On the categorization of ejectives: data from Witsuwit'en

Richard Wright
University of Washington, Seattle
rawright@u.washington.edu

Sharon Hargus
University of Washington, Seattle
sharon@u.washington.edu

Katharine Davis
Conversational Computing Corporation
kdavis@conversa.com

Results of an acoustic study of root-initial /t t^h t'/ for eleven speakers of Witsuwit'en (Athabaskan) are presented. Values averaged across speaker means suggest that Witsuwit'en ejectives have a brief period of creaky voice and slow rise time at vowel onset relative to the voiceless unaspirated stops. However, there is considerable inter-speaker variation in VOT and f0 perturbation, with negligible correlation between these measures, contrary to the predictions of the ejective typologies proposed by Lindau (1984) and Kingston (1985). Results of a perception study indicate that the ejective-voiceless unaspirated contrast is more difficult than the other laryngeal contrasts even for native listeners. Implications of the Witsuwit'en data for the categorization of ejectives are discussed.

1 Introduction

1.1 Ejective typology

As noted by Lindau (1984), Kingston (1985, to appear) and Warner (1996), ejective stops may differ cross-linguistically in a number of acoustic features. The classic type of ejective, termed 'stiff ejective' by Kingston (to appear), is characterized by a silent period between consonant release and vowel onset, resulting in long voice onset time (VOT), sharp rise in the amplitude of the vowel (fast rise time), relatively long total duration (closure duration + VOT) and relatively small closure duration/VOT ratio (Lindau 1984). Other characteristics of stiff ejectives, as detailed by Kingston (1985, to appear), are relatively intense burst, high f0 at vowel onset and modal or tense voice quality at vowel onset. Stiff ejectives apparently occur in Tigrinya (Kingston 1985, Fre Woldu 1984–1986), Nez Perce (Aoki 1970), Sahaptin (Grossblatt 1997), Montana Salish (Flemming, Ladefoged & Thomason 1994), and Navajo (Lindau 1984, McDonough & Ladefoged 1993).

Another type of ejective, termed 'slack ejective' by Kingston (to appear), has a different set of acoustic characteristics. As noted by Lindau (1984), slack ejectives have relatively short VOT, creaky voice quality at vowel onset, slow rise times, relatively shorter total duration, and relatively larger closure duration/VOT. Other characteristics

Journal of the International Phonetic Association (2002) 32/1
DOI:10.1017/S0025100302000142
© International Phonetic Association
Printed in the United Kingdom
https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

Table 1 Proposed ejective typology (after Lindau (1984), Kingston (1985, to appear)).

	stiff ejective	slack ejective
total duration	long	short
closure duration/VOT	small	large
burst	intense	normal
VOT	long	short
f0	high	low
voice quality	modal or tense	creaky
rise time	fast	slow

of slack ejectives noted by Kingston (1985, to appear) include normal burst and f0 lowering at vowel onset. Slack ejectives apparently occur in Hausa (Lindau 1984, cf. Lindsay, Hayward & Hayuna 1992), Quiché (Kingston 1985) and Gitksan (Ingram & Rigsby 1987). A summary of the Lindau/Kingston classification of stiff and slack ejectives is provided in table 1. Lindau and Kingston proposed that the articulatory basis for the two clusters of acoustic characteristics of ejectives is differences in vocal fold stiffness or longitudinal tension. With stiff ejectives, glottal adduction is achieved through longitudinal tension and medial compression of the vocal folds. If the vocal folds remain stiff and/or tense following the stop release, then voice onset should be delayed, f0 raised when voicing begins, voice quality tense or modal, and the rise to peak energy fast. However, with slack ejectives, medial compression is the main adductor of the vocal folds. Since there is little longitudinal tension, voice onset time should be relatively short, f0 depressed, voice quality creaky and the rise to peak energy slow.

Although the production of ejectives appears to be well characterized by the stiff/slack classification across many languages, Warner (1996) has pointed out that the ejective stops of some languages do not neatly fit the proposed typology. She found that Ingush ejective stops have the relatively short VOT, irregular voicing at vowel onset and slow rise times that are characteristic of slack ejectives, but raise the pitch of the following vowel, like stiff ejectives. Even Kingston (1985) noted speaker-dependent variation in f0 following ejectives in Tigrinya. One speaker exhibited lower f0 at vowel onset following ejectives than voiceless unaspirated stops, whereas for the other two speakers in his study, there was no significant difference in f0 between the ejectives and voiceless unaspirates. However, in an earlier study of Tigrinya ejectives, Kingston (1982) found significantly higher f0 following ejectives than voiceless unaspirates.

1.2 Athabaskan ejectives

Proto-Athabaskan is reconstructed with a series of ejective stops and affricates (Sapir 1915, Krauss 1964, Leer 1999). The phonetic characteristics of the ejective stops and affricates have been instrumentally investigated in several of the Athabaskan languages, most thoroughly in Navajo. Lindau (1984) and McDonough & Ladefoged (1993) reported that Navajo ejectives have long VOT, fast rise time and modal voice quality at vowel onset, all characteristics of stiff ejectives. However, McDonough & Ladefoged (1993) also noted that [tɬ'] is followed by a brief period of creaky voice for some speakers, perhaps due to the properties of the lateral release, which resembled approximants. Ladefoged & Maddieson (1996: 79) presented data from Navajo in which the 'glottal closure is released into creaky voice which continues for several periods into the beginning of the vowel'. Both Lindau (1984) and McDonough & Ladefoged (1993) note that Navajo ejectives have a consistent closure duration/VOT

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

ratio (Lindau) or VOT/closure + VOT ratio (McDonough & Ladefoged). With respect to this ration, Navajo ejectives pattern with voiceless aspirated stops and affricates and not with the voiceless unaspirated stops and affricates.

Ejectives in other Athabaskan languages have been less well studied. The data presented by Hogan (1976) for one speaker of Chipewyan indicate a substantial post-release silent period, characteristic of stiff ejectives. Hogan also noted that '30% of the vowels begin with a very low fundamental frequency for two to four glottal pulses' (1976: 278). In Hupa, Gordon (1995) observed that the ejective stops and affricates have relatively long VOT, comparable to the voiceless aspirated stops in the language. In the Minto dialect of Lower Tanana, Tuttle (1997, 1998) found that the ejectives and affricates had relatively long VOT but there were some productions where the difference between ejective and voiceless unaspirated stops was signalled primarily through f0 modifications of the following vowel.

This article reports the results of studies of the production and perception of the ejectives in another Athabaskan language, Witsuwit'en, a dialect of the Bulkley Valley-Lakes District language. (The language is better known in academic circles as Babine, but some Witsuwit'en people consider that name offensive.) Witsuwit'en is spoken in New Hazelton, Moricetown, Smithers, Houston, Broman Lake and surrounding areas in British Columbia, Canada. While the phonology of Witsuwit'en indicates that the language has a three-way contrast between voiceless unaspirated, voiceless aspirated and ejective stops and affricates, the phonetic contrast between the ejective and the voiceless unaspirated stops and affricates is subtle to both fieldworkers and second language learners.

2 Phonological background

Previous researchers on Witsuwit'en or related dialects (Morice 1932, Kari 1975, Story 1984) have posited a phonemic series of ejectives in this language. In table 2, we provide a consonant inventory for Witsuwit'en (rare consonants are parenthesized).

Table 2 Witsuwit'en consonant phonemes.

	Labial	Alveolar	Palatal	Labial Velar	Uvular	Laryngeal
Plosive	p (p^h)	t t^h	c c^h	k^w k^{wh}	q q^h	ʔ
Affricate		ts ts^h				
Lateral Affricate		tɬ tɬ^h				
Ejective	(p')	t'	c'	k^w'	q'	
Ejective Affricate		ts'				
Ejective Lateral Affricate		tɬ'				
Nasal	m	n				
Fricative		s z	ç	x^w	χ ʁ	h
Lateral Fricative		ɬ				
Approximant			j	w		
Lateral Approximant		l				

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

Table 3 Effect of laryngeal setting on vowel allophony.

t____	/təz/	[təz]	'driftwood'
	/tosɔ/	[tosɔ]	'gunny sack'
	/tɑji/	[tæji]	'appointed chief'
t^h____	/t^həz/	[t^hʌz]	'cane'
	/t^hɔ/	[t^hɔ]	'water'
	-/t^hɑj/	[t^hɑj]	'paternal uncle'
t'____	-/t'əts/	[t'ʌts]	'incisor'
	-/t'ɔts/	[t'ɔts]	'peel, bark'
	/t'ɑts/	[t'ɑts]	'backwards'

Although the phonetic differences between voiceless unaspirated and ejective stops and affricates can be subtle, the phonology of the language nonetheless clearly maintains the contrast, as revealed by a variety of phonological and morphological evidence.

(1) Minimal or near-minimal pairs for the laryngeal contrasts can be found at all combinations of manner of articulation and supralaryngeal place (except labial, where mainly voiceless unaspirated stops are found). Compare -/t'əq/ 'collarbone', -/təq/ '(straight) up', -/t^həχ/ 'among'. Such minimal sets are found mainly before the vowels /ə/ and /e/, for historical reasons, but they can be found at most places of articulation, supporting the idea that there is a synchronic contrast between these three laryngeal states.

(2) As discussed by Kari (1975) and Story (1984), the ejective stops/affricates pattern with the voiceless aspirated stops/affricates and the voiceless fricatives in triggering allophonic raising and lowering processes on the following vowel. Here we focus on processes of vowel allophony involving the vowels /ə o ɑ/. (The underlying vowel inventory is /i e o a ə u/.) Following voiceless unaspirated stops and affricates, voiced fricatives and nasals, /ə o ɑ/ have allophones [ə~ɪ o æ], respectively. However, following voiceless aspirated or ejectives or voiceless fricatives, the allophones of /ə o ɑ/ are [ʌ ɔ ɑ] (Story 1984), as verified in an acoustic study (Hargus 1998). Table 3 presents lexical items which illustrate the allophones of /ə o ɑ/ following the three laryngeal settings for alveolar stops. Cross-linguistic surveys of this sort of phenomenon (interaction of voice quality and vowel quality) can be found in Denning 1989 and Trigo 1991.

(3) The distribution of allomorphs of the first person plural subject prefix is in part determined by consonant laryngeal state (Hargus 1996). The two allomorphs, [ts']- and [z]-, of the first person plural subject prefix can be observed in [pət ʁetsʼɛne] 'we're cooking' vs. [pət ʁezt^hɑneɬ] 'we'll cook'. The [ts']- allomorph is generally used following a consonant within the same word. Note that an epenthetic vowel [ə] (in square brackets) also occurs in the forms in table 4.

Table 4 [ts'] / C___.C.

[nenəx^{w}ts'[ə]t^{h}ɑt'ɛɬ]	'we'll see you (pl.) again'
[weɬts'[ə]tistəts'əq]	'we don't hear each other'
[pəʁents'[ə]t^{h}ɑqəz]	'we'll skin several'

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

Table 5 [z] /ʔ___.C.

[neʔ[ə]z̲tistc'əj], *[neʔts'ətistc'əj]	'we shot ourselves'
[ɬɑʔ[ə]z̲təceʁ]	'we finished it'
[ʔ[ə]z̲tətqes]	'we're scratching ourselves'

However, when the preceding consonant is [ʔ], then the [z] allomorph must be used instead, as exemplified in table 5. The failure to use the regular [ts'] allomorph in the ʔ___C context makes sense if viewed as avoidance of adjacent segments with glottal constriction. Thus, the phonology of Witsuwit'en treats both [ʔ] and [ts'] as glottalized consonants. (Note that the cognate to the Witsuwit'en first person plural subject prefix has changed from ejective to voiceless unaspirated in some Athabaskan languages; e.g. Navajo.)

To summarize, phonological contrasts, vowel allophony and prefix allomorphy indicate that Witsuwit'en has inherited the three-way contrast between voiceless unaspirated, aspirated and ejective stops and affricates from Proto-Athabaskan.

3 Production study

3.1 Research questions

The production study, which builds on studies reported in Davis & Hargus (1994, 1996), was designed to investigate the acoustic correlates of the three-way laryngeal contrast in Witsuwit'en with emphasis on the ejectives. Given the noted crosslinguistic variation in ejectives, we were interested in formulating any generalizations across speakers that could characterize the three laryngeal contrasts in the language as a whole. We were also interested in individual speaker tendencies in the phonetic realization of ejectives. We wondered about the effect of gender on the pronunciation of the ejectives, since the impression is that the ejectives are more easily distinguished from the voiceless unaspirates when pronounced by women. The questions below were designed to elucidate the acoustic correlates of the laryngeal contrast in Witsuwit'en.

(1) Are there systematic differences between the voiceless unaspirated, voiceless aspirated and ejective consonants with respect to any of the following measures: VOT, f0 perturbation, jitter, and vowel rise time?
(2) Do any of the acoustic traits cluster together, or can they vary independently from each other? For example, does VOT covary with f0 such that the longer the VOT (i.e. the more ejective-like the stop is) the higher the f0?
(3) To what degree does language internal variation, both within speakers, across speakers, and across groups of speakers, echo the cross-linguistic variation noted in previous work? How many types of ejectives are there?

3.2 Methods

Data were collected in the field from eleven adult native speakers (five male, six female) ranging in age from approximately 50 years to 75 years. All speakers who participated in the study spoke English as well, although competence in English varied from clearly non-native to native-like.

The data measured in the production study are a subset of a larger database of recordings of words spoken in isolation illustrating the contrast between voiceless

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

Table 6 Sample of production word list (consonants of interest are underlined).

phonemic	phonetic	
/t̲ɑji/	[t̲æji]	'appointed chief, leader'
/t̲'ɑc^{h}o/	[t̲'ɑc^{h}o]	'mallard'
/t̲^{h}ɑni/	[t̲^{h}ɑni]	'three people'

unaspirated, voiceless aspirated, and ejective stops at four places of articulation and in a variety of segmental environments. All consonants measured were root initial consonants since the attraction of stress to roots has been noted in many Athabaskan languages (e.g. Ahtna, Kari 1990). The current study analyzes the voiceless unaspirated, voiceless aspirated, and ejective alveolar stops. All but one of the tokens occurred in word initial position. Some sample words are provided in table 6. The complete word list for alveolar initial words is provided in appendix 1.

To facilitate elicitation, the word list was organized semantically, which effectively randomized the list phonologically. The speakers varied in their ability to read Witsuwit'en orthography. Those who could read with ease read each word twice. The others were prompted orally and repeated each word at least twice.

Field recordings were made in the quietest available rooms using a professional field analog cassette recorder (Sony WM-D6C) and a table top microphone with flat resonse up to 20 kHz (Sony ECM-929LT). The recordings were later digitized at 16 kHz, 16 bits, using a Kay CSL setup. Measurements were made using Kay's Multispeech program.

Following an initial qualitative auditory analysis, a variety of quantitative measurements were made. The auditory analysis indicated a wide range of phonetic realizations of /t'/, from ejective to post-creaky (i.e. voiceless but followed by creaky voicing in the vowel onset) to voiceless unaspirated. Because the ejectives varied in complex ways we decided on a variety of measures, based on past studies of ejective consonants, that would allow us to characterize the ejectives as well as the other two laryngeal settings for stops along a variety of dimensions.

VOT (voice onset time) is a well established measure for examining aspiration contrasts in stops (e.g. Lisker & Abramson 1964) and, as discussed above, it has been used to compare ejectives with other laryngeal contrasts. VOT measurements in this study were taken from a waveform with an accompanying spectrogram for reference. VOT was taken to be the interval between the onset of the release burst to the first positive or negative movement away from zero at the onset of periodicity.

Much of the information that distinguishes ejective from voiceless unaspirated consonants occurs in the form of voice-quality perturbations at the onset of the following vowel. Past voice quality studies (e.g. Ladefoged, Maddieson & Jackson 1988) have established a set of measures that can be used for quantifying differences between creaky, modal, and breathy voicing. Relative to modal voicing, creaky voicing has lower f0, increased jitter (irregularity in the glottal cycle), lower airflow accompanied by a concomitant decrease in signal intensity, and more gradual dropoff of energy in the higher frequencies (decreased spectral tilt).

Thus, in addition to VOT, we chose three measures that are appropriate for the onset of the vowel: 1) f0, 2) jitter and 3) vowel rise time. Average f0 and jitter were obtained from 30 ms windows at onset of voicing and vowel midpoint. Voiced period marks were generated automatically using a Multispeech 2.4 algorithm and inspected visually for errors. Outlier values were flagged and reinspected by a team of two researchers. Where there was agreement, the period marks were hand corrected. The harmonic mean of the pitch was calculated by computing the inverse of the time

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

Table 7 Formulas for normalized measures.

f0 perturbation	(mean f0 at onset of voicing) – (mean f0 at vowel midpoint)
jitter perturbation	(mean jitter at onset of voicing) – (mean jitter at vowel midpoint)
rise time	(RMS energy over vowel peak) – (RMS energy at 30 ms from vowel onset)

between each voiced period mark. The mean percent jitter ratio was calculated by taking the mean absolute difference between sequential voiced periods, dividing it by the mean vocal period of the window, and multiplying it by 100 (Koike 1973). Energy was obtained from instantaneous measures at vowel onset and 30 ms from vowel onset, and from a 30 ms window centered over the vowel's amplitude peak. Energy RMS values were calculated from the absolute amplitude values in a frame of data divided by the duration of the frame. Decibels of Sound Pressure (dB SPL) were computed from 20 times the log (base 10) of RMS energy values.

f0, jitter and rise time (energy) were then internally normalized as summarized in table 7. f0 perturbation was calculated by subtracting mean f0 at vowel midpoint from mean f0 at vowel onset. Jitter perturbation was calculated in analogous fashion. Rise time was calculated by subtracting energy at 30 ms from vowel onset from energy at the vowel peak (see also figure 6). Because f0 perturbation, jitter perturbation and rise time are normalized measures, they allow us to make generalizations across speakers without having to compensate for individual differences in pitch, voice quality, or speech intensity.

Spectral tilt is missing from our study because it is affected by differences in the first formant (Ladefoged, Maddieson & Jackson 1988). Any difference in VOT would result in measurements at different points in the F1 transition, thereby creating a confound for spectral tilt in our data. Moreover, the allophonic differences in vowel height following the voiceless aspirated and ejective stops and affricates on one hand and the voiceless aspirated and unaspirated consonants on the other would mean that small but systematic differences in F1 could not be controlled for.

The result of each dependent measure (VOT, pitch, jitter, rise time) was submitted to two statistical tests, each with alpha set at .05. (1) To investigate the effect of the three laryngeal types (voiceless unaspirated, ejective, aspirated) for the group, a repeated-measures ANOVA (laryngeal type × speaker) was performed, using each speakers's mean value on each measure. The interaction of laryngeal type with the independent variable of gender was also investigated in the repeated measures design. This ANOVA seeks to avoid the possibility of type 1 error through inflated *n* by using each speaker's mean as the dependent variable. However, within-speaker variability, which we had also wanted to investigate, is not captured by this repeated measures design. (2) Therefore, a second, one-factor ANOVA was also performed for each speaker, using all tokens produced by that speaker. Post-hoc comparison of mean levels of laryngeal type was also performed via Fischer's PLSD.

3 Group results

3.3.1 Voice onset time

Unlike the aspirated stops, the period between the release burst and the onset of voicing was silent for the ejectives. This typical ejective signature is illustrated in figure 1, a waveform of the word [t'ɛs] 'charcoal' (with 92 ms VOT).

Across subjects, /t^{h}/ has the longest VOT (average 59 ms, s = 10.9), followed by /t'/

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

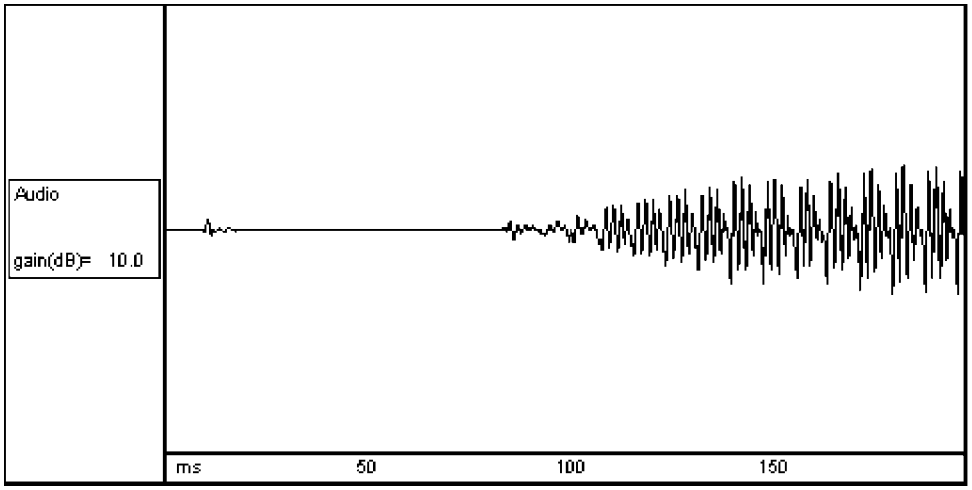


Figure 1 Waveform of [tʼɛs] 'charcoal' (speaker SM).

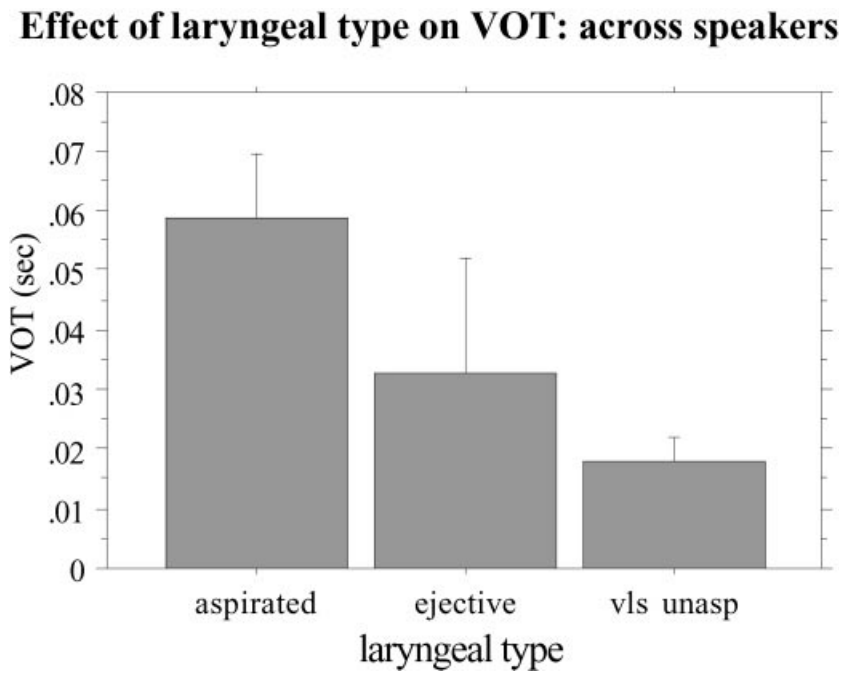


Figure 2 Bar chart with means and standard deviations illustrating the effect of laryngeal type on VOT duration across speakers.

(33 ms, s = 19.4) and /t/ (18 ms, s = 4.3), as shown in figure 2. Laryngeal type has a significant effect on VOT duration ($F[2,20] = 42.799$, $p < .0001$). The ejectives produced by male speakers had longer VOT (36 ms, s = 23.0) than those of female speakers (30 ms, s = 17.7), but there was no significant interaction between gender and laryngeal type.

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

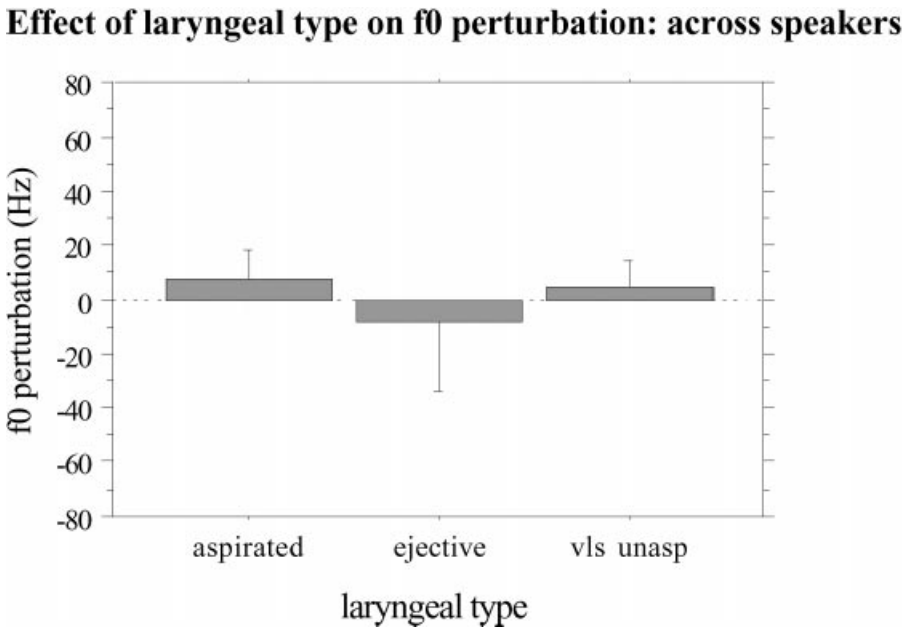


Figure 3 Bar chart showing the mean and standard deviation of f0 perturbation (Hz) for each laryngeal type: across speakers.

3.3.2 f0 perturbation

Across speakers, /t'/ is on average followed by a period of lowered f0 (−9 Hz, s = 25.3) relative to the vowel mid-point, while f0 at the onset of voicing is raised relative to the vowel mid-point for both /t^h/ (8 Hz, s = 10.5) and /t/ (4 Hz, s = 10.2), as shown in figure 3. Across speakers, laryngeal type did not have a significant effect on pitch perturbation. However, we found a significant interaction between laryngeal type and gender ($F[2,18] = 11.473$, $p = .0006$): the ejectives produced by men generally began with a period of raised f0 (8 Hz, s = 7.0), whereas those produced by women began with a period of lowered f0 (−22 Hz, s = 27.1) relative to the vowel's medial f0, as shown in figure 4.

3.3.3 Jitter perturbation

Jitter perturbation was greater following ejectives (6.9, s = 3.20) than aspirated stops (1.8, s = 1.19) or voiceless unaspirated stops (1.6, s = 1.45), as shown in figure 5. Laryngeal type had a significant effect on jitter perturbation ($F[2, 20] = 98.226$, $p < .0001$). However, there was no significant interaction of gender with laryngeal type.

3.3.4 Rise time

Energy was measured at three points in the vowel: at the onset of voicing, at 30 ms into the vowel, and at the vowel's peak. Figure 6, a plot of the raw RMS energy values against the measurement point, gives a rough portrait of the effect of the three laryngeal types on the increase in amplitude following the onset of voicing. The energy at onset and especially at the 30 ms point best separates the ejective from the non-ejectives. The slope is shallower between 0 and 30 ms for the ejectives than for the other kinds of stops, but steeper between 30 ms and the peak. Rise time, as we are defining it here, is the difference between the peak energy and the energy at 30 ms. It is effectively determined by the energy level at the 30 ms point: the lower the energy at that point,

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

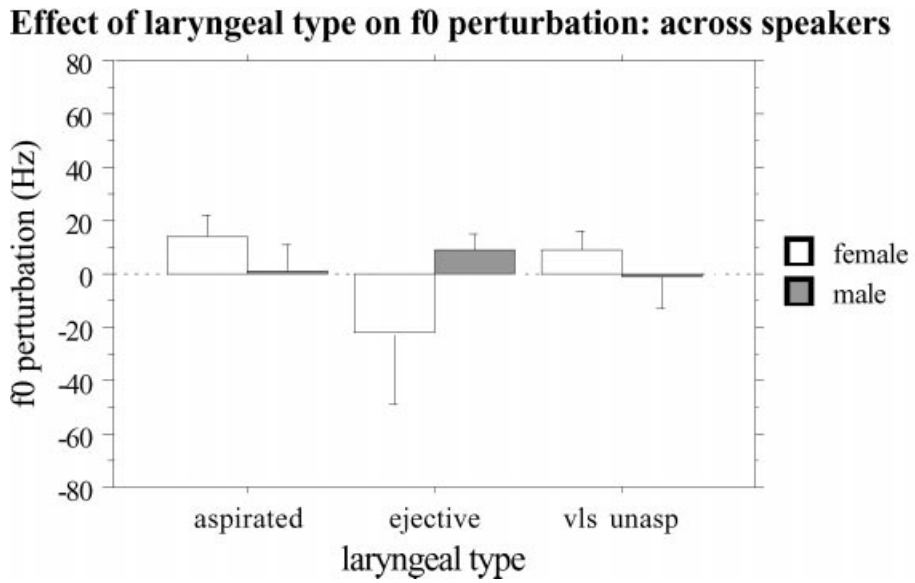


Figure 4 Bar chart showing the mean and standard deviation of f0 perturbation (Hz) for laryngeal type * gender.

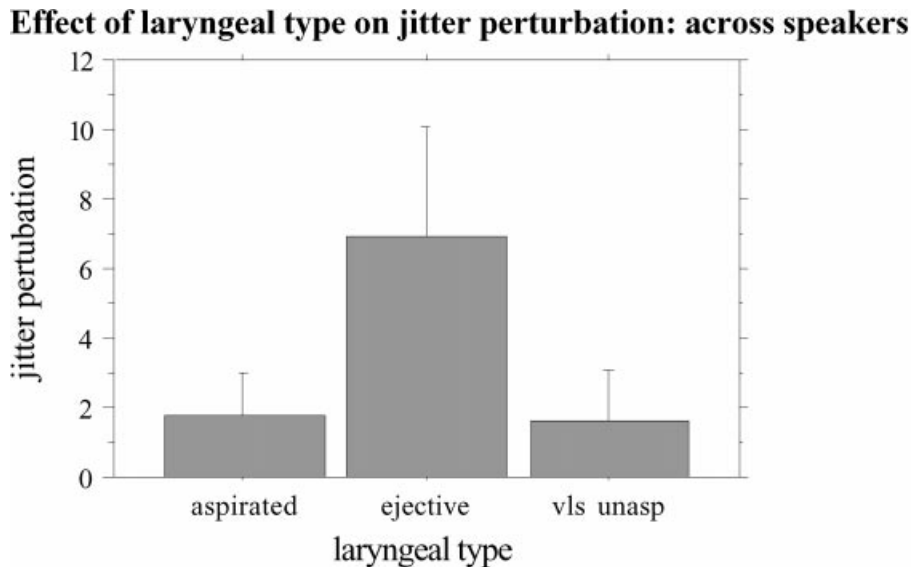


Figure 5 Bar chart with the mean and standard deviation of jitter perturbation for each laryngeal type: across speakers.

the shallower the slope overall and the greater the difference between the vowel midpoint and the 30 ms measure.

As shown in figure 7, the ejective had a slower rise time, or a greater difference between the 30 ms point (5 dB, s = 1.5) and the peak energy of the vowel, than either the aspirated (2 dB, s = 1.3) or voiceless unaspirated stops (1 dB, s = .9). Laryngeal type

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

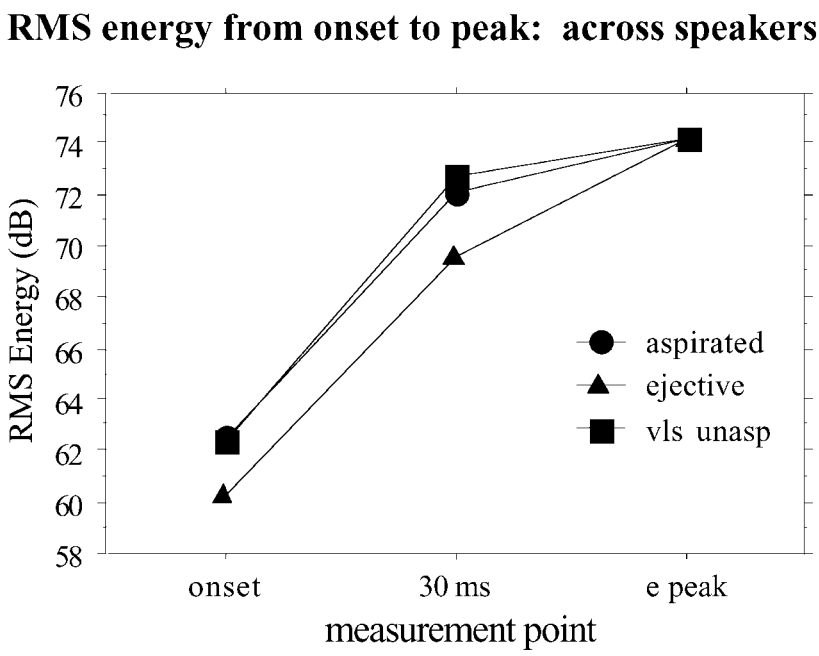


Figure 6 Line chart with the mean RMS energy values in dB plotted against the measurement points: across speakers.

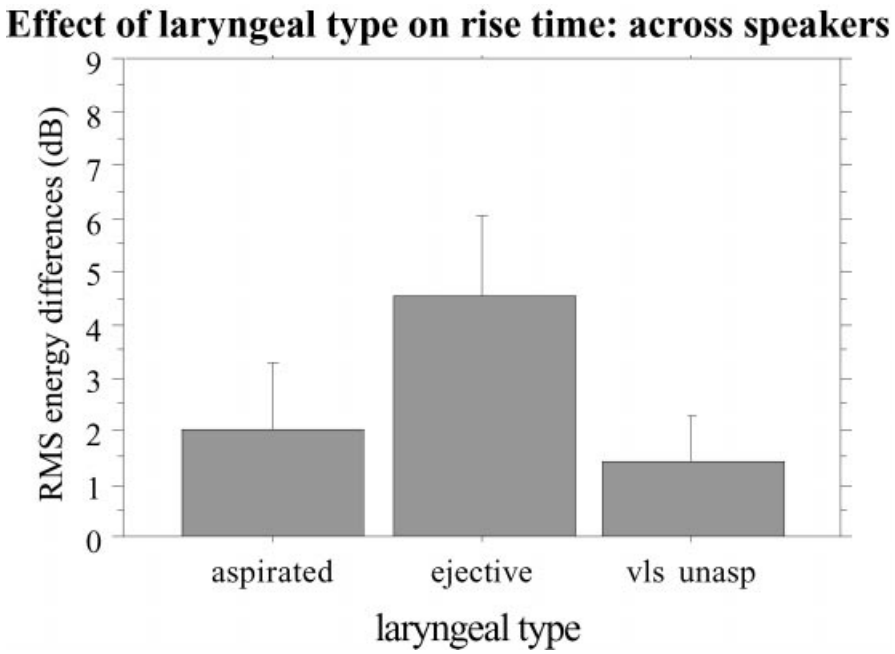


Figure 7 Bar chart showing the means and standard deviations of rise time for the three laryngeal types: across speakers.

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

Table 8 Summary of characteristics of /t'/ relative to other laryngeal types: across speakers. Bold type indicates significant effect.

	VOT	f0 perturbation	jitter perturbation	rise time
group	**intermediate**	depressed	**more**	**slow**
males	**intermediate**	raised	**more**	**slow**
females	**intermediate**	**depressed**	**more**	**slow**

had a significant effect on rise time ($F[2,20] = 31.824$, $p < .0001$). The interaction of gender with laryngeal type was not significant.

3.3.5 Summary

Table 8 summarizes the characteristics of ejectives with respect to the other laryngeal types for the group and for subgroups defined by gender. Significant effects are in bold. In the VOT column of table 8, 'intermediate' means that the VOT was intermediate between voiceless aspirated and voiceless unaspirated stops. In the f0 column, 'raised' means that there was a brief period of raised f0 at the vowel onset, and 'depressed', a brief period of lowered f0. In the jitter column, 'more' means that at vowel onset following an ejective there was more jitter than at the vowel onset following a voiceless unaspirated stop. In the rise time column, 'slow' means that rise time was slower than that found after voiceless unaspirated stops. In general, table 8 suggests uniform, 'slack' characteristics of /t'/ for the group and the gender-defined subgroups for all measures. However, f0 represents a major exception to the 'slack' characterization: laryngeal type had no significant effect on f0 perturbation for the group, and male and female speakers exhibited opposite tendencies. The uniformity suggested by table 8 is also contradicted in the production characteristics of individuals, detailed in the next section.

3.4 Individual results

3.4.1 Voice onset time

As expected, for each speaker, aspirated stops had significantly longer VOT than the voiceless unaspirated stops. However, the VOT of the ejectives was not consistently longer than that of the voiceless unaspirated stops. Speakers divided into three groups. For the largest ('short VOT') group, consisting of seven speakers, the VOT of /t'/ for each speaker was significantly shorter than their /t^h/ but not significantly different from their /t/. The next largest ('long VOT') group, consisting of three speakers, exhibited VOT for /t'/ that was significantly longer than /t/ but not significantly shorter than /t^h/. The final ('intermediate VOT') group consisted of one speaker, whose /t'/ VOT was significantly shorter than /t^h/ but significantly longer than /t/.

Figure 8 illustrates VOT (mean and standard deviation) of the three laryngeal types for a speaker representative of the short VOT group. For this speaker, laryngeal type has a significant effect on VOT duration ($F[2,29] = 14.748$, $p < .0001$.). Post hoc analysis confirms that the VOT of this speaker's /t'/ (25 ms, $s = 18.4$) is significantly shorter ($p = .0001$) than that of her /t^h/ (61 ms, $s = 24.8$) but not her /t/ (20 ms, $s = 8.6$).

Figure 9 illustrates means and standard deviations of VOT for the three laryngeal types of a speaker belonging to the long VOT group. For this speaker, laryngeal type has a significant effect on VOT duration ($F[2,21] = 15.326$, $p < .0001$). Her /t'/ has significantly ($p < .0001$) longer VOT at 48 ms ($s = 34.8$) than her /t/ (13 ms, $s = 3.4$) but is not significantly shorter than her /t^h/ (66 ms, $s = 19.8$).

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

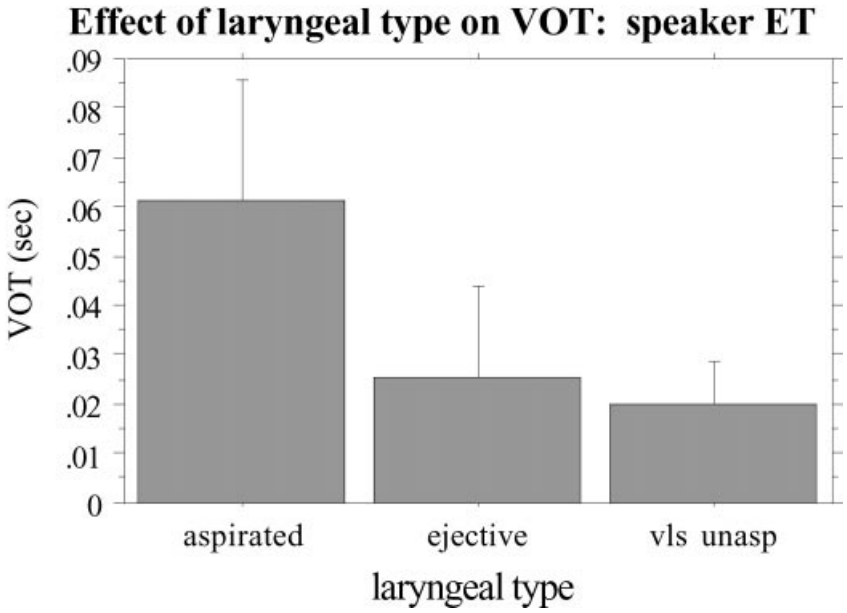


Figure 8 Bar chart with means and standard deviations illustrating the effect of laryngeal type on VOT for ET, a short VOT speaker.

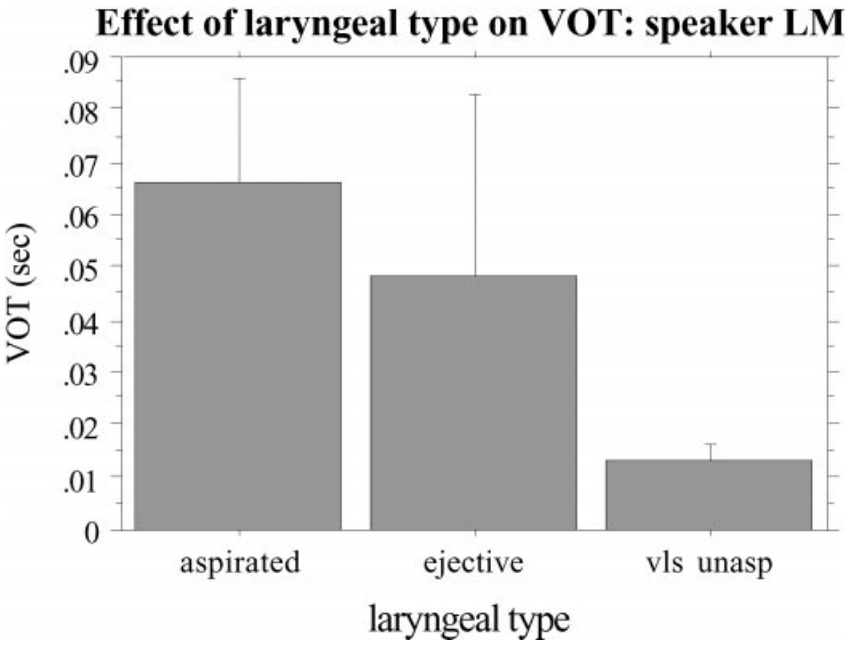


Figure 9 Bar chart with means and standard deviations illustrating the effect of laryngeal type on VOT duration for LM, a long VOT speaker.

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

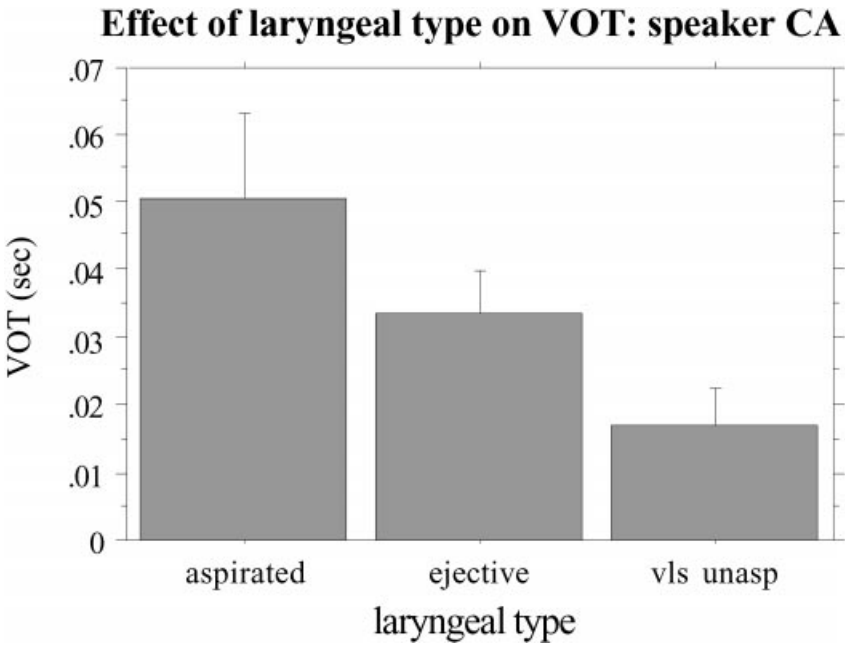


Figure 10 Bar chart with means and standard deviations illustrating the effect of laryngeal type on VOT duration for CA (intermediate group).

The one speaker in the intermediate group had a VOT for /t'/ that was comparable to the group pattern, in which the ejective VOT was intermediate between the voiceless aspirated and unaspirated stops. Figure 10 illustrates VOT means and standard deviations for this speaker's three laryngeal types. For this speaker, laryngeal type has a significant effect on VOT duration ($F[2,13] = 17.733$, $p = .0002$). Post-hoc analysis confirms that this speaker's /t'/ is significantly longer ($p < .05$) at 33 ms ($s = 6.4$) than his /t/ (17 ms, $s = 5.6$) as well as significantly shorter than his /t^h/ (50 ms, $s = 13.0$) ($p < .05$).

Although our speakers can be divided into three groups on the basis of the significance of their VOT contrast between /t'/ and /t/, VOT of the ejective actually varied continuously from relatively short, around the average for /t/, to relatively long, around the average for /t^h/, with all values in between. As shown in figure 11, a histogram of VOT for all /t'/ productions in our sample, there is a single peak around the mean (32 ms, $s = 25.7$) but with considerable rightward skew.

3.4.2 f0 perturbation

Our speakers again divided into three groups with regards to f0 perturbation in the production of the ejective. One group, the f0 depressors, consisted of four speakers for whom the f0 perturbation of /t'/ was significantly lower than that of /t/. The f0 neutral group consisted of six speakers for whom the f0 perturbation of /t'/ did not differ significantly from that of /t/. The third group, f0 raisers, consisted of one speaker for whom the f0 perturbation of /t'/ was significantly higher than that of /t/.

Figure 12 illustrates f0 perturbation for a speaker representative of the f0 depressor group. For this speaker, /t'/ is followed by a brief period of lowered pitch (average −27 Hz, $s = 49.1$) at vowel onset, whereas /t/ and /t^h/ have slightly raised pitch perturbation (/t/, 19 Hz, $s = 8.2$; /t^h/, 11 Hz, $s = 15.8$). The pitch perturbation following /t'/ is significantly lower than that of /t/ ($p = .0062$) and /t^h/ ($p = .0188$), but the difference between /t/ and /t^h/ is not significant.

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

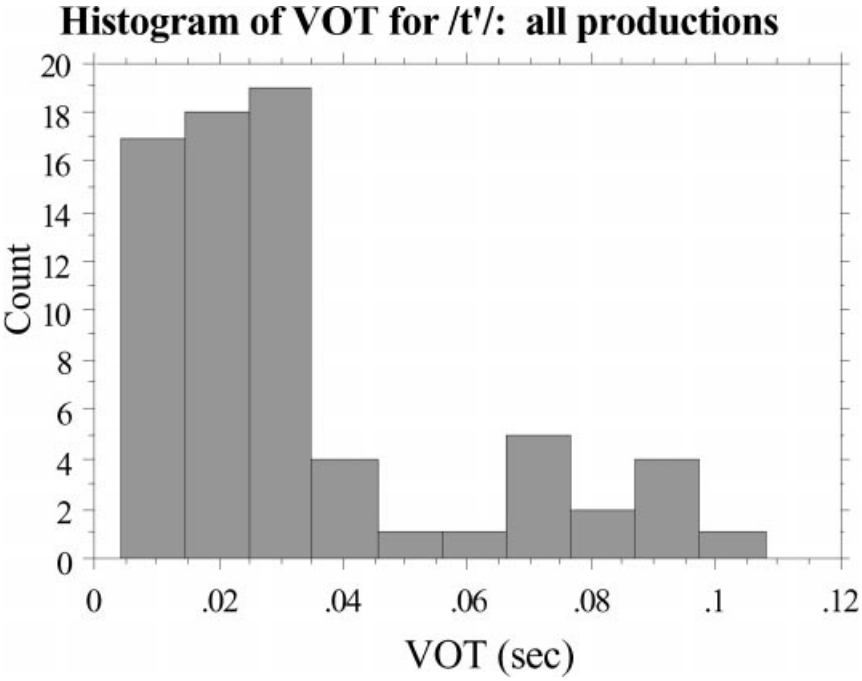


Figure 11 Histogram of VOT of all /t'/ productions in the sample.

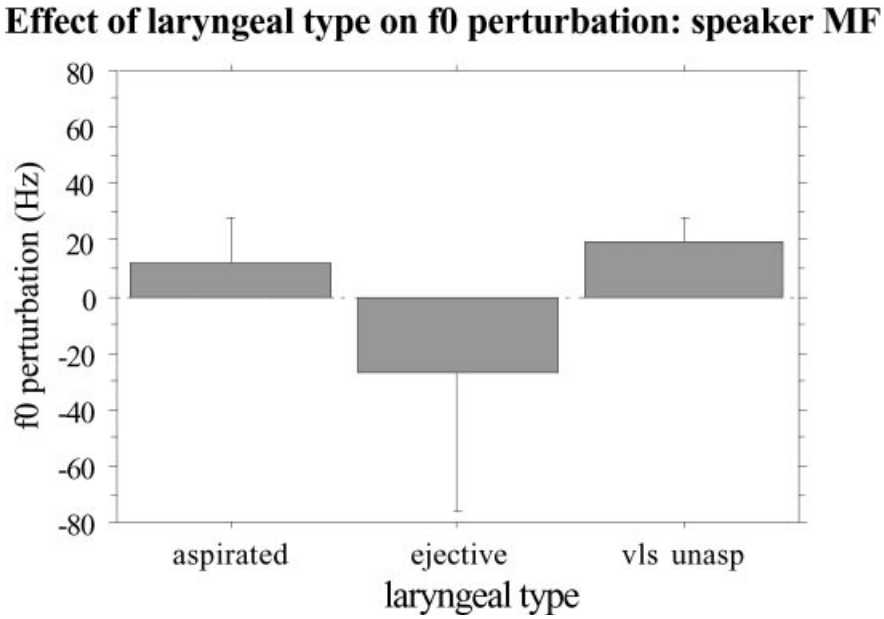


Figure 12 Bar chart with the mean and standard deviation of f0 perturbation (Hz) for each laryngeal type for MF (f0 depressor group).

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

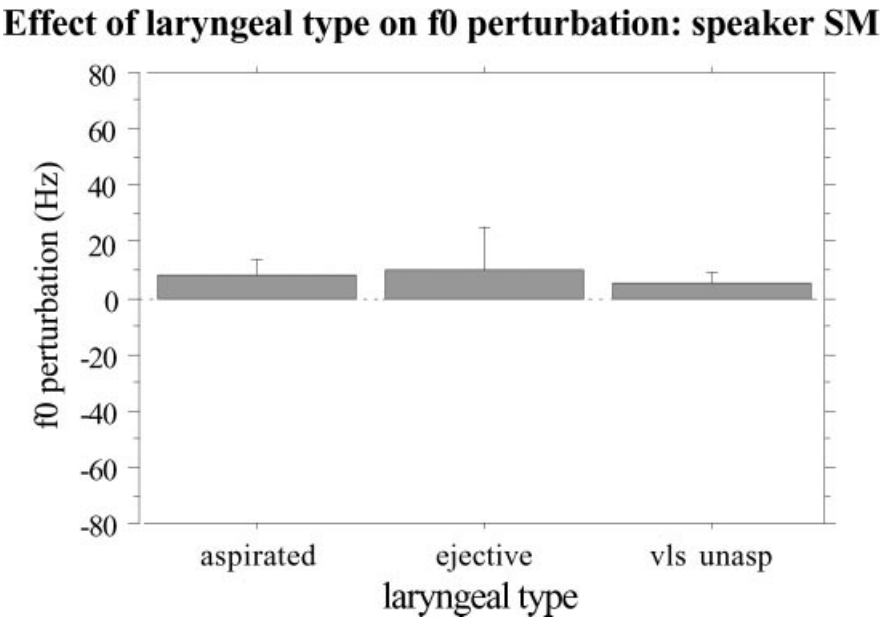


Figure 13 Bar chart with the mean and standard deviation of f0 perturbation (Hz) for each laryngeal type for SM (representative of f0 neutral group).

Six of the eleven speakers in our sample were f0 neutral: for these speakers, /t'/ was followed by an f0 at vowel onset that did not differ significantly from that of /t/. Figure 13 illustrates f0 perturbation for a representative f0 neutral speaker. For the f0 neutral group, the effect of laryngeal type on f0 depression is not significant.

One speaker, AJ, belonged to the remaining f0 raiser group. As shown in figure 14, this speaker's /t'/ was followed by an f0 at the onset of voicing that was higher than the medial f0. For this speaker, analysis of variance reveals no significant effect of laryngeal type on f0 perturbation. However, Fisher's PLSD indicates weakly significant ($p = .04$) pitch raising following /t'/ (17 Hz, $s = 30.9$) than /t/ (-5 Hz, $s = 16.3$). /t^h/ is effectively neutral with respect to pitch raising for this speaker (0 Hz, $s = 8.7$).

As with VOT, we found considerable variation in the f0 perturbation of /t'/ in individual productions. A histogram of f0 perturbation is provided in Figure 15, which shows a mean centered around -8 Hz and standard deviation of 39.4.

3.4.3 Jitter perturbation

At the individual level, two groups of individuals can be distinguished, greater jitter and neutral jitter. For three individuals in the greater jitter group, there was significantly more jitter following /t'/ than /t/. Figure 16 presents results for a typical speaker of the greater jitter group. This speaker's /t'/ shows a jitter of 6.1 ($s = 4.16$), compared to 1.6 ($s = 1.06$) for her voiceless unaspirated stop and 1.5 ($s = 1.73$) for her voiceless aspirated stop. Laryngeal type has a significant effect on jitter perturbation for this speaker ($F[2,22] = 8.211$, $p = .0022$). The jitter following /t'/ is significantly greater ($p = .0030$) than that following /t/ for this speaker.

For the remaining eight individuals, /t'/ was followed by more jitter than /t/ but the difference was not significant. Figure 17 shows jitter perturbation results for a speaker typical of the neutral jitter group.

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

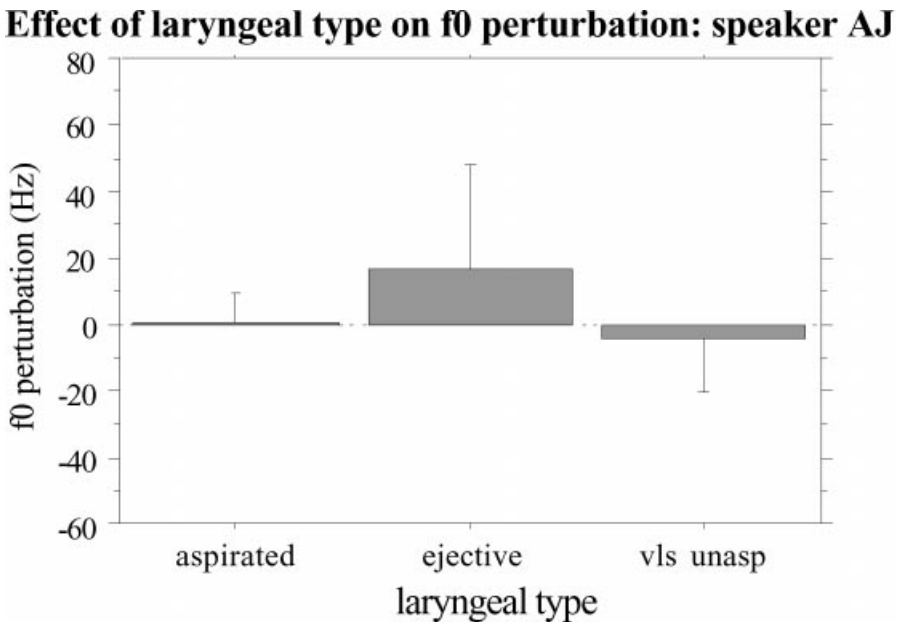


Figure 14 Bar chart with the mean and standard deviation of f0 perturbation (Hz) for each laryngeal type for speaker AJ (f0 raiser group).

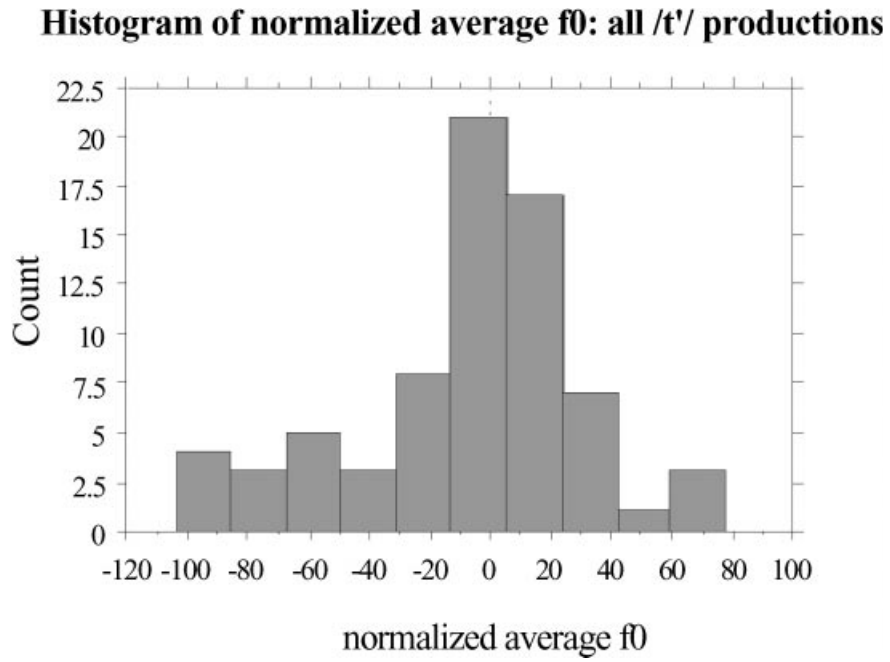


Figure 15 Histogram of f0 perturbation of all /t'/ productions in the sample.

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

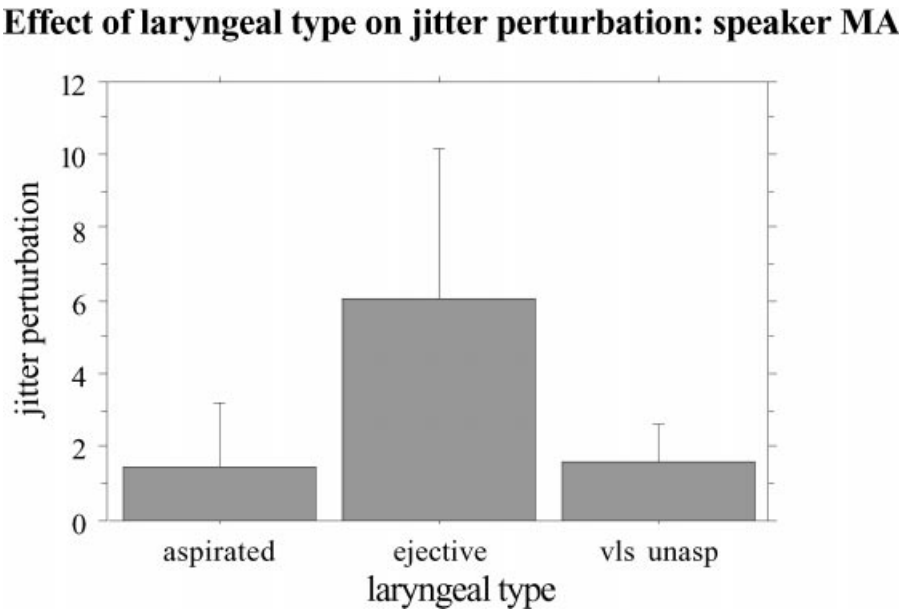


Figure 16 Bar chart with the mean and standard deviation of jitter perturbation for each laryngeal type for speaker MA (greater jitter group).

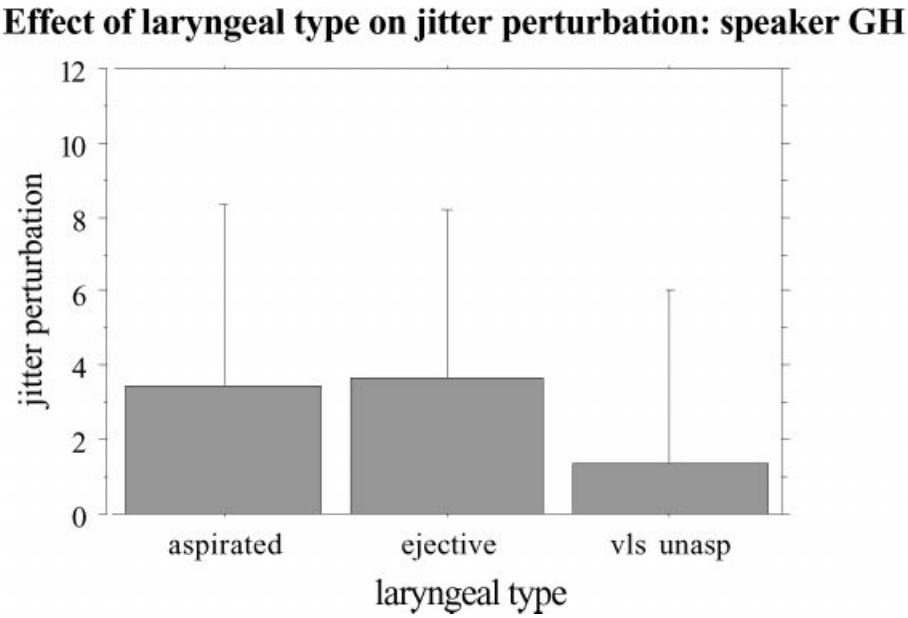


Figure 17 Bar chart with the mean and standard deviation of jitter perturbation for each laryngeal type for speaker GH (neutral jitter group).

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

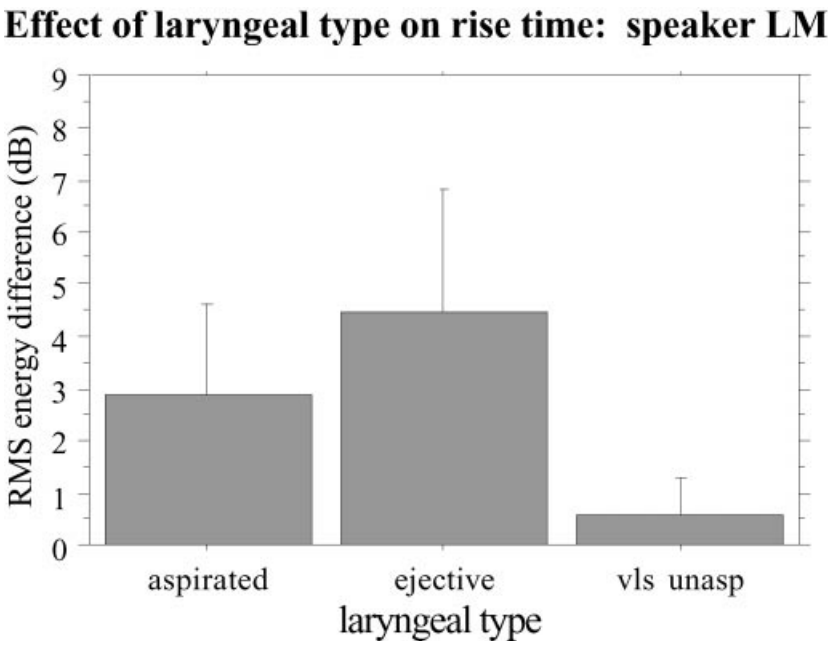


Figure 18 Bar chart showing the means and standard deviations of rise time for the three laryngeal types: speaker LM (slow rise time group).

3.4.4 Rise time

At the individual level, our speakers also divided into two rise time groups. Nine of the eleven speakers had significantly slower rise times following /t'/ than /t/. Figure 18 presents rise time results for a speaker representative of the slow rise time group. For this speaker, laryngeal type had a significant effect on rise time ($F[2,22] = 4.696$, $p = .0200$). /t'/ exhibited significantly ($p = .0001$) slower rise time (4.4 dB, $s = 2.39$) than the voiceless unaspirated stop (2.9 dB, $s = 1.73$).

For the two speakers of the normal rise time group, the rise time for the ejectives was not significantly different from that of their voiceless unaspirated stops. Figure 19 presents rise time results which are representative of the normal rise time group.

3.4.5 Summary

Table 9 provides a summary of the characteristics of /t'/ relative to /t/ for each individual. In table 9, 'short' means that the VOT was comparable to that of a voiceless unaspirated stop; 'long' means that the VOT was comparable to that of a voiceless aspirated stop. Other labels have the definitions used in table 8. Two points from table 9 are noteworthy: (1) there are quite consistent differences in jitter and rise time between ejectives and voiceless unaspirated stops, but speaker-specific patterns in VOT and f0; and (2) some speakers exhibit more significant differences between the ejectives and voiceless unaspirated stops than others.

3.5 Discussion

Returning to the questions we posed at the outset of the production study, we can now conclude that there are systematic differences in VOT, f0 perturbation, jitter, and vowel rise time between the consonants which differ in laryngeal type. The average /t'/ in

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

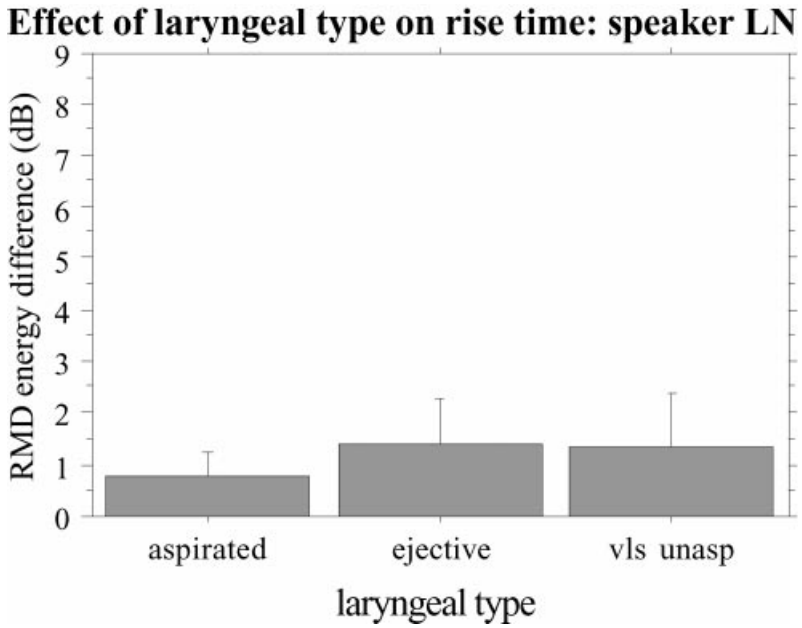


Figure 19 Bar chart showing the means and standard deviations of rise time for the three laryngeal types: speaker LN (normal rise time group).

Table 9 Summary of characteristics of /t'/ relative to /t/: individuals. Bold type indicates a significant difference between /t/ and /t'/.

	VOT	f0 perturbation	jitter perturbation	rise time
AJ	short	**raised**	more	**slow**
CA	**intermediate**	raised	more	**slow**
ET	short	raised	more	slow
GH	short	raised	more	**slow**
LM	**long**	**depressed**	**more**	**slow**
LN	short	depressed	more	slow
MA	short	**depressed**	**more**	**slow**
MF	**long**	**depressed**	more	**slow**
PN	short	**depressed**	more	**slow**
RG	short	flat	**more**	**slow**
SM	**long**	raised	more	**slow**

Witsuwit'en has VOT intermediate between the voiceless unaspirate and the voiceless aspirate, a period of depressed f0 and increased jitter at vowel onset, and relatively low amplitude at vowel onset. However, given the considerable individual variation in the production of the ejectives, the notion 'average Witsuwit'en /t'/' is problematic for our sample.

At the outset of this study we raised the question of how the ejective stops of Witsuwit'en compare to those of other languages. As summarized in table 1, in the models of Lindau and Kingston, there are primarily two types of ejectives, with opposite values of certain acoustic features. Some aspects of this model describe our

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

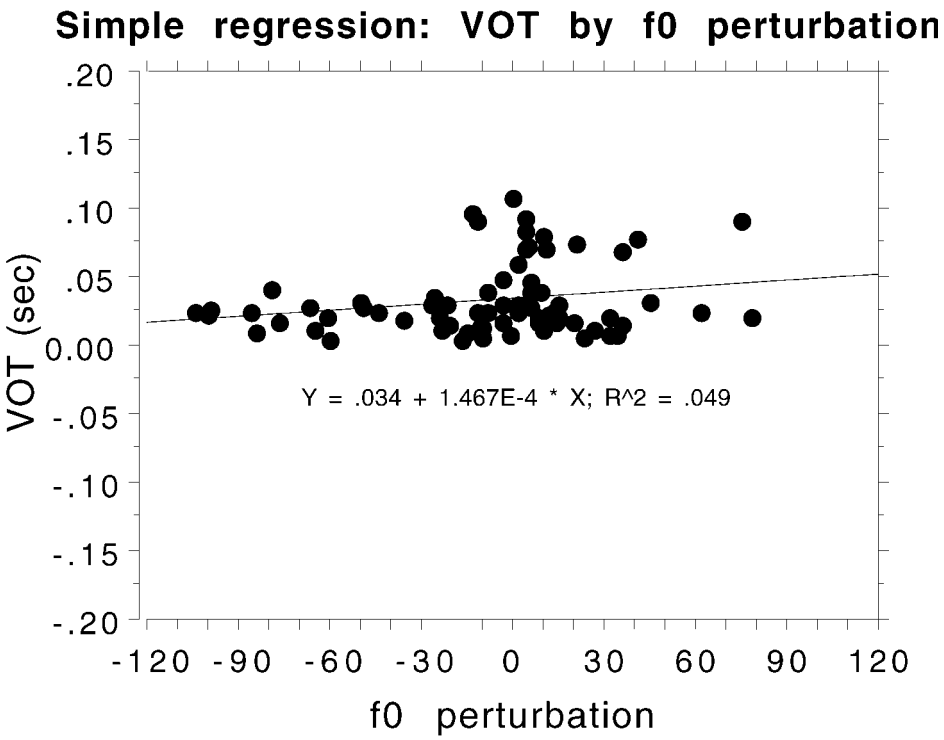


Figure 20 Scattergram with a simple regression line illustrating the relationship between variation in VOT and f0 perturbation for the ejective.

data well: both for the group and for individuals, the Witsuwit'en ejectives have relatively slow rise times and more creaky voice quality, indicative of the slack ejective class. At the group level, f0 is typically depressed at vowel onset, again characteristic of slack ejectives, but at the individual level, there is considerable variation in the amount of f0 depression. Most women were f0 depressors, but most men were f0 raisers, a gender effect for which we have no explanation. VOT is intermediate for the group, again with considerable variation among individuals.

Given this variation, the models of Lindau and Kingston would be useful if they could predict the way in which f0 perturbation varied with VOT. That is, despite the variation in our data, we might find that ejectives with relatively long VOT had raised f0, whereas those with relatively short VOT might have depressed f0. Testing this prediction against our data, however, we found very low ($r^2 = .049$) correlation between f0 perturbation at vowel onset and VOT duration. As the scatter plot and regression line in figure 20 show, long and short VOT can vary nearly independently of f0.

Thus, in our data, we find all possible combinations of long and short VOT vs. lowered f0, raised f0 and essentially flat f0 at vowel onset. In figure 21 through figure 26, we provide examples of each of these six types. Figure 21 is a spectrogram illustrating a long VOT accompanied by f0 raising. The example utterance is -[t'ɔc^{h}o] 'big nest', pronounced by speaker AJ with a VOT of 90 ms and f0 perturbation of 75 Hz (compare the group average for ejectives: 33 ms VOT and −8 Hz f0 perturbation). Figure 22 illustrates the combination of short VOT with f0 raising. The spectrogram is of the word [t'ɑc^{h}o] 'mallard', uttered by speaker RG, with a VOT of 7 ms and a 34 Hz perturbation. Figure 23 is a spectrogram illustrating long VOT combined with

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

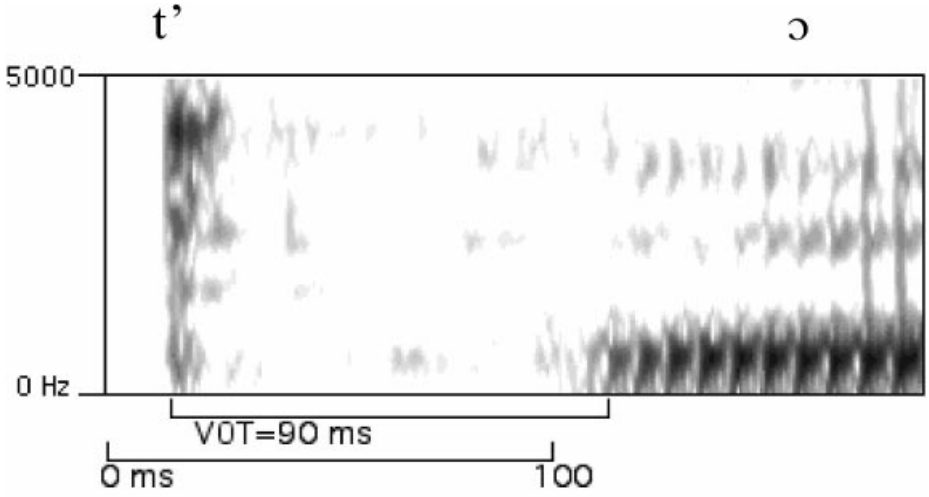


Figure 21 Spectrogram illustrating long VOT with raised f0: speaker AJ.

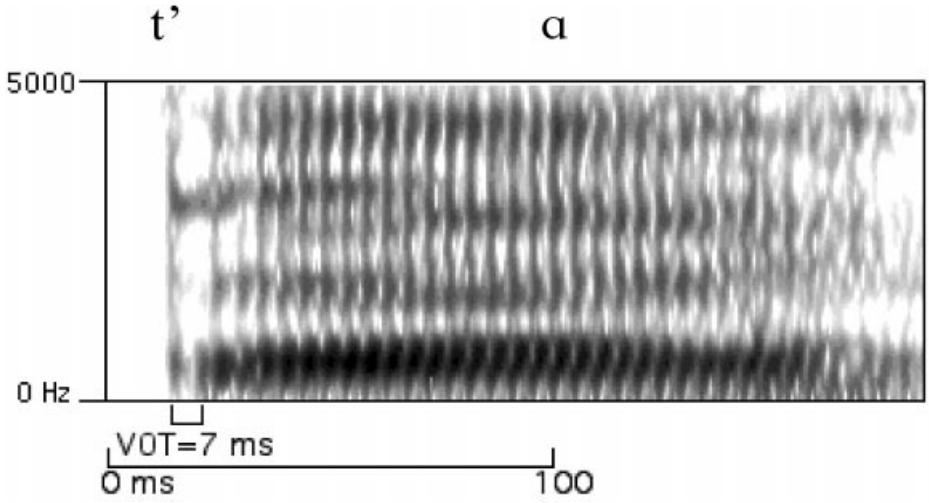


Figure 22 Spectrogram illustrating short VOT with raised f0: speaker RG.

depressed f0. The example is [t'ɛs] 'charcoal', pronounced by speaker MF with a VOT of 41 ms and −79 Hz perturbation. Figure 24 illustrates short VOT combined with depressed f0. The spectrogram is of the word [t'ɑq'ɛt] 'autumn', produced by speaker ET, with 9 ms VOT and −84 Hz f0 perturbation. Figure 25 is a spectrogram illustrating the combination of long VOT with flat f0. The example utterance is [t'ɛs] 'charcoal', pronounced by speaker SM with 108 ms VOT and 0 Hz f0 perturbation. Finally, figure 26 illustrates short VOT combined with flat f0. The spectrogram is of the word [t'ɑq'ɛt] 'autumn', produced by speaker RG. With a VOT of 7 ms and a −1 Hz f0 perturbation, this production of [t'] is essentially not distinguishable from a voiceless unaspirated stop.

Some aspects of the Lindau and Kingston models are supported by our data. Some of the characteristics of Witsuwit'en /t'/ that we have examined here (slow rise time,

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

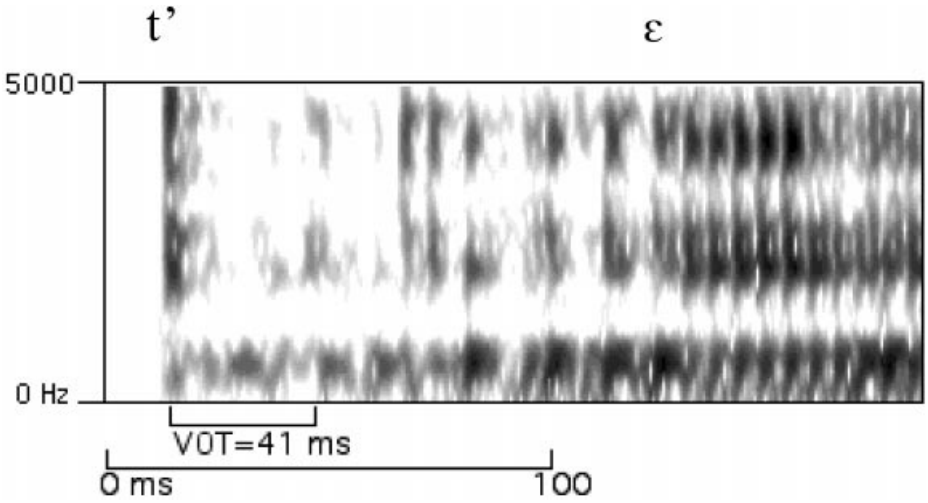


Figure 23 Spectrogram illustrating long VOT with depressed f0: speaker MF.

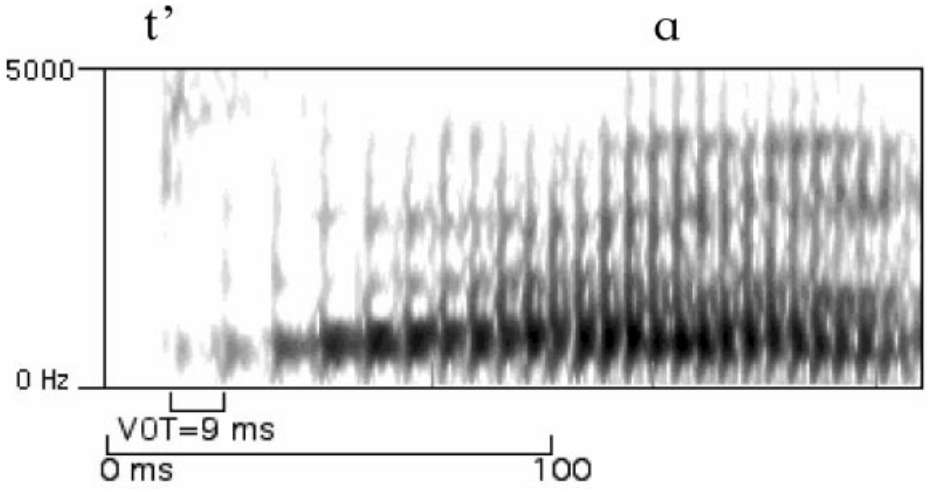


Figure 24 Spectrogram illustrating short VOT with depressed f0: speaker ET.

depressed f0, creaky voice quality) suggest that this sound is a slack ejective. However, the predicted correlation between short VOT and lowered f0 does not hold either on the group or individual levels for our sample of alveolar ejectives. It remains to be seen whether the results obtained here hold for Witsuwit'en ejectives of other places of articulation and in other prosodic contexts.

4 Perception study

4.1 Research questions

Our production study revealed the basis for the fieldworker's impression that the distinction between ejective and voiceless unaspirated consonants can be difficult to

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

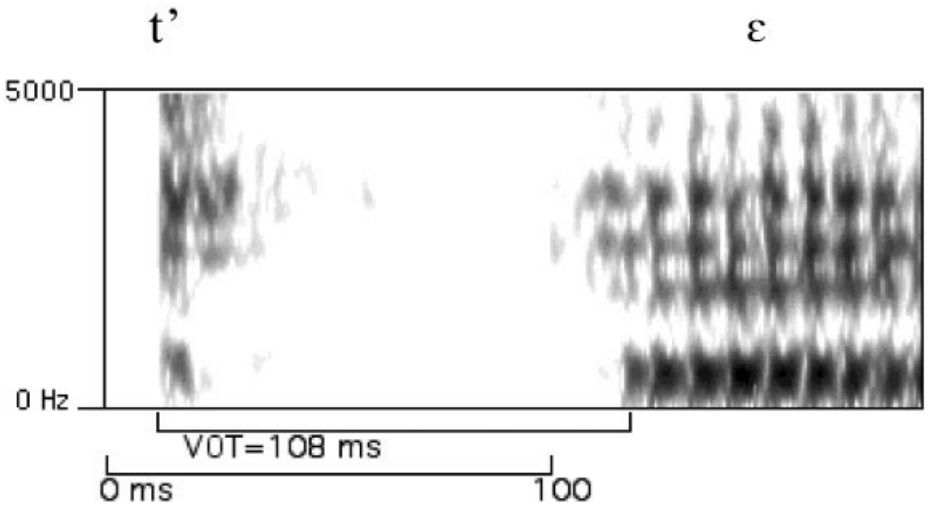


Figure 25 Spectrogram illustrating long VOT with flat f0: speaker SM.

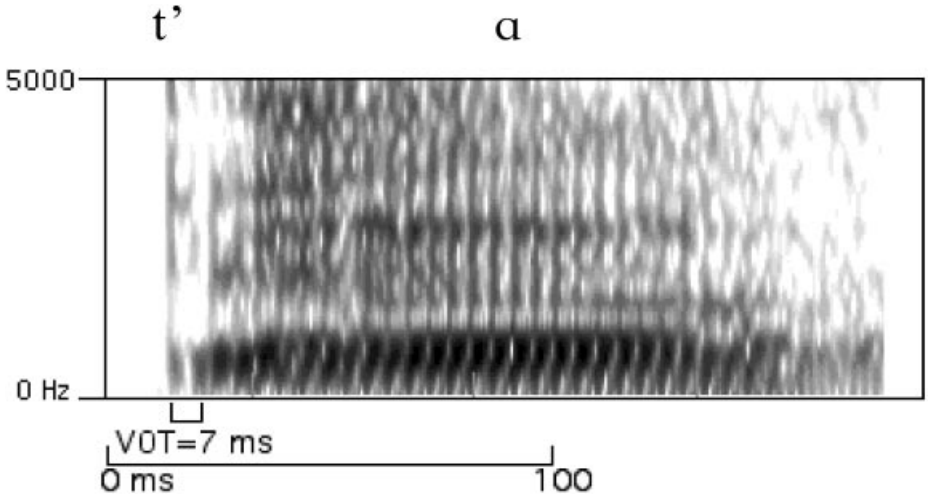


Figure 26 Spectrogram illustrating short VOT with flat f0: speaker RG.

hear: for some speakers, there is little or no difference between the production of ejective and voiceless unaspirated stops. Moreover, when there is a difference, it may be signalled only by a brief period of lowered f0 or slower rise time at vowel onset.

One of the questions resulting from the production study is whether or not the dimensions that we measured sufficiently characterize the ejectives of Witsuwit'en. The fact that we found consistent differences in jitter and rise time across speakers suggests that we have captured at least some of the properties that distinguish Witsuwit'en ejectives from other kinds of stops in this language. To investigate this topic further, we conducted a forced choice identification experiment in which native speakers identified ejective and voiceless unaspirated stops produced by two different speakers, Speaker A, whose ejective/voiceless unaspirated contrast is signalled along all or at least more than one of the measured dimensions, and Speaker B, whose contrast is signalled by few of

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

the measured dimensions. We hypothesized that if the dimensions are good at characterizing the contrast, then the ejective vs. voiceless unaspirated contrast would be more reliably encoded for Speaker A than Speaker B, and listeners would perform well when identifying Speaker A's stops and perform poorly when identifying Speaker B's stops. If, on the other hand, Witsuwit'en speakers are using some unmeasured dimension to signal the ejective/voiceless unaspirated contrast, then listeners would perform equally well when identifying the stops of both speakers.

4.2 Methods

A simple perception study, originally reported on in Davis & Hargus (1998), was designed and carried out in the field. A word list containing eight minimal pairs (sixteen tokens) contrasting ejective and voiceless unaspirated stops or affricates was constructed. A control list of eighteen minimal pairs contrasting other consonantal features was also constructed. (See appendix 2.) Test tokens and control tokens were interspersed in random order.

Two such randomizations were created and recorded with two different speakers, GH (male) and LM (female). The speakers who were chosen for the perception study were located at the extremes in terms of hypothesized intelligibility of the laryngeal contrast: in the production study, the ejectives of GH were significantly different from the voiceless unaspirated stops in rise time only, whereas those of LM differed in all measured dimensions. GH was the first speaker on the tape recording, followed by LM.

Listeners were asked to identify which word they heard by circling one of two words on a forced-choice answer sheet. Words were presented in the local orthography. Each listener heard each word only once.

4.3 Results

A dichotomous inter-subject correlation test was performed on correct/incorrect scores for each token. Scores from two subjects, who had been tested under unavoidably noisy conditions, were excluded from analysis. All ten remaining subjects showed high correlation ($r^2 > .7$).

A pair of tokens (test or control) could receive a maximum correct score of 20. A test pair was considered poorly perceived if at least one member received an identification score at chance levels. Two of GH's test pairs were poorly perceived, whereas all eight of LM's test pairs were well-perceived. The average score for test pairs was 88%, and for control pairs, 98%. (Results for each token are provided in appendix 3.) Tokens produced by both speakers showed similar perceptibility in the controls (GH: 97%, LM: 99%) but among the test pairs GH (85%) was harder to perceive than LM (92%).

In order to investigate acoustic cues to ejective perception, each member of the ejective/non-ejective test pairs was measured for closure duration (non-initial consonants only), VOT, normalized f0, normalized jitter and F1 (first formant) of the following vowel. (A rise time measure comparable to that of the production study could not be obtained because the vowels of the stimulus words were generally too short.) The ratio closure duration/VOT, which Lindau (1984) had found to be significantly different in Navajo vs. Hausa, was calculated. Average values and standard deviations of each speaker's stimuli are given in table 10. The ejective and non-ejective stops and affricates produced by GH were not significantly different by any of these acoustic measures. However, LM's ejective and non-ejective stimuli differed significantly in VOT ($p = .0064$), closure duration/VOT ($p = .02$), f0 perturbation ($p = .0005$) and jitter perturbation ($p = .0163$).

A regression test investigated the hypothesis that the magnitude of the difference between the ejective and non-ejective test pairs on any one of these measures might

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

Table 10 Acoustic values of stimuli (average and standard deviation).

	speaker	F1	VOT	closure duration/ VOT	f0 perturbation	jitter perturbation
ejective	GH	432 (70.5)	24 (16.5)	11.0 (7.96)	−24 (25.5)	−2.3 (2.97)
non-ejective		393 (104.6)	16 (10.5)	9.3 (3.72)	−9 (14.1)	−1.6 (1.54)
ejective	LM	649 (129.9)	50 (30.9)	7.0 (6.52)	33 (30.6)	−12.5 (12.08)
non-ejective		605 (120.6)	12.9 (10.9)	18.7 (6.85)	−22 (16.1)	−.6 (2.04)

Table 11 Correlation (r^2) between % correct and difference between ejective and non-ejective stimulus pairs for five measures.

VOT	f0 perturbation	F1	jitter perturbation	closure duration/VOT
.229	.285	.058	.085	.343

correlate with the percent correct score for the pair. Results are shown in table 11. In general, the correlations between score and magnitude of the phonetic difference are lower than normally seen in correlation studies. However, it is important to keep in mind that each measure represents a subset of the cues that listeners might use to identify the laryngeal setting.

The f0 perturbation and closure duration/VOT ratio differences had the highest correlations with percent correct. Note that f0 perturbation difference and jitter perturbation difference do not have the same correlational value in table 11: jitter perturbation difference, in contrast to f0 perturbation difference, exhibited the second-worst correlation with percent correct. This is perhaps not surprising since in the production study f0 perturbation and jitter did not always co-occur. (Significant f0 perturbation differences between ejectives and voiceless unaspirates were found for five speakers, but of these five, only two also had significantly greater jitter perturbation following ejectives; see table 9.)

Closure duration/VOT difference and VOT difference did not have the same correlational value either. Closure duration/VOT ratio was noted by Lindau (1984) to be a good indicator of inter-language (Navajo vs. Hausa) variation. In the case of Witsuwit'en, closure duration/VOT ratio may also be a good indicator of inter-speaker differences. If this is the case, closure duration/VOT difference may reflect the well-known ability of listeners to calibrate to specific talkers (e.g. Ladefoged & Broadbent 1957). Although VOT difference was among the better correlates of percent correct, particularly for LM (it was also significant for her in the production study), VOT difference was not as good a correlate of percent correct as closure duration/VOT difference. We suggest that VOT alone may be only a subtle, even if statistically reliable, measure even for LM. Closure duration/VOT ratio may be a better cue to the laryngeal setting of ejectives than VOT because it enhances VOT. However, one must be careful in reaching any conclusions comparing VOT and closure duration/VOT ratio since the latter includes the former (i.e. a part-whole comparison). It may be that other composite measures would also show a higher correlation than their individual components.

Finally, we note that the correlation between F1 difference and percent correct was smallest. We believe that the low correlation is the result of two factors: 1) the stimuli

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

were selected to minimize segmental differences between pairs other than the laryngeal contrast, and 2) the allophonic differences in vowel quality may be obscured by certain final consonants, such as uvulars, as are found in many of the minimal pair stimuli.

4.4 Discussion

Our initial hypothesis was that the contrast between ejective and voiceless unaspirated stops and affricates would be more reliably encoded for speakers whose ejective and non-ejective consonants differed along more than one dimension. This hypothesis was borne out in that LM's laryngeal contrast was identified at a higher rate (92%) than GH's (85%). As summarized above in table 10, LM's stimulus pairs exhibited a larger phonetic difference along all of the measures. Moreover, all of the measures except F1 were significantly different for LM, whereas none of the measures were significantly different for GH. Alternatively, it may be that listeners had an easier time with LM because the presentation of stimuli was blocked by speaker (GH before LM) so that listeners might have become more habituated to the task. However, it seems more likely that the order of presentation did not affect intelligibility scores because the control words produced by both speakers were identified at nearly identical rates. Therefore, any differences between speakers can be attributed to differences in their production of the ejective consonant stimuli.

Another interesting finding is that for both speakers the ejective-voiceless unaspirated pairs had lower identification scores than the control stimuli. The poorer identification scores indicate that even for the more intelligible of the two speakers, the ejective-voiceless unaspirated contrast is more difficult to perceive than the other phonetic contrasts in Witsuwit'en.

The main goal of the perception study was to determine whether or not the phonetic dimensions selected for measurement in the production study – VOT, f0 perturbation, jitter perturbation, and rise time – adequately characterize the laryngeal contrast in Witsuwit'en. All of the correlations between percent correct and difference along the selected measures (all but rise time, for reasons discussed above) were positive, as predicted. However, the relatively weak correlations indicate that this conclusion must be considered tentative. It may be that no single measure is an exclusive cue to the laryngeal contrast. That is, the cues may have an additive or even multiplicative interaction.

5 Conclusion

5.1 Witsuwit'en and Athabaskan ejectives

Results of our eleven-speaker production study showed considerable inter-speaker variation in the production of alveolar ejective stops. While averaging the means across speakers might suggest that the typical alveolar ejective in Witsuwit'en has (relative to voiceless unaspirates) intermediate VOT, depressed f0, more jitter, and slow rise time at vowel onset, no one speaker exactly matched the group profile. The notion 'average ejective stop' is thus problematic for Witsuwit'en.

The perception study showed that the ejective-voiceless unaspirated contrast is more difficult than other consonantal contrasts for even native listeners. The perception study also indicated that ejectives which differ in multiple dimensions are easier to perceive than those which differ in only one dimension. In light of this finding, a future study which manipulates cues independently of each other in a fully crossed design might shed light on the relative weighting of the cues and their interactions. Such a

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

Table 12 Characteristics of ejectives in Athabaskan languages.

	source	# speakers	VOT	f0	voice quality	rise time
Navajo	L84	9	long		modal	fast
	ML93	7	long		modal	fast
	LM96				some creaky	
Hupa	G95	3	long			
Chipewyan	H76	1	long	30% low	30% creaky	fast
L. Tanana	T97, T98	3	variable but mostly long		some creaky	
Witsuwit'en		11	intermediate	depressed	creaky	slow

study would rely on synthetic rather than natural speech stimuli so that the experimenter would have independent control over each of the dimensions. Ideally, the experiment would include discrimination and identification tasks and reaction times would be collected. It would also be valuable to observe cue interactions in a variety of word-position and noise conditions since cues are known to interact with these factors (Wright 2001).

The evidence from other instrumental studies of Athabaskan languages is summarized in table 12. (The sources in table 12 are abbreviations of the following references: L84 = Lindau 1984, ML93 = McDonough & Ladefoged 1993, LM 96 = Ladefoged & Maddieson 1996, G95 = Gordon 1995, H76 = Hogan 1976, T97 = Tuttle 1997, T98 = Tuttle 1998.) The comparative evidence indicates that Athabaskan ejectives are usually realized with long VOT, modal voice and fast rise time, to the extent that these features are reported. But even within the one-speaker study of Chipewyan, variable voice quality has been noted, as it has in Navajo, Lower Tanana and Witsuwit'en. With new data from Witsuwit'en, we see that there is also variability in VOT and rise time in the pronunciation of ejectives in the languages of the family. Despite possibly original variability in production, the contrast between voiceless unaspirated, voiceless aspirated and ejective stops and affricates has been historically stable in the estimated 2000–2500 year time-depth of the Athabaskan family: no language has merged all of the voiceless unaspirated stops and affricates with the ejectives, although some languages of the Pacific Coast Athabaskan subgroup have merged voiceless aspirates and unaspirates of certain places of articulation (Krauss 1973, 1979).

The variable but historically stable nature of this contrast may pose an interesting problem for theories of sound change which predict that contrasts which are variable and confusable tend to be lost over time (e.g. Ohala 1981). We have seen that in Witsuwit'en, the ejective/voiceless unaspirated contrast is more difficult to perceive than other phonetic contrasts. An unreliable perceptual distinction may eventually be lost, but doesn't neccesarily disappear very quickly – consider the perceptually difficult contrast between [v] and [ð] in English that has nevertheless endured for centuries. Being difficult to hear/easy to mistake is only one of the factors at work in loss of contrast.

5.2 Implications for ejective typology

In our study of the production of /t'/, we found negligible correlation between f0 depression and VOT, contrary to the predictions of the models proposed by Lindau (1984) and Kingston (1985, to appear). More generally, we found the typology to be inadequate to predict the range of variation in Witsuwit'en ejective production.

The typology is based on widely held assumptions about vocal fold tension and compression that may underestimate the complexities of laryngeal muscular adjust-

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

ment, aerostatic-myoelastic interactions, and timing relationships involved in the production of ejectives. Stevens 1998, a proponent of the stiff-slack dichotomy in consonantal features, notes that changes in f0 are achieved primarily through manipulations of STIFFNESS of the vocal folds and secondarily through manipulations of the MASS of the vocal folds. Increased stiffness is achieved mainly by stretching of the vocal cords. In pitch adjustments the vocal folds are stretched via longitudinal tension primarily through contraction of the cricothyroid muscle (CT). Another intrinsic muscle implicated in increased longitudinal tension with adduction is the vocalis (V), which makes up part of the mass of the vocal folds. Activity of extrinsic muscles such as the thyroid, the stylohyoid, and the anterior and posterior bellies of the digastric (whose contractions may raise the larynx thereby increasing longitudinal tension) may also contribute to vocal fold stiffness. However, vocal fold stiffness alone is insufficient to cause a cessation of voicing or the tight closure necessary for the pressure build-up seen in ejectives. Stiffness must be accompanied by an increase in medial compression that is greater than that seen in normal phonation. Contraction of the interarytenoid (IA), lateral cricoarytenoid (LCA), and lateralis (L) cause adduction of vocal cords and apply medial compression. The lateral cricoarytenoid is also implicated in control of f0 and is activated in the onset of phonation. It is worth noting that muscular activity is applied in concert and that in the absence of tight coordination with other muscles, an adductor muscle such as the interarytenoid might actually pull the vocal folds apart (Zemlin 1998).

The complex situation suggests a variety of possible articulatory mechanisms for the combinations that we have observed in our data. Consider, for example, long VOT, low f0, and creaky voicing. It may be that the supression of voicing was achieved exactly as predicted in the Lindau and Kingston models: through a combination of stiff vocal folds (longitudinal tension) and medial compression. But the timing of the various forces at work decrease longitudinal tension (CT, V) prior to relaxation of medial compression (IA, LCA, L) resulting in shortened, thickened, vocal folds with a tight posterior closure but a relatively slack tension. This could result in an ejective with depressed f0 and creaky voice quality despite stiff vocal folds during the compression stage. Alternatively, and contra the assumptions in Lindau's and Kingston's models, it may be that strong medial compression due to the constriction of the adductor muscles (IA, LCA, L) in the absence of longitudinal tension is sufficient to inhibit the onset of voicing, resulting in a long VOT with accompanying low f0 and creaky voicing. Finally, differing degrees of compression, larynx raising, or differences in timing between larynx raising and oral release could also produce some of the variability seen in our data.

Kingston (to appear) discusses the factors at work in Athabaskan tonogenesis, proposing that the phonetics of Athabaskan ejectives can be understood in part through instrumental studies of other consonants with complex laryngeal configurations such as the Korean fortis consonants, Danish *stød*, and glottal stops. The differences between fortis consonants, *stød*, glottal stops, and ejectives are achieved through subtle variations in one or a combination of articulatory dimensions: longitudinal tension, medial compression, pharyngeal tension, and larynx raising. This point of view is intriguing given the range of variability seen in our study of Witsuwit'en ejectives. Our data, combined with cross-linguistic data, point out a problem for the study and classification of ejectives: they may fall along a continuum in at least three dimensions, the endpoints of which form other speech sounds. Figure 27 illustrates the possible endpoints of the three dimensions (longitudinal tension, medial compression, and larynx raising) together with some possible variants on ejectives seen in our Witsuwit'en sample.

To produce a classic stiff ejective (the filled triangle), with raised f0, long VOT, sharp rise time, and tense to modal voicing, speakers use extreme larynx-raising and

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

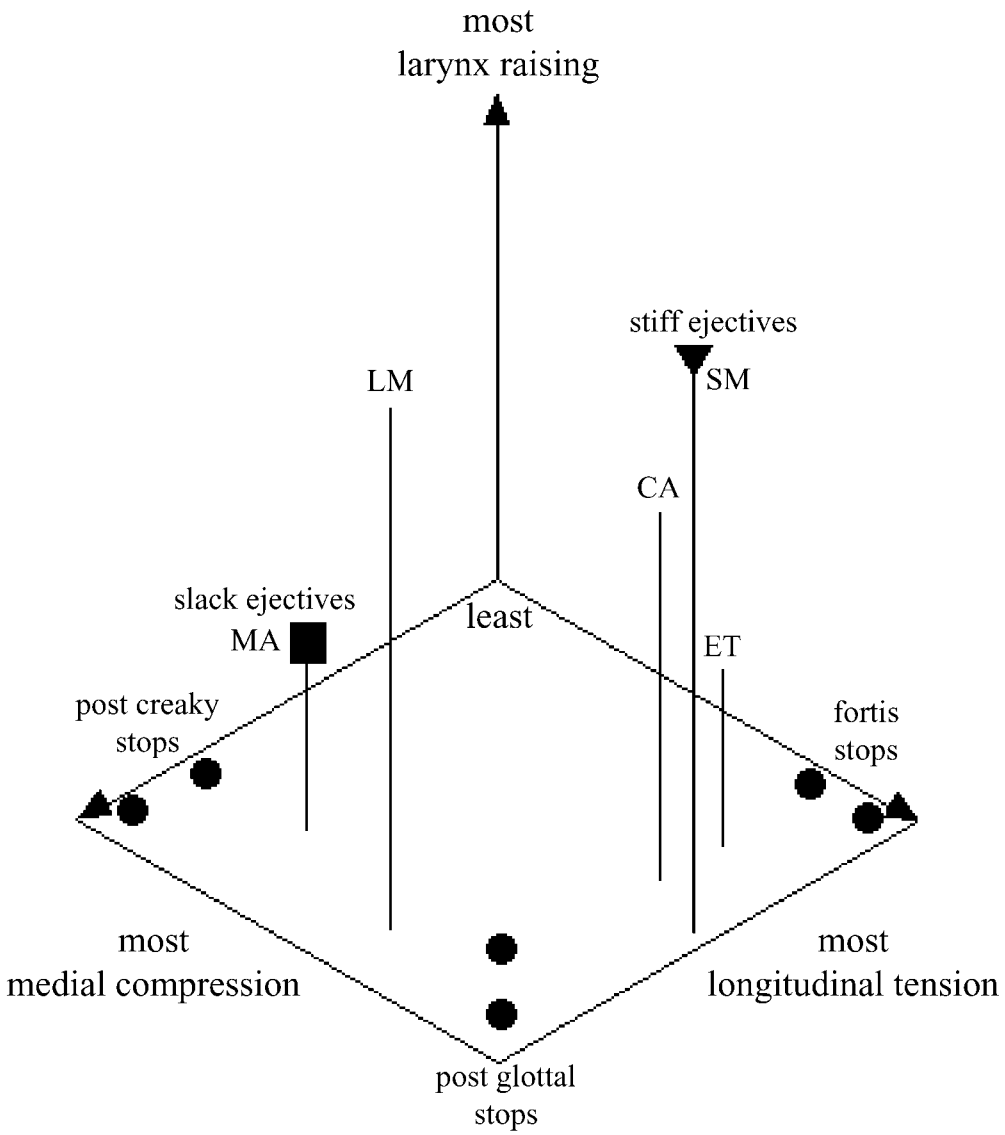


Figure 27 Some of the possible ejective variants given continuous adjustments along three dimensions: larynx raising, medial compression, and longitudinal tension. Examples from the current study are marked with speaker initials. The stiff ejective is marked with a triangle, and the slack ejective is marked with a square.

longitudinal-tension with moderate medial compression. On the other hand, to produce a slack ejective (filled square), with depressed f0, short VOT, slow rise time, and creaky voicing, speakers use extreme medial compression but little longitudinal tension or larynx raising. Five example variants from our data (means for speakers), which are in between these extremes, are also plotted with speaker initials: 1) long VOT and raised f0: SM, 2) intermediate VOT and raised f0: CA, 3) short VOT and raised f0: ET, 4) long VOT and depressed f0: LM, 5) short VOT and depressed f0: MA.

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

Acknowledgements

Funding for this study was provided by NSF grant #OPP-9307704 to Sharon Hargus. We are deeply grateful to the Witsuwit'en speakers who participated in both studies. We note with sadness that three speakers, CA, ET and PN, have passed away since the data was first collected.

The comments of reviewers Siri Tuttle, Joyce McDonough, John Kingston, and editor Peter Ladefoged have greatly improved the article. Alicia Beckford Wassink and Ellen Kaisse also provided useful comments and advice. We are thankful to everyone for their help.

Appendices

Appendix 1: Production study word list

Table 13 Voiceless unaspirates.

/t̲ete/	[t̲ete]	'sickness'
/t̲æji/	[t̲æji]	'chief'
/t̲əz/	[t̲əz]	'driftwood'
/t̲oso/	[t̲osɔ]	'gunny sack'

Table 14 Voiceless aspirates.

/t̲^{h}iɬ/	[t̲^{h}eɬ]	'basket'
/t̲^{h}ɑni/	[t̲^{h}ɑni]	'3 people'
/t̲^{h}əz/	[t̲^{h}ʌz]	'cane'
/t̲^{h}o c^{h}o/	[t̲^{h}ɔc^{h}o]	'big body of water'

Table 15 Ejectives.

/t̲'es/	[t̲'ɛs]	'charcoal'
/t̲'ɑq'et/	[t̲'ɑq'ɛt]	'autumn'
/t̲'ɑc^{h}c/	[t̲'ɑc^{h}o]	'mallard'
-/t̲'o c^{h}o/	[t̲'ɔc^{h}o]	'large nest'

Appendix 2: Perception study word list

Table 16 Test pairs: [t] vs. [t'].

/wejitəs/	[wejitəs]	'he/she didn't twist it'
/wejit'əs/	[wejit'ʌs]	'he/she didn't cut it'
/ntəq/	[ntʌq]	'straight up'
/nt'əq/	[nt'ʌq]	'your collarbone'
/pənin wenitək^{w}/	[pənin wenitək^{w}]	'his/her face isn't round'
/pənin wenit'ək^{w}/	[pənin wenit'ʌk^{w}]	'his/her face isn't short'

Table 17 Test pairs: [tɬ] vs. [tɬ'].

/sətɬet/	[sətɬet]	'it's licking me'
/sətɬ'it/	[sətɬ'et]	'he/she farted'

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

Table 18 Test pairs: [ts] vs. [ts'].

/tsəχ/	[tsʌχ]	'wrong'
/ts'əχ/	[ts'ʌχ]	'hat'
/pətsəq/	[pətsʌq]	'his/her outer ear'
/pəts'əq/	[pəts'ʌq]	'his/her little finger'

Table 19 Test pairs: [q] vs. [q'].

/qəχ/	[qʌχ]	'rabbit'
/q'əχ/	[q'ʌχ]	'backwards'
/c'əqəz/	[c'əqəz]	'Broman Lk. area'
/c'əq'əz/	[c'əq'ʌz]	'half, side, .50'

Table 20 Control pairs: aspirated vs. voiceless unaspirated stops.

/cəs/	[cəs]	'hook'
/c^{h}əs/	[c^{h}ʌs]	'down feathers'
/cen/	[cen]	'here'
/c^{h}en/	[c^{h}ɛn]	'summer'
/təz/	[təz]	'driftwood'
/t^{h}əz/	[t^{h}ʌz]	'cane'

Table 21 Control pairs: voiced vs. voiceless fricatives.

/t^{h}ezt$\downarrow$i/	/t^{h}ɛzt$\downarrow$i/	'outlet of body of water'
/t^{h}est$\downarrow$i/	/t^{h}ɛst$\downarrow$i/	'steelhead'

Table 22 Control pairs: final consonant vs. no final consonant.

/soχ/	[sɔχ]	'robin'
/so/	[sɔ]	'frost'

Table 23 Control pairs: stop/affricate vs. fricative.

/q^{h}e/	[q^{h}ɛ]	'footwear'
/χe/	[χɛ]	'grease'
/ts^{h}e/	[ts^{h}ɛ]	'rock'
/se/	[sɛ]	'belt'

Table 24 Control pairs: place of articulation.

/pəq^{h}ɑq/	[pəq^{h}ɑq]	'his/her surface'
/pəc^{h}ɑq/	[pəc^{h}ɑq]	'his/her rib'
/pəjih/	[pəjih]	'his/her throat'
/pəʁih/	[pəʁih]	'its roots'

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

Appendix 3: Perception study results

Table 25 Results for each test token.

speaker	GH		LM	
token	#correct	pair score	#correct	pair score
/wejit'əs/	9		10	
/wejitəs/	10	19	10	20
/nt'əq/	9		8	
/ntəq/	9	18	10	18
/wenit'əkʷ/	7		10	
/wenitəkʷ/	10	17	10	20
/sətl'it/	10		9	
/sətlet/	10	20	9	18
/ts'əχ/	10		10	
/tsəχ/	7	17	9	19
/pəts'əq/	9		9	
/pətsəq/	8	17	8	17
/q'əχ/	5		8	
/qəχ/	10	15	10	18
/c'əq'əz/	8		9	
/c'əqəz/	4	12	8	17

References

Aoki, Haruo (1970). A note on glottalized consonants. *Phonetica* **21**, 65-74.

Davis, Katharine & Hargus, Sharon (1994). Ejectives in Babine-Witsuwit'en. *Journal of the Acoustical Society of America* **96** (5, pt. 2), 3230.

Davis, Katharine & Hargus, Sharon (1996). The glottal contrast in Babine-Witsuwit'en. *Journal of the Acoustical Society of America* **100** (4, pt. 2), 2690.

Davis, Katharine & Hargus, Sharon (1998). Perception of glottalized consonants in Babine-Witsuwit'en. *Journal of the Acoustical Society of America* **103** (5, pt. 2), 3087.

Denning, Keith (1989). *The Diachronic Development of Phonological Voice Quality, With Special Reference to Dinka and the other Nilotic Languages.* Ph.D. dissertation, Department of Linguistics, Stanford University.

Flemming, Edward, Ladefoged, Peter & Thomason, Sarah (1994). Phonetic structures of Montana Salish. *UCLA Working Papers in Phonetics* **87** (Fieldwork Studies of Targeted Languages II), 1–33.

Fre Woldu, Kiros (1984–1986). Evidence of auditory similarity between Tigrinya ejective /t/ and Arabic emphatic /t/. *Orientalia Suecana* **23–25**, 123–138.

Gordon, Matthew (1995). The phonetic structures of Hupa. *UCLA Working Papers in Phonetics* **93** (Fieldwork Studies of Targeted Languages IV), 1–24.

Grossblatt, Ben (1997). *Sahaptin Ejection and the Fortis/Lenis Model.* M.A. thesis, Department of Linguistics, University of Washington, Seattle.

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

Hargus, Sharon (1996). The first person plural subject prefix in Witsuwit'en. Ms., Department of Linguistics, University of Washington, Seattle.
Hargus, Sharon (1998). An acoustic study of Witsuwit'en vowel quality. Paper read at Athapaskan Languages Conference, University of Calgary, Alberta.
Hogan, John T. (1976). An analysis of the temporal features of ejective consonants. *Phonetica* **33**, 275–284.
Ingram, John & Rigsby, Bruce (1987). Glottalic stops in Gitksan: an acoustic analysis. *Proceedings of the XIth International Congress of Phonetic Sciences* **2**, 134–137. Tallin, Estonia: Academy of Sciences of the Estonian SSR.
Kari, James (1975). Babine, a new linguistic grouping. Ms., Alaska Native Language Center, Fairbanks.
Kari, James (1990). *Ahtna Athabaskan Dictionary*. Fairbanks: Alaska Native Language Center.
Kingston, John (1982). Why tones don't come from ejectives. Ms., University of California, Berkeley.
Kingston, John (1985). *The Phonetics and Phonology of the Timing of Oral and Glottal Events*. Ph.D. dissertation, Department of Linguistics, University of California, Berkeley.
Kingston, John (to appear). The phonetics of Athabaskan tonogenesis. In Hargus, Sharon & Rice, Keren (eds.), *Athabaskan Prosody*. Amsterdam: John Benjamins.
Koike, Yasuo (1973). Application of some acoustic measures for the evaluation of laryngeal dysfunction. *Studia Phonologica* **7**, 17–23.
Krauss, Michael (1964). Proto-Athapaskan-Eyak and the problem of Na-Dene I: the phonology. *International Journal of American Linguistics* **30**, 118–131.
Krauss, Michael (1973). Na-Dene. In Sebeok, Thomas A. (ed.), *Linguistics in North America*, 903–978. The Hague & Paris: Mouton.
Krauss, Michael (1979). Na-Dene and Eskimo-Aleut. In Campbell, Lyle & Mithun, Marianne (eds.), *The Languages of Native America: An Historical and Comparative Assessment*, 803–901. Austin, TX & London: University of Texas Press.
Ladefoged, Peter & Broadbent, David E. (1957). Information conveyed by vowels. *Journal of the Acoustical Society of America* **29** (1), 98–104.
Ladefoged, Peter & Maddieson, Ian (1996). *Sounds of the World's Languages*. Oxford: Blackwell.
Ladefoged, Peter, Maddieson, Ian & Jackson, Michel (1988). Investigating phonation types in different languages. In Fujimura, Osamu (ed.), *Vocal Physiology: Voice Production, Mechanisms and Functions*, 297–317. New York: Raven Press.
Leer, Jeff (1999). Tonogenesis in Athabaskan. In Kaji, Shigeki (ed.), *Cross-Linguistic Studies of Tonal Phenomena: Tonogenesis, Typology, and Related Topics*, 37–66. Tokyo: Institute for the Study of Languages and Cultures of Africa and Asia, Tokyo University of Foreign Studies.
Lindau, Mona (1984). Phonetic differences in glottalic consonants. *Journal of Phonetics* **12**, 147–155.
Lindsey, Geoff, Hayward, Katrina & Haruna, Andrew (1992). Hausa glottalic consonants: a laryngographic study. *Bulletin of the School of Oriental and African Studies, University of London* **55**, 511–527.
Lisker, Leigh & Abramson, Arthur (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word* **20**, 384–422.
McDonough, Joyce & Ladefoged, Peter (1993). Navajo stops. *UCLA Working Papers in Phonetics* **87** (Fieldwork Studies of Targeted Languages), 151–164.
Morice, Adrian G. (1932). *The Carrier Language: A Grammar and Dictionary Combined*. St. Gabriel-Mödling bei Wien, St. Gabriel, Austria: Anthropos.
Ohala, John (1981). The listener as a source of sound change. In Masek, Carrie S., Hendrick, Roberta A. & Miller, Mary Frances (eds.), *Chicago Linguistic Society 17: Papers from the Parasession on Language and Behavior*, 178–203. Chicago: Chicago Linguistic Society.
Sapir, Edward (1915). The Na-Dene languages: a preliminary report. *American Anthropologist* **17**, 534–558.
Stevens, Kenneth N. (1998). *Acoustic Phonetics*. Cambridge, MA: MIT Press.

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press

Story, Gillian (1984). *Babine and Carrier Phonology: A Historically Oriented Study*. Arlington: Summer Institute of Linguistics.

Trigo, Loren (1991). On pharynx-larynx interactions. *Phonology* **8**, 113–136.

Tuttle, Siri G. (1997). Acoustic realizations of laryngeal contrasts in Minto Athabaskan obstruents. Paper read at Athabaskan Languages Conference, University of Oregon, Eugene, OR.

Tuttle, Siri G. (1998). Ejectives in Tanana Athabaskan. *Journal of the Acoustical Society of America* **103** (5, pt. 2), 3086.

Warner, Natasha (1996). Acoustic characteristics of ejectives in Ingush. In *Proceedings of the International Conference on Spoken Language Processing, October 3–6, 1996, Philadelphia*, 1525–1528. New York: Institute of Electrical and Electronics Engineers.

Wright, Richard (2001). Perceptual cues in contrast maintenance. In Hume, Elizabeth & Johnson, Keith (eds.), *The Role of Speech Perception Phenomena in Phonology*, 251–277. San Diego, CA: Academic Press.

Zemlin, Willard R. (1998). *Speech and Hearing Science: Anatomy and Physiology*. Boston, MA: Allyn and Bacon.

https://doi.org/10.1017/S0025100302000142 Published online by Cambridge University Press