An identity bias in phonotactics: Evidence from Cochabamba Quechua

Gillian Gallagher, New York University

January, 2014

Abstract

Speakers of Cochabamba Quechua (CQ) participated in two tasks involving phonotactically illegal nonce forms with pairs of identical (e.g., [p'ap'u) and non-identical ejectives (e.g., [k'ap'u]). In a repetition task, speakers were more accurate on identical than non-identical ejective pairs, though no asymmetry was found an ABX discrimination task, nor in acoustic analysis of nonce roots with identical and non-identical ejective pairs. The latent preference for identical ejectives is unexpected given the phonotactics of CQ, which categorically disallows both identical and non-identical ejective pairs. The asymmetry is in accord with the typology, however. Many languages systematically exempt identical segments from a phonotactic restriction that applies to non-identical segments. It is argued that this cross-linguistic identity preference has its roots in a synchronic bias in favor of identical segments.

0. Introduction

Cross-linguistically, it is common for cooccurrence restrictions to apply only to non-identical pairs of segments, while allowing identical segments to co-occur freely, a phenomenon known as the 'identity effect' (MacEachern 1999). This pattern is frequently found in cooccurrence restrictions on ejectives and other laryngeally marked segments (MacEachern 1999), as in Bolivian Aymara, where pairs of identical ejectives may co-occur in a root, e.g., [t'ant'a] 'bread', but pairs of non-identical ejectives are categorically absent, *[t'ank'a]. This paper presents evidence of a latent preference for nonce roots with identical pairs of ejectives (e.g., [p'ap'u]) over non-identical pairs (e.g., [k'ap'u]) in speakers of Cochabamba Quechua (CQ), a language in which roots with both identical and non-identical pairs of ejectives are categorically unattested.

The current study contributes to the body of research documenting a correlation between typology and synchronic grammars (Greenberg & Jenkins 1964; Broselow & Finer 1991; Broselow, Chen & Wang 1998; Berent et al. 2007; Berent & Lennertz 2010), and investigating the sources of such effects in detailed generalization from the lexicon (Daland et al. 2011) and/or learned interpretation of phonetic properties (Davidson & Shaw 2012). Accounts of the current findings based on either lexical generalization or phonetic preferences are presented, though the data do not conclusively choose between these explanations. Further experiments to tease apart the contribution of lexical and phonetic information are sketched.

The paper is organized as follows. Section 1 lays out the identity effect as a typological phenomenon and Section 2 lays out the relevant details of Cochabamba Quechua. Section 3 reports on the repetition study, followed by the discrimination study in Section 4. A general discussion of the findings is found in Section 5, and Section 6 concludes.

1. The identity effect in cooccurrence restrictions

1.1. The pattern

Many languages distinguish between identical and non-identical segments with respect to some phonotactic restriction, allowing identical pairs of segments but disallowing non-identical pairs (MacEachern 1999). In (1), data from Bolivian Aymara illustrate the pattern; the data are from

de Lucca's (1987) dictionary and are also discussed in MacEachern (1999) and Gallagher (2010a,b). Disyllabic roots in Bolivian Aymara may contain a single ejective, as in (1a), while roots with two non-identical ejectives are unattested (1b). Roots with two identical ejectives, however, are attested (1c).[1]

(1)   a. tʃʼaka 'bone'   b. *tʃʼakʼa   c. tʃʼatʃʼa 'to soak'

           k'apa 'cartilage'   *k'ap'a   k'aʎk'u 'acidic'

           p'eqe 'head'   *p'eq'e   p'ap'i 'type of fish'

           q'olti 'drink'   *q'olt'i   q'oq'e 'to catch fire'

           t'aku 'calm'   *t'ak'u   t'ant'a 'bread'

The identity effect is found in many languages that restrict the cooccurrence of laryngeally marked consonants like ejectives, aspirates and implosives. MacEachern (1999) identifies the identity effect in laryngeal cooccurrence restrictions in Tz'utujil (Mayan), Hausa (Afro-Asiatic), Peruvian Aymara and Gojri (Indo-Aryan), and the pattern is also found for ejectives in Chol (Mayan) (Gallagher & Coon 2009) and Yucatec Mayan (Straight 1976), and for implosives in Muna (Austronesian) (van den Berg & Sidu 1996; Coetzee & Pater 2008).

    Languages with laryngeal cooccurrence restrictions that do not distinguish between identical and non-identical segments are also common. This pattern is found for ejectives in Bolivian Quechua, Shuswap (Salish) (MacEachern 1999) and Yapese (Austronesian) (Jensen 1977), and for aspirates in Quechua, Ofo (Siouan), Sanskrit (Indo-Aryan) and Souletin Basque (isolate) (MacEachern 1999).

    Cooccurrence restrictions on major place of articulation may also show the identity effect. In many languages, pairs of non-identical homorganic consonants are under-attested or unattested, though pairs of identical consonants are well-attested. This pattern is well-documented in Muna (Coetzee & Pater 2008), Semitic languages (McCarthy 1986; Berent & Shimron 1997; Frisch et al. 2004; Frisch & Zawaydeh 2001) and in Ngbaka (Thomas 1963; Mester 1986). In Semitic languages, identical consonant pairs may only occur in certain positions within the root and are frequently analyzed as corresponding to a single underlying consonant (McCarthy 1986, 1988), or as containing reduplicated structures (Gafos 1998).

1.2 Identity and the grammar

The existence of the identity effect requires the phonological grammar to include some mechanism for distinguishing between identical and non-identical segments, either in phonological representations, as with double linking in Autosegmental Phonology (Leben 1973; Goldsmith 1976), or in the structure of grammatical statements, as in Coetzee & Pater's (2008) formulation of the Obligatory Contour Principle (OCP) (McCarthy 1986, 1988; Suzuki 1998;

---

[1] Like CQ, Bolivian Aymara has primarily disyllabic roots, though no counterexamples to the generalizations here are found in the few trisyllabic roots. Suffixes may contain ejectives in Bolivian Aymara, resulting in words with multiple ejectives across a root+suffix boundary (de Lucca 1987).

Coetzee 2009) to refer to identical or non-identical classes of segments. Both approaches have also been formalized in inductive models of phonotactic learning. Colavin et al. (2011) present a modification of the UCLA phonotactic learner (Hayes & Wilson 2008), that enriches the representations given to the learner. In this model, identical segment pairs are represented with a single fully specified feature matrix and a copy, a featurally emtpy segment. Phonotactic restrictions on pairs of identical and non-identical segments must then be learned separately. Berent et al. (2012) propose a modification to the UCLA phonotactic learner that allows the learner to build constraints that penalize only the identical members of some class, like labial consonants. The proposed model learns constraints on the distribution of identical segments, constraints of the form *$[+F]_i[+F]_i$, and could presumably be further modified to learn constraints on the distribution of non-identical segments, *$[+F]_i[+F]_j$.

In addition to the existence of languages that distinguish identical and non-identical classes of segments in their phonologies, further evidence for this distinction comes from experimental work with both adults and infants. Gallagher (2013a) finds that adult speakers of American English learn a pattern in which made-up words in an artificial language with identical consonants behave differently than those with non-identical consonants. Participants learn this identity based pattern better than an arbitrary distinction between consonant pairs, and generalize the identity based pattern to novel identical and non-identical pairs of consonants, indicating that the identity based generalization is represented by learners as such. While Gallagher (2013a) shows that identity behaves like other natural classes in being targeted by generalizations, Endress et al. (2007) and Gervain et al. (2008) find evidence that identity based patterns may be priveleged compared to other systematic patterns. Endress et al. (2007) found that a tonal pattern based on identity is learned better by adult Italian speakers than a systematic pattern that does not contain an identity relation. In an optical brain-imaging experiment, Gervain et al. (2008) played newborns in Italian speaking families auditory stimuli containing syllable repetitions, e.g., *mubaba*, and found that the newborns distinguished the stimuli with repetitions from control stimuli after very few trials. They interpret these results as evidence for a mechanism in the perceptual system to automatically detect repetitions. These results corroborate the typological data in supporting explicit reference to the identity relation in grammatical generalizations.

1.3 Connecting typology and synchronic grammars
The identity effect is quite well attested cross-linguistically, raising the question of what makes identical segments grammatically favorable. Typological patterns have been argued to arise from universal biases about the shape of grammars, or the types of grammatical restrictions that can be posited or exploited in synchronic grammars (Chomsky & Halle 1968; Jakobson 1941; Trubetzkoy 1958). Moreton (2008) argues for analytic biases that favor constraints with certain formal structure, e.g., constraints that reference a single autosegmental tier, that may be active within a theory of inductive learning (Hayes 1999; Albright & Hayes 2003; Hayes & Wilson 2008). Moreton & Pater (2012a,b) further establish evidence for a role for the formal simplicity of generalizations in how easily generalizations are learned.

Alternatively, cross-linguistic patterns may arise from substantive biases in learning phonological patterns. Patterns that favor phonetically preferable forms, those that are perceptually distinctive or articulatorily easy, are easier to learn and are predicted to be more common (Steriade 1997; Hayes 1999; Hayes, Kirchner & Steriade 2004). Blevins (2004) also argues that phonetic factors influence the shape of phonological grammars, but only indirectly, through sound change. Phonetically difficult structures are more likely to undergo change, resulting in phonetically advantageous grammars.

If identical ejectives, or identical segments more generally, are favored cross-linguistically because of a synchronic bias (formal, substantive, or both), then we expect to see the asymmetry between identical and non-identical ejectives reflected in the behavior speakers of languages *without* the identity effect. In a language where all pairs of ejectives are restricted, like Cochabamba Quechua, speakers may still find identical ejectives easier to produce, perceive or process than non-identical ejectives.

2. Background on Cochabamba Quechua

CQ, like other Quechua varieties spoken in Southern Peru and Bolivia (e.g., Cuzco Quechua (Parker & Weber 1996; Parker 1997; MacEachern 1999)), contrasts ejectives, aspirates and plain stops/affricates at five places of articulation. The phonemic inventory of CQ is given in Table 1.

	labial	dental	palatoalveolar	palatal	velar	uvular	glottal
plain	p	t	tʃ		k	q	ʔ
aspirate	pʰ	tʰ	tʃʰ		kʰ	qʰ	
ejective	p'	t'	tʃ'		k'	q'	
fricative		s	ʃ				h
nasal	m	n		ɲ			
liquid		l ɾ		ʎ			
glide	w			j			

Table 1: Cochabamba Quechua consonant inventory.

Ejectives in CQ occur only in onset position in roots, which are primarily CV(C)CV, and are prohibited from co-occurring in pairs. While ejectives may occur in both initial and medial position in a root (2a), and identical and homorganic stops are generally attested (2b), both identical and non-identical ejective pairs are categorically absent from the language (2c).

(2) a. k'inti 'a pair'   sut'i 'clear'   c. *k'int'i   *k'unk'a
        p'atʃa 'clothes'   satʃ'a 'tree'   *p'atʃ'a   *t'ant'a

    b. k'anka 'rooster'   kunka 'neck'
        t'anta 'bread'   tanta 'gather'

Laryngeally marked consonants are subject to several additional restrictions in CQ, which are not central to this paper. Ejectives are found in medial position only in roots with an initial fricative or sonorant consonant, e.g., [mat'i] 'forehead', but are absent from roots with an initial plain stop, e.g., *[pat'i]. Further, ejectives may not occur in vowel initial words, which are analyzed as glottal stop initial (Parker & Weber 1996; Parker 1997; MacEachern 1999; Gallagher 2011; Bennet 2013), e.g., *[ʔap'i]. Aspirates show the same patterns as ejectives: they may not co-occur in pairs (e.g., *[pʰatʰi], *[pʰapʰi]), nor may an aspirate follow a plain stop in the root (e.g., *[patʰi]. Aspirates may also not co-occur with the glottal fricative, e.g., *[hapʰi]. Finally, ejectives and aspirates may not co-occur with one another, e.g., *[pʰat'i], *[p'atʰi].

Gallagher (2013b) provides experimental evidence that CQ speakers are sensitive to (some of) these static phonotactic restriction in their language. In this study, CQ speakers were asked to repeat nonce words with a medial ejective that were either phonotactically legal controls (e.g., [sap'i]) or were phonotactically illegal items that violated the cooccurrence (e.g., [k'ap'u]) or the ordering restriction (e.g., [kap'u]) on ejectives. Participants made more errors on illegal items than control items, often repairing the phonotactic violation and producing a phonotactically legal form. These results are interpreted as evidence that speakers of CQ have learned these phonotactic restrictions in their language, as has been found for other non-local restrictions on consonant pairs in Arabic (Frisch & Zawaydeh 2001) as well as Amharic and Chaha (Rose & King 2007).

The current study follows up on the results of Gallagher (2013b), exploring the representation of the cooccurrence restriction in more detail by comparing identical and non-identical ejective pairs. Furthermore, the current study pairs a repetition task with a perception task, allowing for a better understanding of the mechanisms involved in repetition errors.

3. Experiment 1 – nonce word repetition

Experiment 1 presented CQ speakers with nonce words, some of which were phonotactically legal (e.g., [huk'a]) and some of which contained a phonotactically illegal pair of ejectives (e.g., [k'up'a]). Accurate repetition was compared between legal and illegal nonce words, and also between illegal forms with identical and non-identical pairs of ejectives. Additionally, the acoustic properties of responses were analyzed.

3.1. Method

3.1.1. Participants

The experiment was conducted in Cochabamba, Bolivia with 22 native speakers of Quechua from the Cochabamba area. The participants were 4 men and 18 women, all in their 20s and 30s. Further demographic information (e.g., education level) about the participants was not collected, though the majority were recruited through contacts at the major university in Cochabamba (Universidad Mayor de San Simón). All the participants were bilingual in Quechua and Spanish, and Spanish was used as the metalanguage to conduct the experiment. Participants were paid the equivalent of about $7 dollars for their participation.

Data from one participant was removed from analysis because this participant showed a pattern that was opposite from the pattern shown by all other subjects. This participant produced all target ejectives accurately, including in the phonotactically illegal categories, but also misproduced target plain stops in filler items as ejectives 2/3 of the time. While interesting, this participant's data were not included in either the analysis of accuracy or the acoustic analysis because they showed an outlier pattern.

3.1.2. Stimuli

The stimuli for the experiment were $C_1V_1C_2V_2$ nonsense words, classified into three categories of target items and three corresponding filler categories. All target items had an ejective stop, [p'] [tʃˀ] or [k'], in C_2, and were further divided into three categories according to C_1. In the 'control' items, C_1 was a fricative or sonorant [s h ɲ ʎ] and the resulting stimulus item was phonotactically legal, e.g., [hitʃˀu], [ɲak'u].[2] In the 'identical' items, C_1 was an ejective identical to C_2, resulting in a phonotactically illegal nonce word, e.g., [k'ak'u], [tʃˀitʃˀu]. Finally, in the 'non-identical' items, C_1 was an ejective stop that was non-identical to C_2, also resulting in a phonotactically illegal nonce word, e.g., [p'atʃˀi], [k'up'i]. The filler items were the same as the target items, except that C_2 was the corresponding plain stop or affricate [p tʃ k]. Stimuli were balanced for place of articulation of C_1 and C_2; that is, each unique C_1-C_2 combination in each category appeared in an equal number of stimuli.

The vowels in the stimuli were always non-identical, and were roughly balanced between categories. Perfect balancing for vowel patterns was impossible since many potential filler items are actual, existing words. The stimuli were designed to also be used for the perception study presented below in §4, and thus it was crucial that targets be matched precisely with fillers.

There were 12 items in each category, for a total of 72 stimuli, shown in Table 2 below. It was not possible to make 12 unique items for the 'identical' and 'non-identical' categories without using attested, meaningful words as fillers, so there are 6 items that were repeated. Repeated items are given in italics in the table below, they are: [k'ak'u] (target identical), [k'aku] (filler identical), [p'atʃˀi] (target non-identical], [p'atʃi] (filler non-identical), [tʃˀup'i] (target non-identical) and [tʃˀupi] (filler non-identical).

---

[2] The palatal nasal and lateral [ɲ, ʎ] were used in the control stimuli instead of their cross-linguistically less marked dental counterparts because the palatals are much more common in CQ than the dentals, which are actually somewhat rare.

target control		target identical		target non-identical	
suk'i	ʎuk'a	p'ap'u	tʃʼutʃʼi	p'atʃʼi	tʃʼup'i
sitʃʼa	ʎatʃʼu	p'up'a	tʃʼitʃʼa	*p'atʃʼi*	*tʃʼup'i*
sap'i	ʎup'i	p'up'i	k'ak'u	p'uk'i	k'ap'u
huk'a	ɲak'u	p'ip'a	*k'ak'u*	p'uk'a	k'up'i
hitʃʼu	ɲutʃʼi	tʃʼatʃʼi	k'uk'a	tʃʼak'u	k'atʃʼu
hap'u	ɲup'a	tʃʼutʃʼa	k'uk'i	tʃʼuk'a	k'itʃʼa
filler control		**filler identical**		**filler non-identical**	
suki	ʎuka	p'apu	tʃʼutʃi	p'atʃi	tʃʼupi
sitʃa	ʎatʃu	p'upa	tʃʼitʃa	*p'atʃi*	*tʃʼupi*
sapi	ʎupi	p'upi	k'aku	p'uki	k'apu
huka	ɲaku	p'ipa	*k'aku*	p'uka	k'upi
hitʃu	ɲutʃi	tʃʼatʃi	k'uka	tʃʼaku	k'atʃu
hapu	ɲupa	tʃʼutʃa	k'uki	tʃʼuka	k'itʃa

Table 2: Stimulus items for the repetition experiment. Italics indicate tokens that were repeated.

The stimuli focused on three of the five places of articulation, excluding the dentals and uvulars [t t' q q']. These three places of articulation were chosen largely to simplify the design of the experiment by limiting the variety in the stimuli. As the acoustic and articulatory properties of ejection differ by place of articulation (McDonough & Ladefoged 1993; Gordon 1996; Gordon & Applebaum 2006; Flemming et al. 2008; Gallagher 2011), place must be taken into account in the analysis. Having only three places of articulation to compare, as opposed to five, simplifies the analysis and makes the results easier to interpret. Furthermore, previous, informal work with CQ speakers has found that the labial and dental ejectives are often confused with one another, as are the velar and uvular ejectives. The labial and velars were chosen from these pairs because they are more common in the Quechua lexicon (as estimated from a root list taken from Laime Ajacopa 2007).

The stimuli were made by splicing initial and medial syllables together from productions of a native Quechua speaker reading phonotactically legal nonce words. The speaker was a 28 year old female from the city of Cochabamba, Bolivia, with training and experience as a Quechua teacher. Since all stimulus items have a stop or affricate in C_2, the stimuli were spliced together by making a cut during the silent closure of C_2. For example, the stimulus [k'ak'u] was made by splicing the initial CV of [k'aku] and the final CV of [ɲak'u], making the cut during the silent closure of the medial [k']. All stimuli, whether phonotactically legal or illegal, were spliced in this fashion. The stimuli were normalized for amplitude, using the 'scale peak' function in Praat (Boersma & Weenink 2011), and ejectives were additionally normalized for a Voice Onset Time of 130 ms by adding or removing silence during the VOT period. VOT is the main cue, along with burst amplitude, for ejection in CQ, and thus all ejectives in the stimuli were equally well cued along this dimension, though natural differences in VOT by place of articulation were obscured in the stimuli. As will be seen in the acoustic measurements presented below in §3.2.2, 130 ms is at the long end of the VOT range for all three ejectives.

3.1.3. Equipment

The stimuli were presented aurally through the experimental software Psyscope X, running on a Macbook Air, via Audio Technica noise cancelling headphones. Participants' repetitions of the stimuli were recorded with a Marantz PMD560 digital recorder at a sampling rate of 44 kHz and an Audio Technica 831b lapel microphone.

3.1.4. Procedure

The experiment was conducted in a hotel room in Cochabamba, Bolivia. The stimuli were presented aurally in random order to each participant.[3] On each trial, a single stimulus was repeated twice, 400 ms apart, and then the participant was asked to repeat what they heard as precisely as possible. Participants were told that the words had no meaning in Quechua or in any other language, though they contained sounds that were used in existing Quechua words. The instructions were given in Spanish, both in written form on the computer screen, and orally by the experimenter, and are included in the Appendix. Participants were allowed to practice with as many trials as they wanted (about 5 trials for most people) until they were comfortable with the task. Participants had as long as they needed to respond; once they had repeated the stimulus item, they pressed the space bar on the keyboard to move on to the next trial. The experiment took between 4 and 7 minutes depending on the pace of the individual participant.

3.1.5. Analysis
3.1.5.1. Transcription of responses

Participants' repetitions were transcribed based on listening to the sound file and examining the spectrogram. Responses were coded for accuracy and type of error, if any. 25% of the data were transcribed by a second transcriber; inter-transcriber reliability was 95%.

The vast majority of errors were productions of a target ejective as a plain stop, for example, a participant heard [k'up'i] and repeated it as [k'upi]. These misproductions of ejectives as plain stops are the error of main interest, since this error occurs most often when it changes a phonotactically ill-formed item to a well-formed item. In a smaller number of cases, target plain stops were produced as ejectives, e.g., target [huka] repeated as [huk'a]. In 6 tokens, the initial stop was de-ejectivized as opposed to the target, medial stop (e.g., target [p'uk'a] repeated as [puk'a]). In 1 token, both stops were de-ejectivized, target [tʃ'utʃ'a] was repeated as [tʃutʃa].

To determine whether a given stop or affricate was ejective or plain, impressionistic categorization was supplemented by examination of the spectrogram and waveform, as well as acoustic measurements. The principle cues to ejectives in Cochabamba Quechua are VOT and burst amplitude (Gallagher 2010a, 2011). To gauge whether a given stop in the phonotactically illegal categories is an ejective or a plain stop, the VOT and burst measures were compared to

---

[3] Stimuli were presented aurally as opposed to orthographically because comfort and familiarity with the standard orthography for Bolivian Quechua differs greatly between individuals. While aural presentation allows for misperception, orthographic presentation would allow for mis-reading, in addition to severely limiting the pool of possible participants. The potential effects of misperception in the repetition task were further quantified with the perception study in Experiment 2.

these same measures for (a) target control items, which have a phonotactically legal medial ejectives and (b) filler items, which have a (phonotactically legal) medial plain stop. In the majority of cases, the transcription of a stop or affricate as an ejective or plain stop was quite clear. In cases that were less clear, the acoustic measurements were closely attended to.

In addition to mismatches between ejectives and plain stops/affricates, a small number of other errors were also found, including change in a target vowel, e.g., [huka] produced as [hika], change in place of articulation of a target stop, e.g., [hup'a] produced as [hut'a], or change of a plain stop or ejective to an aspirate, e.g., [hup'a] produced as [hup^{h}a]. Vowel errors were ignored in the analysis (i.e., responses with vowel errors were counted as 'correct'). Errors in place of articulation were also ignored, unless they changed the category of the stimulus. For example, [k'up'a] produced as [k'ut'a] was ignored, because both [k'up'a] and [k'ut'a] have two, non-identical ejectives, but [p'up'a] produced as [p'ut'a] changes a stimulus with two identical ejectives to one with two non-identical ejectives. A total of 13 tokens were removed from analysis, 10 because of an error in place of articulation that changed the category of the stimulus, and 3 because of an ejective produced as an aspirate. In 6 of the 10 place of articulation errors, there is also de-ejectivization of the second consonant. The results presented in the next section do not change if these excluded tokens are included in the analysis as incorrect repetitions.

3.1.5.2 Acoustic analysis
Acoustic measurements of participants' responses to the repetition task were taken in order to quantify the properties of both correctly and incorrectly produced ejectives. For all responses, the following measurements were taken: duration of V_1, closure duration of C_2, burst intensity of C_2 and VOT of C_2.[4] A sample spectrogram and waveform is given in Figure 1, showing the segmentation of an utterance for these measurements. The duration of V_1 was delimited by the onset and offset of the second formant and the closure duration of C_2 was delimited by the offset of F2 in V_1 and the release burst. VOT was measured from the beginning of the release burst until the onset of a periodic waveform in the following vowel. To quantify burst amplitude, the difference between the maximum intensity in the burst itself and the minimum intensity in the period between the burst and the vowel was taken. For ejectives, this difference is usually quite large, because the burst itself has a high intensity and the silent period following the burst has a low intensity. For plain stops, the difference is small or zero.

---

[4] Ejectives are often associated with creaky phonation in the following vowel (Kingston 1985; Lindau 1984; Ham 2008; Wright, Hargus & Davis 2002; Vicenik 2010; Shosted & Rose 2011). H1-H2 measurements were taken using a Praat script by Chad Vicenik (available at: http://www.linguistics.ucla.edu/faciliti/facilities/acoustic/PraatVoiceSauceImitator.txt), to assess phonation quality (Ladefoged 2003; Vicenik 2010; Garellek & Keating 2011), but were not found to differ between ejectives and plain stops. H1-H2 was previously found to not distinguish ejectives and plain stops in Cochabamba Quechua in Gallagher (2011).
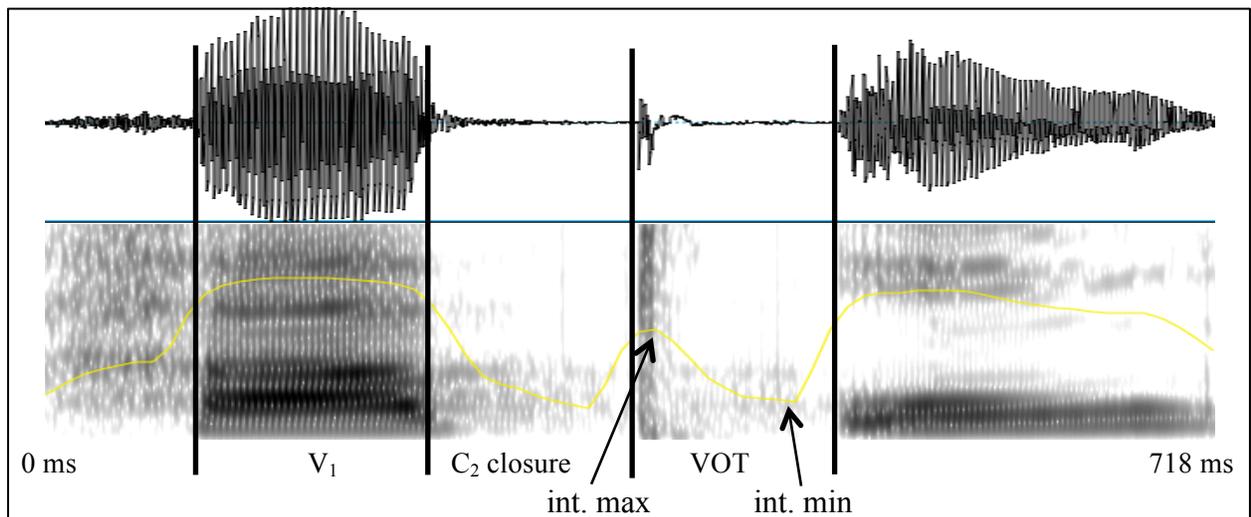
Figure 1: Spectrogram and waveform of [hap'u], showing landmarks for duration of V_1, C_2 closure and VOT, and burst intensity measurements. Intensity window is 40-100 dB.

Responses from participant 2 were not included in the acoustic analysis, due to excessive noisiness in the recordings (other participants talking loudly from the waiting area). Additionally, analysis of incorrect productions could only be done for participants that made some errors. Participants 3 and 4 were 100% correct across the board, so these participants were not included in the analyses of incorrect productions.

3.2.   Results

3.2.1.   Transcription results

For target items with a medial ejective (fillers with a medial plain stop are discussed below), participants performed nearly at ceiling (99% accurate) at accurately producing the phonotactically legal ejectives in control items (e.g., producing [huk'a] as [huk'a], not [huka]). For phonotactically illegal target items in the identical and non-identical categories, participants made a substantial number of errors, repeating the target medial ejective as a plain stop. More errors were made on pairs of non-identical ejectives (e.g., [k'up'a] produced as [k'upa]) than pairs of identical ejectives (e.g., [p'up'a] produced as [p'upa]); accuracy was 49% on the non-identical target items and 61% on the identical target items. The results are shown in Figure 2.
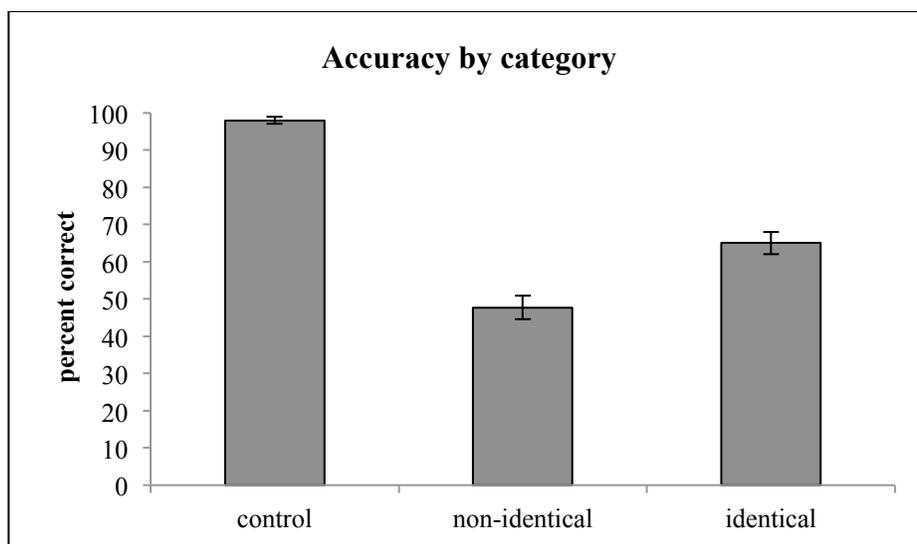
Figure 2: Accuracy on target items with medial ejectives in the phonotactically legal control category and the phonotactically illegal identical and non-identical categories. Error bars indicate standard error. Control category *n=250*, non-identical *n=245*, identical *n=252*.

In Figure 3, the results are broken down by place of articulation of the target ejective: labial, palatoalveolar or velar. As can be seen, the same pattern holds for all three places of articulation, though overall accuracy for the non-identical and identical categories is lower for the palatoalveolars than for the labials and velars.
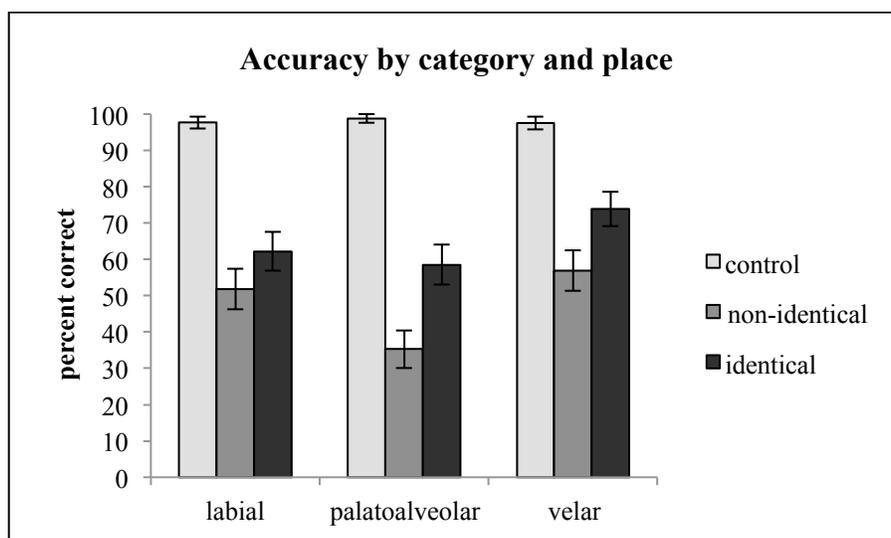

Figure 3: Accuracy on target items with medial ejectives, by category and place of articulation. Error bars indicate standard error. Labials: control *n=86*, non-identical *n=81*, identical *n=82*, palatoalveolars: control *n=84*, non-identical *n=85*, identical *n=82*, velars: control *n=80*, non-identical *n=79*, identical *n=88*.

A Mixed Logit Model was fit to test for an effect of category and place of articulation of the target stimulus on the accuracy of responses. The binomial, dependent variable was 'correct' or

'incorrect'. The model included a Helmert coded predictor of category, which compared the control category to the identical and non-identical categories as a group, and the identical and non-identical categories to one another, and also a Helmert coded predictor of place, which compared the palatoalveolars to the labials and velars as a group, and the labials and velars to one another. All interactions between category and place were also included. The model was fit using the *lmer*() function in the lme4 package (Bates & Maechler 2010) in the R software. All predictors were centered and random intercepts for participant and item were included, as well as random slopes for all predictors by participant. Model comparison was used to assess significance of predictors; predictors were removed one-by-one until removing a predictor resulted in a significant difference in the fit of the model.

The final model has three significant predictors. Both predictors of category are significant, showing that accuracy on the control category is significantly higher than accuracy on the identical and non-identical categories, and that accuracy on the identical category is significantly higher than accuracy on the non-identical category. The predictor of place that compares the palatoalveolars to the labials and velars as a group is also significant, showing that accuracy on the labials and velars is higher than on the palatoalveolars. The comparison between the labials and velars did not reach significance, nor did any of the interactions between place and category. Table 3 gives the final model.

	Estimate	Standard Error	Wald's z	p
intercept	2.80	0.37	7.61	< .0001
control vs. identical & non-identical	-6.13	0.89	-6.87	< .0001
identical vs. non-identical	-1.51	0.34	-4.42	< .0001
palatoalvear vs. labial & velar	1.50	0.34	4.40	< .0001

Table 3: Results of Mixed Logit Model with binomial dependent variable of accuracy and predictors of stimulus category and place of articulation for target items.

Of the 21 participants, 15 of the participants are consistent with the overall pattern, with higher accuracy on stimuli with identical ejectives than stimuli with non-identical ejectives. 4 participants show the opposite pattern, with higher accuracy on non-identical ejective pairs than identical pairs, and 2 show no difference.

Errors on filler items, with medial plain stops, were not very common, though the error rate did differ by category. Participants were 97% accurate on filler control items (an error like [huka] produced as [huk'a] was made 3% of the time), and 97% accurate on filler non-identical items (an error like [k'upa] produced as [k'up'a] was made 3% of the time), but only 89% accurate on filler identical items (an error like [p'upa] produced as [p'up'a] was made 11% of the time). The significance of these differences in error rate on filler items was assessed with a Mixed Logit Model with two predictors of category, one comparing identical filler items to the other filler items as a group, and the other comparing non-identical filler items and control filler items, and two predictors of place, one comparing palatoalveolars to labials and velars, and the

other comparing labials and velars to one another. All predictors were centered and the model also included random slopes for participant and random intercepts for participant and target. The only significant predictor is the comparison between identical filler items and the fillers in the other two categories, showing that the error rate of identical filler items is substantially higher than for control or non-identical filler items. The final model is given in Table 4.

	Estimate	Standard Error	Wald's z	p
intercept	4.75	0.46	10.42	< .0001
identical vs. control & non-identical	2.78	0.64	4.37	< .0001

Table 4: Results of Mixed Logit Model with binomial dependent variable of accuracy and predictors of stimulus category and place of articulation, for fillers.

The increased number of errors on the filler identical items is consistent with the errors on target ejectives: participants show a preference for pairs of identical ejectives over pairs of non-identical ejectives both in their accurate repetition of these forms and in their addition of ejectives to filler items.

3.2.2. *Acoustic results*

The analysis of acoustic measurements tests for three classes of effects. First, differences in the duration of V_1 and/or C_2 closure duration were compared between correct productions of the identical and non-identical categories, and between these two categories and controls. These measures quantify the temporal distance between ejectives, which may be actively manipulated by participants, as was found in Anonymous (under review) for pairs of ejectives across word boundaries in CQ. Second, differences in the VOT and/or burst amplitude of an ejective in C_2 were assessed between correct productions of the identical, non-identical and control categories. Third, differences in VOT and/or burst amplitude between incorrect productions of target ejectives in the identical and non-identical categories and correct productions of target plain stops in the filler categories. These measures of the VOT and burst properties of stops are meant to look for a small effects that might have been missed in transcription, e.g., reduction or hyperarticulation of ejectives in correctly produced phonotactically illegal roots, or relics of attempted in ejection in incorrect repetitions. Previous studies that have looked at the phonetic details of errors in experimental tasks have found acoustic and articulatory effects that were not reflected in transcription (Frisch & Wright 2002; Pouplier & Goldstein 2005; Goldrick & Blumstein 2006). The results pertaining to each of these effects are discussed in turn. All statistical models include maximal random effects structure, and significance is assessed via model comparison.

First, effects of the duration of V_1 or C_2 closure were assessed. For V_1 duration, the low vowels are generally longer than the high vowels, but no other big differences are apparent. The average durations are shown in Figure 4.
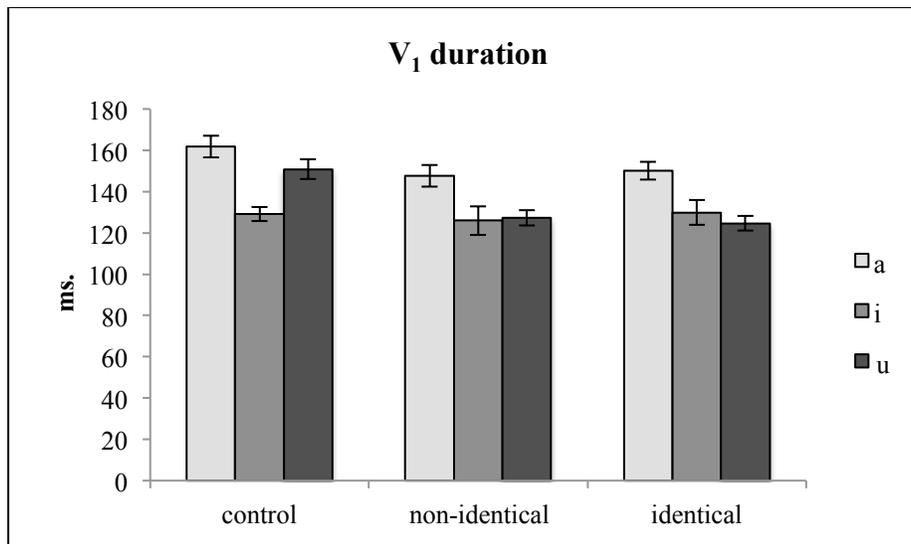
V₁ duration

Figure 4: Duration of V_1 in correct productions of ejectives in the non-identical, identical and control categories, by vowel. Error bars indicate Standard Error. Control [a] *n=58*, [i] *n=37*, [u] *n=82*; non-identical [a] *n=37*, [i] *n=11*, [u] *n=69*; identical [a] *n=60*, [i] *n=24*, [u] *n=82*.

To test for effects of vowel and category on the duration of V_1, a Linear Mixed Model was fit with V_1 duration as the dependent variable. There were two predictors of category: one predictor compared correct productions of the identical and non-identical categories as a group to the control category, and a second predictor compared correct productions of the identical and non-identical categories to each other. Additionally, the model included two predictors of vowel: one comparing the high vowels to the low vowel and the other comparing the two high vowels to one another. All predictors were centered, and all interactions between category and vowel were included. Only the predictor of high vs. low vowels was significant, due to the longer duration of low vowels. The final model is given in Table 5.

	Estimate	Standard Error	*t*
intercept	140.37	5.39	26.03
high vs. low vowels	23.63	6.32	3.74

Table 5: Results of Linear Mixed Model with dependent variable of V_1 duration and predictors of stimulus category and vowel.

The average C_2 closure durations are shown in Figure 5. Labial stops tend to have a slightly longer closure than velars, which in turn have a slightly longer closure than palatoalveolars. No clear differences between categories are apparent.
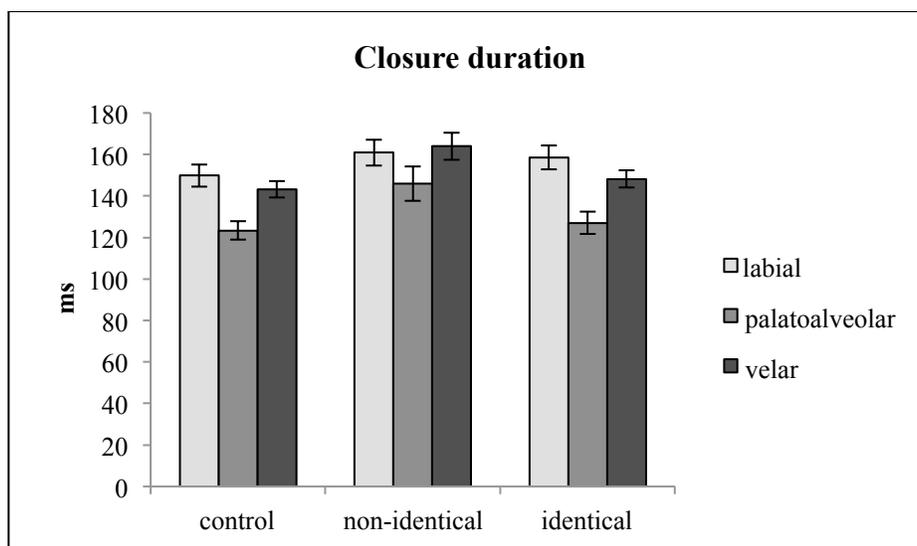
14

Figure 5: Closure of C_2 in correct productions of the non-identical, identical and control categories, by vowel. Error bars indicate Standard Error. Control labial n=63, palatoalveolar n=61, velar n=62; non-identical labial n=41, palatoalveolar n=30, velar n=45; identical labial n=51, palatoalveolar n=48, velar n=66.

A Linear Mixed Model was fit with C_2 closure duration as the dependent variable. There were two predictors of category: one predictor compared correct productions of the identical and non-identical categories as a group to the control category, and a second predictor compared correct productions of the identical and non-identical categories to each other. Additionally, the model included two predictors of place: one comparing the labials to the palatoalveolars and velars as a group and the other comparing palatoalveolars and velars. All predictors were centered, and all interactions between category and place were included. None of the predictors emerges as significant, indicating that closure duration is not significantly affected by either place of articulation or category.

Next, the VOT and burst amplitude of correctly produced ejectives were compared between the control, identical and non-identical categories. Average VOT values are shown in Figure 6, where it can be seen that VOT is longest in palatoalveolars, followed by velars and then labials. Additionally, VOT is shorter in non-identical and identical categories, where the ejectives are phonotactically illegal, than in the control category, where the ejectives are phonotactically legal.
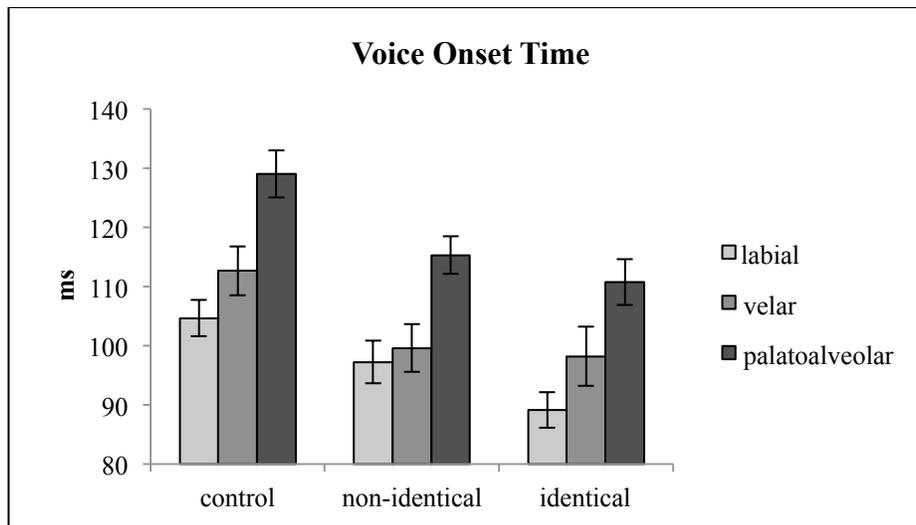
Voice Onset Time

Figure 6: Duration of VOT in correct productions of ejectives in the non-identical, identical and control categories, by place of articulation. Error bars indicate Standard Error. Control labial *n=79*, palatoalveolar *n=75*, velar *n=77*; non-identical labial *n=40*, palatoalveolar *n=29*, velar *n=46*; identical labial *n=50*, palatoalveolar *n=49*, velar *n=62*.

A Linear Mixed Model was fit with a dependent measure of VOT. There were two predictors of category, comparing the identical and non-identical categories as a group to the control, and comparing the identical and non-identical categories to one another, as well as two predictors of place, one comparing palatoalveolars to labials and velars as a group, and the other comparing labials and velars to one another. All predictors were centered, and the model included all interactions. Both predictors of place are significant, as well as the predictor comparing the control category to the identical and non-identical categories as a group. No other predictors are significant. The final model is given in Table 6.

	Estimate	Standard Error	*t*
intercept	105.53	3.33	31.67
control vs. identical & non-identical	-15.86	3.03	-5.23
palatoalveolar vs. velar & labial	-20.89	2.56	-8.17
velar vs. labial	-9.81	2.80	-3.50

Table 6: Results of Linear Mixed Model with dependent variable of VOT and predictors of stimulus category and place of articulation.

Average intensity difference in the burst of ejectives is shown in Figure 7. There are large differences among places of articulation, with palatoalveolars having the largest burst intensity difference, followed by velars and then labials. Intensity difference is also slightly higher in the control condition than in the identical and non-identical conditions. This is the same pattern as seen for VOT duration.
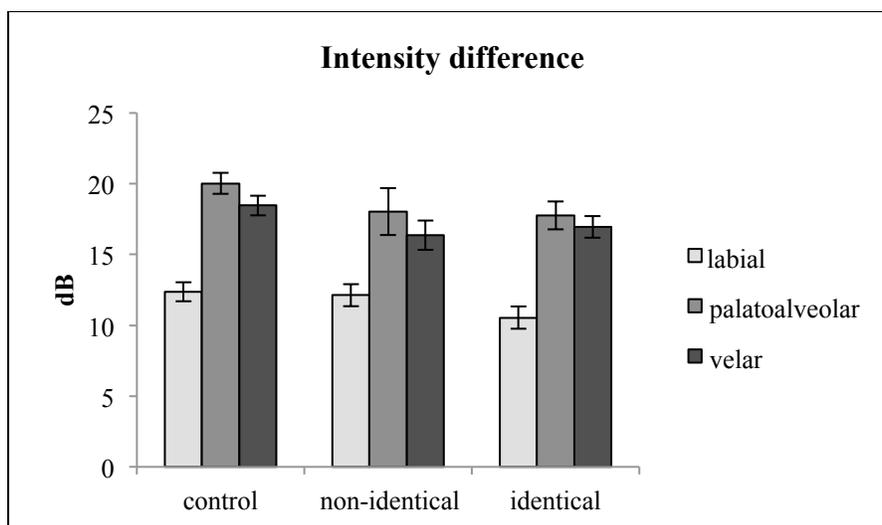
Figure 7: Intensity difference during burst and VOT period in correct productions of ejectives in the non-identical, identical and control categories, by place of articulation. Error bars indicate Standard Error. Control labial *n=79*, palatoalveolar *n=75*, velar *n=77*; non-identical labial *n=40*, palatoalveolar *n=29*, velar *n=46*; identical labial *n=50*, palatoalveolar *n=49*, velar *n=62*.

A Linear Mixed Model was fit with a dependent measure of intensity difference to assess the significance of differences based on place and category. The model had the same predictors as the VOT model. Both predictors of place are significant, as well as the predictor comparing the control category to the identical and non-identical categories as a group. Additionally, the difference between the control category and the identical and non-identical categories is significant. The final model is given in Table 7.

	Estimate	Standard Error	*t*
intercept	16.10	0.70	23.03
control vs. identical & non-identical	-2.55	0.70	-3.66
palatoalveolar vs. velar & labial	-3.78	0.83	-4.55
velar vs. labial	-6.13	0.93	-6.58

Table 7: Results of Linear Mixed Model with dependent variable of intensity difference and predictors of stimulus category and place of articulation.

Finally, no acoustic differences were found between correctly produced plain stops in the filler categories and target ejectives misproduced as plain stops. When participants produce a target with two ejectives with only a single ejective, e.g., [k'up'a] produced as [k'upa], the resulting plain stop has the same VOT and burst amplitude as an accurately produced plain stop in a filler item, e.g., [k'upa] produced as [k'upa].

3.3. *Discussion*

Two major asymmetries in accuracy on the repetition task are found. First, participants make many more errors on target items that contain a phonotactically illegal structure than on those

17

which are phonotactically legal. Target items with pairs of ejectives, whether identical or non-identical, are often repaired to contain only a single ejective, e.g., [k'ap'u] is often repeated as [k'apu]. Second, participants are significantly more accurate in repeating nonce words with pairs of identical ejectives than those with non-identical ejectives, e.g., [p'ap'u] is repeated accurately more often than [k'ap'u].

The higher error rate on pairs of ejectives relative to control (and filler) items reflects the phonotactic pattern of CQ, which contains no roots with multiple ejectives. This pattern replicates the finding of Gallagher (2013b), in which Quechua speakers were also asked to produce phonotactically illegal nonce roots in a repetition task. The asymmetry between target items with identical and non-identical ejective pairs, however, receives no support from the phonotactic patterns of attested roots in CQ, since both identical and non-identical ejective pairs are categorically absent. This finding suggests that the typological preference for identical ejectives seen in languages with the identity effect has its source in a synchronic asymmetry at the level of the individual speaker.

The asymmetry in accuracy between the control category and the identical and non-identical categories was mirrored in the acoustic properties of ejectives in these categories. Both VOT and intensity difference were significantly higher in the control category than in the phonotactically illegal categories, e.g., VOT is longer and intensity difference greater in [p'] in [hap'u] than in [k'ap'u] or [p'ap'u]. This pattern indicates reduction of the second of two ejectives, which may be driven by the phonotactic restriction. The second of two ejectives may be slightly reduced because this second ejective is phonotactically illegal, or it may generally be the case that repeated glottal gestures, as required in ejection, are subject to reduction. A comparison with a language with phonotactically legal pairs of ejectives is necessary to determine the source of this effect. Unlike the comparison between legal and illegal categories, the difference in accuracy between the identical and non-identical categories has no correlate in the acoustic analysis; the acoustic properties of an ejective preceded by an identical and a non-identical ejective are comparable. If the difference in accuracy between identical and non-identical ejective pairs has roots or reflexes in articulation, such articulatory differences do not have acoustic consequences.

In addition to these effects of category, place of articulation was also found to affect both accuracy and acoustic measures. Participants were overall least accurate on palatoalveolar affricates, though the palatoalveolar affricates have the longest VOT and highest intensity difference. One possible explanation for this effect is that while palatoalveolar ejective affricates have particularly long VOTs and high intensity bursts compared to other ejectives, the difference in VOT and intensity between an ejective and plain affricate is smaller, or simply less perceptible, than for stops.

The role of perception in accounting for the place of articulation effect, as well as errors on the repetition task in general, is assessed via an ABX discrimination task in Experiment 2. The repetition task involves both perception and production, as well as the mapping between two, and thus errors could stem from the phonotactic grammar, perceptual difficulties, production difficulties, or a combination of these factors. A perception study was run using the same stimuli

as in the repetition experiment in order to isolate perceptual difficulty associated with pairs of ejectives, and to determine whether the distinction between identical and non-identical ejective pairs found in the repetition experiment could have a source in perceptual asymmetries.

4. Experiment 2 – ABX discrimination

The discrimination study tested participants' ability to distinguish pairs of ejectives from forms with a single ejective, e.g., [k'ap'u] and [k'apu], the most common error found in the repetition experiment.

4.1. Methods

4.1.1. Participants

16 native speakers of Cochabamba Quechua participated in the experiment, which was conducted at the Quechua Indigenous University of Bolivia (Universidad Indígena Boliviana Quechua) in Chimore, a town in the Chapare region of the Cochabamba department of Bolivia. All participants were in their 20s and 30s. The participants for the discrimination task did not overlap with the participants for the repetition task, no individual completed both experiments. All the participants were bilingual in Quechua and Spanish, and Spanish was used as the metalanguage to conduct the experiment. Participants were paid the equivalent of about $7 dollars for their participation.

Data from three participants were removed from analysis, one for giving the same response on every trial, and two for performing at or below chance on the control trials.

4.1.2. Stimuli

The stimuli were the same $C_1V_1C_2V_2$ tokens used in the repetition experiment. For each ABX trial, two stimulus items that differed only in whether C_2 was ejective or plain were presented as stimuli A and B; the third stimulus, X, matched either A or B. There were three types of trials. In control trials, C_1 was a fricative or sonorant, so both a plain stop/affricate or an ejective in C_2 resulted in a phonotactically legal nonce word. In the identical and non-identical trials, phonotactically illegal nonce words with two ejectives, either identical or non-identical, were paired with a phonotactically legal nonce word with a single ejective in C_1. All stimulus pairs are shown in Table 8. Each pair corresponds to one target item and one filler item from the repetition experiment. As in the repetition task, three pairs, given in italics, are repeated twice to reach a total of 12 pairs in each category.

control	identical	non-identical
suk'i - suki	p'apʃ'u - p'apu	p'atʃ'i - p'atʃi
sitʃ'a - sitʃa	p'up'a - p'upa	*p'atʃ'i - p'atʃi*
sap'i - sapi	p'up'i - p'upi	p'uk'i - p'uki
huk'a - huka	p'ip'a - p'ipa	p'uk'a - p'uka
hitʃ'u - hitʃu	tʃ'atʃ'i - tʃ'atʃi	tʃ'ak'u - tʃ'aku
hap'u - hapu	tʃ'utʃ'a - tʃ'ut'a	tʃ'uk'a - tʃ'uka
ʎuk'a - ʎuka	tʃ'utʃ'i - tʃ'utʃi	tʃ'up'i - tʃ'upi
ʎatʃ'u - ʎatʃu	tʃ'itʃ'a - tʃ'itʃa	*tʃ'up'i - tʃ'upi*
ʎup'i - ʎupi	k'ak'u - k'aku	k'ap'u - k'apu
ɲak'u - ɲaku	*k'ak'u - k'aku*	k'up'i - k'upi
ɲutʃ'i - ɲutʃi	k'uk'a - k'uka	k'atʃ'u - k'atʃu
ɲup'a - ɲupa	k'uk'i - k'uki	k'itʃ'a - k'itʃa

Table 8: Stimulus items for the ABX discrimination experiment. Italics indicate repeated items.

Each stimulus pair was used in four trials, evenly balancing which item was A and which was B, and whether X matched A or B. For example, for the stimulus pair *huk'a-huka* the four trials were *huk'a-huka-huka, huk'a-huka-huk'a, huka-huk'a-huk'a* and *huka-huk'a-huka*. There were a total of 144 trials.

4.1.3. Equipment

The stimuli were presented aurally through the experimental software Psyscope X, running on a Macbook Air, via Audio Technica headphones. Participants indicated their response by pressing one of two keys on the keyboard; for all participants, the "1" key indicated a response of X=A and "0" indicated a response of X=B.

4.1.4. Procedure

The experiment was conducted in a quiet room at the university. The stimuli were presented in random order to each participant. On each trial, three stimuli were presented, with a 300 ms inter-stimulus interval. After the third stimulus, a green line appeared on the screen indicating the response period. Participants were instructed to press the "1" key if the third item they heard matched the first item, and the "0" key (marked with a sticker saying "2") if the third item matched the second item. Participants were told that the words they would hear had no meaning in Quechua or in any other language, though they contained sounds that were used in existing Quechua words. Additionally, participants were told that the first and second items they heard would always be different, and that the third item would match either the first or the second. The instructions were given in Spanish, both in written form on the computer screen, and orally by the experimenter. The instructions are given in the appendix.[5] Participants were allowed to

---

[5] Please note that the written instructions do not include the information that the first and second words were always different. This information was given orally, as it was deemed crucial to explaining the task after the first couple of participants.

practice with as many trials as they wanted (about 5 trials for most people) until they were comfortable with the task. Participants had 2 seconds to respond after each trial; if they did not respond within this timeframe, the experiment continued automatically to the next trial. The experiment took about 10 minutes to complete.

4.2. *Results and Analysis*

D-prime scores (MacMillan and Creeleman 2005; Boley & Lester 2009) were calculated for each subject for each trial type (control, identical and non-identical).[6] The results are shown in Figure 8. D-prime is much higher in the control category than in the identical and non-identical categories, indicating that participants can generally distinguish medial ejectives from medial plain stops in phonotactically legal nonce words (e.g., [huk'a] from [huka]) but that distinguishing a phonotactically illegal item with two ejectives from a form with just a single ejective (e.g., [k'up'a] from [k'upa]) is more difficult. There is no difference between the identical and non-identical categories; a form with two non-identical ejectives, e.g., [k'up'a], is just as confusable with a form with a single ejective, e.g., [k'upa], as a form with two identical ejectives is, e.g., [p'up'a] and [p'upa]. Accuracy quantified as percent correct shows the same pattern: participants are 80% accurate on control trials compared to 69% and 71% accurate on identical and non-identical trials, respectively.
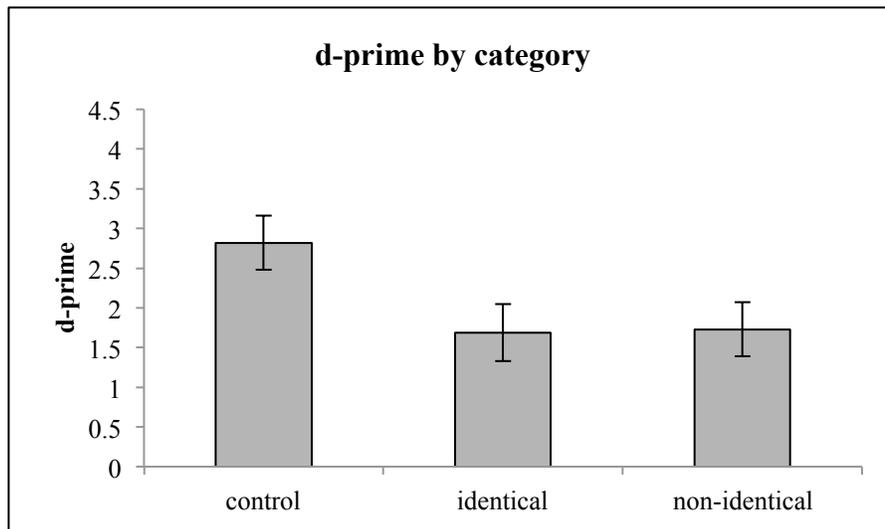

Figure 8: D-prime scores on control, identical and non-identical trials in the ABX experiment. Error bars indicate Standard Error.

In Figure 9, the results are broken down by place of articulation of the target ejective: labial, palatoalveolar or velar. The overall pattern holds for the velar place of articulation, but the

---

[6] D-prime is calculated by taking the z score of the proportion of hits and false alarms, and then subtracting false alarms from hits: d' = $z(H)$-$z(F)$. In the ABX task, a "hit" was defined as a correct "A" response and a "false alarm" as an incorrect "A" response. In other words, accuracy on ABA trials was relativized to errors on ABB trials. Proportions with the values of 0 and 1 were converted to 0.01 and 0.99, respectively, yielding a maximum d prime score of 4.5 (MacMillan & Creelman 2005).

palatoalveolar and labial places of articulation both show differences between the identical and non-identical categories that go in opposite directions.
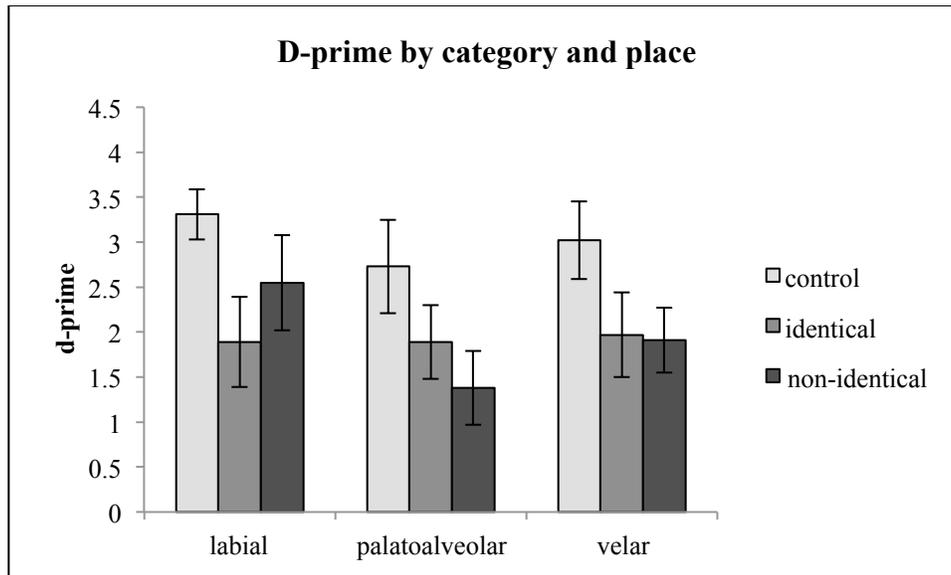

Figure 9: D-prime scores on control, identical and non-identical trials in the ABX experiment, by place of articulation. Error bars indicate Standard Error.

A Linear Mixed Model was fit to the d-prime scores with two predictors of category, one comparing the control category to the identical and non-identical categories as a group, and the other comparing the identical and non-identical categories to one another. There were also two predictors of place in the model, one comparing the velar place of articulation to the palatoalveolars and labials as a group, and the other comparing the palatoalveolars and the labials to one another. All interactions were included, and all predictors were centered. Random slopes and intercepts were included for participant. Significance of predictors was assessed via model comparison. In the final model, given in Table 9, the comparison between the control category and the identical and non-identical categories as a group is significant, as is the comparison between the palatoalveolar and the labial categories. D-prime is higher for the control category than the other two categories, and is higher for labials than for palatoalveolars. The interaction between palatoalveolar vs. labial and identical vs. non-identical is also significant: for labials, d-prime is higher on the non-identical category than on the identical category, while for palatoalveolars, the opposite pattern holds, d-prime is higher on the identical category than on the non-identical category. The comparison between identical and non-identical categories is not significant, as assessed by a *t* score that is within the ±2 range (Gelman & Hill 2006), but it is included in the model because it participates in a significant interaction.

	Estimate	Standard Error	_t_
intercept	2.26	0.27	8.5
control vs. identical & non-identical	-1.08	0.18	-5.89
identical vs. non-identical	0.03	0.14	0.27
palatoalveolar vs. labial	0.59	0.17	3.39
ident vs. non : pal vs. lab	1.12	0.52	2.15

Table 9: Results of Linear Mixed Model with dependent variable of d-prime and predictor of control vs. identical & non-identical categories.

4.3. Discussion

The contrast between an ejective and a plain stop or affricate is more perceptible for Quechua speakers in forms with an initial fricative or sonorant, where both ejectives and plain stops/affricates are phonotactically legal, than in forms with an initial ejective, where ejectives are phonotactically illegal. This difference mirrors the results of the repetition task, where it was found that forms with two ejectives were often repeated with only a single, initial ejective, and reflects the grammatical restriction against pairs of ejectives. This result suggests that at least some of the errors in the repetition task may be due to misperception: a form like [k'up'a], with two ejectives, may be repeated as [k'upa], with a single ejective, because it is misperceived with only a single, initial ejective.

The distinction between pairs of identical ejectives and pairs of non-identical ejectives, however, is found in the repetition task but not in the discrimination task. While forms with two non-identical ejectives are more likely to be misproduced with a single ejective than forms with two identical ejectives, forms with non-identical ejectives are not more likely to be misperceived than pairs of identical ejectives. There is thus no evidence that the asymmetry found between identical and non-identical ejective pairs in the repetition task is due to a difference in the perceptual properties of these two types of forms.

Misperception in the identical and non-identical categories was highest for the palatoalveolars, reflecting the differences in accuracy on the repetition task, where palatoalveolars were also the least accurate. The differences in overall accuracy among the three places of articulation, then, is likely due to differences in the perceptibility of the ejective-plain contrast.

5. General discussion

The results of the repetition task show two asymmetries. First, phonotactically illegal nonce roots with pairs of ejectives are often repeated with only a single ejective, while controls are repeated accurately (as in Gallagher 2013b). This asymmetry between phonotactically legal and illegal forms is also found in the perception task, where forms with two ejectives are confusable with forms with a single ejective. These results suggest that the grammatical restriction is internalized by CQ speakers, and pervades production and perception, as found in much previous work (Werker & Tees 1984; Hallé et al. 1998; Pitt 1998; Dupoux et al. 1999; Hallé & Best 2007). The

consistent effects of phonotactic legality across tasks is unsurprising given the categorical nature of the restriction against forms with multiple ejectives, as well as previous findings that pairs of ejectives are both generally articulatorily and perceptually challenging (Ohala 1981, 1993; Gallagher 2010a,b; Anonymous under review).

The second, more interesting asymmetry, is between phonotactically ill-formed structures. CQ speakers repeat nonce roots with identical ejectives more accurately than nonce roots with non-identical ejective pairs. This result raises two questions that are addressed in the remainder of this section. First, given that both identical and non-identical ejective pairs are categorically unattested in CQ, where does this preference for identical ejectives come from? And second, how does the preference for identical ejectives in CQ relate to the typological preference for identical ejectives, seen in the many languages where only non-identical pairs are prohibited?

5.1.   *Sources of the identity preference*

The identity preference in CQ speakers found here is consistent with previous work that shows a correlation between typology and synchronic asymmetries in individual speakers, particularly work examining sonority profiles in onsets (Greenberg & Jenkins 1964; Scholes 1966; Broselow & Finer 1991; Coleman & Pierrehumbert 1997; Moreton 2002; Davidson 2006, 2010; Berent et al. 2007; Berent & Lennertz 2010; Daland et al. 2011). Proposals to explain asymmetries among unattested onsets, and unattested structures more generally, have argued that these asymmetries arise (i) from universal grammatical principles, such as the sonority hierarchy (Selkirk 1984; Clements 1990), (ii) from the formal properties of phonological learning that allow abstraction from the lexicon, and/or (iii) from detailed knowledge of perceptual and articulatory structures. Several researchers have established asymmetries among unattested onsets that favor onsets with a greater sonority rise and argued that the grammatical capacity of all speakers includes the sonority hierarchy and a markedness scale favoring greater sonority rises in onsets (Broselow & Finer 1991; Moreton 2002; Berent et al. 2007; Berent & Lennertz 2010). Others have proposed that sonority preferences arise not from a universal markedness scale, but rather by generalizing over features or natural classes in attested onsets (Albright 2009; Daland et al. 2011). In these models, a novel stimulus is more or less acceptable depending on how similar it is to existing clusters. Asymmetries in unattested onsets may also arise from phonetic distinctions between different types of clusters, as well as language specific information about what phonetic cues are informative (Davidson 2010; Davidson & Shaw 2012). Arguments for these latter two lines of explanation for the identity preference are evaluated here. First, it will be shown in §5.1.1 that a preference for identical ejectives could in principle be learned from the lexicon of CQ, and that greater accuracy on nonce roots with identical ejectives could stem from the phonotactic grammer. Second, identical pairs of ejectives may be preferred to non-identical ejectives from an articulatory planning perspective, and the greater accuracy in repeating nonce roots with identical ejectives may have its roots in production.

5.1.1 Evidence for an identity preference from attested CQ roots

One source of evidence for an identity preference in CQ comes from the relative size of the classes of identical and non-identical ejective pairs. CQ has five ejective consonants, [p' t' tʃ' k' q'], yeilding 20 potential non-identical pairs of ejectives and only 5 identical pairs. Looking in more detail at the lexical statistics of CQ, based on counts of the 2290 roots in the Laime Ajacopa et al. (2007) dictionary, non-identical ejective pairs are expected to occur 35.8 times by chance, and identical ejective pairs 8.91 times. By both counts, non-identical ejective pairs are predicted to be about 4 times as frequent as identical ejective pairs. Learning models typically include a measure of confidence or reliability in statistical generalizations, putting greater weight in generalizations that are based on more data (Tenenbaum & Griffiths 2001; Albright & Hayes 2003). From this perspective, a constraint against non-identical ejective pairs has more evidence supporting it than a constraint against identical pairs; the greater number of expected non-identical pairs makes it less likely that their total absence is accidental. This asymmetry in the evidence predicts that speakers of CQ should learn a stronger grammatical restriction against non-identical ejective pairs than against identical pairs. This argumentation holds for CQ in particular, but should also apply to many other languages. In any language with more than two ejectives (or segments in any other natural class), there are more possible pairs of non-identical ejectives than identical ejectives, and thus there is more evidence for the grammatical status of non-identical ejectives than identical ejectives.

For the difference in evidence for restrictions on identical and non-identical ejectives to be relevant to CQ speakers, it must first be assumed that CQ speakers assess the evidence for restrictions on identical and non-identical ejectives separately. Instead of a single phonotactic generalization that rules out all pairs of ejectives, *[+cg][+cg] (where [+cg] stands for [+constricted glottis], the feature that distinguishes ejectives from plain stops), CQ speakers must postulate two generalizations: one against identical ejectives *$[+cg]_j[+cg]_j$ and one against non-identical ejectives *$[+cg]_j[+cg]_k$.

An assumed universal distinction in the representation of identical and non-identical structures has precedent in the phonological literature, formalized as double linking within Autosegmental Phonology (Leben 1973; Goldsmith 1976; McCarthy 1986, 1988) and as distinct OCP constraints in Coetzee & Pater (2008). Empirical support for the idea that all learners may consider identical and non-identical segments as different classes comes from Endress et al. (2007) and Gervain et al. (2008), discussed in Section 2, who find that identity based patterns are learned particularly easily.

A second potential source for the identity preference in CQ comes from the frequency of identical pairs of non-ejective consonants in the language. The phonotactics of CQ disallow all pairs of identical ejectives and identical aspirates, but identical pairs of all other consonants are well-attested, or highly over-attested. For those consonants that are not phonotactically precluded from cooccurring, the observed-over-expected (O/E) ratios (Pierrehumbert 1993; Frisch et al. 2004) of all but one pair of identical consonants in CQ is greater than 1 (ranging from 0.84 for [p-p] to 8.04 for [l-l]), indicating that identical consonants are more likely to occur than expected

by chance.[7] The overattestation of identical pairs of consonants could lead to CQ speakers representing $C_iV(C)C_iV$ as a kind of template or positive constraint over root types. While negative prohibitions on identical ejectives and aspirates must be stronger than this positive generalization, the frequency of phonotactically legal roots with identical consonants may still result in a small preference for nonce words of this form even when they are phonotactically illegal.

While the lexical statistics of CQ do provide evidence that the preference for identical ejectives could be learned, these data aren't informative about the specific representation of this knowledge. Models of phonotactic learning typically assume that generalizations are stated over abstract phonological features (Hayes & Wilson 2008; Albright 2009), but it is also probable that learners generalize over phonetically detailed representations, like articulatory plans (Bybee 2001). The proposal in this section is therefore not about the precise representational content of the identity preference – specifically whether it is phonetically contentful or not – but simply that this preference can be learned by CQ speakers from their experience with attested roots in their language.

5.1.2   An identity preference in speech planning

In addition to, or instead of, a learned preference for identical ejectives based on experience with CQ, the identity preference seen on the repetition task may have its sources in articulatory planning. Under this view, greater accuracy on the repetition task does not result from any high-level, stored representation of a preference for identical ejectives, but instead arises only in the construction of a speech plan for an intended utterance.

Research on speech errors has found that similar segments are particularly likely to be subject to error, and that segments that are similar but not identical are likely to become identical (MacKay 1970; Fromkin 1971; Shattuck-Hufnagel & Klatt 1979; Stemberger 1982; Frisch 1996; Vousden et al. 2000; Walker 2007). The prevalence of errors involving similar segments or structures is accounted for in spreading activation models of language production and processing (Dell & Reich 1980; McClelland & Rumelhart 1981; Dell 1984, 1986; Stemberger 1985; MacKay 1987; Frisch 2004), in which featurally or gesturally similar sounds cause activation of many of the same processing nodes. In such a model, a form with identical ejectives like [p'up'a] would be easier to process than a form like [k'up'a], because in the former each incidence of [p'] activates the same set of abstract nodes, reinforcing the speech plan, and leading to accurate production of both ejectives. Furthermore, in forms like [k'up'a] or [p'upa], which contain consonants sharing some but not all of their features, segments are likely to interfere with one another, leading to errors. The role of speech planning in the repetition task must be in preferentially blocking errors on targets with identical ejectives, which are driven by a general

---

[7] The O/E ratio is calculated by multiplying the probability of C_1 as the word-initial consonant by the probability of C_2 as the medial syllable onset by the total number of roots.

26

phonotactic restriction on pairs of ejectives, and driving errors on filler items with homorganic segments, like [p'upa].

A more specific proposal for an articulatory preference for identity comes from studies of gestural phasing (Pouplier & Goldstein 2005; Goldstein et al. 2007; Pouplier 2008). When participants are asked to repeat phrases like *top cop*, which involve alternating alveolar and velar onsets in otherwise matching syllables, articulatory data show intrusion errors where a velar gesture is produced during the alveolar gesture, changing the alternating phasing of the alveolar and velar gestures to be simultaneous. Pouplier (2008) finds that intrusion errors resulting in synchonous gestures are more likely the more material is shared between the interacting structures; a shared coda consonant in phrases like *top cop* increases error rate compared to phrases with only a shared vowel, like *taa kaa*. Applying these results to the data here, identical pairs of ejectives may be prefered to non-identical ejective pairs or pairs of an ejective and a plain homorganic stop because of the synchronous relationship between the place and glottal onset gestures in [p'up'a], as compared to the alternating place or glottal gestures required in [p'upa] or [p'uk'a].

While in general terms the results of the repetition task are consistent with effects seen in speech error research, the specifics are less clear. The most major problem is that the vast majority of speech error research, and particularly articulatory and acoustic studies of elicited speech errors in experimental tasks, are done on English. There is no guarantee that any of the effects found on these tasks generalize to other languages, particularly a language like CQ which differs substantially from English in both phonemic inventory and word structure. Gestural coupling for tongue body and tongue tip gestures in English may not generalize straightfowardly to coupling between the multiple gestures required in an ejective, namely closely and raising of the glottis with an oral constriction of the tongue body, tongue tip or lips. To pursue an explanation for the current results in speech planning, work needs to be done that explicitly addresses speech planning considerations for ejectives and plain stops in CQ and establishes a production preference for repeated, simultaneous gestures. A task could be designed that more closely resembles the conditions that triggered errors in *top cop* phrases in Pouplier & Goldstein (2005), Goldstein et al. (2007) and Pouplier (2008). In those experiments, participants were asked to repeat the same phrase with alternating onsets over and over. If gestural coupling works similarly in CQ and English, than if CQ speakers are asked to repeat a nonce root with alternating onsets, e.g., *p'upa p'upa p'upa*, the expectation is that errors with repeated ejectives would occur, e.g., *p'upa p'upa p'up'a*. Parallel to Pouplier (2008), more errors would be expected if the onsets were followed by identical vowels, or identical vowel+coda strings, as in *p'apa* or *p'aspas*.

5.1.3 Assessing the role of the grammar and production

The major difference between the two proposals above is whether the identity preference results from stored grammatical knowledge, learned from making generalizations over experience with attested forms, or whether the identity preference results from articulatory preferences that are

active in production only. If a preference for identical ejective pairs is a part of the learned phonological competence of CQ speakers, there should be evidence for this preference in various tasks that involve manipulation of phonological generalizations, not only tasks that involve production.

The perception study provides one piece of evidence that the identity preference is particularly linked to production. The perception study involved only listening, and revealed no differences between identical and non-identical pairs of ejectives. The null result on the perception study could be a result of the identity preference being only a production effect, or it could simply be that the preference for identity isn't reflected in an ABX task. While some grammatical distinctions have been shown to influence perception (Werker & Tees 1984; Hallé et al. 1998; Pitt 1998; Dupoux et al. 1999; Moreton 2002; Hallé & Best 2007), other work has shown that perception is not equally influenced by all grammatical distinctions (Best, McRoberts & Sithole 1988; Kabak & Idsardi 2007). A more varied range of tasks is needed to determine whether the identity preference is active only in production.

In addition to studies of speech planning in CQ, as sketched above, which would explicitly establish the planning considerations relevant to pairs of ejectives, further studies that involve phonological encoding and processing, but not production, could help assess the role of the grammar in accounting for the identity preference. Possible tasks include those that involve memory and storage of novel forms, but no production of these forms. For example, if CQ speakers were taught nonce words as names for novel objects, grammatical wellformedness should influence their learning of these new words, as assessed by non-production tasks like matching names to objects or judging whether a name-object pair is correct. If the identity preference is in the grammar, the expectation would be that nonce words with identical ejectives would be learned faster and better than nonce words with non-identical ejectives.

With the data currently available, the source of the identity preference in phonological knowledge, speech production, or both, cannot be conclusively determined. More study is needed to determine both the pervasiveness of the identity preference in a variety of tasks, and to explore the dynamics of speech planning in CQ.

5.2. *Connections to typology*

The identity preference in CQ speakers is of interest because it is the same preference seen in a categorical form in the phonotactics of many languages with the identity effect. This correlation between the typology and the behavior of CQ speakers suggests that the cross-linguistic preference for identical ejectives is driven by a bias in the synchronic knowledge of individual speakers. The details of the connection between typology and the effects here differ depending on whether the identity preference in CQ speakers is a learned grammatical preference or is isolated in production.

If the identity preference is learned from the lexical data of CQ, then the identity effect may be common cross-linguistically because it is easier to learn than the opposite, anti-identity effect pattern, where only pairs of non-identical ejectives are permitted. For any language with more

than two ejectives, the number of possible identical pairs is smaller than the number of possible non-identical pairs. In the absence of any roots with pairs of ejectives, like CQ, this difference in the size of the classes leads to a greater restriction on non-identical ejective pairs: the class of identical ejectives is smaller, so the chances that the absence of identical ejectives is accidental is higher than for non-identical ejectives. On the other hand, if we consider a language where some pairs of ejectives are attested, the size of the class of identical ejectives is an advantage for learning that only identical pairs of ejectives are attested. The chances that a random sample of n roots contains only identical pairs of ejectives is smaller than the chances that it contains only non-identical pairs. The learner then requires fewer data points to learn that only identical pairs of ejectives are attested than that non-identical pairs of ejectives are attested. This asymmetry in learning identity-based generalizations was found in an artificial grammar paradigm in Linzen & Gallagher (2013). When 50% of the training tokens participants saw contained identical consonants, they showed evidence of having extracted a generalization and rated novel pairs of identical consonants as acceptable. When participants saw no identical consonants pairs, however, they didn't learn a comparable dispreference for identical consonant pairs. The difference in the size of identical and non-identical classes, then, could underlie the typological preference for the identity preference by making the identity effect pattern learnable with relatively little data.

If identical ejectives are favored by the production system, the greater difficulty associated with producing non-identical ejectives should support learning a generalization against this relatively difficult structure (Hayes 1999; Hayes, Kirchner & Steriade 2004). Learners may be better able to discover a grammatical restriction if it is supported both by the statistical properties of their language, and is phonetically grounded.

6. Conclusion

This paper has presented evidence for a bias in favor of identical ejective pairs over non-identical ejectives in speakers of Cochabamba Quechua, a language in which both identical and non-identical pairs of ejectives are categorically unattested. This latent preference for identical ejectives is consistent with the typological preference for identical ejectives, commonly seen in languages that exhibit the identity effect.

The repetition and discrimination studies presented here do not conclusively determine the source of the asymmetry between identical and non-identical ejectives in CQ. The preference for identical ejectives may have its source in the grammar, learned from the statistics of attested CQ roots, or it may have its source in production, where repeated, coupled gestures are prefered. Further exploration of the pervasiveness of the identity effect on a variety of tasks is needed, along with explicit investigation of speech planning effects in CQ, to achieve a greater understanding of the latent identity preference found here.

Appendix

The instructions for the repetition experiment were written in Spanish as follows:

> Escuchará una palabra, pronuncida dos veces, y debe repetir lo que escucha lo más exactamente que puede. Las palabras serán secuencias de sonidos sin sentido - no son palabras reales de ningún idioma. Despues de repetir la palabra, presione "espacio" para continuar. Antes de empezar el experimento, hay algunas palabras para practicar.

Translated into English:

> You will hear a word, pronounced twice, and you must repeat what you hear as exactly as you can. The words are sequences of sounds without meaning – they aren't real words in any language. After repeating the word, press "space" to continue. Before beginning the experiment, there are a few words for practice.

The instructions for the discrimination experiment were written in Spanish as follows:

> Escuchará tres palabras y debe decidir si la tercera y la primera son iguales o si la tercera y la segunda son iguales. Las palabras serán secuencias de sonidos sin sentido - no son palabras reales de ningún idioma. Despues de escuchar las tres palabras, aparecerá una línea verde. Cuando aparezca la línea, presione "1" si piensa que la tercera palabra y la primera palabra son iguales y presione "2" si piensa que la tercera palabra y la segunda palabra son iguales. Antes de empezar el experimento, hay algunas palabras para practicar.

Translated into English:

> You will hear three words, and must decide if the third and the first are the same or if the third and the second are the same. The words are sequences of sounds without meaning – they aren't real words in any language. After listening to the three words, a green line will appear. When this line appears, press "1" if you think the third word and the second word are the same and press "2" if you think the third word and the second word are the same. Before beginning the experiment, there are a few words for practice.

References

Laime Ajacopa, Teofilo, et al. 2007. *Diccionario Bilingüe, Iskay Simipi Yuyayk'ancha: Quechua – Castellano Castellano –Quechua*. La Paz, Bolivia.

Albright, Adam and Bruce Hayes. 2003. Rules vs. analogy in English past tenses: a computational/experimental study. *Cognition* 90:119-161.

Albright, Adam. 2009. Feature-based generalization as a source of gradient acceptability. *Phonology* 26:9-41.

Anonymous. under review. Details omitted due to double blind reviewing.

Bates, Doug and Martin Maechler. 2010. lme4: Linear mixed effects models using S4 classes. http://CRAN.Rproject.org/package=lme4.

Bennett, William. 2013. Dissimilation, consonant harmony, and surface correspondence. Doctoral Dissertation, Rutgers University.

Berent, Iris & Joseph Shimron. 1997. The representation of Hebrew words: evidence from the obligatory contour principle. *Cognition 64*:39-72.

Berent, Iris & Tracy Lennertz. 2010. Universal constraints on the sound structure of language: Phonological or acoustic? *Journal of Experimental Psychology: Human Perception and Performance* 36(1):212-223.

Berent, Iris, Donca Steriade, Tracy Lennertz and Vered Vaknin. 2007. What we know about what we have never heard: Evidence from perceptual illusions. *Cognition* 104:591-630.

Berent, Iris, Colin Wilson, Gary Marcus and David Bemis. 2012. On the role of variables in Phonology: Remarks on Hayes and Wilson 2008. *Linguistic Inquiry* 43:97-119.

Best, C., T. McRoberts & N. Sithole. 1988. The phonological basis of perceptual loss for non-native contrasts: Maintenance of discrimination among Zulu clicks by English speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14:345-360.

Blevins, Juliette. 2004. *Evolutionary Phonology*. Cambridge: Cambridge University Press.

Boersma, Paul and David Weenink. 2011. Praat: doing phonetics by computer. Version 5.1.30, available at www.praat.org.

Boley, Jon & Michael Lester. 2009. Statistical analysis of ABX results using signal detection theory. *Audio Engineering Society Convention Papers*.

Broselow, E., Chen, S., & Wang, C. 1998. The emergence of the unmarked in second language phonology. Studies in Second Language Acquisition, 20, 261–280.

Broselow, E., & Finer, D. 1991. Parameter setting in second language phonology and syntax. Second Language Research, 7(1), 35–59.

Bybee, Joan. 2001. Phonology and Language Use. Cambridge: Cambridge University Press.

Chomsky, Noam & Morris Halle. 1968. *The Sound Pattern of English*. Cambridge: MIT Press.

Clements, G. N. 1990. The role of the sonority cycle in core syllabification. In M. Beckman (ed.), *Papers in laboratory phonology I: Between the grammar and physics of speech*, 282-333. Cambridge: Cambridge University Press.

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.

Coetzee, Andries W. 2009. Grammar is both categorical and gradient. In Steven Parker, ed. *Phonological Argumentation*. London: Equinox.

Colavin, Rebecca, Roger Levy and Sharon Rose. 2010. Modeling OCP-Place in Amharic with the Maximum Entropy phonotactic learner. *Papers from the 46th Chicago Linguistic Society*.

Coleman, John and Janet Pierrehumbert. 1997. Stochastic phonological grammars and acceptability. *Proceedings of the third meeting of the Association for Computational Linguistics special interest group in computational phonology*. Somerset, NJ: Association for Computational Linguistics.

Daland, Robert, Bruce Hayes, James White, Marc Garellek, Andrea Davis & Ingrid Norrman. 2011. Explaining sonority projection effects. *Phonology* 28:197-234.

Davidson, Lisa & Jason Shaw. 2012. Sources of illusion in consonant cluster perception. *Journal of Phonetics* 40:234-248.

Davidson, Lisa. 2006. Phonology, phonetics of frequency: Influences on the production of non-native sequences. *Journal of Phonetics* 34(1):104-137.

Davidson, Lisa. 2010. Phonetic bases of similarities in cross-language production: Evidence from English and Catalan. *Journal of Phonetics* 38:272-288.

de Lacy, Paul. 2002. The formal expression of markedness. Doctoral Dissertation, University of Massachusetts, Amherst.

de Lucca, Manuel. 1987. *Diccionario practico aymara-castellano castellano-aymara*. Cochabamba:Los Amigos del Libro.

Dell, Gary & Peter Reich. 1981. Stages in sentence production: An analysis of speech error data. *Journal of Verbal and Learning Behavior* 20:611-629.

Dell, Gary. 1984. Representation of the serial order of speech: Evidence from the repeated phoneme effect in speech errors. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 10: 222-233.

Dell, Gary. 1986. A spreading-activation theory of retrieval in sentence production. *Psychological Review* 92:283-321.

Dupoux, Emmanuel, Yuki Hirose, Kazuhiko Kakehi, Christophe Pallier and Jacque Mehler. 1999. Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25:1568-1578.

Endress, Ansgar, Ghislaine Dehaene-Lambertz and Jacque Mehler. 2007. Perceptual constraints and the learnability of simple grammars. *Cognition* 105:577-614.

Flemming, Edward, Peter Ladefoged & Sarah Thomason. 2008. The phonetic structures of Montana Salish. *Journal of Phonetics* 36:465-491.

Frisch, Stefan and Bushra Zawaydeh. 2001. The psychological reality of OCP-place in Arabic. *Language 77*:91-106.

Frisch, Stefan & Richard Wright. 2002. The phonetics of phonological speech errors: An acoustic analysis of slips of the tongue. *Journal of Phonetics* 30:139-162.

Frisch, Stefan, Janet Pierrehumbert, and Michael Broe. 2004. Similarity avoidance and the OCP. *Natural Language and Linguistic Theory* 22: 179-228.

Frisch, Stefan. 1996. Similarity and frequency in phonology. Doctoral Dissertation, Northwestern University.

Frisch, Stefan. 2004. Language processing and segmental OCP effects. In B. Hayes, R. Kirchner and D. Steriade (eds.), *Phonetically based phonology*: 346-371. Cambridge: Cambridge University Press.

Fromkin, Victoria. 1971. The non-anamalous nature of anomalous utterances. *Language* 47: 27-52.

Gafos, Diamandis. 1998. Eliminating long-distance consonantal spreading. *Natural Language and Linguistic Theory* 16:223-278.

Gallagher, Gillian and Jessica Coon. 2009. Distinguishing total and partial identity: evidence from Chol. *Natural Language and Linguistic Theory.*

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. 2013a. Learning the identity effect as an artificial language: bias and generalization. *Phonology* 30:1-43.

Gallagher, Gillian. 2013b. Speaker awareness of non-local laryngeal phonotactics in Cochabamba Quechua. *Natural Language and Linguistic Theory* 31:1067-1099.

Garellek, Marc & Patricia Keating. 2011. The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. *Journal of the International Phonetic Association* 41:185-205.

Gelman, Andrew & Jennifer Hill. 2006. Data analysis using regression and multilevel/hierarchical models. Cambridge: Cambridge University Press.

Gervain, Judit, Francesco Macagno, Silvia Cogoi, Marcela Peña and Jacques Mehler. 2008. The neonate brain detecets speech structure. *Proceedings of the National Academy of Sciences* 105(37), 14222.

Goldrick, Matt & Sheila Blumstein. 2006. Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters.

Goldsmith, John. 1976. *Autosegmental Phonology*. Doctoral Dissertation, Massachusetts Institute of Technology.

Goldstein, Louis, Marianne Pouplier, Larissa Chen, Elliot Saltzman and Dani Byrd. 2007. Dynamic action units slip in speech production errors. *Cognition* 103:386-412.

Gordon, Matthew and Alya Applebaum. 2006. Phonetics structures of Turkish Kabardian. *Journal of the International Phonetic Association* 36:159-186.

Gordon, Matthew. 1996. The phonetic structures of Hupa. *UCLA Working Papers in Phonetics* 93:164-187.

Greenberg, Joseph & James Jenkins. 1964. Studies in the psychological correlates of the sound system of American English. *Word* 20:157-177.

Greenberg, Joseph. 1978. Some generalizations concerning initial and final consonant clusters. In E. A. Moravcsik (ed.), *Universals of human language,* Vol 2: 243-279. Stanford, CA: Stanford University Press.

Hallé, Pierre & Catherine Best. 2007. Dental-to-velar perceptual assimilation: A cross-linguistic study of the perception of dental stop+/l/ clusters. *Journal of the Acoustical Society of America* 121:2899-2914.

Hallé, Pierre, Juan Segui, Uli Frauenfelder & Christine Meunier. 1998. Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance* 24(2):592-608.

Ham, SooYoun. 2004. *Tsilhqut'in ejectives: a descriptive phonetic study*. MA. Thesis, University of Victoria.

Hayes, Bruce & Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39:379-440.

Hayes, Bruce. 1999. Phonetically-driven phonology: the role of Optimality Theory and inductive grounding. In M. Darnell, E. Moravscik, M. Noonan, F. Newmeye and K. Wheatly, eds., *Functionalism and Formalism in Linguistics Volume I: General Papers*. John Benjamins: Amsterdam, 243-285.

Hayes, Bruce, Robert Kirchner & Donca Steriade. 2004. *Phonetically based phonology*. Cambridge: Cambridge University Press.

Jakobson, Roman. 1941. *Kindersprache, Aphasie und allgemeine Lautgesetze [Child language, aphasia, and phonological universals]*. Frankfurt: Suhrkamp.

Jensen, John Thayer. 1977. *Yapese Reference Grammar*. Honolulu: University Press of Hawaii.

Kabak, Baris & William Idsardi. 2007. Speech perception is not isomorphic to phonology: The case of perceptual epenthesis. *Language and Speech* 50:23-52.

Kingston, John. 1985. The phonetics and phonology of the timing of oral and glottal events. Doctoral Dissertation, University of California, Berkeley.

Ladefoged, Peter. 2003. *Phonetic Data Analysis*. Oxford: Blackwell Publishing.

Leben, W. 1973. Suprasegmental Phonology. Doctoral Dissertation, Massacusetts Institute of Technology.

Lindau, Mona. 1984. Phonetic differences in glottalic consonants. *Journal of Phonetics* 12: 147-155.

MacEachern, Margaret. 1999. *Laryngeal co-occurrence restrictions*. New York: Garland.

MacKay, Donald. 1970. Spoonerisms: The structure of errors in the serial order of speech. *Neuropsychologia* 8:323-350.

MacKay, Donald. 1987. *The organization of perception and action: A theory for language and other cognitive skills*. New York: Springer.

Macmillan, Neil and Douglas Creelman. 2004. *Detection Theory: A user's guide*. 2nd ed. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

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.

McClelland, J. & D. Rumelhart. 1981. An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review* 88: 375-407.

McDonough, Joyce and Peter Ladefoged. 1993. Navajo Stops. *UCLA Working Papers in Phonetics* 87:151-164.

Mester, R. Armin. 1986. Studies in tier structure. Doctoral disseration, University of Massachusetts, Amherst.

Moreton, Elliott & Joe Pater. 2012a. Structure and substance in artificial-phonology learning. Part 1: Structure, Part II: Substance. *Language and Linguistics Compass* 6:686-718.

Moreton, Elliott & Joe Pater. 2012b. Structurally biased phonology: Complexity in learning and typology. In K.G. Vijayakrishnan, ed., *Journal of the English and Foreign Languages University Special Issue*, Hyderabad, 1-44.

Moreton, Elliott. 2002. Structural constraints in the perception of English stop-sonorant clusters. *Cognition* 84:55-71.

Moreton, Elliott. 2008. Analytic bias and phonological typology. *Phonology* 25:83-127.

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. 1997. An OT account of laryngealization in Cuzco Quechua. *Working Papers of the Summer Institute of Linguistics, University of North Dakota Session* 41:1-11.

Pierrehumbert, Janet. 1993. Dissimilarity in Arabic verbal roots. In A. Schafer (ed.), *NELS 23:Proceedings of the North East Linguistic Society*, 367-381. GLSA: Amherst.

Pitt, Mark. 1998. Phonological processes and the perception of phonotactically illegal consonant clusters. *Perception & Psychophysics* 60:941-951.

Pouplier, Marianne & Louis Goldstein. 2005. Asymmetries in the perception of speech production errors. *Journal of Phonetics* 33:47-75.

Pouplier, Marianne. 2008. The role of a coda consonant as error trigger in repetition tasks. *Journal of Phonetics* 2008: 114-140.

Rose, Sharon and Lisa King. 2007. Speech error elicitation and cooccurrence restrictions in two Ethiopian Semitic languages. *Language and Speech 50*:451-504.

Scholes, Robert. 1966. *Phonotactic Grammaticality*. The Hague: Mouton.

Selkirk, Elisabeth. 1984. On the major class features and syllable theory. In M. Aronoff & R. Oehrle (eds.), *Language sound structure*. Cambridge: MA: MIT Press.

Shattuck-Hufnagel, Stephanie & Dennis Klatt. 1979. The limited use of distinctive features and markedness in speech production: Evidence from speech error data. *Journal of Verbal Learning and Verbal Behavior* 18:41-55.

Shosted, Ryan & Sharon Rose. 2011. Affricating ejective fricatives: The case of Tigrinya. *Journal of the International Phonetic Association* 41:41-65.

Stemberger, J. 1982. The nature of segments in the lexicon. Evidence from speech errors. *Lingua* 56:235-259.

Stemberger, J. 1985. An interactive activation model of language production. In A. W. Ellis (ed.), *Progress in the psychology of language,* Vol. 1, 143-186. London: Lawrence Erlbaum.

Steriade, Donca. 1997. Phonetics in phonology: the case of laryngeal neutralization. Ms. UCLA.

Straight, H. Stephen. 1976. *The acquisition of Maya phonology: Variation in Yucatec child language*. New York: Garland.

Suzuki, Keiichiro. 1998. A typological investigation of dissimilation. Doctoral dissertation, University of Arizona, Tucson.

Tenenbaum, Joshua & Thomas Griffiths. 2001. Generalization, similarity and Bayesian inference. *Behavioral and Brain Sciences* 24:629-640.

Thomas, Jacqueline M.C. 1963. *Le parler Ngbaka de Bokanga: Phonologie, morphologie, syntaxe*. Paris: Mouton.

Trubetzkoy, N. 1958. *Grundzuge der Phonologie [Foundation of phonology]*. Gottingen: Vandenhoeck & Ruprecht.

van den Berg, Rene & Ode Sidu. 1996. *Muna-English Dictionary*. Leiden:KITLV Press.

Vicenik, Chad. 2010. An acoustic study of Georgian stop consonants. *Journal of the International Phonetic Association* 40:59-92.

Vousden, Janet, Gordon Brown & Trevor Harley. 2000. Serial control of phonology in speech production: A hierarchical model. *Cognitive Psychology* 41:101-175.

Walker, Rachel. 2007. Nasal and oral consonant similarity in speech errors: Exploring parallels with nasal consonant harmony. *Language and Cognitive Processes* 22(7): 1073-1113.

Werker, Janet and Richard Tees. 1984. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7:49-63.

Wright, Richard, Sharon Hargus & Katherine Davis. 2002. On the categorization of ejectives: Data from Witsuwit'en. *Journal of the International Phonetic Association* 32:43-77.