A preliminary investigation of some acoustic characteristics of ejectives in Waima'a: VOT and closure duration

Mary Stevens & John Hajek
Department of French, Italian and Spanish Studies
University of Melbourne
Australia
mes@unimelb.edu.au; j.hajek@unimelb.edu.au

Abstract

Waima'a is a little known language spoken in East Timor. From a typological perspective, its stop system is unusual for an Austronesian language: it has a four-way stop system which includes a set of voiceless ejectives. In this paper, we report on the results of our first experimental investigation of Waima'a, in which ejectives and pulmonic voiceless aspirated and unaspirated stops are compared. After first presenting some qualitative observations about ejectives, we examine voice onset time (VOT) duration, closure duration and overall duration i.e. closure + VOT. We then turn to a cross-linguistic comparison, looking in particular at the behaviour of ejectives relative to other voiceless stop types, and at possible place of articulation dependencies concerning VOT and closure duration.

1. Introduction

This paper presents preliminary results from a first acoustic phonetic investigation into ejective stops in Waima'a (or Waimoa), an Austronesian language spoken in East Timor. Although ejectives occur frequently in different parts of the world, e.g. Africa and the Caucasus, they are extremely rare in Austronesian languages. The only other Austronesian language known to have ejective stops is Yapese, spoken some distance away in Micronesia. Ejectives have arisen quite independently in both languages.

To date only a limited amount of experimental evidence for ejectives in Yapese has been made available (Maddieson 2001, Cho & Ladefoged 1999). The only published source on ejectives in Waima'a is a short note by Hajek and Bowden (2002) who provide a brief articulatory description.

While there have been a number of recent experimental studies of ejectives in North American languages in particular, e.g. Gordon et al. 2001, Billerey-Mosier (2003), more detailed experimental data on ejectives from the Asia-Pacific region would be a useful contribution towards a better understanding of the phonetics of ejective stops in general. In this pilot study, we look specifically at a range of interrelated issues that have been examined in earlier phonetic investigations of ejectives: (i) voice onset time (VOT); (ii) closure duration; (iii) overall duration; (iv) possible interactions between each of these with place of articulation, and (v) how ejectives compare with other voiceless stop types with regard to (i-iv).

One particular advantage of investigating ejectives in Waima'a is that they occur at three different places, including in bilabial position. In many languages with ejectives, bilabial /p'/ is missing, and experimental studies and observations are, as a result, restricted to non-bilabial ejectives (see, for instance, Gordon et al. 2001 on Western Apache). One issue that remains uncertain is the extent to which VOT in ejectives follows the same general place-governed pattern reported in non-ejective, pulmonic stops, i.e. VOT is much greater in velars than in bilabials, with intermediate values typical for coronals (Maddieson 2001). In their survey of VOT in 18 languages, Cho and Ladefoged (1999) included 6 languages with ejectives (5 of these North American). They found no regular patterning of VOT according to place in ejectives. However, as Maddieson (2001) points out, their results are problematic: only 2 languages had bilabial ejectives, and measurement criteria do not appear to have been applied uniformly across the 6 languages.

1.1. Ejectives

Ejectives are distinguished from other voiceless, non-glottalized, stop types by having a closed glottis as well as a supralaryngeal closure gesture. As the larynx is raised the air between the glottal and oral closures is compressed. This increased pressure in the oral cavity results in a strong, distinctive burst at release of the oral closure. The glottis remains closed at the point of oral release, and is not opened until some time after (see e.g. Ladefoged & Maddieson 1996:78).

1.2. Waima'a stop series

Waima'a has a four-way stop system at three places of articulation, as shown in Table 1. Stops are one of four kinds: (a) voiced; (b) voiceless unaspirated; (c)

Proceedings of the 10th Australian International Conference on Speech Science & Technology
Macquarie University, Sydney, December 8 to 10, 2004.

Accepted after abstract only review

voiceless aspirated; and (d) voiceless ejective. Lexical sets contrasting the four stop types are easily found: /gama/ 'shark', /kama/ 'bed', /k'ama/ 'scratch' and /kha ma/ 'eat already'. Voiceless coronals /t t^h t'/ are dental in Waima'a. The status of the voiceless aspirate /p/ is less secure than that of other stops – it occurs only in recent loans from Portuguese and Tetum, and tends to merge with its aspirated counterpart /p^h/. Voiceless glottal /'/ also exists but has an extremely restricted distribution, with a tendency to weakening and deletion. Because of their unusual characteristics, no further reference is made to /p/ and /'/ in this study.

Table 1. Stop contrasts in Waima'a

Voiceless unaspirated	(p)	t	k
Voiceless aspirated	p^h	t^h	k^h
Voiceless ejective	p'	t'	k'
Voiced plain	b	d	g

Waima'a has a relatively simple phonotactic structure, with a strong tendency towards CVCV forms. Unlike other stops, ejectives never appear as the first member of a cluster, and can occur only once in a word. They appear most often in word-initial position.

2. Aims

This study aims to provide the first acoustic analysis of ejectives in Waima'a. After making some qualitative observations about ejectives in Waima'a, we then focus specifically on the duration of VOT and of the closure period, as well as on overall durations (VOT + closure) for ejective stops. We then consider results for each of these measurements and compare them with those for the other stop types under examination. Existing experimental studies of ejectives in other languages have concentrated primarily on their VOT properties, where the duration of the VOT has been shown to be useful in distinguishing ejectives from other stop types (e.g. Warner 1996 for Ingush; Billerey-Mosier 2003 for Kiowa; Wright et al. 2002 for Witsuwit'en). Very little information is available as to the role of closure duration in distinguishing ejectives from other stop types, primarily because most studies have been based upon ejectives in utterance-initial position. As ejectives were recorded in post-vocalic position in the Waima'a data, we are able to contribute some evidence of the duration of closure for ejective stops.

We compare ejectives specifically with voiceless unaspirated and aspirated stops in our Waima'a data, as well as with ejectives in other languages. Following Warner (1996), voiced stops /b d g/ were excluded from this initial study, since the negative VOT associated with voiced stops quite clearly separates them from ejective stops in both acoustic and perceptual terms.

3. Methods

3.1. The data

The data consist of a series of target words read in a simple carrier sentence by one adult male speaker of Waima'a, recorded in a laboratory setting. While the investigation is limited by the use of only one speaker subject at this stage, we note that other experimental studies also involve single speaker studies (e.g. Warner 1996; Billerey-Mosier 2003). The frame *ehe*_____ "say_____" was used so that the stop was always in post-vocalic position and closure duration could be measured. All target consonants used for this preliminary study precede /a/ except the 8 /t'/ tokens which precede /i/ in the word /t'iba/, as no corresponding word-initial /t'a/ tokens could be found at the time of recording. Up to 5 repetitions were made for each word, giving 71 tokens in total which were distributed as follows:

Table 2. Number of tokens examined and their distribution across place of articulation and stop type.

	Bilabial	Coronal	Velar	All
Aspirated	8	8	12	28
Unaspirated	-	8	13	21
Ejective	10	8	4	22
All	18	24	29	71

Using spectrographic and waveform displays within the PCQuirer program, durations of the closure period and of VOT were measured for each token. The quality of the onset of the post-stop vowel was also noted, since creak has sometimes been reported to occur after ejectives in other languages. Statistical tests were, unless otherwise indicated, conducted on the duration measurements using single-factor ANOVA tests within Excel.

3.2. Measurement criteria

In all cases the closure period was defined as the interval between the offset of the vowel and the onset of the oral release burst. For all three stop types examined (voiceless aspirated, unaspirated and ejective) the VOT was measured from the start of the oral release burst to the first glottal pulse associated with the vowel.

Regarding the ejective stops, existing experimental studies differ as to whether VOT was measured from the oral release to the glottal release (e.g. Cho & Ladefoged 1999:222), or to the actual onset of voicing associated with the following vowel (e.g. Billerey-Mosier 2003; Maddieson 2001; Wright et al. 2002). In this study VOT is measured to the onset of voicing, rather than to the glottal release. This is shown in Figure 1 where the

Accepted after abstract only review

arrow points to the glottal release following the oral release for the ejective stop /k’/:

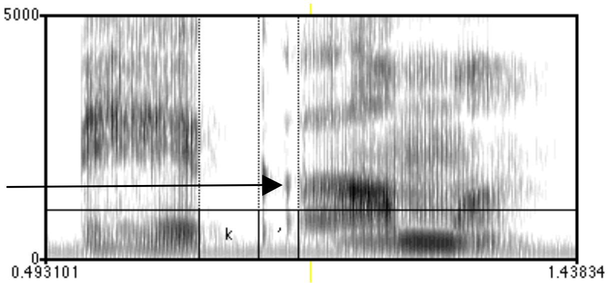

Figure 1. *The ejective velar stop in /k’ama/, showing measurement of VOT from the oral burst to the onset of voicing rather than to the glottal release shown by the arrow.*

In cases where the onset of voicing immediately follows the glottal release there will be no difference between the two measurement criteria. However, a distinguishing feature for ejectives is the silence that typically occurs between the glottal release and the delayed onset of voicing (see e.g. Warner 1996:1525; Gordon et al. 2001:422). Whether this period is included in the overall VOT is an important consideration, particularly in terms of the comparison of duration measurements for ejectives across languages (see Maddieson 2001:823; Billerey-Mosier 2003 for discussion). Given that we compare ejectives with aspirated and unaspirated stops in this study, we considered it most appropriate to measure VOT according to the same criteria for all stop categories i.e. to the actual onset of voicing associated with the vowel. Though differing from Cho & Ladefoged’s (1999) definition of VOT for ejectives, this is consistent with the measurement of VOT for ejectives in other studies with which we compare the Waima’a data (Billerey-Mosier 2003; Maddieson 2001; Gordon et al. 2001; Wright et al. 2002).

Vowel quality after each stop was also noted with primary reference to the spectrogram, where creaky voice was distinguished from modal voiced by its irregular vertical striations – the result of a more adducted glottis with respect to that required for modal voice.

4. Results

4.1. Qualitative observations

The acoustic appearance of the burst and VOT differs across all three stop types. Figure 1 shows an unaspirated velar stop /k/ with a short VOT consisting only of the burst at release of closure in the oral tract:

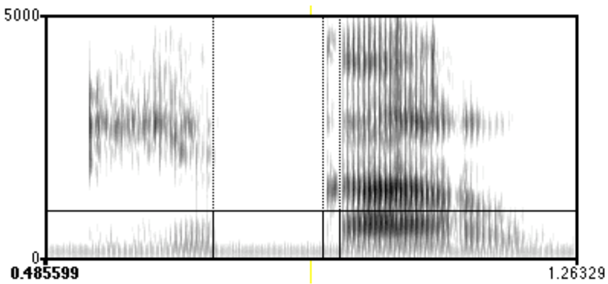

Figure 2. *The unaspirated velar voiceless stop in /karu/, showing the measurement of closure duration and VOT.*

Unaspirated stops are clearly distinguished from aspirated stops, which have a much longer VOT which is filled with aspiration noise:

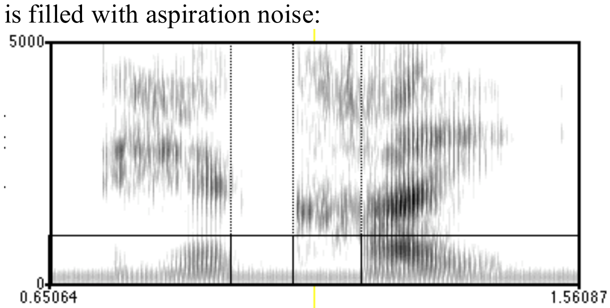

Figure 3. *The aspirated voiceless velar stop in /khai/, showing inclusion of the aspiration noise in the VOT duration.*

The VOT for ejective tokens typically involved a clearly visible oral release, a weaker glottal release and a slightly delayed onset of voicing (as in Figure 1). In some cases the release of the glottal closure was not distinguishable on the spectrogram or waveform displays from the onset of voicing, as below:

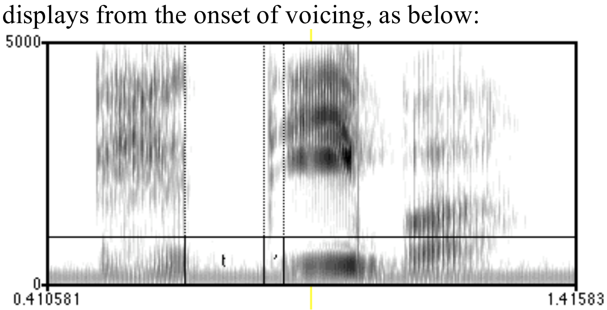

Figure 4. *The coronal ejective in /t’ibu/ which does not have the expected silence after the release of the oral burst because glottal closure is released immediately after.*

As seen in the onset of the vowel following /k’/, below, the first few periods of the following vowel can be characterized by creaky voice, as the glottis returns to the normal position for modal phonation following complete adduction:

Proceedings of the 10th Australian International Conference on Speech Science & Technology
Macquarie University, Sydney, December 8 to 10, 2004.

Accepted after abstract only review

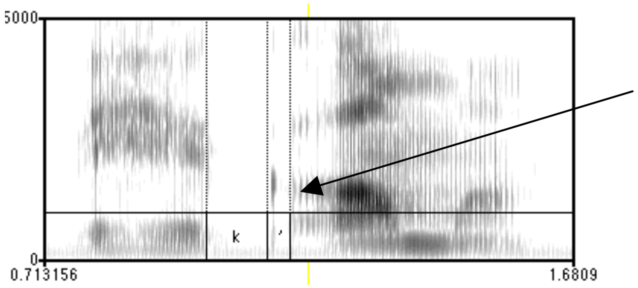

Figure 5. *The velar ejective stop in /k'ama/ showing creak in the onset of the following vowel.*

This period of creak can provide a secondary cue to distinguish ejective stops (see e.g. Ladefoged & Maddieson 1996:79).[1] However, creak was seen in the vowel following only two of the four velar /k'/ tokens, but none of the 8 /t'/ tokens nor any of the 10 /p'/ tokens. Post-release creak is overall not frequent, and in this respect ejectives in Waima'a are similar to reported descriptions of ejectives in Western Apache (Gordon et al. 2001) for which some contextual creak is reported to occur occasionally.

4.2. VOT duration

The duration of the VOT for ejectives in Waima'a is intermediate between that recorded for aspirated and unaspirated stops across all places of articulation:

Table 3. Average VOT values across stop places of articulation and type; standard deviations in parentheses.

	p	t	k
Aspirated	60 (6)	96 (15)	103 (13)
Ejective	33 (6)	30 (6)	66 (19)
Unaspirated	-	22 (4)	22 (3)

The difference between the VOT for aspirated and ejective stops is statistically significant at all three places of articulation ($p < 0.001$ in all cases). Similarly, VOT differences between ejective and unaspirated stops are statistically significant at both places for which data were available: velar /k/ v. /k'/ $p = 0.001$; coronal /t/ v. /t'/ $p = 0.006$.
Amongst the ejective stops themselves, VOT is much longer in velar (66ms.) than in bilabial or coronal stops. The difference is strongly significant, $p < 0.001$ in both cases. Coronal and bilabial ejectives show similar VOT durations, and the small difference between them is not statistically significant ($p = 0.345$). Within the aspirated series, VOT durations are significantly different between /k^h/ and /p^h/ and between /t^h/ and /p^h/ to the same degree ($p = 0.001$) but are not significant between /k^h/ and /t^h/ ($p = 0.293$). No effect was found in the plain unaspirated series between /t/ and /k/ ($p = 0.911$).

4.3. Closure duration

Comparing the closure durations for ejectives with those of other stop types, our Waima'a data show no consistent patterns across place of articulation or stop type:

Table 4. Average closure durations & standard deviations in ms. according to stop place of articulation and type.

	p	t	k
Aspirated	164 (62)	121 (4)	107 (7)
Ejective	131 (14)	136 (6)	113 (3)
Unaspirated	-	119 (8)	139 (21)

Although average closure duration is longer for bilabial aspirated stops /p^h/ than bilabial ejectives /p'/, this difference is not statistically significant ($p = 0.128$). As for coronal stops, the order of significance is reversed: ejective /t'/ has a significantly longer closure duration than aspirated /t^h/ ($p < 0.001$). There is no significant difference in the closure duration of the velar aspirated stop /k^h/ and its ejective counterpart /k'/ ($p = 0.123$).
When we compare ejectives with the unaspirated stops, closure duration is shorter for ejective /k'/ than its unaspirated counterpart /k/. This difference in duration is statistically significant ($p = 0.033$). For coronal stops the order is once again reversed: ejective /t'/ has a longer closure duration than unaspirated /t/, and the difference is highly significant ($p < 0.001$). Combined ANOVAs for whole stop series did not distinguish between the closure durations of aspirated and ejective stops ($p = 0.767$), unaspirated and ejective stops ($p = 0.820$) or aspirated and unaspirated stops ($p = 0.691$).
In terms of place effects within each stop series, we report first on the ejectives. Closure duration is longest for coronal and bilabial ejectives, and is shortest for the velar ejective stop. The difference between the closure duration for velar /k'/ and coronal /t'/ is statistically significant ($p < 0.001$), as is the difference between bilabial /p'/ and velar /k'/, albeit to a lesser degree ($p = 0.030$). There was no significant difference between bilabial /p'/ and coronal /t'/ ($p = 0.367$). With respect to the pulmonic stops, closure duration were strongly differentiated in parts of the aspirated series: /p^h/ v. /k^h/ $p = 0.006$, /t^h/ v. /k^h/ $p < 0.001$, albeit not between /p^h/ and /t^h/ where $p = 0.072$. The difference between closure durations is less marked but still significant in the unaspirated series, with only two members: /t/ v. /k/ $p = 0.021$.

[1] Creaky periods were included in vowel durations.

4.4. Overall stop duration

For each stop type, the VOT and closure duration measurements given in §4.2 and §4.3 were added to calculate the overall duration of the stop at each place of articulation:

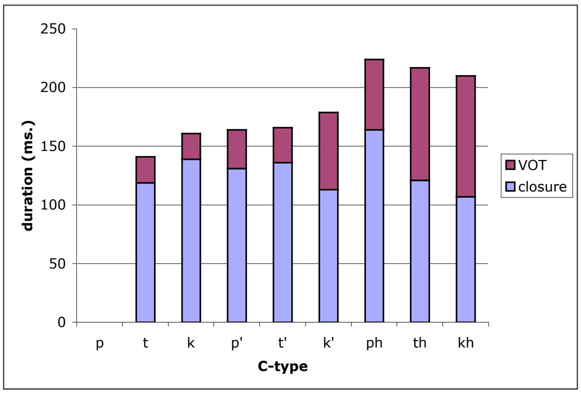

Figure 6. *Overall stop durations (VOT & closure) according to type (voiceless unaspirated, ejective, and aspirated) and place of articulation.*

In Figure 6 we see that aspirated stops have the longest overall durations, ejectives have intermediate overall durations, and unaspirated stops are the shortest. ANOVAs comparing overall durations (closure + VOT) show that the differences in stop type are statistically highly significant: unaspirated v. aspirated $p < 0.001$; ejective v. aspirated $p < 0.001$; ejective v. unaspirated $p = 0.005$.

With regard to specific place interactions with overall duration, no significant effect was found for any pairwise comparison in either the aspirated or ejective series: /p^h/ v. /k^h/ p = 0.493; /k^h/ v. /t^h/ p = 0.408; /t^h/ v. /ph/ = 0.775; /p'/ v. /t'/ p = 0.660; /t'/ v. /k'/ p = 0.086; /k'/ v. /p'/ p = 0.085. However, a relatively small place effect was found in the unaspirated set between velar /k/ and coronal /t/ (p= 0.022).

5. Discussion

Our Waima'a data pattern well with reported findings for ejective stops in terms of VOT duration in a number of other languages with the same three-way series of voiceless stops, e.g. Witsuwit'en (Wright et al. 2002) and Western Apache, Navajo and Tlingit (Gordon et al. 2001). In all of these cases the duration of the VOT for ejectives is intermediate between that of aspirated and unaspirated stops. In Kiowa, much higher VOT values are reported for ejectives than for voiceless aspirates. It is not clear why this is the case, although Billerey-Mosier (2003) suggests it may reflect the different segmentation technique he used in contrast to earlier studies such as Cho and Ladefoged (1999). However, we have followed Billerey-Mosier in measuring VOT to the onset of voicing rather than to glottal release, and have found no difference between Waima'a and most other languages with respect to the relative ordering of stop VOT values. It appears then that Kiowa is an exception amongst languages with ejectives.

The duration of the VOT for velar ejectives is considerably longer than at other places of articulation in our Waima'a data, in line with available duration measurements for ejectives in other languages e.g. Western Apache (Gordon et al. 2001), and Kiowa (Billerey-Mosier 2003). Our Waima'a data pattern particularly well with measurements reported for ejectives in Yapese (Maddieson 2001), the only other Austronesian language to have ejective stops:

Table 5. VOT durations for ejective stops in Yapese (Maddieson 2001) and Waima'a.

	p'	t'	k'
Yapese	29	35	61
Waima'a	33	30	66

It is also known more generally that across languages VOT tends to be longest at velar than at other places of articulation, particularly bilabial, regardless of stop type (e.g. Maddieson 1999, 2001; Lisker & Abramson 1964). As we have already seen in §4.2, the effect of velar place on VOT duration is highly significant in Waima'a across all three stop types. A similar effect is reported for Western Apache (Gordon et al. 2001) and Kiowa (Billerey-Mosier 2003).

While the closure duration for ejective stops is reported to be intermediate between aspirated and unaspirated stops in other languages for which data are available (e.g. Kiowa), effects are not always great enough to be statistically significant (e.g. Western Apache). Our Waima'a data do not show a consistent pattern across stop types, and no significant pairwise effect between stop series was uncovered. This appears to be the result of place-governed internal inconsistencies within each series: within bilabials closure duration is longer for aspirated than for ejective stops, but for coronals the opposite is true. In this case vowel height might have an effect, as the ejective /t'/ preceded the high vowel /i/ whereas all other stop consonants in our Waima'a data preceded /a/. This could also explain the longer closure durations for coronal ejectives than for unaspirated coronal stops (i.e. /t'i/ v. /ta/). Further investigation of Waima'a on this point is clearly needed. We note only at this stage that Gordon et al. (2001) who discuss possible vowel height effects between /i/ and /a/ only as they pertain to VOT, did not find any effect. Unfortunately, they do not discuss vowel height in the context of closure duration.

When VOT and closure durations are summed in Waima'a (cf. Figure 6), we see that the overall duration of ejective stops is intermediate between that of aspirated and unaspirated stops. The difference in duration between the three stop types is found to be

Accepted after abstract only review

highly significant. Ejectives and voiceless aspirates are similar in that they show a general inverse correlation between closure and VOT durations across place, as has often been noted in stops across languages (Maddieson 1999). Such an effect is particularly evident when bilabials and velars are compared. The absence of bilabial /p/ means that the same effect cannot be observed in the unaspirated series.

Place of articulation has no effect on overall stop durations in Waima'a, as opposed to their component parts (closure & VOT), within either the ejective or voiceless aspirated series. This result reflects the frequently observed trade-off between VOT and closure durations at different places of articulation (see Maddieson 1999, 2001 and especially Weismer 1980 for a possible explanation of such an effect). However, a small effect of place on overall durations is observable between /t/ and /k/ in the voiceless unaspirated series in Waima'a.

With respect to the general properties of voiceless stops, our Waima'a data show that ejectives are clearly distinguished from other voiceless stop types by differences in relative VOT and in overall duration. In both cases, they fall between voiceless unaspirated and aspirated stops. Our results lend support to the suggestion by Gordon et al. (2001) that VOT may serve as a perceptual cue in differentiating ejectives from other stops. Based on our results, we extrapolate that overall duration could also function with the same perceptual effect. Closure duration alone, on the other hand, does not appear to differ according to manner of articulation.

Specific to possible place of articulation dependencies, our results show strong effects of place on VOT and closure durations: velar ejectives are significantly different on both measures relative to other places of articulation. An identical pattern is evident in voiceless aspirates. However, place has no effect on overall durations within the ejective series, a pattern once again mirrored in aspirated stops. Unaspirated pulmonic stops /t k/ behave somewhat differently: VOT values are the same, but differences in closure duration and overall duration differences are significant, albeit not at high levels. Given the absence of bilabial /p/ from the tested set, less weight should be given to generalizations about place effects among unaspirated stops based on our reduced data.

There is ample scope for further investigation of ejectives in Waima'a. But even at this stage, our results show more consistent articulatory patterning within the ejective series and between ejectives and other stops than Cho & Ladefoged (1999) found in their study. Our results also show that ejectives fit in well amongst other stop types along specific duration-based phonetic continua. As a result, we must agree with Maddieson (2001) that "ejectives have a similar pattern to plosives and … a unified explanation for all three types of stop should be sought".

6. References

Billerey-Mosier, R. (2003). *An acoustic analysis of Kiowa stops*. Unpublished manuscript, UCLA. Available at http://www.linguistics.ucla.edu/people/grads/billerey/billerey/htm.

Cho, T. & P. Ladefoged (1999). Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics* 27: 207-229.

Gordon, M. (2001). Linguistic aspects of voice quality with special reference to Athabaskan. *Proceedings of the 2001 Athabaskan Languages Conference*: 163-178.

Gordon, M., Potter, B., Dawson, J., de Reuse, W. & P. Ladefoged (2001). Some phonetic structures of Western Apache. *International Journal of American Linguistics* 67: 415-448.

Hajek, J. & J. Bowden (2002). A phonological oddity in the Austronesian area: Ejectives in Waimoa, *Oceanic Linguistics* 41: 222-224.

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

Lisker, L. & A. Abramson (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word* 20: 384-422.

Maddieson, I. (1999). Phonetic universals. In W. J. Hardcastle & J. Laver (Eds.) *The Handbook of Phonetic Sciences*. Oxford, Blackwell: 619-639.

Maddieson, I. (2001). Good timing: place dependent VOT in ejective stops. *Proceedings of Eurospeech 2001*: 823-826.

Warner, N. (1996). Acoustic characteristics of ejectives in Ingush. *Proceedings of the International Conference on Spoken Language Processing,* October 3-6 1996, Philadelphia: 1525-1528.

Weismer, G. (1980). Control of the voicing distinction for intervocalic stops and fricatives: some data and theoretical considerations. *Journal of Phonetics* 8: 427-438.

Wright, R., Hargus, S. & K. Davis (2002). On the categorization of ejectives: data from Witsuwit'en. *Journal of the International Phonetic Association* 32: 43-77.

Accepted after abstract only review