data sets have equal means).

Table 11. Comparison of formant frequencies and durations between [a] and [ã] (N = number of unique tokens; SD = standard deviation from the mean; * indicates rejection of H_0).

Group	Context	N	F_1			F_2			Duration	
			Mean	SD	*p*-value	Mean	SD	*p*-value	Mean	SD
1	/kah/	10	756	42.5	0.0001236*	1777	107.4	0.8387	80	16.5
	/kãh/	10	864	54.2		1768	88.4		80	13.4
2	/kah/-/sah/	20	759	38.7	8.968e-12*	1781	88.7	0.07304	81.8	13.3
	/kãh/-/tãh/	33	874	55.2		1734	92.1		80	15.6

For each group in table 11, the results are similar in that H_0 is only rejected for the first, and not the second, formant. For F_2, the *p*-values are smaller than the significance level ($p < 0.05$), which means that the null hypothesis is not rejected. For F_1, the *p*-values are greater than 0.05, meaning the null hypothesis is rejected. Thus, there is no sufficient evidence to suggest that F_2 is different between the oral [a] and the nasal [ã]. However, the two vowels differ significantly in F_1. This difference is illustrated in figure 10.

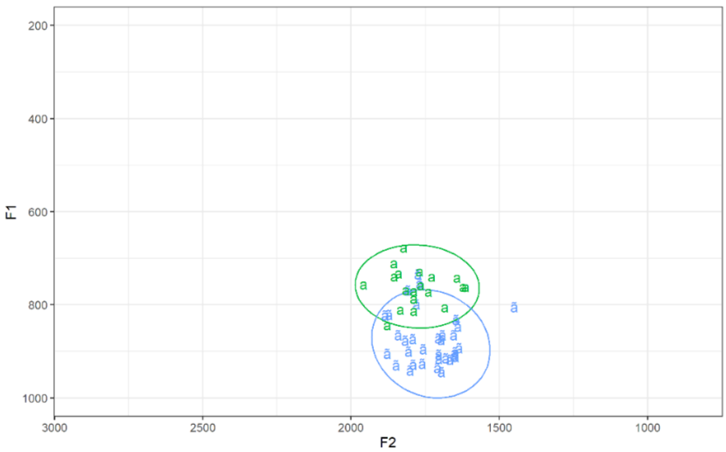


Figure 10. Comparison of [a] and [ã] depicting all vowel tokens (ellipses are based on bivariate *t*-distribution).

6. Conclusion and Discussion

6.1 General findings

The acoustic space for oral vowels in Siona shows six distinctive vocalic categories, similar to the phonemic vowels described by Bruil (2014). The acoustic measurements for the vowels /i, ɨ, u, e, o, a/ suggest that they are realized as [i, ɨː, ʊ, e, ɔ, a] respectively. Furthermore, after [t, s], the vowel /o/ is not only lowered but is also more front, suggesting that it is articulated with an advanced tongue root and can be transcribed as [ɔ̘]. Elsewhere, i.e., following the velar stop [k] and in the syllable [oh] ~ [ʔoh], it is realized as [ɔ]. The reason for the more advanced articulation can be attributed to the distance that the tongue must travel from the denti-alveolar area in [t, s] to the dorsal area in [ɔ̘]. Similarly, Recasens (1991) reports that coarticulatory effects on back vowels, including [ɔ], are greater in dental and alveolar contexts due to tongue dorsum fronting and raising.

The acoustic space for nasal vowels shows overlaps for all vowels except /ĩ/, namely /ɨ̃, ũ, ẽ, õ, ã/, indicating a general shrinking effect of vocalic contrasts. The results suggest that the vowels /ĩ, ɨ̃, ẽ, õ, ã/ are phonetically realized as [ĩː, ə̃, ɛ̃, ɔ̃, ã], and the vowel /ũ/ is realized as [ʉ̃] in the verbal root *ũhku* 'drink' and as [œ̃] in the verbal root *pũ'pu* 'smoke'. Aside from lip rounding (which was not measured in this study), the vowel [œ̃] has the same quality as [ɛ̃] but with a lot more variation. The back vowels /õ/ and /ũ/, realized as [ɔ̃] and [ʉ̃] ~ [œ̃] respectively, and the central vowel /ɨ̃/, realized as [ə̃], show the greatest variation among nasal vowels, which can be attributed to their high susceptibility to coarticulation (cf. Recasens, 1991).

Apart from [ʉ̃], all nasal vowels appear lower in the vowel space than their oral counterparts. As a case in point, a statistical analysis on the differences between [a] and [ã] shows that the two vowels only differ with regards to F_1 such that the nasal vowel appears lower. Overall, then, vowel nasalization is associated with increased F_1 in Siona. For high and mid vowels, the nasalization effect observed along the height dimension is in line with the general cross-linguistic patterns from Beddor (1983, 1993). For low vowels, however, the results are the opposite of Beddor's cross-linguistic observations, namely [ã] is lowered and not raised compared to [a]. Among the 75 languages examined by Beddor (1983), only in Inuit, the nasalization of vowels has a lowering effect for all vowels, including /a, aː/ (although in Inuit vowel nasalization is not contrastive as it is in Siona). Hence, Siona potentially presents a typologically interesting case of the acoustic effects on the height parameter in nasal vowels.

The nasalization effect on F_2 is a lot less uniform. The most striking shift along the front–back dimension, namely an increase in F_2, occurs from the oral [ʊ] to the nasal [ʉ̃] ~ [œ̃]. A slight increase in F_2 is also apparent from the high front oral [i] to the nasal [ĩː], and a slight decrease – from the oral vowels [e] and [ɨː] to their nasal counterparts [ɛ̃] and [ə̃] respectively. Lastly, no significant difference in F_2 is observed between the oral [a, ɔ] and the nasal [ã, ɔ̃].[18]

Similar to the findings reported by Wright (1986) on the perceptual vowel space, the acoustic results in this study demonstrate centralization of the nasal vowels along the height dimension, except for the vowel [ã], which is not raised but lowered. A centralization effect along the front-back dimension is found for /ẽ, ɨ̃/, which are both, considerably lowered and slightly more retracted than their oral counterparts, and are, thus, realized as [ɛ̃, ə̃], as well as for the back rounded /ũ/, which is centralized to either [ʉ̃] or [œ̃]. It appears then that the second formant shows more variability, or in other words, less consistency, than the first formant, especially when it comes to back vowels, which is congruent with the findings on the perceptual vowel space by Wright (1986).

[18] Note that vowel duration likely contributed to the quality of some vowels, such as the phonetically long [ɨː, ĩː] and the reduced nasal [ə̃]; durational results are discussed further in this section.

The reported differences in the acoustic quality between oral and nasal vowels can be broadly attributed to the velum lowering or the coupling of oral and nasal cavity; however, such a conclusion is rather speculative and needs to be followed up by an articulatory study (such as Carignan, 2014). Next, I discuss the durational results and the effects of vowel length on vowel quality.

The results suggest that vowel length contributes to the realization of vowels, namely longer vowels reach their targets better, while shorter vowels appear more reduced (as described by Lindblom, 1963). The clearest examples come from the vowels [ɨː, ĩː], which are more peripheral in the vowel space than their short counterparts [ə̃, i]. The difference is especially striking in the case of the high central vowels /ɨ, ɨ̃/, realized as [ɨː, ə̃]: the duration of the oral vowel is 121 ms, allowing it to fully reach the target, whereas the duration of the nasal vowel is only 50 ms, reducing it to a schwa. The durational difference between the high front vowels [i] and [ĩː] is not as striking, namely 53 ms and 94 ms respectively, which is also reflected in the acoustic space, where [ĩː] appears slightly more peripheral than [i].

Another vowel that has a markedly short duration is [ɔ], whose average length is 60 ms. Of the oral vowels, it is the second shortest vowel after [i], and it is considerably shorter than [ɔ̃], whose average length is 83 ms. While [ɔ̃] appears more peripheral, specifically lower, than [ɔ], it is difficult to attribute this difference to the vowel duration since ultimately all nasal vowels appear on average lower than their oral counterparts. Aside from the aforementioned cases, the results of vowel duration are generally in line with the physical properties of vowel articulation: vowel duration increases from high vowels to low vowels as the oral cavity becomes more open. Lastly, nasal vowels are not consistently longer than oral vowels, indicating that duration does not contribute to the acoustic contrast between the two, as for example, was described for perceptual contrast between nasal and oral vowels in English and French (Beddor, 1993).

6.2 Lowering of back vowels

On average, all back vowels appear lower in the acoustic vowel space than their phonemic values as described by Bruil (2014). The lowering of /u, o, ũ, õ/ is not unexpected since it has been noted before by Bruil (2014) and Bruil & Stewart (2022), but no conditions were described for this process. It then becomes unclear when the back vowels are realized as [u, o, ũ, õ].

The target vowel /u/ only appears in the syllable [tʊh] in verbal and nominal roots; thus, the lowering of /u/ can only be generalized for this context. The vowel /o/ appears in a variety of contexts and in all instances, it is lowered to [ɔ]: in syllables [tɔh], [kɔh], [sɔh], [ɔh]~[ʔɔh], which can be a part of a verb, or a noun used in a sentence or a word pair. This suggests that the high-mid /o/ in Siona may actually be a low-mid /ɔ/. The vowel /õ/ appears only four times in the data in the word *sõhkɨ̃ɨ̃* 'tree'. On average, it appears lower in vowel space but with a great amount of variation (see table 9 and figure 8). Because of the lack of data with this vowel it is difficult to judge whether /õ/ is consistently lower in Siona. The realization of the back nasal /ũ/ is unusual compared to other back vowels in that it is consistently fronted, but similar in that it is on average lowered. The vowel /ũ/ appears in the verbal root *ũhku* 'drink', where it was reported by Bruil (2014: 120) to be occasionally lowered. While there are a couple of instances where it appears slightly lower in *ũhku*, it is on average not lower but instead fronted to [ʉ̃]. The vowel /ũ/ also appears in the verbal root *pũ'pu* 'smoke', where it is both considerably lower and more front, and in the vowel space, it appears in place of [œ̃] (see figures 7 and 8 for realization of /ũ/).

In conclusion, the results suggest that back vowels may be inherently lower than previously described; therefore, a case could be made for representing them as /ʊ, ɔ, ʊ̃, ɔ̃/. To resolve this question, further research can measure these vowels in clear speech, that is when vowels are over-articulated and are better reaching their targets, including more speakers of different demographics.

6.3 Long vowels

In the analyzed data, the oral high central /ɨ/ appears as phonetically long [ɨː], and the nasal high front /ĩ/ is also realized as long [ĩː]. These realizations cannot be accounted for by any of the processes described in section 2.3.2. In other words, the vowel length in these cases cannot be attributed to the effect of vowel coalescence or to the morphological context. Therefore, I explore other possible explanations for the occurrence of these long vowels, such as prosodic effects and a possibility of underlyingly long vowels.

The vowel /ɨ/ always occurs in an open syllable, in the suffix *-kɨ* (a classifier denoting a masculine, animate entity; Bruil, 2014: 138) in carrier sentences, such as *Jure kahkasih**kɨ**bi daha'i* 'He came after entering today', where the verb root *kahka* is variable. The first condition for vowel lengthening that comes to mind is prosodically motivated, namely that it appears in an open syllable. However, if it were the case, such vowel lengthening would have been observed in other open syllables. Instead, the target vowel /e/ is not realized as long even though it appears in open syllables. There is also a possibility that /ɨ/ is inherently long, which would enhance the contrast between hypothetical /ɨː/ and /i/, as well as between hypothetical /ɨː/ and /ɨ̃/.

The target vowel /ĩ/, on the other hand, always occurs in a closed syllable, in a monomorphemic word *ĩhjo*, roughly translated as 'here', at the beginning of sentences such as *Ĩhjo horo ba'ihi* 'There is a flower here', where the noun *horo* is variable. The vowel length in this case may be prosodically conditioned as it appears at the beginning of an utterance. As such, the vowel appears at an intonational phrase boundary, and thus, in a prosodically prominent position. It is possible that the vowel /ĩ/ becomes phonetically long due to its prosodically stronger position. Alternatively, it may be underlyingly long, which would enhance the contrast between hypothetical /ĩː/ and /ɨ̃/, as well as between hypothetical /ĩː/ and /i/.

If the vowels /ɨ, ĩ/ turn out to be underlyingly long, this will create an additional feature used to contrast vowels in Siona, namely length. This topic can be further investigated by comparing the vowel durations between the hypothesized short and long vowel pairs produced by different speakers in different contexts, controlled for prosodic effects and effects of phonological environment.

References

Beddor, P. S. (1983). *Phonological and Phonetic Effects of Nasalization on Vowel Height.* Bloomington, Ind.: Indiana University Linguistics Club.

Beddor, P. S. (1993). The perception of nasal vowels. In M. K. Huffman & R. A. Krakow (Eds.), *Phonetics and Phonology. Vol. 5: Nasals, Nasalization, and the Velum*, 171–196. San Diego, CA. [etc.]: Academic Press.

Boersma, P., & Weenink, D. (2021). Praat: doing phonetics by computer [Computer program]. Version 6.1.55, retrieved 25 October 2021 from http://www.praat.org/.

Botma, E. D. (2004). *Phonological Aspects of Nasality: An Element-based Dependency Approach.* [Doctoral dissertation, Universiteit van Amsterdam]. LOT. https://hdl.handle.net/11245/1.220655.

Bruil, M. (2012). *Documentation of Ecuadorian Siona.* Endangered Languages Archive. Handle: http://hdl.handle.net/2196/00-0000-0000-000D-EA53-3. Accessed 14 April 2022.

Bruil, M. (2014*). Clause-typing and Evidentiality in Ecuadorian Siona* (Publication No. 350) [Doctoral dissertation, Leiden University]. LOT. https://www.lotpublications.nl/clause-typing-andevidentiality-in-ecuadorian-siona-clause-typing-and-evidentiality-in-ecuadorian-siona.

Bruil, M. (2015). When evidentials are not evidentials: The case of the Ecuadorian Siona reportative. *Linguistic typology, 19*(3), 385–423. https://doi-org.proxy.uba.uva.nl/10.1515/lingty-2015-0012.

Bruil, M. (2018). The development of the portmanteau verbal morphology in Ecuadorian Siona: A story of the formal merger of linguistic categories. *Journal of Historical Linguistics 8*(1), 128–167. https://doi.org/10.1075/jhl.17003.bru.

Bruil, M., & Stewart, J. (2022). Phonetics and phonology of nasality in Ecuadorian Siona. *Phonological Data & Analysis, 4*(3), 1–34. https://doi.org/10.3765/pda.v4art3.44.

Carignan, C. (2014). An acoustic and articulatory examination of the “oral” in “nasal”: The oral articulations of French nasal vowels are not arbitrary. *Journal of Phonetics*, *46*, 23–33. https://doi.org/10.1016/j.wocn.2014.05.001.

Chacon, T. C. (2012). *The Phonology and Morphology of Kubeo: The Documentation, Theory, and Description of an Amazonian Language* (Order No. 3520641) [Doctoral dissertation, University of Hawai'i, Mānoa]. ProQuest Dissertations Publishing.

Chacon, T. (2014). A revised proposal of proto-Tukanoan consonants and Tukanoan family classification. *International Journal of American Linguistics, 80*(3), 275–322. https://doi.org/10.1086/676393.

Choi, J. D. (1991). An acoustic study of Kabardian vowels. *Journal of the International Phonetic Association 21*(1), 4–12. https://doi.org/10.1017/S0025100300005958.

Cohn, A. C. (1990). *Phonetic and Phonological Rules of Nasalization* (Publication No. 76) [Doctoral dissertation, UCLA]. UCLA Working Papers in Phonetics.

Cook, D. M., & Criswell, L. L. (1993). *El Idioma Koreguaje (Tucano Occidental).* Santafé de Bogotá: Asociación Instituto Lingüístico de Verano.

Deterding, D. H. (1990). *Speaker Normalization for Automatic Speech Recognition* [Doctoral dissertation, Cambridge University]. https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.359822.

Deterding, D. (1997). The formants of monophthong vowels in standard Southern British English pronunciation. *Journal of the International Phonetic Association 27*, 47–55. https://doi.org/10.1017/S0025100300005417.

Farnetani, E., & Recasens, D. (2010). Coarticulation and connected speech processes. In W. J. Hardcastle, J. Laver & F. E. Gibbon (Eds.), *The Handbook of Phonetic Sciences 2*, 316–352. Singapore: Blackwell Publishing.

Feng, L. (2009). *A Phonetic Study of the Vowel System in Suzhou Chinese* [Doctoral dissertation, City University of Hong Kong]. CityU Scholars. https://scholars.cityu.edu.hk/en/theses/theses(52b089ed-b109-4b7f-968d-5e9a96e6f9b1).html.

Feng, G., & Castelli, E. (1996). Some acoustic features of nasal and nasalized vowels: A target for vowel nasalization. *The Journal of the Acoustical Society of America, 99*(6), 3694–3706. https://doi.org/10.1121/1.414967.

Ferragne, E., & Pellegrino, F. (2010). Formant frequencies of vowels in 13 accents of the British Isles. *Journal of the International Phonetic Association, 40*(1), 1–34. https://doi.org/10.1017/S0025100309990247.

Gómez-Imbert, E. (2004). Fonología de dos idiomas tukano del Piraparaná: barasana y tatuyo. *Amerindia, 29/30*, 43–80.

Gómez-Imbert, E. (1997). *Morphologie et Phonologie Barasana: Approche Non-linéaire* [Doctoral dissertation, Université de Paris]. Université de Paris 8.

Gordon, M., Munro, P., & Ladefoged, P. (2000). Some Phonetic Structures of Chickasaw. *Anthropological Linguistics*, *42*(3), 366–400. http://www.jstor.org/stable/30028763.

Gregores, E., & Suárez, J. A. (1967). *A Description of Colloquial Guarani.* Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783111349633.

Hajek, J. (1997). *Universals of Sound Change in Nasalization.* Oxford [etc.]: Blackwell.

Harrington, J. (2010). Acoustic Phonetics. In W. J. Hardcastle, J. Laver & F. E. Gibbon (Eds.), *The Handbook of Phonetic Sciences 2*, 81–129. Singapore: Blackwell Publishing.

Harshman, R., Ladefoged, P., & Goldstein, L. (1977). Factor analysis of tongue shapes. *The Journal of the Acoustical Society of America*, *62*(3), 693–707. https://doi.org/10.1121/1.381581.

Iivonen, A. (1995). Explaining the dispersion of the single-vowel occurrences in an F1/F2 space. *Phonetica*, *52*(3), 221–227. https://doi.org/10.1159/000262174.

Iivonen, A., & Harnud, H. (2005). Acoustical comparison of the monophthong systems in Finnish, Mongolian and Udmurt. *Journal of the International Phonetic Association, 35*(1), 59–71. https://www.jstor.org/stable/44526397.

Johnson, K. (2012). *Acoustic and Auditory Phonetics* (3rd ed.). Chichester [etc.]: Wiley-Blackwell.

Johnson, O. E., & Levinsohn, S. H. (1990). *Gramatica Secoya* (Vol. 11). Quito: Instituto Lingüístico de Verano.

Joos, M. (1948). *Acoustic Phonetics* (Language monograph no. 23). Baltimore: Linguistic Society of America.

Kaye, J. D. (1971). Nasal harmony in Desano. *Linguistic Inquiry, 2*(1), 37–56. https://www.jstor.org/stable/4177608.

Kharlamov, V., & Oberly, S. (2021). Phonetics of Southern Ute vowels. *Journal of the International Phonetic Association*, *51*(1), 36–54. https://doi.org/10.1017/S0025100318000427.

Kirk, P. L., Ladefoged, J., & Ladefoged, P. (1993) Quantifying acoustic properties of modal, breathy and creaky vowels in Jalapa Mazatec. In A. Mattina & T. Montler (Eds.), *American Indian Linguistics and Ethnography in Honor of Laurence C. Thompson*, 435–50. Missoula, MT: University of Montana.

Klein, W., Plomp, R., & Pols, L. C. W. (1970). Vowel spectra, vowel spaces and vowel identification. *The Journal of the Acoustical Society of America, 48*(4), 999–1009. https://doi.org/10.1121/1.1912239.

Labov, W. (1985). Hypercorrection by the lower middle class as a factor in linguistic change. In W. Bright (Ed.), *Sociolinguistics: Proceedings of the UCLA Sociolinguistics Conference, 1964*, 84–113. Berlin, Boston: De Gruyter Mouton. https://doi-org.ezproxy.leidenuniv.nl/10.1515/9783110856507-008.

Ladefoged, P. (1967). *Three Areas of Experimental Phonetics: Stress and Respiratory Activity, the Nature of Vowel Quality, Units in the Perception and Production of Speech.* London: Oxford University Press.

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

Ladefoged, P. (1975). *A Course in Phonetics.* New York: Harcourt Brace Jovanovich.

Ladefoged, P. (2001). *Vowels and Consonants: An Introduction to the Sounds of Languages.* Malden, MA [etc.]: Blackwell Publishers.

Ladefoged, P. (2006). *A Course in Phonetics* (5th ed.). Southbank, Victoria, Australia [etc.]: Wadsworth.

Ladefoged, P., & Maddieson, I. (1996). *The Sounds of the World's Languages.* Oxford [etc.]: Blackwell.

Le Roux, M., & le Roux, J. (2008). An acoustic assessment of Setswana vowels. *South African Journal of African Languages, 28*(2), 156–171. https://doi.org/10.1080/02572117.2008.10587311.

Lindblom, B., (1963) Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America, 35*(11), 1773–1781. https://doi-org.ezproxy.leidenuniv.nl/10.1121/1.1918816.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H & H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech Production and Speech Modelling*, 403–440. Dordrecht: Kluwer.

Lindau, M. (1978). Vowel features. *Language, 54*(3), 541–563. https://doi.org/10.2307/412786.

Lindau, M., Wood, P., & LaFage, P. (1976). Vowel spaces in two Kwa languages. *The Journal of the Acoustical Society of America, 60*(S1), S44–S44. https://doi.org/10.1121/1.2003348.

Maddieson, I. (1984). *Patterns of Sounds*. Cambridge: Cambridge University Press. https://doi-org.ezproxy.leidenuniv.nl/10.1017/CBO9780511753459.

Maddieson, I. (2007). Areal distribution of nasalized vowels. In *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS 2007)*. Saarbrücken: Institute of Phonetics, Saarland University.

Mejeant, L. (2001). Culturas y lenguas indígenas del Ecuador. *Revista Yachaikuna.* Retrieved 10 April 2023 from http://icci.nativeweb.org/yachaikuna/1/mejeant.pdf.

Moon, S., & Lindblom, B. (1994). Interaction between duration, context, and speaking style in english stressed vowels. *The Journal of the Acoustical Society of America, 96*(1), 40–55. https://doi.org/10.1121/1.410492.

Moran, S., & McCloy, D. (Eds.). (2019). PHOIBLE 2.0. Max Planck Institute for the Science of Human History. http://phoible.org.

Muehlbauer, J. (2012). Vowel spaces in Plains Cree. *Journal of the International Phonetic Association*, *42*(1), 91–105. https://doi.org/10.1017/S0025100311000302.

Noske, M. (1995). *The Ternary Use of Distinctive Features: In Defense of The Number Three* (Order No. 9542714) [Doctoral dissertation, University Of Chicago]. ProQuest Dissertations Publishing.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America*, *24*(2), 175–184. https://doi.org/10.1121/1.1906875

Pillai, S., Mohd. Don, Z., Knowles, G., & Tang, J. (2010). Malaysian English: an instrumental analysis of vowel contrasts. *World Englishes*, *29*(2), 159–172. https://doi.org/10.1111/j.1467-971X.2010.01636.x.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Recasens, D. (1991). An electropalatographic and acoustic study of consonant-to-vowel coarticulation. *Journal of Phonetics, 19*(2), 177–192. https://doi.org/10.1016/S0095-4470(19)30216-5.

RStudio Team (2021). RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

Sarvasy, H., Elvin, J., Li, W., & Escudero, P. (2020). An acoustic phonetic description of Nungon vowels. *The Journal of the Acoustical Society of America, 147*(4), 2891–2900. https://doi.org/10.1121/10.0001003.

Serrurier, A., & Badin, P. (2008). A three-dimensional articulatory model of the velum and nasopharyngeal wall based on MRI and CT data. *The Journal of the Acoustical Society of America, 123*(4), 2335–2355. https://doi.org/10.1121/1.2875111.

Silva, W. de L. (2012). *A Descriptive Grammar of Desano* (Order No. 3522100) [Doctoral dissertation, University of Utah]. ProQuest Dissertations Publishing.

Stenzel, K. (2013). *A Reference Grammar of Kotiria (Wanano)*. University of Nebraska Press. https://doi.org/10.2307/j.ctt1ddr99n.

Stewart, J., & Bruil, M. (2018). Toward a description of Intonation Patterns in Ecuadorian Siona. In *Amazonicas VII*.

Szende, T. (1994). Hungarian. *Journal of the International Phonetic Association, 24*(2), 91–94. https://doi.org/10.1017/S0025100300005090.

Tuttle, S. G., Lovick, O., & Núñez-Ortiz, I. (2011). Vowels of Upper Tanana Athabascan. *Journal of the International Phonetic Association*, *41*(3), 283–312. https://doi.org/10.1017/S002510031100034X.

Vallejos, R. (2013). El secoya del Putumayo: Aportes fonológicos para la reconstrucción del ProtoTucano Occidental. *Línguas Indígenas Americanas (LIAMES)* 13, 67–100. https://doi.org/10.20396/liames.v0i13.1532.

Veer, B.M. van 't, Bruil, M., & Damonte-Matveienko, O. (in press). Pre-aspiration in Ecuadorian Siona. In N. Kuznetsova, C. Anderson & S. Easterday (Eds.), *Rarities in Phonetics and Phonology: Evolutionary, Structural, Typological and Social Dimensions*.

Walker, R. (2011). Nasal harmony. In M. van Oostendorp, C. J. Ewen, E. Hume & K. Rice, *The Blackwell Companion to Phonology*. Malden, MA [etc.]: Wiley-Blackwell.

Weenink, D. (2022). *Speech Signal Processing with Praat.* To be published. URL: http://www.fon.hum.uva.nl/david/sspbook/sspbook.pdf.

Wheeler, A. (1967). Grammatical structure in Siona discourse. *Lingua, 19*(1-2), 60–77. https://doi.org/10.1016/0024-3841(69)90108-9.

Wheeler, A. (1970). *Grammar of the Siona language, Colombia, South America* (Order No. 7100865) [Doctoral dissertation, University of California], ProQuest Dissertations Publishing.

Wheeler, A. (1987a). *Gantëya Bain, el Pueblo Siona de río Putumayo Colombia, Tomo I*. Lomalinda: Editorial Townsend.

Wheeler, A. (1987b). *Gantëya Bain, el Pueblo Siona del río Putumayo Colombia, Tomo II*. Lomalinda: Editorial Townsend.

Wheeler, A. (2000). La lengua Siona. In M. S. González de Pérez & M. L. Rodríguez de Montes (Eds.), *Lenguas Indígenas de Colombia: Una Visión Descriptive*, 181–198. Santafé de Bogotá: Instituto Caro y Cuervo.

Wheeler, A., & Wheeler, M. (1975). Fonémica Siona (Tucano Occidental). *Estudios Fonológicos sobre Lenguas Vernáculas del Ecuador*, 56–67. Quito: Ministerio de Educación Pública del Ecuador.

Wright, J. T. (1986) The behavior of nasalized vowels in the perceptual vowel space. In J. J. Ohala & J. J. Jaeger (Eds.), *Experimental Phonology*, 45–67. New York: Academic Press.

Yusuf, Y. Q., Fata, I. A., & Karwinda, S. (2021). Oral monophthong vowel qualities of the Jamee language in Aceh. *Indonesian Journal of Applied Linguistics, 10*(3). https://doi.org/10.17509/IJAL.V10I3.31767.

Appendix A. Data

This appendix includes relevant orthographic conventions (table A1), changes made in the transcriptions (table A2), discarded data (table A3), and two lists of analyzed data for oral vowels and nasal vowels. The utterances and words are presented in original Siona orthography, just as in the archived ELAN annotations (Bruil, 2012). Every utterance and word pair has a file name that corresponds to the file name in the archive. The target syllables are in bold. The orthographic conventions relevant to the target syllables are listed in table A1 below (if not listed, the orthographic characters are identical to IPA, such as <s> is equal to [s]). For orthographic conventions outside of target syllables refer to Bruil (2014: 129–132).

Table A1. Relevant conventions for original Siona orthography.

Orthography	IPA	IPA
j	/h/	[h], [h̃]
’	/ʔ/	[ʔ]
ë	/ɨ/	[ɨ]
b	/p̰/	[β]
V̱	/Ṽ/	[Ṽ]

Table A2. Changes made in the transcription.

Change	Utterance	File name
<’> /ʔ/ → <j> /h/	yë'ë **se̱j**seni ñañë ‘I see a bore.’	20120919elicr004
<’> /ʔ/ → <j> /h/	i̱jño **so̱j**këñë ba'iji ‘There is a tree here.’	20120912elicr007
<’> /ʔ/ → <j> /h/	yë'ë **so̱j**këñëre ñañë ‘I see a tree.’	20120913elicr001
<’> /ʔ/ → <j> /h/	**so̱j**këñë – **so̱j**këñëa ‘tree – trees’	20120919elicr005
<e> /e/ → <e̱> /ẽ/	yë'ë ë̱jë̱ juju'i **se̱j**seni ‘My husband killed a bore with an arrow.’	20120920elicr001
<e’> /eʔ/ → <e̱j> /ẽh/	yë'ë ë̱jë̱ juiji **se̱j**seni. ‘My husband is killing a bore with an arrow.’	20120920elicr001

Table A3. Discarded tokens.

Why discarded?	Utterance/word	File name
Some audio artifact present in the recording; unclear where the vowel ends and begins.	yure ku̱e̱ne**sij**kobi dajko'ë ‘She came after drying herself today’	20120918elicr004
The target vowel is nasalized but does not represent an underlyingly nasal vowel.	tojto – **toj**ña ‘board – boards’	20120919elicr005
The target vowel is creaky throughout.	yure **u̱j**kuko do̱mio ‘The woman is drinking now’	20120917elicr006

ANALYZED ORAL VOWELS:

/sih/

yure jëo**sij**këbi daja'i.	He came after cleaning today.	20120918elicr004
yure sai**sij**këbi daja'i.	He came after going today.	20120918elicr004
yure dai**sij**këbi daja'i.	He came after coming today.	20120918elicr004
yure kajka**sij**këbi daja'i.	He came after entering today.	20120918elicr004
yure ajcha**sij**këbi daja'i.	He came after listening today.	20120918elicr004
yure moo**sij**këbi daja'i.	He came after fishing today.	20120918elicr004
yure pu̲'pu**sij**këbi daja'i.	He came after smoking today.	20120918elicr004
yure ñaa**sij**këbi daja'i.	He came after watching today.	20120918elicr004
yure saa**sij**këbi daja'i.	He came after taking (it with him) today.	20120918elicr004
yure cha'ka**sij**këbi daja'i.	He came after having jumped today.	20120918elicr004
yure sai**sij**kobi dajko'ë.	She came after having gone today.	20120918elicr004
yure u̲jku**sij**kobi dajko'ë.	She came after drinking today.	20120918elicr004
yure go'i**sij**kobi daja'i.	She came after returning today.	20120918elicr004
yure ka̲i**sij**kobi dajko'ë.	She came after sleeping today.	20120918elicr004
yure ajcha**sij**kobi dajko'ë.	She came after listening today.	20120918elicr004
yure guya**sij**kobi dajko'ë.	She came after bathing today.	20120918elicr004
yure choi**sij**kobi dajko'ë.	She came after inviting today.	20120918elicr004
yure nëjka**sij**kobi dajko'ë.	She came after standing today.	20120918elicr004
yure moo**sij**kobi dajko'ë.	She came after fishing today.	20120918elicr004
yure dai**sij**këni ñawë̲.	I saw the one who came today.	20120918elicr005
yure sai**sij**këni ñawë̲.	I saw the one who went today.	20120918elicr005
yure kajka**sij**këni ñawë̲.	I saw the one who entered today.	20120918elicr005
yure guya**sij**këni ñawë̲.	I saw they one who bathed today.	20120918elicr005
yure ka̲i**sij**këni ñawë̲.	I saw the one who slept today.	20120918elicr005
yure aya**sij**këni ñawë̲.	I saw the one filled today.	20120918elicr005
yure ku̲e̲ne**sij**koni ñawë̲.	I saw the one who dried herself today.	20120918elicr005
yure guya**sij**koni ñawë̲.	I saw the one (woman) who bathed today.	20120918elicr005
yure aya**sij**koni ñawë̲.	I saw the one (woman) who filled today.	20120918elicr005
yure u̲jku**sij**koni ñawë̲.	I saw the one (woman) who drank today.	20120918elicr005
yure go'i**sij**koni ñawë̲.	I saw the one (female) who returned today.	20120918elicr005
yure ta̲jte**sij**koni ñawë̲.	I saw the one who sowed.	20120918elicr005

/ki.p̰/

yure jëosij**këb**i daja'i.	He came after cleaning today.	20120918elicr004
yure saisij**këb**i daja'i.	He came after going today.	20120918elicr004
yure daisij**këb**i daja'i.	He came after coming today.	20120918elicr004
yure kajkasij**këb**i daja'i.	He came after entering today.	20120918elicr004
yure ajchasij**këb**i daja'i.	He came after listening today.	20120918elicr004
yure moosij**këb**i daja'i.	He came after fishing today.	20120918elicr004
yure pu̲'pusij**këb**i daja'i.	He came after smoking today.	20120918elicr004
yure ñasij**këb**i daja'i.	He came after watching today.	20120918elicr004
yure saasij**këb**i daja'i.	He came after taking (it with him) today.	20120918elicr004
yure cha'kasij**këb**i daja'i.	He came after having jumped today.	20120918elicr004

/te.k/

sani ta̲j**tek**ëna ñawë̲.	I saw him while was going to sow.	20120914elicr006
yure ta̲j**tek**ëña ë̲më.	The man is sowing now, they say.	20120915elicr001
yure ta̲j**tek**o do̲mio.	The woman is sowing now.	20120917elicr006
yure oj**tek**o do̲mio.	The woman is sowing now.	20120917elicr006
yure ta̲j**tek**oña do̲mio.	The woman is sowing now, they say.	20120917elicr007
yure oj**tek**oña do̲mio.	The woman is sowing now, they say.	20120917elicr007
yure oj**tek**oña do̲mio.	The woman is sowing now, they say.	20120917elicr007
sani ta̲j**tek**ona ñawë̲.	I saw her when she was going to sow.	20120918elicr006
sani oj**tek**ona ñawë̲.	I saw her when she was going to sow.	20120918elicr006

/te.h/		
yure ta̱j**tej**i ë̱më.	The man is sowing now.	20120914elicr005
sani ta̱j**tej**ë̱na ñawë̱	I saw them while they were going to sow.	20120917elicr003
sani oj**tej**ëna ñawë̱	I saw them while they were going to sow.	20120917elicr003
/tuh/		
go'ye mo'se **tuj**këña ë̱më.	The man sat on top (of something) yesterday, they say.	20120914elicr008
më'ë **tuj**këna daë'ë.	When you had been sitting on top (of something), I came.	20120915elicr002
me **tuj**kë ë̱më?	How did the man sit on top?	20120915elicr003
go'ye mo'se **tuj**teña si'awa'i.	Everyone was sitting on top (of something), they say.	20120917elicr001
si'awa'i **tuj**tena daë'ë.	I came after everyone had been sitting on top (of something).	20120917elicr004
go'ye mo'se **tuj**ko'ë do̱mio.	The woman was following yesterday.	20120918elicr001
go'ye mo'se **tuj**koña do̱mio.	The woman was following yesterday, they say.	20120918elicr003
më'ë **tuj**kona daë'ë.	I came when you (fem) had been following	20120918elicr007
go'ye mo'se **tuj**tida'wë yë'ë.	I would have been sitting on top (of something) yesterday.	120919elicr001
i̱jño **tuj**kuñë ba'iji.	There is a red seed tree here.	20120919elicr003
i̱jño **tuj**tu ba'iji.	There is wind here.	20120919elicr003
yë'ë **tuj**kuñëre ñañë.	I see a red seed tree.	20120919elicr004
yë'ë **tuj**ture ñañë.	I see wind.	20120919elicr004
go'ye mo'se juabi ñojkuamena **tuj**kuga̱.	He was stringing a chambira* wire through the red seeds.	cr001
/koh/		
yë'ë **koj**kare ñañë.	I see a word.	20120913elicr001
kojka - kojkaña	word - words	20120919elicr005
kojka - **koj**kaña	word - words	20120919elicr005
/toh/		
i̱jño ui**toj**to ba'iji.	There is a spear here.	20120914elicr004
tojto - tojña	board - boards	20120919elicr005
/soh/		
yë'ë **soj**tore ñawë̱.	I saw clay.	20120914elicr002
ja̱o̱kato **soj**totsiaya ajkuao.	She is from Sototsiaya.	20120919elicr002
/oh/		
i̱jño **oj**ko ba'iji.	There is water here.	20120912elicr007
më'ë **oj**teina daë'ë.	After you sew corn, I came.	20120915elicr002
yure **oj**teyë si'awa'i	Everyone is sowing now.	20120915elicr005
sani **oj**tejëna ñawë̱	I saw them while they were going to sow.	20120917elicr003
yure **oj**teko do̱mio.	The woman is sowing now.	20120917elicr006
yure **oj**tekoña do̱mio.	The woman is sowing now, they say.	20120917elicr007
sani **oj**tekona ñawë̱.	I saw her when she was going to sow.	20120918elicr006
/sah/		
go'ye mo'se **saj**këña ë̱më.	The man went yesterday, they say.	20120914elicr008
më'ë **saj**këna daë'ë.	When you had left, I came.	20120915elicr002
go'ye mo'se **saj**teña si'awa'i.	Everyone went yesterday, they say.	20120917elicr001
ñamina'a **saj**si'i yë'ë.	I am going to go tomorrow.	20120917elicr002
si'awa'i **saj**tena daë'ë.	I came after everyone left.	20120917elicr004
go'ye mo'se **saj**ko'ë do̱mio.	The woman went yesterday.	20120918elicr001
go'ye mo'se **saj**koña do̱mio.	The woman went yesterday, they say.	20120918elicr003
go'ye mo'se **saj**koña do̱mio.	The woman went yesterday, they say.	20120918elicr003
më'ë **saj**kona daë'ë.	I came after you (fem) went.	20120918elicr007
go'ye mo'se **saj**tida'wë yë'ë.	I would have gone yesterday.	20120919elicr001
/kah/		
sani **kaj**kakëna ñawë̱.	I saw him entering.	20120914elicr006
go'ye mo'se **kaj**kabi ë̱më.	The man entered yesterday.	20120914elicr007
go'ye mo'se **kaj**kaëña ë̱më.	The man entered yesterday, they say.	20120914elicr008
më'ë **kaj**kaëna daë'ë.	When you had entered I came.	20120915elicr002
go'ye mo'se **kaj**kareña si'awa'i.	Everyone was entering yesterday, they say.	20120917elicr001
si'awa'i **kaj**karena daë'ë.	I came after everyone had entered.	20120917elicr004

go'ye mo'se **kaj**kao do̲mio.	The woman entered yesterday.	20120918elicr001
go'ye mo'se **kaj**kaoña do̲mio.	The woman entered yesterday, they say.	20120918elicr003
më'ë **kaj**kaona daë'ë.	I came after you (fem) entered.	20120918elicr007
go'ye mo'se **kaj**kada'wë yë'ë.	I would have entered yesterday.	20120919elicr001

ANALYZED NASAL VOWELS:

/ĩh/		
i̲jño tujkuñë ba'iji.	There is a red seed tree here.	20120919elicr003
i̲jño gono ba'iji.	There is banana juice here.	20120919elicr003
i̲jño joro ba'iji.	There is a flower here.	20120919elicr003
i̲jño bijko ba'iji.	There is smoke here.	20120919elicr003
i̲jño se̲jse ba'iji.	There is a pig here.	20120919elicr003
i̲jño wi'yabe ba'iji.	There is fat here.	20120919elicr003
i̲jño tujtu ba'iji.	There is wind here.	20120919elicr003
i̲jño ya'ime ba'iji.	There is tanchi* here.	20120919elicr003
i̲jño tu:bë ba'iji.	There is a treetrunk here.	20120919elicr003
i̲jño gë̲jso ba'iji.	There is a leg here.	20120919elicr003
i̲jño jë̲jña ba'iji.	There are hands here.	20120919elicr003
i̲jño wë ba'iji.	There is a *huatuso here.	20120919elicr003
i̲jño tobë ba'iji.	There is a bag here.	20120919elicr003
i̲jño kuebë ba'iji.	There is a nose here.	20120919elicr003
i̲jño yauba ba'iji.	There is a tucunari* here.	20120919elicr003
i̲jño i̲'si ba'iji.	There is a pineapple here.	20120919elicr003
i̲jño kuejso ba'iji.	There is a capibara here.	20120919elicr003
i̲jño ma'ñoko ba'iji.	There is a star here.	20120919elicr003
i̲jño o'a ba'iji.	There is a bee here.	20120919elicr003
i̲jño pë'përi ba'iji.	There is a condor here.	20120919elicr003
/tĩh/		
i̲jño **të̲j**të ba'iji.	There is a 'trompetero' here.	20120914elicr004
/kĩh/		
go'ye mo'se **kë̲j**këña ë̲më.	The man was digging yesterday, they say.	20120914elicr008
më'ë **kë̲j**këna daë'ë.	After you were digging I came.	20120915elicr002
go'ye mo'se **kë̲j**teña si'awa'i.	Everyone was digging yesterday, they say.	20120917elicr001
si'awa'i **kë̲j**tena daë'ë.	I came after everyone had been digging.	20120917elicr004
go'ye mo'se **kë̲j**ko'ë do̲mio.	The woman was digging yesterday.	20120918elicr001
go'ye mo'se **kë̲j**koña do̲mio.	The woman was digging yesterday, they say.	20120918elicr003
më'ë **kë̲j**kona daë'ë.	I came after you (fem) had been digging.	20120918elicr007
go'ye mo'se **kë̲j**tida'wë yë'ë.	I would have been digging yesterday.	20120919elicr001
/sẽh/		
i̲jño **se̲j**se ba'iji.	There is a pig here.	20120919elicr003
yë'ë **se̲j**seni ñañë. ('→j)	I see a bore.	20120919elicr004
yë'ë ë̲jë̲ juju'i **se̲j**seni. (e→e̲)	My husband killed a bore with an arrow.	20120920elicr001
yë'ë ë̲jë̲ juiji **se̲j**seni. (e'→e̲j)	My husband is killing a bore with an arrow.	20120920elicr001
/sẽʔ/		
mo**se̲'**kewë - mose̲'kewë̲a̲	fish hook - fish hooks	20120919elicr005
mose̲'kewë - mo**se̲'**kewë̲a̲	fish hook - fish hooks	20120919elicr005
/kãh/		
go'ye mo'se **ka̲j**këña ë̲më.	The man slept yesterday, they say.	20120914elicr008
më'ë **ka̲j**këna daë'ë.	When you had slept, I came.	20120915elicr002
go'ye mo'se **ka̲j**teña si'awa'i.	Everyone was slept yesterday, they say.	20120917elicr001
ñamina'a **ka̲j**si'i yë'ë.	I am going to sleep tomorrow.	20120917elicr002
ñamina'a **ka̲j**si'i yë'ë.	I am going to sleep tomorrow.	20120917elicr002
si'awa'i **ka̲j**tena daë'ë.	I came after everyone had slept.	20120917elicr004

go'ye mo'se **ka̱j**ko'ë do̱mio.	The woman slept yesterday.	20120918elicr001
go'ye mo'se **ka̱j**koña do̱mio.	The woman slept yesterday, they say.	20120918elicr003
më'ë **ka̱j**kona daë'ë.	I came after you (fem) slept.	20120918elicr007
go'ye mo'se **ka̱j**tida'wë yë'ë.	I would have slept yesterday.	20120919elicr001
/tãh/		
i̱jño **ta̱j**ke ba'iji.	There is a machin here.	20120912elicr008
yure **ta̱j**teji ë̱më.	The man is sowing now.	20120914elicr005
sani **ta̱j**tekëna ñawë̱.	I saw him while was going to sow.	20120914elicr006
go'ye mo'se **ta̱j**tebi ë̱më.	The man was sowing yesterday.	20120914elicr007
go'ye mo'se **ta̱j**teiña ë̱më.	The man was sowing yesterday, they say.	20120914elicr008
yure **ta̱j**tekëña ë̱më.	The man is sowing now, they say.	20120915elicr001
më'ë **ta̱j**teina daë'ë.	After you were sowing, I came.	20120915elicr002
yure **ta̱j**teyë si'awa'i.	Everyone is sowing now.	20120915elicr005
yure **ta̱j**teyeña si'awa'i.	Everyone is sowing now, they say.	20120916elicr001
go'ye mo'se **ta̱j**tewë si'awa'i.	Everyone was sowing yesterday.	20120916elicr004
go'ye mo'se **ta̱j**tereña si'awa'i.	Everyone was sowing yesterday, they say.	20120917elicr001
ñamina'a **ta̱j**tesi'i yë'ë.	I am going to sow tomorrow.	20120917elicr002
sani **ta̱j**tejë̱na ñawë̱	I saw them while they were going to sow.	20120917elicr003
si'awa'i **ta̱j**terena daë'ë.	I came after everyone had been sowing.	20120917elicr004
yure **ta̱j**teko do̱mio.	The woman is sowing now.	20120917elicr006
yure **ta̱j**tekoña do̱mio.	The woman is sowing now, they say.	20120917elicr007
go'ye mo'se **ta̱j**teo do̱mio.	The woman was sowing yesterday.	20120918elicr001
go'ye mo'se **ta̱j**teoña do̱mio.	The woman was sowing yesterday, they say.	20120918elicr003
yure **ta̱j**tesijkoni ñawë̱.	I saw the one who sowed.	20120918elicr005
sani **ta̱j**tekona ñawë̱.	I saw her when she was going to sow.	20120918elicr006
më'ë **ta̱j**teona daë'ë.	I came after you (fem) had been sowing.	20120918elicr007
yure **ta̱j**teni sajsi'i.	I will go after sowing today.	20120918elicr009
go'ye mo'se **ta̱j**teda'wë yë'ë.	I would have sown yesterday.	20120919elicr001
/ũh/		
yure **u̱j**kukëña ë̱më.	The man is drinking now, they say.	20120915elicr001
më'ë **u̱j**kuna daë'ë.	After you drank I came.	20120915elicr002
yure **u̱j**kuyë si'awa'i.	Everyone is drinking now.	20120915elicr005
yure **u̱j**kuyeña si'awa'i.	Everyone is drinking now, they say.	20120916elicr001
go'ye mo'se **u̱j**kuwë si'awa'i.	Everyone was drinking yesterday.	20120916elicr004
go'ye mo'se **u̱j**kureña si'awa'i.	Everyone was drinking yesterday, they say.	20120917elicr001
sani **u̱j**kujëna ñawë̱	I saw them while they were going to drink.	20120917elicr003
si'awa'i **u̱j**kurena daë'ë.	I came after everyone drank.	20120917elicr004
yure **u̱j**kukoña do̱mio.	The woman is drinking now, they say.	20120917elicr007
/pũʔ/		
yure **pu̱'**puji ë̱më.	The man is smoking now.	20120914elicr005
sani **pu̱'**pukëna ñawë̱.	I saw him going to smoke.	20120914elicr006
go'ye mo'se **pu̱'**pubi ë̱më.	The man smoked yesterday.	20120914elicr007
go'ye mo'se **pu̱'**puña ë̱më.	The man smoked yesterday, they say.	20120914elicr008
yure **pu̱'**pukëña ë̱më.	The man is smoking now, they say.	20120915elicr001
me **pu̱'**pu ë̱më?	How did the man smoke?	20120915elicr003
yure **pu̱'**puyë si'awa'i.	Everyone is smoking now.	20120915elicr005
sani **pu̱'**pujë̱na ñawë̱.	I saw them while they were going to smoke.	20120917elicr003
si'awa'i **pu̱'**purena daë'ë.	I came after everyone had smoked.	20120917elicr004
/sõh/		
i̱jño **so̱j**këñë ba'iji. ('→j)	There is a tree here.	20120912elicr007
yë'ë **so̱j**këñëre ñañë. ('→j)	I see a tree.	20120913elicr001
so̱jkëñë ('→j)	tree	20120919elicr005
so̱jkëñëa ('→j)	trees	20120919elicr005

Appendix B. Praat Script

```
# CHANGE FILE PATH in 3 places and CHANGE SYLLABLE in appendInfoLine

files = Create Strings as file list: “list”, “DATA/oral/kah/*.TextGrid”
writeInfoLine: “syllable, F1, F2, length, file_name”
numberOfFiles = Get number of strings
for fileNumber to numberOfFiles
	selectObject: files
	textgridFile$ = Get string: fileNumber
	textgrid = Read from file: “DATA/oral/kah/” + textgridFile$
	numberOfIntervals = Get number of intervals: 1
	soundFile$ = textgridFile$ - “.TextGrid” + “.wav”
	sound = Read from file: “DATA/oral/kah/” + soundFile$
	selectObject: sound
	formantObject = To Formant (burg): 0, 5, 5500, 0.025, 50
	for interval to numberOfIntervals
		selectObject: textgrid
		text$ = Get label of interval: 1, interval
		if text$ = “v”
			onsetTime = Get start time of interval: 1, interval
			endTime = Get end time of interval: 1, interval
			vowelDuration = endTime – onsetTime
			halfVowelDuration = vowelDuration/2
			midpoint = onsetTime + halfVowelDuration
			targetStart = midpoint – 0.005
			targetEnd = midpoint + 0.005
			selectObject: formantObject
			f1 = Get quantile: 1, targetStart, targetEnd, “hertz”, 0.50
			f2 = Get quantile: 2, targetStart, targetEnd, “hertz”, 0.50
			appendInfoLine: “kah”, “, “,f1,”, “, f2,”, “, vowelDuration,”, “, textgridFile$ -
“.TextGrid”
		endif
	endfor
removeObject: sound, formantObject, textgrid
endfor
```

Appendix C. Praat Script /sih/

```
files = Create Strings as file list: “list”, “DATA/oral/sih/*.TextGrid”
writeInfoLine: “syllable, F1, F2, length, file_name”
numberOfFiles = Get number of strings
for fileNumber to numberOfFiles
	selectObject: files
	textgridFile$ = Get string: fileNumber
	textgrid = Read from file: “DATA/oral/sih/” + textgridFile$
	numberOfIntervals = Get number of intervals: 1
	soundFile$ = textgridFile$ - “.TextGrid” + “.wav”
	sound = Read from file: “DATA/oral/sih/” + soundFile$
	selectObject: sound
	filSoundFile$ = textgridFile$ - “.TextGrid” + “_fil” + “.wav”
	filSound = Filter (formula): “if x>900 and x<2100 then 0 else self fi”
	selectObject: filSound
	Save as WAV file: “DATA/oral/sih_fil/” + filSoundFile$
	formantObject = To Formant (burg): 0, 5, 5500, 0.025, 50
	for interval to numberOfIntervals
		selectObject: textgrid
		text$ = Get label of interval: 1, interval
		if text$ = “v”
			onsetTime = Get start time of interval: 1, interval
			endTime = Get end time of interval: 1, interval
			vowelDuration = endTime – onsetTime
			halfVowelDuration = vowelDuration/2
			midpoint = onsetTime + halfVowelDuration
			targetStart = midpoint – 0.005
			targetEnd = midpoint + 0.005
			selectObject: formantObject
			f1 = Get quantile: 1, targetStart, targetEnd, “hertz”, 0.50
			f2 = Get quantile: 2, targetStart, targetEnd, “hertz”, 0.50
			appendInfoLine: “sih”, “, “,f1,”, “, f2,”, “, vowelDuration,”, “, textgridFile$ -
“.TextGrid”
		endif
	endfor
	removeObject: sound, formantObject, textgrid, filSound
endfor
```

Appendix D. Filtered Data

Here, the tokens that were filtered manually and the corresponding formulas applied in Praat, are listed. To illustatrate, the first formula bans all formants between 1000 Hz and 1600 Hz and returns a filtered .WAV file without the banned formants.

Syllable	*File name*	*Formula*
/sah/	20120917elicr001	“if x>1000 and x<1600 then 0 else self fi”
/sah/	20120917elicr004	“if x>1000 and x<1600 then 0 else self fi”
/te.h/	20120917elicr003_01	“if x>1300 and x<2200 then 0 else self fi”
/sẽh/	20120919elicr004	“if x>1300 and x<2200 then 0 else self fi”
/sẽh/	20120920elicr001_02	“if x>1200 and x<2000 then 0 else self fi”
/tãh/	20120916elicr001	“if x>1200 and x<2000 then 0 else self fi”
/tãh/	20120917elicr002	“if x>1200 and x<2000 then 0 else self fi”
/tãh/	20120917elicr003	“if x>1200 and x<2000 then 0 else self fi”
/tãh/	20120917elicr004	“if x>1200 and x<2000 then 0 else self fi”
/tãh/	20120918elicr001	“if x>1200 and x<2000 then 0 else self fi”
/tãh/	20120918elicr009	“if x>950 and x<1650 then 0 else self fi”
/ĩh/	20120919elicr003_017	“if x>1300 and x<1800 then 0 else self fi”
/ĩh/	20120919elicr003_011	“if x>1400 and x<2400 then 0 else self fi”
/sõh/	20120919elicr005_01	“if x<250 then 0 else self fi”
/ũh/	20120916elicr004	“if x<1700 and x>1500 then 0 else self fi”