

An acoustic study of trans-vocalic ejective pairs in Cochabamba Quechua¹

Gillian Gallagher and James Whang, New York University
March, 2013

1 Introduction

In Cochabamba Quechua, as in other dialects of Quechua spoken in Southern Peru and Bolivia, pairs of ejectives do not cooccur root-internally (e.g., *[k'it'a]), but are attested across word boundaries (e.g., [misk'i t'anta] 'good bread'). This paper presents an acoustic study of these cross-word-boundary, trans-vocalic ejective pairs with the goal of investigating phonetic precursors for the root based restriction. Specifically, acoustic properties of attested pairs of ejectives may provide evidence that this is an articulatorily challenging structure, and thus that it is avoided root-internally to minimize articulatory effort. We begin by presenting the core descriptive facts of Cochabamba Quechua in §1.1, and then discuss previous explanations of the relevant restrictions in §1.2. Our research questions are presented in §1.3, where we also motivate the study discussed in §2 and §3.

1.1 Cochabamba Quechua descriptive facts

Quechua languages are spoken by about 10 million people throughout Ecuador, Peru, Bolivia, Northern Argentina and Chile (Lewis et al. 2013). Cochabamba Quechua (henceforth CQ) falls into the Southern Bolivian Quechua dialect group (2.8 million speakers), within the Southern Quechua branch of the family, also known as Quechua IIC (Adelaar & Muysken 2004). The other dialects of the Southern Quechua branch share the laryngeal phonology reported here for CQ. The description and data in this paper draw from the Ajacopa et al. (2007) dictionary of Bolivian Quechua, as well as the Bills et al. (1971) textbook of Bolivian Quechua, and is heavily supplemented by the first author's fieldwork in Cochabamba, Bolivia. The facts are the same as those described for Cuzco Quechua in MacEachern (1999), Parker & Weber (1996) and Parker (2007).

The consonantal inventory is given in Table 1. Of particular interest is the ternary laryngeal distinction among the stops and affricates: voiceless unaspirated or plain, aspirated and ejective.

¹ We are grateful to Lisa Davidson for detailed comments on this manuscript. The work has also benefited from feedback from audiences at the University of Maryland, College Park, the University of Delaware, Wilmington and McGill University. This work was partially supported by NSF grant BCS-1222700 to Gillian Gallagher. Many thanks also to Gladys Camacho Rios and Neil Myler for help with participant recruitment and data collection.

	labial	alveolar	postalveolar	velar	uvular	glottal
plain	p	t	tʃ	k	q	ʔ
aspirate	p ^h	t ^h	tʃ ^h	k ^h	q ^h	
ejective	p'	t'	tʃ'	k'	q'	
fricative		s				h
nasal	m	n	ɲ			
liquid		l r	ʎ			
glide	w		j			

Table 1: Cochabamba Quechua consonant inventory.

CQ has three phonemic vowels /i a u/ and two allophones [e o], which result from lowering of /i u/ adjacent to uvulars. [ʔ] is restricted to root initial position; there are no vowel initial roots.

Roots in CQ are primarily CV(C)CV, where the optional coda is a fricative, nasal, liquid or glide. Ejectives and aspirates occur only in onset position in roots, and are subject to a range of long-distance restrictions. The present paper is concerned with the absence of roots with multiple ejectives, exemplified in (1e) below. The examples in (1) illustrate the distribution of ejectives in roots: ejectives may occur in initial position when either a stop (1a) or a non-stop is in the second syllable (1b), and may occur in medial position when the initial consonant is a non-stop (1c). Roots with medial ejectives and an initial stop are unattested (1d), as are roots with two ejectives (1e).

- (1) a. k'inti 'a pair' b. p'upu 'jug' c. rit'i 'snow'
 p'atʃa 'clothes' tʃ'iri 'cold' satʃ'a 'tree'
- d. *kint'i e. *k'int'i
 *patʃ'a *p'atʃ'a

In addition to the restrictions in (1), medial ejectives also do not occur if the initial consonant is an aspirate or a glottal stop (*[k^hap'i], *[ʔap'i]). Though the current paper is concerned only with ejectives, it should be noted that aspirates are subject to parallel restrictions. Aspirates may not occur in pairs, and may not appear in medial position if the initial consonant is a plain stop, an ejective, or the glottal fricative [h] (see Gallagher 2010a, 2011 for a full description and analysis).

The cooccurrence of ejectives is prohibited in roots, but pairs of ejectives in adjacent syllables do appear across word boundaries, as shown in (2). These sequences arise when a root with a medial ejective is followed directly by a root with an initial ejective. This situation occurs in compounding, and in noun-phrase internal, attributive adjective-noun phrases. CQ is a highly agglutinative, suffixing language, and thus roots usually appear in words with many suffixes. Suffixes in CQ do not contain ejectives or aspirates, and as a result the cooccurrence restriction on ejectives holds at the root level as well as the word level.

(2) misk'i t'anta	'good bread'
hutʃʷuj k'anika	'small rooster'
hunt'a q'ontʃa	'full oven'

Phrases like those in (2) are not reported in the previous literature to undergo any modification, though it is not stated explicitly that they are unmodified (Bills et al. 1971; Parker & Weber 1996; MacEachern 1999; Adelaar & Muysken 2004), nor verified impressionistically or quantitatively. The exact nature of trans-vocalic ejective pairs like those in (2) is thus unknown, and is investigated in the current production experiment.

The root-bound restriction on pairs of ejectives in CQ is a common type of restriction cross-linguistically, where pairs of segments that share some feature are prohibited from co-occurring in some domain. This pattern is often analyzed as an effect of the Obligatory Contour Principle (OCP) (Leben 1973; Goldsmith 1976; McCarthy 1986, 1988), and is often referred to as an OCP effect. In phonological theory, the OCP is a formal principle that restricts multiple instances of a given phonological feature within some domain. For example, multiple instances of the feature [+constricted glottis], which distinguishes ejectives from voiceless unaspirated stops, are prohibited within the domain of the root in CQ. Similar restrictions on aspirates, ejectives and implosives are found in a genetically and geographically diverse range of languages, including dialects of Quechua (Quechuan) and Aymara (Aymaran), Shuswap (Salishan), Tz'utujil (Mayan), Old Georgian (Kartvelian), Hausa (Afro-Asiatic), Ofo (Siouan), Gojri (Indo-Aryan), and Souletin Basque (isolate) as described in detail in MacEachern (1999), as well as Sanskrit (Indo-Aryan) and Ancient Greek (Indo-European) ('Grassman's Law'), some North East Bantu languages (Ohala 1981) ('Dahl's Law') and Muna (Austronesian) (van den Berg & Sidu 1996). Pairs of voiced stops are restricted in Japanese (Mester & Ito 1986; Ito & Mester 2003 and references there; Kawahara 2008). Pairs of pre-nasals are restricted in Muna and some North East Bantu languages (Meinhof 1932; Herbert 1977) ('Meinhof's Law' or the 'Ganda Rule'), and restrictions on homorganic segments are well established in Semitic, Muna and many other languages (Greenberg 1960; McCarthy 1986; Frisch, Broe & Pierrehumbert 2004; Coetzee & Pater 2008).

1.2 Articulatory explanations of ejective cooccurrence restrictions

Perhaps the simplest explanation for the existence of a cooccurrence restriction on ejectives is based on articulatory effort. Ejectives are articulatorily complex, and a cooccurrence restriction avoids structures with multiple complex gestures in close proximity to one another. To flesh out exactly what such an account would entail, it must first be established how ejectives are produced.

Ejectives involve a supraglottal constriction, as in any other stop consonant, as well as a constriction and raising of the glottis (Ladefoged 1993; Ladefoged & Maddieson 1996; Catford 2001). The glottal gesture must be timed with respect to the supraglottal gesture, either preceding or coinciding with release of the supraglottal closure (Lindau 1984; Kingston 1985). If the glottal closure precedes release of the supraglottal gesture, the air behind the oral closure is compressed, resulting in a 'strong' ejective with a large burst and a long voice onset time (VOT),

corresponding to glottal closure. If the glottal closure is simultaneous with supraglottal release, the ejective will be ‘weak’, characterized by a relatively weaker burst and a shorter VOT than a strong ejective, but with creaky phonation on a following segment caused by the overlapping glottal closure. While the acoustic and perceptual properties of strong and weak ejectives differ, all ejectives are inherently articulatorily complex in that they involve multiple, coordinated gestures (Maddieson 2011).

Given the relative complexity of ejectives, they may be dispreferred on articulatory grounds because they are effortful. Indeed, many languages lack ejectives; these sounds appear in only 16% (92 out of 567) of the languages surveyed in the World Atlas of Language Structures (Dryer & Haspelmath 2011). Multiple ejectives could be dispreferred simply because multiple complex sounds require more effort than a single complex sound within a fairly small temporal window. Languages with a cooccurrence restriction on ejectives may tolerate a single ejective within some domain (like the root), but multiple ejectives within this domain may cross a threshold of articulatory complexity. Restrictions on other articulatorily complex sounds like aspirates, implosives and pre-nasals could have a similar basis.

It is also possible that multiple ejectives are actually more difficult to produce in close proximity to one another; that is, two ejectives in adjacent syllables are more difficult than the simple sum of the effort required in two ejectives that are further apart. This line of analysis is pursued in Walter (2007), who provides experimental support that repetition of gestures is articulatorily difficult. Walter is concerned with restrictions on homorganic consonants, and argues that repeated use of the same articulator, as is required for multiple, homorganic consonants, is more effortful than use of multiple articulators, as occurs in heterorganic consonants pairs. Walter finds that vowels are lengthened between identical and non-identical homorganic consonants compared to heterorganic consonants in a variety of production tasks involving English nonce words. This effect is interpreted as a method for dealing with the articulatory challenge of repeated gestures by lengthening the interval between them. Relevant to this line of analysis is that co-occurrence restrictions often hold of disyllabic roots, like in Quechua, or of consonants in adjacent syllables of trisyllabic roots (Rose & King 2007; Coetzee & Pater 2008). Thus, the segments that are restricted from co-occurring are separated by few other segments. In Quechua, putative pairs of ejectives would be either transvocalic in a CVCV root, or separated by a vowel and a coda consonant in a CVCCV root.

A cooccurrence restriction on multiple ejectives within a root avoids repetition of an effortful gesture (coordinated constriction and raising of the glottis) in a short temporal window, though ejectives may still cooccur across word boundaries. The boundedness of the restriction to roots makes sense if the temporal coordination of gestures across word boundaries is more flexible than that within words, as has been found in several studies (Hardcastle 1985; Holst & Nolan 1995; Byrd 1998; Byrd et al. 2000; Cho 2001). If multiple ejectives pose an articulatory challenge when produced in close temporal proximity, the relatively inflexible timing of word- or root-internal gestures may underlie the categorical restriction against pairs of ejectives within roots. Across word boundaries, speakers may exploit the relatively more flexible gestural timing to lengthen the temporal distance between ejectives, making this structure less effortful.

If trans-vocalic ejectives are indeed articulatorily challenging, effects of this challenge should be observable in the production of trans-vocalic ejectives across word boundaries in CQ, where this structure is grammatical. This study seeks to quantify any such effects.

1.3 Research questions for the acoustic study

The goal of the study is to examine whether sequences that are categorically disallowed within roots are modified in any way when produced across a word boundary. Evidence of articulatory effort may be found in gradient or categorical modification of trans-vocalic ejective pairs, or in sub-phonemic variation in the realization of these pairs compared to other structures. The first research question is whether pairs of ejectives that occur across word boundaries are really produced as such, or whether de-ejectivization applies some or all of the time.

(3) Question 1

Are pairs of trans-vocalic ejectives produced as such, or is one stop de-ejectivized some or all of the time? E.g., is /...k'a#p'.../ always produced as [...k'a#p'...], or is it sometimes produced as [...k'a#p...] or [...ka#p'...]?

A second question is whether the interval between the two ejectives is lengthened, which would be expected if pairs of ejectives are easier to produce if they are temporally further apart. This second question follows up on Walter's (2007) results of vocalic lengthening between repeated place gestures. Lengthening at the boundary is assessed via two measures, the duration of the boundary vowel, and the duration of the stop closure of the second ejective.

(4) Question 2

- a. Is the vowel between two ejectives longer than a vowel flanked by one or zero ejectives? E.g., is the duration of [a] in [...k'a#p'...], longer than in [...k'a#p...], [...ka#p'...] or [...ka#p...]?
- b. Is the closure of an ejective preceded by another ejective longer than the closure of an ejective preceded by a non-ejective? E.g., is the labial closure in [...k'a#p'...] longer than the labial closure in [...ka#p'...]

The third question investigated in this study is broader than the first two, examining a range of cues to ejection and whether they are affected by the presence of another, preceding or following ejective. Previous work has found that ejectives in CQ are characterized by a high burst amplitude, a long VOT, and depressed pitch in the transition into the following vowel (Gallagher 2010a, 2011). One possibility is that some or all of these cues are minimized in the presence of another ejective, indicating lenition or undershoot of the glottal gesture. Another possibility is that these cues are enhanced, indicating hyperarticulation of a challenging sequence.

(5) Question 3

Are there acoustic differences in burst amplitude, VOT or pitch in the following vowel between an ejective that is trans-vocalically adjacent to another ejective and an ejective that is not? E.g., are there acoustic differences between [k'] in [...k'a#p'...] and [k'] in [...k'a#p...], or between [p'] in [...k'a#p'...] and [p'] in [...ka#p'...].

2 The acoustic study

2.1 Methods

2.1.1 Participants

The participants were thirteen native speakers of Quechua, all bilingual in Spanish. Twelve speakers were from the Cochabamba region of Bolivia, and completed the experiment in the city of Cochabamba, Bolivia, where they currently reside. One additional speaker was from Cuzco, Peru (participant 5) and completed the experiment in New York City, where he has lived for several years. The dialect of Quechua spoken in Cuzco is in the same dialect group as Cochabamba Quechua (Quechua IIC: Adelaar & Muysken 2004), and has the same laryngeal phonology as described in MacEachern (1999) and Parker & Weber (1996).

The participants were 9 females and 3 males, all of whom were university educated (or currently pursuing a university degree). The age range was 20-50 years. One female participant was excluded from analysis because she failed to produce ejectives consistently, even in control contexts.²

2.1.2 Stimuli

The stimuli were adjective-noun phrases, containing unsuffixed adjectival and nominal roots. The target sequence in each phrase was the [...C₁V#C₂...] sequence at the boundary between the adjective and noun. Stimuli were in one of four categories, depending on whether C₁ and C₂ are ejective or not: either both C₁ and C₂ are ejective (ejective-ejective), C₁ is ejective and C₂ is not (ejective-plain), C₁ is non-ejective and C₂ is ejective (plain-ejective), or C₁ and C₂ are both non-ejective (plain-plain). There were ten phrases in each category, shown in Table 2.

² The cause of this speaker's failure to produce ejectives cannot be determined conclusively, but it may be indicative of language attrition.

ejective-ejective		ejective-plain	
ʎusp'a tʃuru	polished shell	ʎusp'a tʃuwa	polished plate
ʎusp'a q'ontʃa	polished oven	lak'a tʃaʎwa	salty fish
lak'a tʃuwa	salty liquid	misk'i tʃaʎwa	good fish
misk'i tʃuwa	good liquid	misk'i tʃuwi	good beans
misk'i t'anta	good bread	hutʃuj senqa	small nose
hutʃuj k'anka	small rooster	hutʃuj manka	small pot
hutʃuj p'anqa	small book	hunt'a sonqo	full heart
hunt'a p'uɲu	full jug	hunt'a manka	full pot
hunt'a q'ontʃa	full oven	hank'a tʃuwi	toasted beans
hank'a t'anta	toasted bread	hank'a muhu	toasted seeds
plain-ejective		plain-plain	
kuʎi tʃuru	purple shell	tʃunka tʃaʎwa	ten fish
kuʎi p'uɲu	purple jug	tʃunka muhu	ten seeds
kinsa q'ontʃa	three ovens	kinsa nina	three fires
kinsa p'anqa	three books	kinsa tʃuwi	three beans
miʎaj ³ tʃuwa	bad liquid	puka manka	red pot
miʎaj q'ontʃa	bad oven	puka nina	red fire
ʎasa t'anta	heavy bread	soɣta kunka	six necks
tuʎu k'anka	skinny rooster	soɣta perqa	six walls
jana tʃuwa	black liquid	jana punku	black door
jana t'anta	black bread	jana qowi	black guinea pig

Table 2: Adjective-noun phrase stimuli. Categories refer to the laryngeal properties of the trans-vocalic consonant pair at the boundary between adjective and noun.

In addition to the laryngeal properties of the consonants at the boundary between adjective and noun, several other properties of the stimuli were also attended to. Lexical items with stops, fricatives or nasals in the relevant positions were favored, to make segmentation of vowels and consonants straightforward. In the plain-ejective category, stops in C_1 were avoided because a plain stop-ejective sequence is disallowed within roots (e.g., there are no roots like *[pak'i]). To serve as a true control for the ejective-ejective stimuli, which do contain a sequence disallowed in roots, the plain-ejective, ejective-plain and plain-plain categories must not contain sequences disallowed in roots.⁴ The same adjectives and nouns are present in multiple categories with the aim of making the categories maximally comparable. Effort was made to balance the phrases in the four categories for place of articulation of ejectives (as the acoustic correlates of ejectives vary somewhat with place of articulation) and identity of vowels. Emphasis was also put on constructing stimuli that were natural (i.e., adjectives and nouns that made sense together) and contained commonly used words (as judged by the first author's experience); these

³ The Ajacopa et al. (2007) dictionary gives [miʎa] for this adjective, but speakers pronounced it as [miʎaj].

⁴ The decision to not include plain stop-ejective pairs means that the categories differ in both the manner of articulation of C_1 - C_2 as well as the laryngeal specifications of these consonants.

considerations of naturalness and familiarity resulted in some deviations from perfectly balancing the stimuli for other properties.

2.1.3 Procedure

Participants were told that they would be asked to produce phrases containing two Quechua words. The experimenter (the first author) pronounced the adjective in Quechua, and the participant was asked to repeat the word. Then the experimenter pronounced the noun in Quechua, and the participant was instructed to say the full adjective-noun phrase. The experimenter thus gave each word in the phrase individually, but never pronounced the whole phrase. The phrase was constructed and produced by the participant; it was not repeated from the experimenter's production. A sample transcript from a trial would be as follows, illustrated with the phrase [jana punku] 'black door':

Experimenter: [jana]

Participant: [jana]

Experimenter: [punku]

Participant: [jana punku]

If participants were unsure of a word or did not repeat it accurately, the Spanish translation of the word was given. In all cases of uncertainty, this allowed the participant to arrive at the intended Quechua word on their own, without the experimenter repeating the word or correcting the participant's production. If the participant paused between the adjective and noun, or otherwise pronounced the phrase disfluently, they were asked to repeat it.

For a single pilot participant (not included in the analysis reported here), a different elicitation methodology was used. For this participant, the experimenter gave Spanish words, and asked the participant to translate the words into Quechua. This procedure resulted in the participant offering different lexical items than those intended, and in some cases mirroring the Spanish word order for a noun phrase, noun-adjective, as opposed to the Quechua adjective-noun order.⁵ Eliciting the intended phrases was thus quite difficult with this method. The repetition procedure described above was favored because it resulted in participants producing the intended phrase quite easily.

2.1.4 Measurements and Analysis

Each target ejective was coded as either "correct", if produced as an ejective, or "incorrect", if produced as a voiceless unaspirated stop. The correct/incorrect status of a production was determined through careful examination of the waveform and spectrogram, looking for the large burst and long VOT characteristic of an ejective in CQ (Gallagher 2010a, 2011). A silent VOT period has been found to be a strong acoustic correlate of ejection in many languages, including Georgian (Vicenik 2010), Kabardian (Gordon & Applebaum 2004), Montana Salish (Flemming

⁵ The noun-adjective sequence is also grammatical in Quechua, but it is a predicative structure as opposed to a modificational one, compare [k'anka hutʃuj] 'the rooster is small' (the present tense 3rd singular copula is null) with [hutʃuj k'anka] 'small rooster'.

2008), Tigrinya (Shosted & Rose 2011) and Witsuwit'en (Wright, Hargus & Davis 2002). In general, the ejective productions were perceptually quite distinct from their non-ejective counterparts. Where perceptual impression and visual examination were insufficient in deciding the status of an ejective, additional acoustic measures of plain stops in the participant's data – VOT, burst amplitude, and the pitch of the following vowel – were taken post-hoc for the purpose of comparison (e.g., a given target [p'] was compared to all instances of target [p]). The study was not designed to have plain stops to serve as controls, and thus such post-hoc measurements were limited in their power.

The majority of errors were productions of target ejectives as voiceless unaspirated stops, but a few other misproductions or disfluencies also occurred that required removing data from analysis. On one occasion, a speaker failed to produce a single fluent production of a phrase and on three occasions a word with a plain stop was produced with an ejective. These four tokens were removed from all analyses.

Acoustic analysis of the recordings was conducted using the Praat software (Boersma & Weenink 2012). For all accurately produced ejectives, VOT and burst amplitude were measured, along with pitch in the following vowel (as an indicator of voice quality). Duration of visible glottalization, identified as irregularly spaced pitch periods (Ladefoged 1993), in the vowel following an ejective was also measured. Glottalization was extremely infrequent, and tokens with glottalization were too few to be meaningfully compared. Of 406 ejectives, only 61 had some visible glottalization, and only two speakers had more than 10 glottalized tokens. This measure is thus not included in the results and discussion. The three measures of burst amplitude, VOT and pitch are the measures found in previous studies to distinguish ejectives from voiceless unaspirated stops, and from aspirated stops (Gallagher 2010a, 2011). The purpose of looking at these measures in this study is to determine if the correlates of ejection are significantly affected by the presence of a preceding/following ejective.

To quantify the properties of the boundary between adjective and noun, the duration of the vowel at the boundary (e.g., [yana punku]) was measured for all phrases, as well as the closure duration for initial ejectives (e.g., [yana t'anta]). Only the vowel [a] was measured, as the other vowels were infrequent in the stimuli and not evenly represented in all four categories.

VOT was measured from the beginning of the burst to the onset of periodicity in the following vowel, including the period of frication in affricate ejectives. The recordings were not done in a soundproof room, and thus reverberations from the preceding vowel often obscured the beginning of consonant closure. In such cases, the beginning of closure was segmented where the spectrogram showed an abrupt loss of energy between the vowel and the following consonant burst, as shown in Figure 1 below. The amplitude of the burst was measured by taking the difference between the maximum and minimum intensity during the period between the burst and the onset of the following vowel (the VOT period); intensity values were taken from the “get intensity” function in the intensity menu in Praat. Average pitch was taken in the first 30 ms of the vowel following an ejective, by selecting the relevant period and using the “get average pitch” function in the pitch menu in Praat.

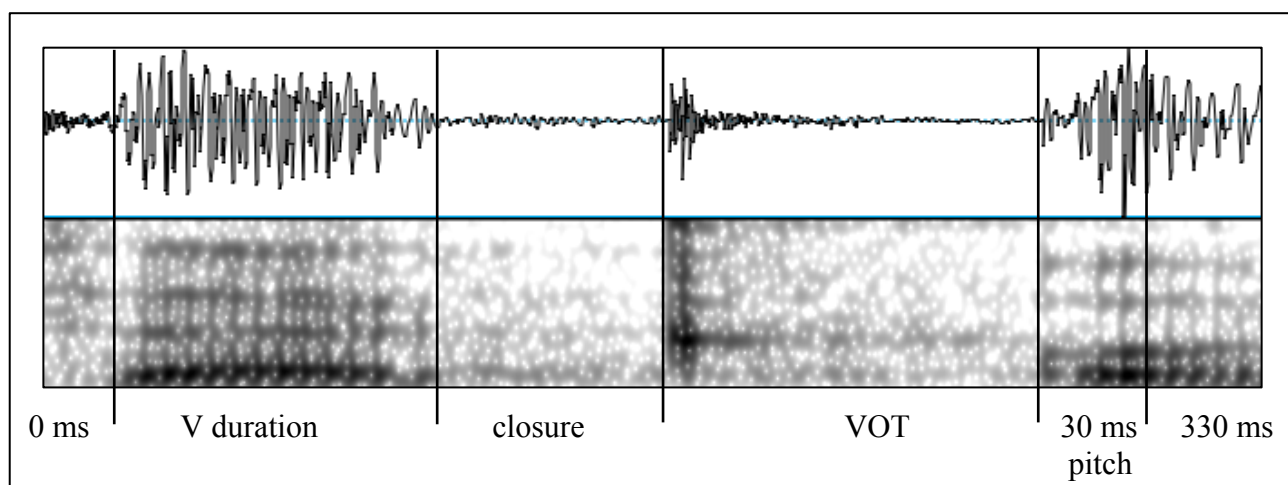


Figure 1: Spectrogram and waveform from [kinsa q'ontʃa], showing landmarks for vowel duration, closure duration, VOT period, and the first 30 ms of the vowel during which pitch was taken.

2.2 Results

The results were analyzed for accuracy, as well as the acoustic properties of ejectives and the boundary between adjective and noun. The accuracy analysis is a binomial analysis of the correct/incorrect production of target ejectives as ejectives (correct) or plain stops (incorrect). The acoustic analysis only includes data from correct productions. We present the binomial analysis of accuracy first in §2.5.1, followed by the analysis of acoustic detail in §2.5.2.

2.5.1 Binomial analysis of accuracy

When analyzing errors in producing ejectives, only the three phrasal categories with ejectives are compared (the ejective-ejective, ejective-plain and plain-ejective categories). As can be seen in Figure 2, accurate production of ejectives was higher when there was a single ejective, as in the ejective-plain and plain-ejective categories, than when there were two ejectives, as in the ejective-ejective category.

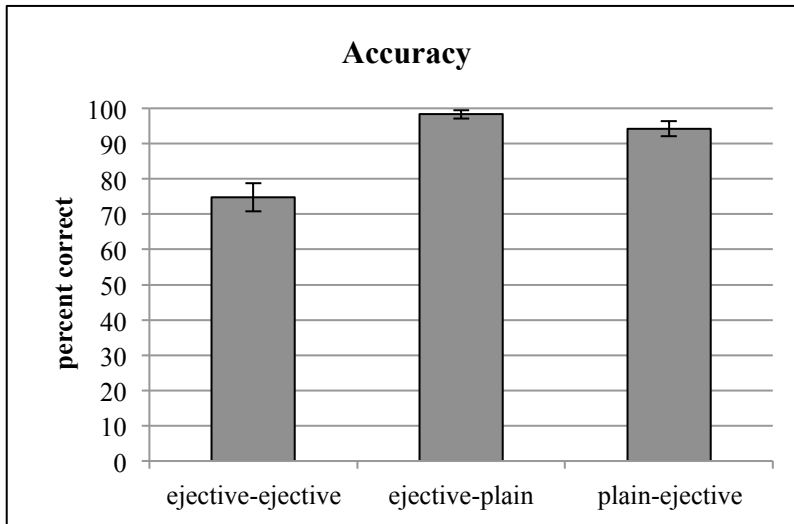


Figure 2: Accuracy on categories with ejectives. Error bars indicate Standard Error.

Inaccurate responses in the ejective-ejective category consist of de-ejectivization of one of the two target ejectives. De-ejectivization was more common on the second ejective than the first: accuracy on the first ejective in the ejective-ejective category was 96%, compared to 79% for the second ejective.

A Mixed Logit Model was fit using the *lmer()* function in the lme4 package (Bates & Maechler 2010) for the R software (<http://www.R-project.org>). The dependent variable was whether the response was correct or not. There was a three-level predictor of category, with the ejective-ejective category as the baseline. The ejective-ejective category was selected as the baseline to directly test the hypothesis that accuracy on pairs of ejectives differs from overall accuracy on ejectives, as represented in the ejective-plain and plain-ejective categories. Random intercepts and slopes by participant were included in the model. Re-leveling the predictor of category with ejective-plain as the baseline reveals no significant difference between ejective-plain and plain-ejective categories.

	Estimate	Standard Error	Wald's <i>z</i>	<i>p</i>
intercept	1.79	0.61	2.92	< .01
<i>ejective-plain</i>	2.29	0.92	2.49	< .02
<i>plain-ejective</i>	1.46	0.53	2.74	< .01

Table 3: Results of a Mixed Logit Model testing for a difference in accuracy between baseline category ejective-ejective and the ejective-plain and plain-ejective categories.

The results and analysis presented above are aggregated across all participants. While there is substantial variation in overall accuracy between participants, as shown in Figure 3, when de-ejectivization occurs it is most likely to occur in the ejective-ejective category for most participants. Eight of the 12 participants have lower accuracy on the ejective-ejective category than the other two categories. For the other participants, subjects 1 and 8 are completely accurate on all categories, while subjects 3 and 5 show a different pattern (both participants make just a

single error, on the ejective-plain category for subject 3 and on the plain-ejective category for subject 5).

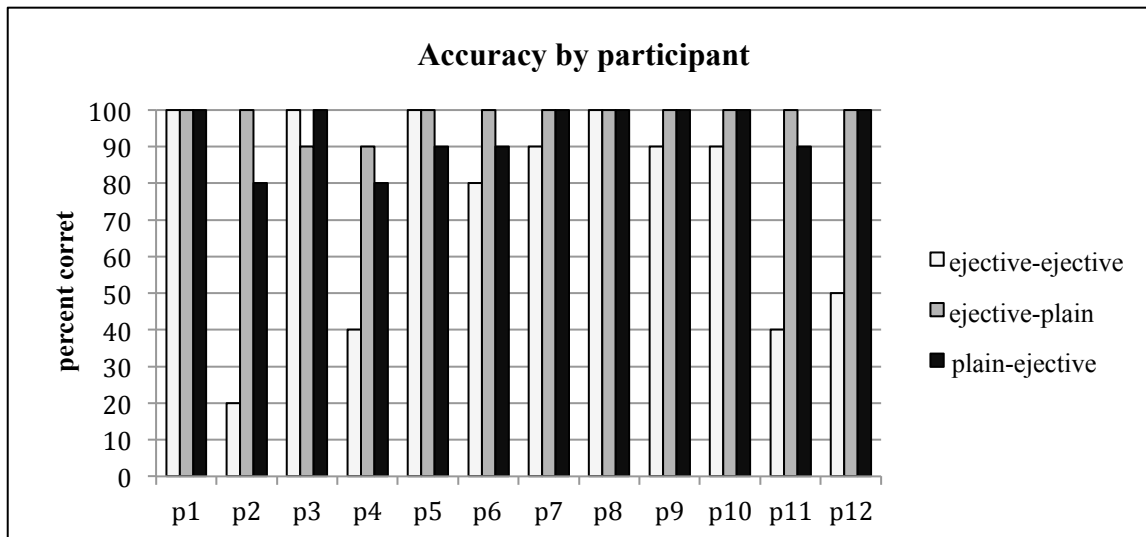


Figure 3: Accuracy on categories with ejectives, by participant.

2.5.2 Acoustic analysis

The acoustic properties measured for all correct productions of ejectives in the target sequences are divided into two categories: boundary measures and ejection measures. Boundary measures examine whether the interval between two ejectives is lengthened, assessed via the duration of the boundary vowel between two ejectives as compared to the other categories (e.g., duration of [a] in [ʌusp'a q'ontʃa] compared to [kinsa q'ontʃa], [hank'a muhu] and [tʃunka muhu]), and the stop closure of initial ejectives in ejective-ejective and plain-ejective categories (e.g., closure of [q'] in [ʌusp'a q'ontʃa] and [kinsa q'ontʃa]).

Ejection measures examine whether the primary acoustic correlates of ejection in CQ – VOT, intensity difference and pitch in the following vowel – are affected by the presence of a preceding or following ejective. Hyperarticulation would be indicated by longer VOT, a greater intensity difference and lower pitch in an ejective preceded/followed by another ejective compared to an ejective preceded/followed by non-ejective consonants (e.g., [misk'i t'anta] vs. [misk'i tʃuwa] and [hutʃ'uj k'anaka] vs. [tuʌu k'anaka]). Undershoot or lenition would be indicated by shorter VOT, a smaller intensity difference and higher pitch. For all measures, the analyses look for a difference between a context with two ejectives and contexts with only a single ejective.

The data from participants 2, 4, 11 and 12 were removed from the acoustic analysis because these speakers had accuracies of 50% or less on the ejective-ejective category, and thus comparison between the few correct productions in the ejective-ejective category and the other categories would be based on very few tokens.

Linear mixed models were fit for each acoustic measure, using the *lmer()* function in the *lme4* package (Bates & Maechler 2010) in the R software. The significance of predictors was determined through model comparison. All models included a random intercept and slope for

participant. The results of the analysis of boundary measures are presented first, followed by the ejection measures.

2.5.2.1 Boundary measures - vowel and closure duration

For vowel duration, all four categories are compared (ejective-ejective, plain-ejective, ejective-plain, and plain-plain). Analysis was done only for the vowel [a], as diphthongs and high vowels were rare in the stimulus set and were not evenly distributed among the four categories. The analysis had a single, two-leveled factor comparing the ejective-ejective category to the other three categories as a group. The duration of the vowel between two ejectives is longer than in the other three categories, as can be seen in Figure 4. The effect of this predictor reaches significance with a t value of ± 2 (Gelman & Hill 2006), as shown in Table 4, and a model with this predictor is significantly different from a model without the predictor ($p < .05$).

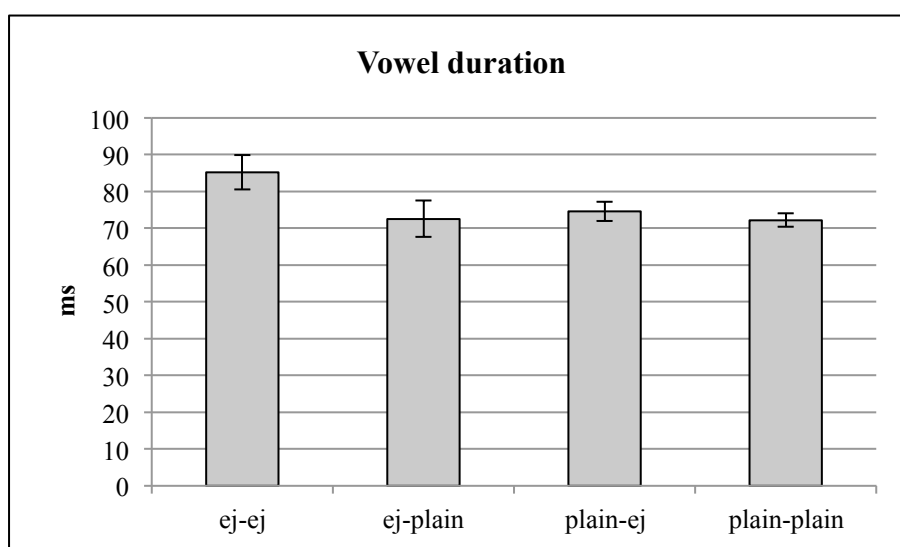


Figure 4: Duration of [a] at the word boundary in the four stimulus categories. Error bars represent Standard Error.

	Estimate	Standard Error	t
intercept	76.26	4.99	15.28
<i>ejective-ejective vs. others</i>	15.02	6.88	2.18

Table 4: Results of a Linear Mixed Model testing for a difference in vowel duration between the ejective-ejective category and the other categories.

Figure 5 below shows the comparison by category for each participant. Participants 1 and 10 show strong effects consistent with the overall pattern of longer vowels in the ejective-ejective category. Participants 3, 5 and 9 show small effects in this direction, while participant 7 goes in the opposite direction, with the shortest vowels in the ejective-ejective category. For participants 6 and 8, vowel length in the ejective-ejective category is neither shorter nor longer than in all the other categories.

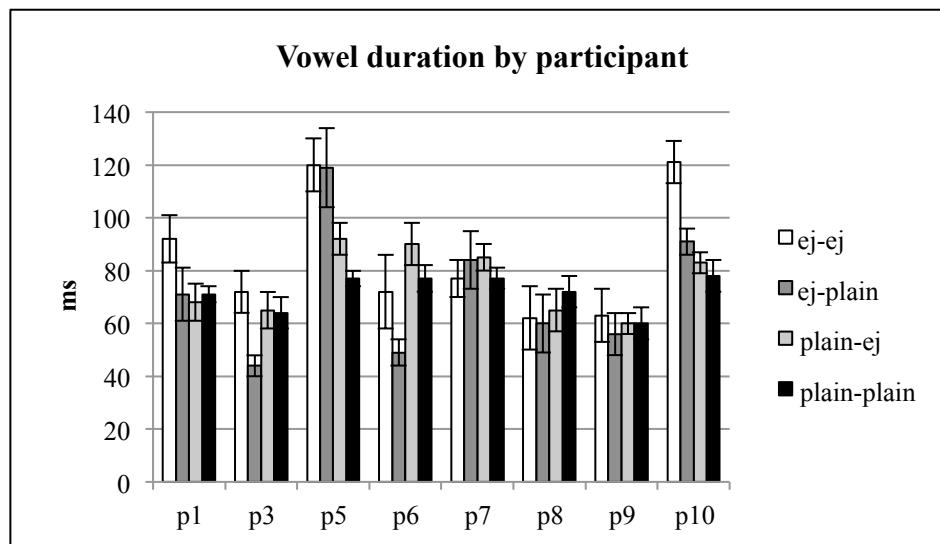


Figure 5: Vowel duration by category and participant. Error bars represent standard error.

For the analysis of closure duration, a comparison is made between the ejective-ejective and plain-ejective categories to test for an effect of a preceding ejective on closure duration, e.g., is the closure of [q'] different in [ʎusp'a q'ontʃa] compared to [kinsa q'ontʃa]. These two categories were compared because C₂ in these categories was always a stop, and thus closure duration could be measured. In the ejective-plain and plain-plain categories, C₂ could be a plain stop, a nasal or a fricative. The comparison between the two categories is significant ($p < .01$ compared to a model without the predictor), with a longer closure duration in the ejective-ejective category than the plain-ejective category. The cumulative results across all participants are shown in Figure 6, and the statistical results are given in Table 5.

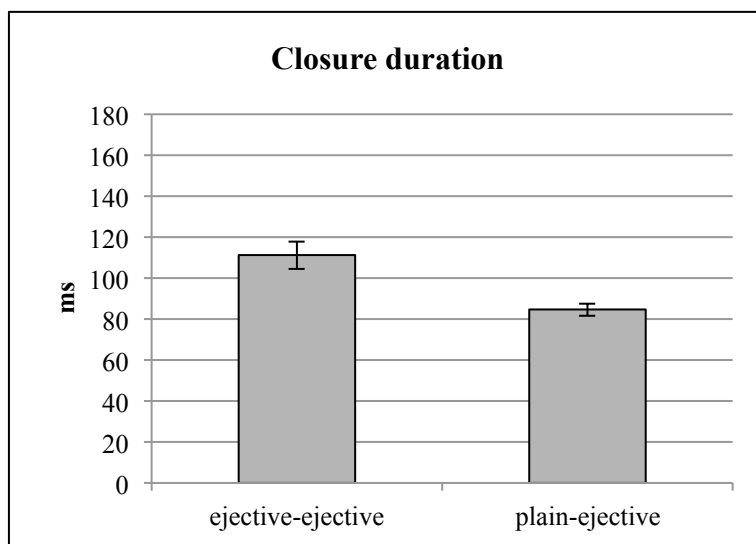


Figure 6: Closure duration for ejective C₂ by category. Error bars represent standard error.

	Estimate	Standard Error	<i>t</i>
intercept	84.78	7.09	11.97
<i>ej-ej vs. plain-ej</i>	36.53	7.73	3.43

Table 5: Results of a Linear Mixed Model testing for effects of category on closure duration of ejective C_2 .

Figure 7 shows that all individual participants are consistent with the overall effect, though participants 3, 5 and 6 show very small differences.

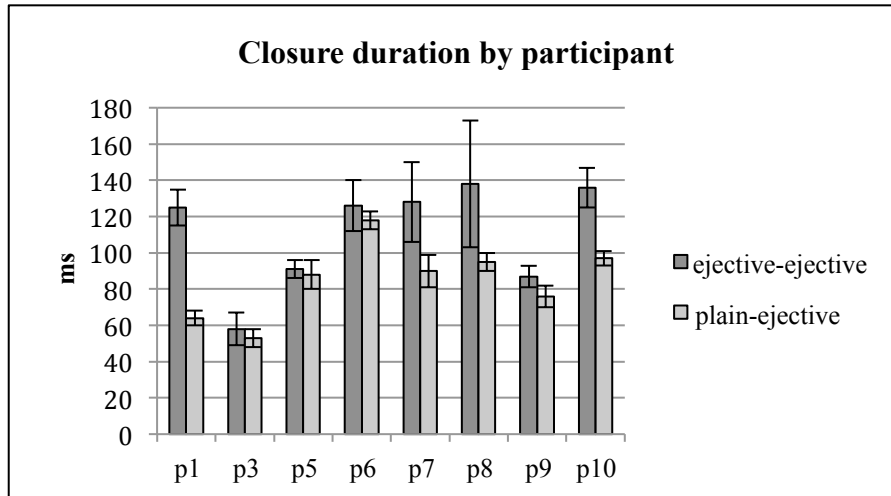


Figure 7: Closure duration for ejective C_2 by category and participant. Error bars represent standard error.

2.5.2.2 Ejection measures – intensity difference, VOT and pitch

For each ejection measure, the analysis looked for effects of position, context and their interaction. The predictor of position compared ejectives in C_1 (e.g., [k'] in [misk'i tʃaɫwa]) to ejectives in C_2 (e.g., [k'] in [tuɫu k'anka]); the predictor of context compared ejectives with a preceding or following ejective (e.g., [k'] and [t'] in [misk'i t'anta]), referred to as pre- and post-ejective, to those with a preceding or following plain consonant (e.g., [k'] and [t'] in [misk'i tʃaɫwa] and [ɫasa t'anta]), referred to as pre-plain and post-plain.

For intensity difference, there is a significant effect of position ($p < .03$ compared to a model without position as a predictor), but no effect of context or interaction between position and context. Intensity difference is bigger for ejectives in C_2 (word-initial position) than in C_1 (word-medial position), indicating larger bursts, but is unaffected by the presence of a preceding or following ejective. The results are graphed in Figure 8, and the statistical results are given in Table 6.

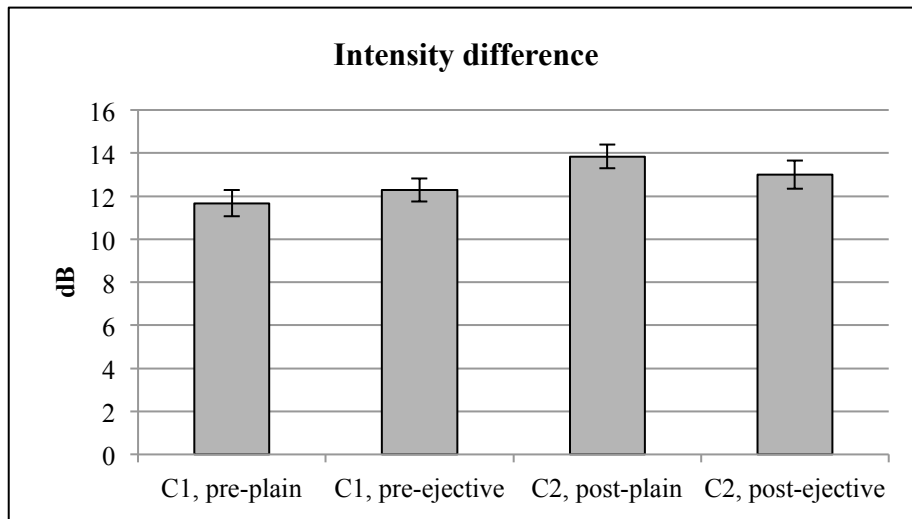


Figure 8: Intensity difference by context and position. Error bars represent Standard Error.

	Estimate	Standard Error	<i>t</i>
intercept	13.21	0.89	14.82
<i>position</i>	1.55	.51	3.03

Table 6: Results of a Linear Mixed Model testing for effects of context and position on intensity difference for ejectives.

Intensity difference is shown in Figure 9 for each individual participant. Participants 3 and 8 show large effects in the direction of the overall pattern. The remaining participants show mixed effects, with small difference between categories, largely overlapping in values. Participant 7 has the smallest intensity difference for post-ejective ejectives in C₂, contrary to the overall pattern of a greater intensity difference in C₂.

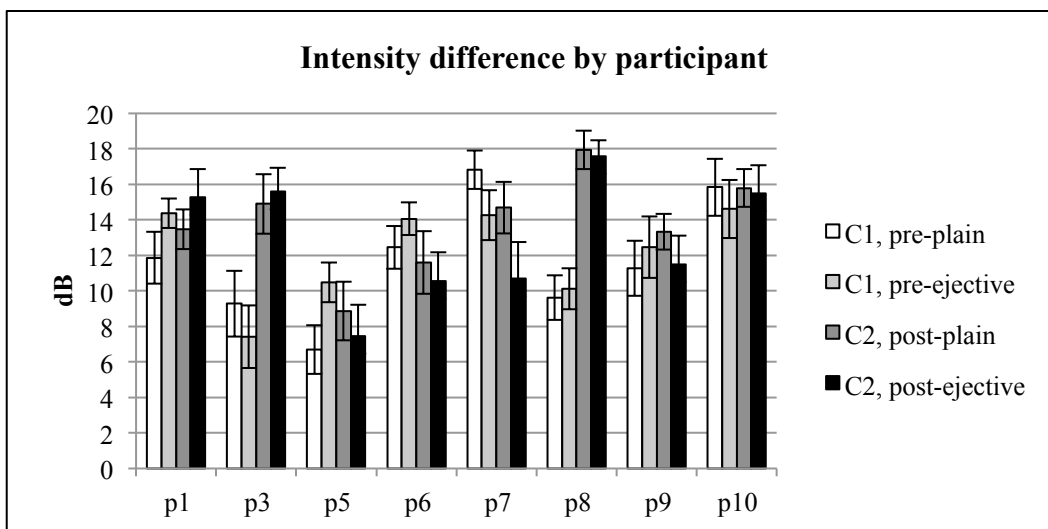


Figure 9: Intensity difference by context, position and participant. Error bars represent Standard Error.

The measure of VOT shows a similar pattern to the measure of intensity difference: there is a significant main effect of position ($p < .05$ compared to a model without a predictor of position), but no effect of context or interaction between position and context. VOT is longer for ejectives in C_2 (which is word-initial position) than in C_1 (which is word-medial position). The results are graphed in Figure 10 and the statistical results are given in Table 7.

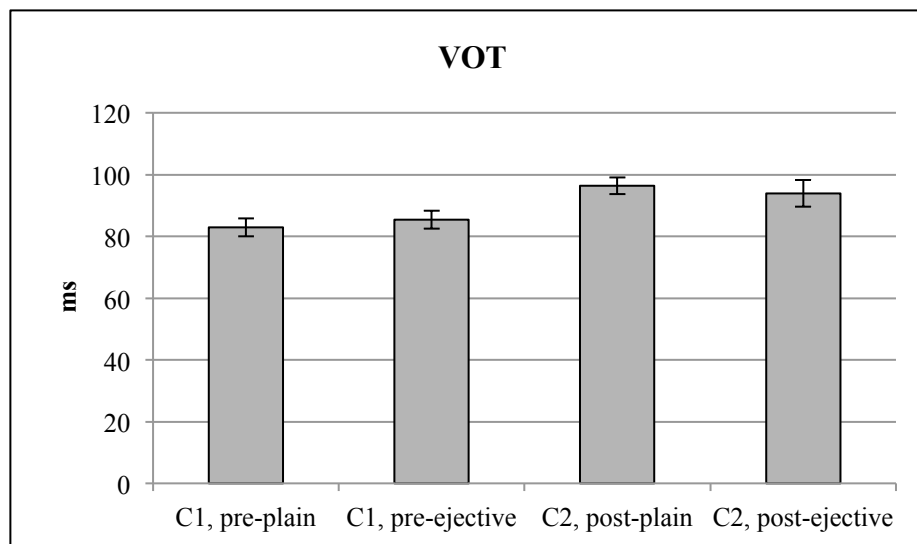


Figure 10: VOT by context and position. Error bars represent Standard Error.

	Estimate	Standard Error	<i>t</i>
intercept	91.39	6.19	14.76
<i>position</i>	11.37	3.35	3.40

Table 7: Results of a Linear Mixed Model testing for effects of context and position on ejective VOT.

The results for the measure of VOT by participant are given in Figure 11. Participants 7, 8, 9 and 10 are consistent with the overall result, while the other participants show small, varied differences between categories. For participant 6, VOT is shortest for post-ejective ejectives in C_2 , contrary to the overall pattern for longer VOT in C_2 .

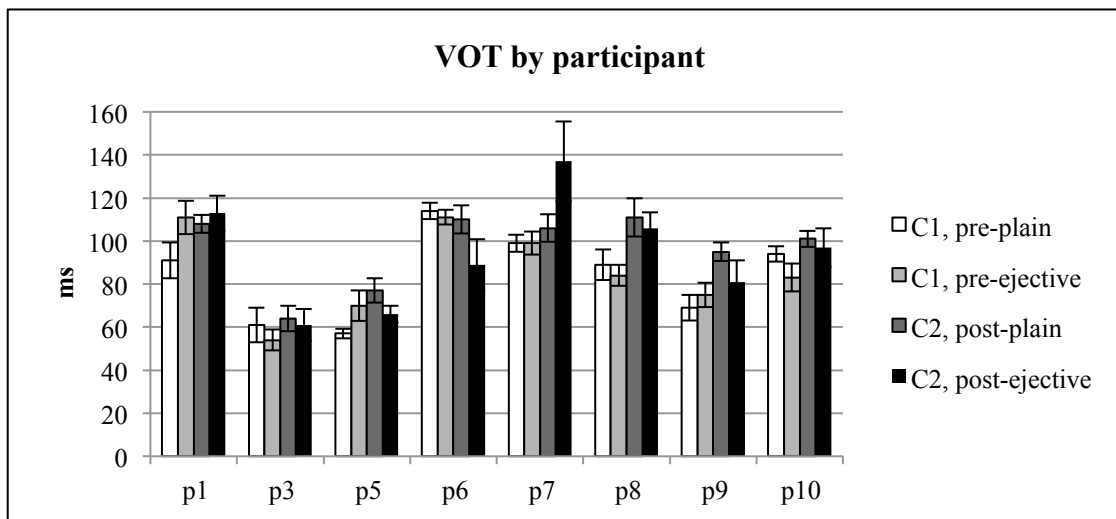


Figure 11: VOT by context, position and participant. Error bars represent Standard Error.

For pitch in the vowel following an ejective, there were no significant effects of context, position, or their interaction. The effect of an ejective on the phonation quality of the following vowel is comparable between ejectives that are preceded by another ejective and those that are preceded by a non-ejective consonant (e.g., [a] in [hutʃʷuj kʰanka] vs. [tuʌu kʰanka]), as well as between ejectives that are followed by another ejective and those that are followed by a non-ejective consonant (e.g., [a] in [hankʰa tʰanta] vs. [hankʰa muhu]). If the production of an ejective were substantially altered by the presence of another, preceding or following ejective, the alignment between the glottal and oral closures could be affected, which in turn may lead to different degrees of influence of the ejective on a following vowel. The acoustic measures, both of pitch and of VOT and intensity difference, suggest that the articulation of an ejective is unaffected by the presence of a preceding or following ejective.

3 Discussion

We began this study with three research questions, (1) are pairs of ejectives subject to de-ejectivization? (2) is the boundary between two ejectives lengthened? and (3) are the acoustic properties of ejectives different in the context of another ejective? We found evidence of both de-ejectivization and boundary lengthening, leading to affirmative answers to questions (1) and (2), but we found no reliable indications of either lenition or hyperarticulation in ejectives that are preceded or followed by another ejective, leading to a negative answer to question (3).

Participants are more likely to de-ejectivize a target ejective in the ejective-ejective category, where there are two ejectives, than in the ejective-plain or plain-ejective categories, where there is only one ejective. Furthermore, in the cases where de-ejectivization occurred in the ejective-ejective category, it is more likely for the second than the first ejective to surface as plain, e.g., /miskʰi tʰanta/ is more likely to be misproduced as [miskʰi tanta] than as [miski tʰanta]. This preference for preserving the ejective in C₁ is also seen in the higher accuracy rate of the ejective-plain category than the plain-ejective category, although the difference between the two

does not reach significance. While de-ejectivization occurs at a significant rate overall, examination of individual participants reveals that de-ejectivization is only prevalent (de-ejectivization occurs on 5 or more of the 10 target items) for four speakers, but rare or unattested for the others (de-ejectivization occurs on 0, 1 or 2 of the 10 target items). All speakers produced some sequences of two ejectives, which is consistent with the description that the cooccurrence restriction on ejectives holds within the root. Across word boundaries, trans-vocalic ejectives are optionally subject to de-ejectivization.

Participants who produce pairs of ejectives relatively accurately generally lengthen both the vowel between two ejectives and the closure duration of the second ejective, relative to other categories. The finding of lengthening of the vowel and closure duration between ejectives is consistent with Walter's (2007) finding that vowels were lengthened between repeated place gestures, and consistent with the hypothesis that cooccurrence restrictions for both ejectives and place of articulation have their phonetic roots in the articulatory difficulty associated with repeating certain gestures within a small temporal frame. When producing multiple ejectives, or multiple, identical place gestures, speakers modify their productions to lengthen the interval between the repeated gestures. The increase in closure duration may result from a lengthening of the stop closure gesture, but it is also conceivable that this period includes a pause between the adjective and noun in the adjective-noun phrase. In a vowel-stop sequence, closure duration and pause are acoustically indistinguishable, and thus it is possible that closure duration measures included short pauses with no articulated stop closure.

The acoustic analysis of ejectives themselves did not find any significant results. When participants produce pairs of ejectives, those ejectives have the same burst, VOT and pitch in the following vowel as other ejectives. There is no evidence that ejectives induce hyperarticulation or lenition of other, surrounding ejectives. Rather, ejectives that are produced as such are unaffected by the surrounding segmental context.

While the results here are only directly relevant to an articulatory hypothesis about the basis of cooccurrence restrictions, they do not contradict previous proposals that dissimilatory cooccurrence restrictions are perceptually grounded (Ohala 1981, 1993; Gallagher 2010a,b, 2012). Perceptually based accounts of dissimilation assume that the distinction between a form with one ejective and a form with two ejectives is relatively indistinct (e.g., that [k'ap'i] is confusable with [k'api]), and thus that languages eliminate one of these forms in order to render existing forms more distinctive. Evidence for this analysis has come from perception studies with English speakers (Gallagher 2010a,b, 2012), but the perceptual properties of ejective pairs have not been studied with speakers of a language with phonemic ejectives. Ongoing work is investigating Quechua speakers' perception of pairs of ejectives, both within hypothetical roots and across word boundaries. It may be the case that in addition to being articulatorily challenging, pairs of ejectives are perceptually challenging, even for speakers with a phonemic distinction between ejectives and plain stops.

The finding that the boundary between two ejectives is lengthened by increasing the duration of both the vowel and the ejective closure suggests that the temporal distance between the two ejectives is directly related to how difficult a transvocalic ejective pair is to produce. The importance of temporal proximity may provide an insight into why cooccurrence restrictions

may hold within roots, but not across morpheme or word boundaries, as is the case in many languages. In Articulatory Phonology (Browman & Goldstein 1986, 1992, et seq.), the degree of overlap between gestures is directly specified and manipulated. Multiple studies have found that timing relations between adjacent gestures within a word or morpheme are more stable than relations between gestures that cross a prosodic boundary (Hardcastle 1985; Holst & Nolan 1995; Byrd 1998; Byrd et al. 2000; Cho 2001). Pairs of ejectives, or other effortful structures, may be grammatical across boundaries just because speakers can exploit the more variable timing relations between gestures, and actively increase the duration between effortful gestures. For the particular case at hand, this line of analysis would predict that the timing between V_1 and C_2 in a $C_1V_1C_2V_2$ root is more stable, and more overlapped, than the timing between V_1 and C_2 in a $C_1V_1\#C_2$ string that straddles a word boundary. If true, this would explain why speakers cannot recruit vowel and closure duration lengthening root-internally to make pairs of ejectives less effortful within roots.

There is an aspect of the CQ cooccurrence restriction that complicates a simple explanation based on articulatory difficulty. Roots in CQ may contain a coda consonant in the first syllable, and thus the prohibited structure is not just trans-vocalic ejectives, as in *[k'ap'i], but also ejectives separated by both a consonant and a vowel, as in *[k'asp'i]. The effect of vowel lengthening found here is an increase of 12 ms (the average duration of [a] in the ejective-ejective category is 85 ms compared to 73 ms in the other categories), and the increase in closure duration is 26 ms (the average closure duration in the ejective-ejective category is 111 ms. compared to 85 ms. in the plain-ejective category). This increase in the duration of the boundary between two ejectives, a total of 34 ms., is smaller than the increase if a coda is introduced. In a previous study, Gallagher (to appear) asked CQ speakers to repeat nonce roots with pairs of ejectives, like *[k'asp'i] and *[k'ap'i]. In correct productions, the average vowel duration in CVCV roots was 129 ms. and the vowel+coda duration in CVCCV roots was 276 ms., an increase of 149 ms. Despite this increase in duration between the two target ejectives, participants were not overall more accurate on stimuli with a coda than those without, i.e., in this task, an increase in temporal distance between two ejectives did not correlate with an increase in accuracy. The results of the repetition task and the production task presented here suggest that the relationship between articulatory difficulty and grammaticality is not direct. Rather, CQ speakers have a categorical restriction on pairs of ejectives within a root, that is unaffected by the phonetic, sub-phonemic properties of the root. Across word boundaries, where there is no categorical restriction on pairs of ejectives, phonetic factors are actively manipulated and reveal a potential source for the grammaticalized avoidance of multiple ejectives within roots.

References

- Adelaar, William with Pieter Muysken. 2004. *The languages of the Andes*. Cambridge: Cambridge University Press.
- Ajacopa, Teofila Laime, et al. 2007. *Diccionario Bilingüe, Iskay Simipi Yuyayk'anacha: Quechua – Castellano Castellano – Quechua*. La Paz, Bolivia.
- Bates, Doug and Martin Maechler. 2010. lme4: Linear mixed effects models using S4 classes. <http://CRAN.Rproject.org/package=lme4>.

- Bills, Garland and Rudolph C Troike; Bernardo Vallejo. 1971. *An Introduction to Spoken Bolivian Quechua*. Austin:University of Texas Press.
- Boersma, Paul and David Weenink. 2012. Praat: doing phonetics by computer. Version 5.1.30, available at www.praat.org.
- Browman, C., & Goldstein, L. 1986. Towards an articulatory phonology. *Phonology Yearbook*, 3, 219–252.
- Browman, C., & Goldstein, L. 1992. Articulatory phonology: An overview. *Phonetica*, 49, 155–180.
- Byrd, Dani. 1998. Intergestural coordination adjacent to multiple boundaries. *Proceedings of the 16th International Congress of Acoustics and the 135th Meeting of the Acoustical Society of America*: 1255-1256.
- Byrd, Dani, Abigail Kaun, S. Narayanan & E. Saltzman. 2000. Phrasal signatures in articulation. In Michael Broe and Janet Pierrehumbert (eds.), *Papers in Laboratory Phonology V: Language acquisition and the lexicon*: 70-87. Cambridge: Cambridge University Press.
- Catford, J.C. 2001. *A practical introduction to phonetics*, 2nd ed. New York: Oxford University Press.
- Cho, Taehong. 2001. Effects of morpheme boundaries on inter-gestural timing: evidence from Korean. *Phonetica* 58: 129-162.
- Coetzee, Andries and Joe Pater. 2008. Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic. *Natural Language and Linguistic Theory* 26:289-337.
- Dryer, Matthew S. & Martin Haspelmath (eds.). 2011. *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library. Available online at <http://wals.info/>
- Frisch, Stefan, Michael Broe and Janet Pierrehumbert. 2004. Similarity avoidance and the OCP. *Natural Language and Linguistic Theory* 22: 179-228.
- Gallagher, Gillian. 2010a. The perceptual basis of long-distance laryngeal restrictions. Doctoral Dissertation, Massachusetts Institute of Technology.
- Gallagher, Gillian. 2010b. Perceptual distinctness and long-distance laryngeal restrictions. *Phonology* 27:435-480.
- Gallagher, Gillian. 2011. Auditory features in Phonology – The case for [long VOT]. *The Linguistic Review* 28:281-313.
- Gallagher, Gillian. to appear. Speaker awareness of non-local ejective phonotactics in Cochabamba Quechua. *Natural Language and Linguistic Theory*.
- Gelman, Andrew & Jennifer Hill. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.
- Goldsmith, John. 1976. *Autosegmental Phonology*. Doctoral Dissertation, Massachusetts Institute of Technology.
- Greenberg, Joseph. 1960. The patterning of root morphemes in Semitic. *Word* 6:162-181.
- Hardcastle, W.J. 1985. Some phonetic and syntactic constraints on lingual coarticulation during /kl/ sequences. *Speech Communication* 4:247-263.
- Herbert, R.K. 1977. Phonetic analysis in phonological description: prenasalized consonants and Meinhof's Rule. *Lingua* 43:339-373.
- Holst, T. & F. Nolan. 1995. The influence of syntactic structure on [s] and [ʃ] assimilation. In Arvaniti Connell (ed.), *Papers in Laboratory Phonology IV: Phonology and phonetic evidence*: 315-333. Cambridge: Cambridge University Press.
- Ito, Junko & Armin Mester. 2003. *Japanese morphophonemics: markedness and word structure*. Cambridge, MA: MIT Press.
- Kawahara, Shigeto. 2008. Phonetic naturalness and unnaturalness in Japanese loanword phonology. *Journal of East Asian Linguistics* 17:317-330.

- Kingston, John. 1985. The phonetics and phonology of the timing of oral and glottal events. Doctoral Dissertation, University of California, Berkeley.
- Ladefoged, Peter. 1993. *A course in phonetics*, 3rd ed. Fort Worth, TX: Harcourt Brace College.
- Ladefoged, Peter and Ian Maddieson. 1996. *The Sounds of the World's Languages*. Oxford: Blackwell.
- Leben, W. 1973. Suprasegmental Phonology. Doctoral Dissertation, Massachusetts Institute of Technology.
- Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.). 2013. *Ethnologue: Languages of the World, 17th edition*. Dallas, TX: SIL International. Online at www.ethnologue.com.
- Lindau, Mona. 1984. Phonetic differences in glottalic consonants. *Journal of Phonetics* 12: 147-155.
- MacEachern, Margaret. 1999. *Laryngeal co-occurrence restrictions*. New York: Garland.
- Maddieson, Ian. 2011. Glottalized consonants. In Matthew Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*, chapter 7. Munich: Max Planck Digital Library. Available online at <http://wals.info/chapter/7>
- McCarthy, John. 1986. OCP effects: gemination and anti-gemination. *Linguistic Inquiry* 17:207-63.
- McCarthy, John. 1988. Feature geometry and dependency: A review. *Phonetica* 43: 84-108.
- Meinhof, C. 1932. *Introduction to the phonology of the Bantu languages*. Berlin: Dietrick Reimer/Ernst Vohsen.
- Mester, Armin and Junko Itô. 1986. The phonology of voicing in Japanese. *Linguistic Inquiry* 17:49-73.
- Ohala, John J. 1981. The listener as a source of sound change. In C. S. Masek, R. A. Hendrick and M. F. Miller (eds.), *Papers from the Parasession on Language and Behavior*, 178-203. Chicago Linguistics Society: Chicago.
- Ohala, John J. 1993. The phonetics of sound change. In Charles Jones, ed. *Historical Linguistics: Problems and Perspectives*, 237-278. London: Longman.
- Parker, Steve and David Weber. 1996. Glottalized and aspirated stops in Cuzco Quechua. *International Journal of American Linguistics* 62:70-85.
- Parker, Steve. 2007. Un análisis acústico del quechua del Cusco. *GIALens* 3.
- Rose, Sharon and Lisa King. 2007. Speech error elicitation and co-occurrence restrictions in two Ethiopian Semitic languages. *Language and Speech* 50:451-504.
- van den Berg, Rene & Ode Sidu. 1996. *Muna-English Dictionary*. Leiden: KITLV Press.
- Walter, Mary Ann. 2007. *Repetition avoidance in human language*. Doctoral dissertation, Massachusetts Institute of Technology.
- Vicenič, Chad. 2010. An acoustic study of Georgian stop consonants. *Journal of the International Phonetic Association* 40:59-92.
- Gordon, Matthew & Alya Applebaum. 2004. Phonetic structures of Turkish Kabardian. *Journal of the International Phonetic Association* 36:159-186.
- Flemming, Edward, Peter Ladefoged & Sarah Thomason. 2008. Phonetic structures of Montana Salish. *Journal of Phonetics* 36:465-491.
- Shosted, Ryan & Sharon Rose. 2011. Affricating ejective fricatives: The case of Tigrinya. *Journal of the International Phonetic Association* 41:41-65.
- Wright, R., S. Hargus and K. Davis. 2002. On the categorization of ejectives: data from Witsuwit'en. *Journal of the International Phonetic Association* 32:43-77.