

ACOUSTIC CHARACTERISTICS OF EJECTIVES IN INGUSH

Natasha Warner

Department of Linguistics
University of California, Berkeley

ABSTRACT

The purpose of this paper is to investigate acoustic characteristics which distinguish ejectives from pulmonic stops in Ingush (a Northeast Caucasian language), and to compare Ingush ejectives to those of other languages. The articulation of ejectives is relatively well understood, but their acoustic effects are less clear. Working with a native speaker of Ingush, ejectives and pulmonic voiceless stops were compared for VOT, closure duration, and post-burst power; and pitch, amplitude, and voice quality of the following vowel. Ingush results are compared to published descriptions of ejectives in several other languages. Ingush ejectives do not have all the same acoustic features as any other language studied. The characteristics of ejectives vary with each language, and do not pattern together to form just two types of ejectives, as has been claimed based on binary comparisons.

1. INTRODUCTION

Ejectives, which occur in approximately 16% of the world's languages, are very different from the more common pulmonic stops in their articulation: the speaker makes an oral closure while simultaneously closing the glottis, then raises the larynx. This action causes an unusually high oral air pressure. The oral closure is then released, and the air rushes out of the mouth. At some point after the release of the oral closure, the glottal closure is also released. However, the acoustic cues for ejectives are less often discussed. Larynx raising and high oral pressure, the articulatorily distinctive aspects of an ejective, cannot be heard, but ejectives are auditorily quite distinct from other types of stops, so what is it that listeners hear in them?

One would expect that the burst of an ejective would have different characteristics from that of a pulmonic voiceless stop, because the glottis is closed during the burst of an ejective, and the air in the mouth is under higher than usual pressure. After the burst, an ejective should have almost no noise, since the glottis is still closed, preventing flow of air from the lungs. The action of the vocal cords (phonation type) at the beginning of the vowel after an ejective should be different from usual, since the vocal cords have been held tightly together, whereas they are apart during a pulmonic stop. Because we know that vocal cord action is likely to be different, we might also predict a difference in the pitch of the vowel following the two stop types.

Past work on ejectives [1, 2, 3, 4, 6, 7] has identified several possible acoustic characteristics of ejectives and established that these features are not uniform across different languages. However, detailed acoustic phonetic descriptions of ejectives exist for only a few languages (such as Hausa, Navajo, Tigrinya, Quiché, and Montana Salish), and most cross-linguistic comparisons have been of just two languages. It is important at this stage to examine ejectives from as many languages as possible in order to see how the acoustic

characteristics pattern in different languages. Ingush contributes to this pool of data, and shows that the acoustic characteristics do not group together to form simply two kinds of ejectives, as has sometimes been concluded in the past based on binary comparisons.

2. DATA

	bilabial	alveolar	velar	uvular
voiceless	p	t	k	q
voiced	b	d	g	--
ejective	(p')	t'	k'	q'

Table 1: Ingush stop system. /p'/ is rare. There is no voiced uvular stop.

Working with a single native speaker of Ingush (who is bilingual in Russian, but Ingush dominant), a list of Ingush words was compiled including every stop except /p'/ in word-initial position before each of three vowel qualities, high front ([i] or [j]), low or mid-central ([a] or [ə]), and back high or mid ([o], [ɔ], or [w]). For /p'/, only three words could be found, so all three were included in the list, although all have a high or mid front vowel following /p'/. Ejectives occur almost exclusively in word-initial position in Ingush, so intervocalic stops were obtained by placing the word [si] 'my', before each word. Because Ingush has a large number of segments and many possible clusters, true minimal pairs are rare, so only the quality of the vowels surrounding the ejective was controlled.

In a recording booth, the same speaker was recorded producing six repetitions of each word (three in isolation and three in the intervocalic frame), each time in a different random order. He was prompted with words written in Cyrillic orthography on index cards. The data was digitized at 48,000 Hz and analyzed using ESPS/Waves.

3. ANALYSES PERFORMED

Since a large quantity of data was recorded, it is not necessary to use all of it for most analyses. Therefore, a selection of 38 tokens (19 pulmonic voiceless, 19 ejective) was used, which includes one token for each pulmonic voiceless or ejective stop before each vowel quality, in both initial and medial environments (less one token for which data was missing).

3.1. Measurements within the Stops

Duration. The average VOT of ejectives in this sample is 26.2 msec., for pulmonic voiceless stops it is 45.1 msec. This difference is statistically significant ($F=17.65$, $F(1,30)=4.17$ at $\alpha=.05$), but a difference in VOT of 19 milliseconds may not be enough to serve as a major perceptual cue. As is usual cross-linguistically, VOT is longer for velars than for bilabials and alveolars, and longer before high vowels than before low

vowels. The voiced stops are fully voiced, and are clearly distinguished from the other two stop types by negative VOT. Therefore, I compared only ejectives and pulmonic voiceless stops in the following analyses. Closure duration (for the intervocalic tokens) shows almost no difference: the ejective average is 56.8 msec., and that for pulmonic voiceless stops is 54.4 msec. This difference is not significant ($F=0.10$ ($F(1,16)=4.49$ at $\alpha=.05$).

Power. To test for the predicted lack of aspiration noise (due to closed glottis) after the burst in ejectives, the average power during the period from the end of the burst to just before the onset of voicing was calculated. The average values were 162,321 for ejectives and 385,193 for pulmonic voiceless stops. This difference is significant: $F=9.83$ ($F(1,30)=4.17$ at $\alpha=.05$).

Ingush ejectives often sound as if they have very weak bursts, but if ejectives have high oral pressure, one would predict them to have strong bursts, so the power at the peak of the burst was also measured. In cases of double bursts (velars), the louder burst was used. Ejectives had an average peak burst power of 1,136,175, while voiceless pulmonic stops had 1,459,551. Although the ejectives did have weaker bursts, this difference was not significant, with $F=0.94$ ($F(1,30)=4.17$ at $\alpha=.05$). The ranges of the values overlap almost completely. Any difference in burst power appears to be inconsistent and small.

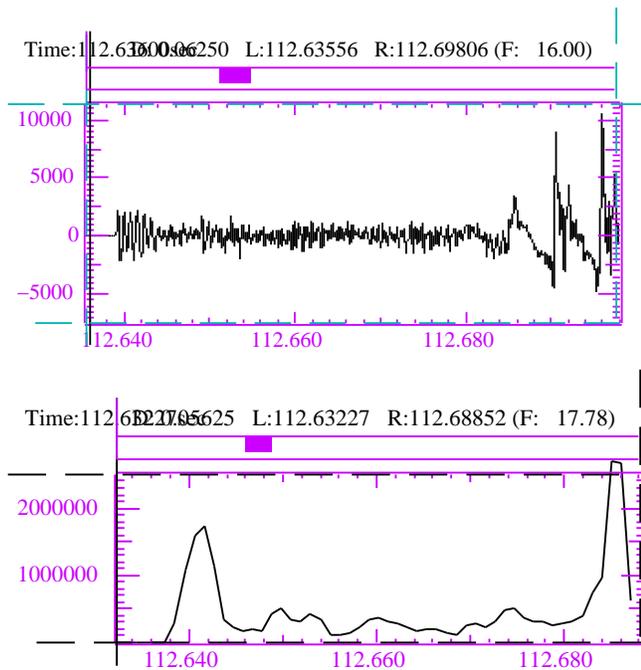


Fig. 1a: Waveform and power for a pulmonic voiceless stop showing aspiration noise and elevated power after the burst. Token is [qat] 'food.'

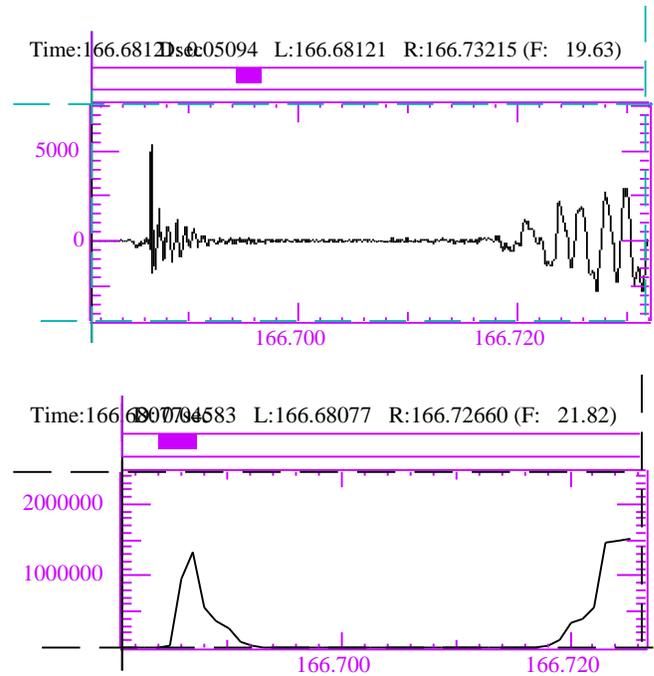


Fig. 1b: Corresponding ejective in [q'əmirk] 'throat' with almost no noise and very low power after the burst.

Spectral characteristics of bursts. Characteristics of ejective and pulmonic voiceless bursts were also preliminarily investigated through examination of spectrograms and both FFT and LPC spectra. Ejectives seem to make little or no difference in the frequency of the prominences in the burst, and no consistent patterns could be found.

3.2. The Following Vowel

Pitch. It is very likely that some acoustic cues for stop type appear in the vowel following the stop, not during the stop itself. Ingush stops have a strong influence on the pitch of the following vowel: pitch is audibly higher after ejectives than after pulmonic voiceless stops. Average pitch at the onset of voicing was 221.4 Hz after ejectives, but only 195.2 Hz after voiceless stops. This difference is both significant ($F=30.43$, ($F(1,36)=4.17$ at $\alpha=.05$) and highly consistent. Fig. 2 shows representative pitch tracks.

To confirm the validity of this finding, I also measured pitch at a point 5-7 periods into the vowel. At that point, the post-ejective vowels still have higher pitch (average 217.4 Hz, pulmonic average 200.2 Hz), but the difference is not as great, showing that the post-ejective pitch is starting to return toward normal. This difference is also significant, at $F=19.65$ ($F(1,36)=4.17$ at $\alpha=.05$).

Thus, pitch is higher after ejectives, both immediately at the onset of voicing and several periods into the vowel. Pitch tends to fall after ejectives, reflecting a return to normal after the elevated pitch, while for pulmonic voiceless stops pitch rises toward the middle of the vowel.

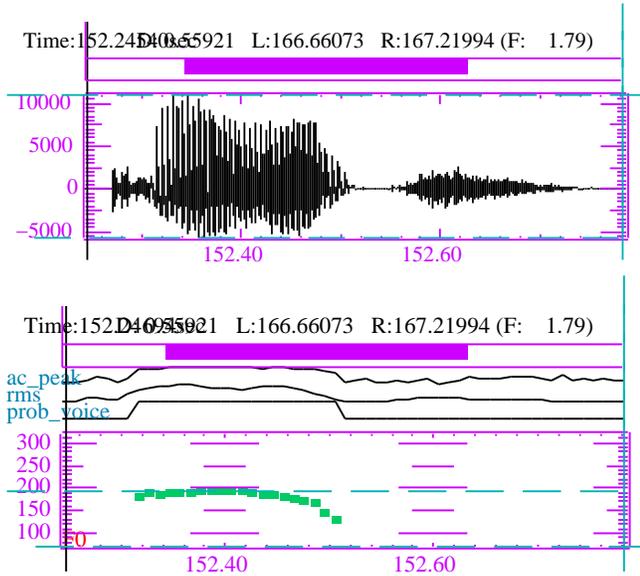


Fig. 2a: Pitch track of [qat] 'food.' Initial pitch of 182 Hz, followed by a rise.

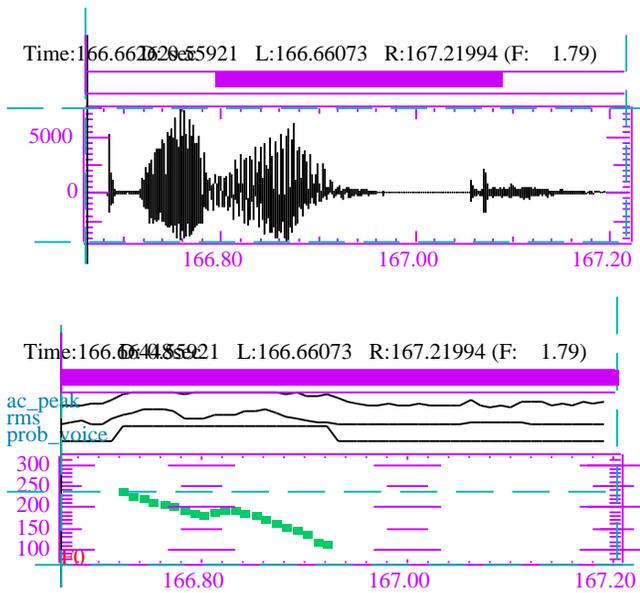


Fig. 2b: Pitch track of [q'əmirkʰ] 'throat.' Initial pitch of 240 Hz, followed by a fall.

Quality of voicing. Ejectives frequently influence the quality of voicing in the first several periods of the following vowel in one or more ways. One such effect is a predominance of F2 and F3 over F1 (effect on spectral tilt), shown in Fig. 3. Detailed measurement of the waveform confirms that the high frequency peaks within each period represent component waves at the frequency of the second and third formants, as judged by comparison with the spectrogram. This distribution of energy is what one would expect after an ejective. Ejectives appear most likely to have this effect on the following vowel when

that vowel is [a] or [ə]. This occurs rarely or never with [i] or [u, w]. This effect is never found after pulmonic stops, only after ejectives. Paradoxically, however, some ejective tokens appear to lack higher frequency energy in the first several periods of the vowel. This effect occurs only when the vowel is [o, ɔ, w].

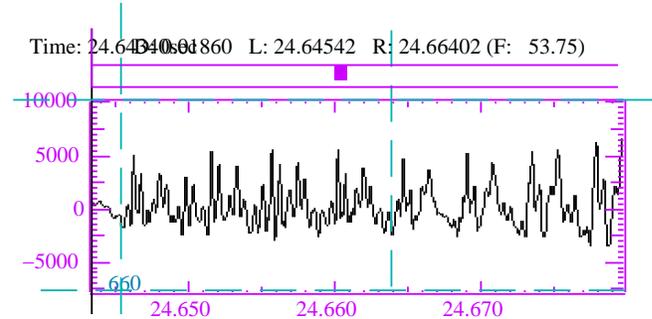


Fig. 3: F2 and F3 predominating over F1 after an ejective in [k'a] 'wheat'. The area between the vertical lines has more energy in the higher frequencies. After this, F1 becomes stronger.

A very common effect of ejectives on phonation type is a slow rise in the following vowel's amplitude (Fig. 4). To quantify this, I calculated the power of the waveform, then the first order difference of the power curve. Although calculating the rise time of the power curve would have been simpler, that method was precluded by the degree of variation in the peak power of the vowels. The first order difference of the power gives a measure of how fast the power is rising at the beginning of the vowel, independent of what point it eventually reaches.

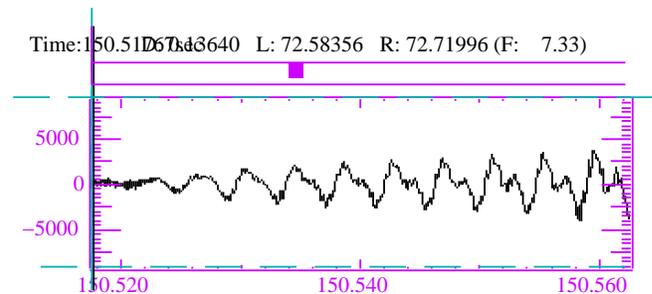


Figure 4: Waveform showing an ejective followed by several periods of very low amplitude voicing, in [t'ixxk] 'bone.'

For this power calculation, a frame length of approximately 14.6 msec and a frame step of approximately 0.1 msec, with a Hamming window, were used. The long frame length and short frame step were necessary to provide a smooth curve, so that the first order difference could be accurately calculated. The use of a Hamming window, however, should minimize the effect of power at the edge of the window on the point being calculated. I then found the average of the first order difference over the first 20 msec. of voicing after the stop. For this measure, the vowels after pulmonic voiceless stops had an average of 21,955, while the vowels after ejectives averaged 6,503. This difference is significant, at $F=27.83$ ($F(1,30)=4.17$ at $\alpha=.05$), and clearly shows that the amplitude of the vowel rises more slowly after ejectives than after pulmonic stops. This effect on phonation

type occurs very frequently in ejective tokens, and even more often with [i, j] than with other vowel qualities. This low amplitude voicing can also occur together with other voice quality effects.

Finally, some ejective tokens show an aperiodic component to the voicing for the first several periods of the following vowel. It is hoped that further research using inverse filtering will elucidate the effects of ejectives on phonation type, bringing together the various component effects discussed here. At this stage, it is only clear that almost all ejective tokens demonstrate at least one of the voice quality effects discussed. This might serve as an important perceptual cue for ejectives.

4. OTHER LANGUAGES

Table 2 allows for comparison of acoustic features of ejectives among Ingush, Hausa, Quiché, Navajo, and Tigrinya.

	Ingush	Hausa	Quiché	Navajo	Tigrinya
Pitch-following vowel	higher		lower		higher
approx. VOT-velar ejective	50 ms	25 ms	50 ms	80 ms	~ 80 ms
ratio of closure dur. to VOT	1.94	~ 2		~ 1.5	
post-burst noise amplitude	low	low	low	low	silence
rise to full amplitude of vowel	slow	normal	slow	very fast	fast
voice quality-onset of vowel	aperiodic	aperiodic	creaky	modal	modal

Table 2: Comparison of acoustic features of ejectives across multiple languages. No data is available for cells left blank. Sources: Tigrinya [2, 3, 4], Quiché [4], Navajo and Hausa [6].

In addition, information on VOT is available for Chechen, Montana Salish, and K'ekchi. In Chechen, bilabial ejectives have a VOT around zero, and ejectives in general have a low VOT, 7-33 msec. less than pulmonic voiceless stops [5]. No VOT data is available for velar ejectives, so Chechen cannot be easily compared with the languages above. Velar ejectives have an average VOT of 86 msec. in Montana Salish, and 97 msec. in K'ekchi [1].

To summarize, pitch following an ejective is higher than usual in Ingush and Tigrinya, but lower in Quiché. VOT is very short for Chechen, short for Hausa, long for Navajo, Tigrinya, Montana Salish, and K'ekchi, and in between for Ingush and Quiché. Amplitude of noise after an ejective burst is low for all languages. The rise time of the following vowel ranges from slow to very fast. The voice quality after an ejective can be aperiodic, creaky, or normal.

Kingston [4], comparing Tigrinya and Quiché, concludes that there are "tense" (Tigrinya) and "lax" (Quiché) ejectives, and that the glottal release is at the onset of voicing for tense ejectives, but with or near the oral release for lax ejectives. Lindau [6] draws a similar conclusion based on Navajo (glottal release at onset of voicing) and Hausa (earlier glottal release). Ingush ejectives are more similar to Quiché or Hausa ejectives than to those of Tigrinya and Navajo, so an early glottal release (at or near the oral release) would be predicted. Waveforms of Ingush ejectives, however, often show a rather late glottal release, at or near the beginning of voicing, as judged by a sudden onset of noise after near silence. This implies that timing of the glottal release may not pattern with amplitude rise or voice quality of the vowel. Further research on Ingush using oral airflow may help to confirm this.

Most previous cross-linguistic comparisons involved two languages each, and this obviously allowed for a two-way division of ejectives into tense and lax or strong and weak, as in [4, 6]. However, a comparison of the several languages above shows that no two languages pattern together for all the characteristics shown in the table. Tigrinya and Navajo do have similarities, but a binary division of ejectives is not sufficient to describe the data. Preliminary investigation of the timing of the glottal release makes a binary division even less likely. Ingush is closer to Hausa and Quiche than to Tigrinya and Navajo, but it does not pattern consistently with any of the other languages.

REFERENCES

1. Flemming, E., Ladefoged, P., and Thomason, S. "Phonetic structures of Montana Salish." *Fieldwork Studies of Targeted Languages II. UCLA Working Papers in Phonetics*, 87, 1-34, 1994.
2. Fre Woldu, K. "Evidence of auditory similarity between Tigrinya ejective /t/ and Arabic emphatic /t/," *Orientalia Suecana*, 23-25, 123-138, 1984-86.
3. Fre Woldu, K. *The perception and production of Tigrinya stops. Reports from Uppsala University Department of Linguistics* 13, 1985.
4. Kingston, J. C. *The phonetics and phonology of the timing of oral and glottal events.* University of California, Berkeley dissertation, 1984.
5. Kingston, J. C., and Nichols, J. "Pharyngealization in Chechen: Its implications for phonetics and phonology." Unpublished manuscript, Cornell University and University of California, Berkeley.
6. Lindau, M. "Phonetic differences in glottalic consonants." *Journal of Phonetics*, 12, 147-155, 1984.
7. McDonough, J., and Ladefoged, P. "Navajo stops." *UCLA Working Papers in Phonetics*, 84, 151-164, 1993.