PHONETIC VARIABILITY IN THE REALIZATION OF GLOTTALIZED STOPS IN USPANTEKO (MAYAN)

Ryan Bennett[a], Robert Henderson[b], Meg Harvey[b]

University of California, Santa Cruz[a] and University of Arizona[b]
rbennett@ucsc.edu, rhenderson@arizona.edu, mharvey3@arizona.edu

ABSTRACT

Glottalized *b' q'* often vary between ejective and implosive realizations in Mayan languages. Some of this variation has been attributed to phonetic differences between languages or dialects, but most descriptions are compatible with the alternative view that individual speakers of the same language/dialect differ in their inclinations to produce these sounds as ejective or implosive. We investigate the phonetic realization of *b' q'* in Uspanteko, a Guatemalan Mayan language. We find that labial *b'* is produced as an implosive more often than uvular *q'*; that word-final position favors ejective realizations; and that speakers of the same dialect differ in their tendencies to produce *b' q'* as implosive vs. ejective.

Keywords: ejectives, implosives, Mayan languages, allophony, inter-speaker variation

1. ALLOPHONIC REALIZATIONS OF GLOTTALIZED STOPS IN MAYAN

All Mayan languages contrast plain voiceless stops and affricates like /p t k t͡s t͡ʃ/ with glottalized counterparts at the same place of articulation [1].

(1) Glottalization contrasts in Uspanteko [2–4]
 a. *ka'n* /kaʔn/ 'animal'
 b. *k'a'n* /kˀaʔn/ 'angry'

The phonetic realization of these glottalized stops varies by language and place of articulation, among other factors. Bilabial *b'* is typically an implosive [ɓ] or [ɓ̥], though in some languages it may be preglottalized [ˀb]. The canonical realization of other glottalized plosives tends to be ejective (1).

However, glottalized stops in Mayan languages sometimes show extensive phonetic variability. Ejective [pˀ] allophones are commonly reported for glottalized *b'*, especially in word-final position [1, 4]. Variation between ejective and implosive realizations is often reported for uvular *q'* = [qˀ] ∼ [ʛ̥], and to a lesser extent coronal *t'* = [tˀ] ∼ [ɗ].[1]

Descriptive sources often report that this variation is conditioned by segmental context [1, 5, 6]. For example, [2] states that word-final *b'* in Uspanteko is realized as an unreleased implosive [ɓ̥̚], or an ejective [pˀ]; it is otherwise described as a voiceless implosive. [6] found that one speaker of Q'anjob'al produced *b'* as ejective [pˀ] at word edges, and as implosive [ɓ] word-medially, especially before [a].

Even so, in at least some Mayan languages, individual speakers may produce glottalized *b' q'* as either implosive or ejective in exactly the same environments, and even in the same words [1, 4–7]. Additionally, [1, 4] note that speakers seem to differ in their tendencies to produce ejective or implosive variants of these sounds. So while the segmental context may bias *b' q'* toward ejective vs. implosive realizations, it can also leave room for variability.

The phonetic variability observed for *b' q'* seems to involve a choice between qualitatively distinct allophones. Ejectives are produced with laryngeal raising, and implosives with laryngeal lowering (even if slight); these two vertical movements require the contraction of different muscles [5, 6, 8–12]. In that sense, the varied realization of *b' q'* as ejective vs. implosive implicates variation between categorically distinct phonetic outputs (cf. [6]).

In important early work addressing the phonetics of *b' q'* in Mayan with instrumental methods, [5] characterizes this variability in terms of differences between languages and/or dialects (based on data from the K'ichean languages Q'eqchi', Poqomchi', Kaqchikel, K'iche', and Tz'utujil). However, [5] presents results from just one speaker of each language/dialect, and no quantitative analysis. Consequently, it is unclear whether the phonetic variation observed by [5] reflects systematic differences between speech communities, or simply differences between individuals (as [5] herself notes; see also [7]). The same point can be made for [6, 13], who each report aerodynamic data from one speaker of Q'anjob'al. [7] is the only study of ejective vs. implosive variation in a Mayan language (Mam) which combines instrumental methods, quantitative analysis, and data from multiple speakers; still, [7] does not investigate contextual factors. Here, we provide a novel quantitative analysis of contextual and inter-speaker variability between ejectives and implosives in another Mayan language, Uspanteko.

Data from Uspanteko could be useful for distinguishing the effects of dialect and language on the realization of *b’ q’* from simple inter-speaker variation. Uspanteko is a fairly small language, with up to 6000 speakers; in comparison, there are over 70,000 speakers of Tz’utujil, over 150,000 speakers of Q’anjob’al, over 400,000 speakers of Kaqchikel, and over 1 million speakers of both K’iche’ and Q’eqchi’ (according to the most recent Guatemalan census; `https://www.censopoblacion.gt/`). Uspanteko is spoken in a compact region, and dialect variation does not occur on the same scale found for larger Mayan languages. Salient dialect differences do exist between the hamlet of Las Pacayas and other locations [2,4,14,15], but none of the speakers reported on here are from Las Pacayas.

In this paper we investigate variability in the realization of *b’ k’ q’* in a wordlist corpus [4]. We code each token of these sounds as either ejective or implosive based on their acoustic properties, and investigate what factors might influence this categorical variation in their phonetic form. We find that labial *b’* is implosive more often than uvular *q’*, and that speakers differ in their tendencies to produce ejectives vs. implosives. We also confirm that word-final position favors ejective realizations.

2. DATA COLLECTION

Audio recordings were collected from 9 native speakers of Uspanteko in 2018 (3 male; 23-50 years old, mean 35). Eight speakers were from the town of San Miguel Uspantán, and one from the nearby village of La Lagunita. The speakers translated target words written on index cards in Spanish to Uspanteko. The words were produced in the frame sentence [ˈjaχ ˈtekˀ ___ tiχ.ˈɓiχ] ‘Diego says ___’. Most target items were 1-2 syllables long, but some were longer (3-5 syllables). Recordings were made in a quiet room with a headset microphone (Audio-Technica ATM73a) and solid-state recorder (Zoom H5) at a 48 kHz sampling rate. See [4] for details.

3. DATA ANNOTATION

Ejective vs. implosive realizations of *b’ q’* and velar *k’* were annotated using TextGrids in Praat [16]. Implosives can be acoustically distinguished from ejectives and plain stops by the lack of a clear, salient release burst, particularly at CV transitions [4,5,7,9,17,18]. Though a release burst is sometimes present for implosives, it is typically very weak, and sometimes appears on the waveform as an initial negative pressure impulse at release, possibly reflecting ingressive airflow. Implosives, unlike ejectives, may also be produced with voicing during closure, especially just before or after the release of the oral constriction.

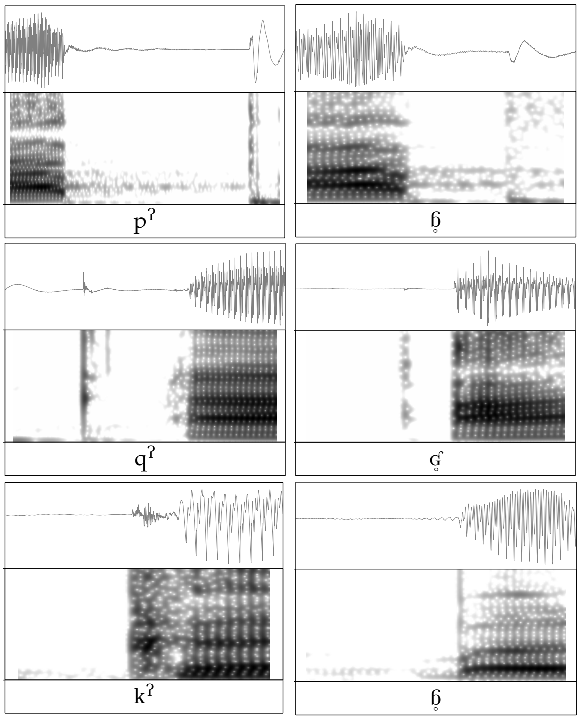


Figure 1: Some coding decisions for *b’ k’ q’*. Top row: /aC#/ context; bottom rows: /Ca/ context.

We thus annotated according to the following algorithm: (i) code as implosive if the release burst is very weak or absent, or if there is a negative impulse at release (Fig. 1, right); (ii) as ejective if there is a clear, strong release burst accompanied by a positive impulse on the waveform (Fig. 1, left); and (iii) if there is only a weak positive burst at release, look for other diagnostics, primarily voicing during closure or just prior to the oral release (Fig. 1, bottom right).

While most tokens were unambiguously ejective or implosive, some were difficult to annotate. Ejective tokens sometimes seemed to be ‘slack’, with weaker bursts [6, 7, 17, 19–21]. These could be difficult to distinguish from implosives, and so we had to rely on criteria other than the salience of the release burst to make coding decisions (e.g. voicing prior to release, direction of the initial impulse at release, etc.). Ambiguous tokens were most common with labial *b’*, likely because labials have weak release bursts in general, and so often lack the strong burst associated with ejectives [22].

We observed 14 tokens that appeared to be plain (non-glottalized) stops, and 28 unreleased stops in word-final position that could not be reliably classified. We excluded these tokens from analysis.

Each annotation was checked multiple times, by at least two co-authors, to ensure that annotation guidelines were consistently followed. We did not compute any measures of inter-annotator agreement, but after joint review, it was clear our classification practices were quite consistent.

4. STATISTICAL ANALYSIS

Statistical analyses were carried out using `R` [23]. This dataset includes 312 ejective and 225 implosive tokens (Fig. 2). As expected, glottalized *b'* was mostly commonly produced as an implosive, and *q'* most commonly produced as an ejective, though ejective and implosive tokens did occur frequently for both sounds. Also as expected, velar *k'* was strictly realized as ejective, and never implosive except for one anomalous token.

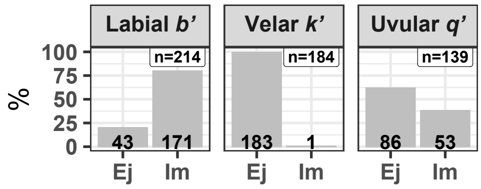
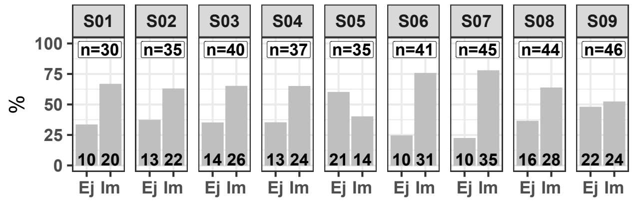


Figure 2: *b' k' q'* tokens by allophone. 'Ej' = ejective, 'Im' = implosive. Numbers in bars indicate token counts.

Since velar *k'* shows essentially no allophonic variability, it is not informative about the conditions favoring implosive vs. ejective variants of glottalized stops. Indeed, the effect of PLACE = VELAR cannot be estimated in the logistic regression presented below, because *k'* = [kˀ] is basically invariant. We thus exclude velar *k'* from further discussion. This leaves 129 ejective and 224 implosive tokens of *b' q'* (= 353 total) for analysis.

Speakers apparently differ in their propensities to produce ejectives vs. implosives for *b' q'* (Fig. 3). For example, speaker 7 seems to favor implosives, while speaker 5 seems to favor ejectives; both speakers are women from San Miguel Uspantán, though they differ in age.

We coded the segmental environment in terms of any neighboring consonants (= C), word-boundaries (= #), and vowels (= V), distinguishing stressed from unstressed vowels (= [±str], Fig. 4; stress may be final or penultimate in Uspanteko, and is predictable [4]). We treated lexical consonant clusters (e.g. *xq'aax*) and clusters produced by vowel deletion (e.g. *xáj(a)b'*) as equivalent (cf. [24]). With this coding, we observe more ejectives following stressed vowels and in word-final position, but no other obvious contextual effects.



Speaker	1	2	3	4	5	6	7	8	9
Age	30	34	30	27	29	50	44	23	47
Gender	F	M	F	M	F	M	F	F	F
Town	U	U	U	L	U	U	U	U	U

Figure 3: *b' q'* tokens by speaker.

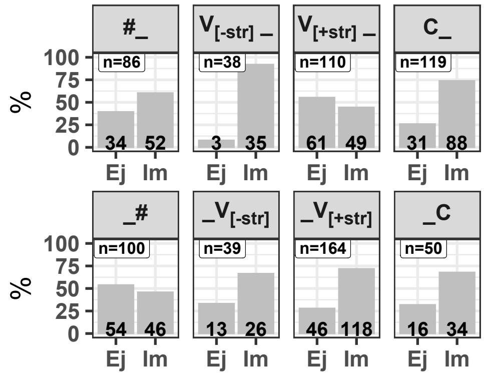


Figure 4: *b' q'* tokens by context. '_' = position.

The distribution of implosive and ejective variants was modeled with a mixed-effects binomial logistic regression (2) using the `glmer` function in `lme4` [25]. We included SPEAKER as a fixed effect rather than a random effect because we are interested in inter-speaker variation in ejective vs. implosive productions of glottalized stops, which is more straightforwardly assessed if SPEAKER is treated as a fixed effect. A random intercept for WORD was also included. All predictors were sum coded to reduce collinearity between simple predictors and interactions; this means that each level of a variable is compared to the grand mean across all levels of that variable. We chose speaker 4 as the reference level for SPEAKER because the estimated effect for speaker 4 was close to the overall sample mean. Fixed effects included SPEAKER, C PLACE, PRECEDING SEGMENT TYPE, and FOLLOWING SEGMENT TYPE, along with a two-way interaction between SPEAKER and C PLACE (2a). Models including other interactions involving C PLACE failed to converge, likely because some environments don't include enough tokens of both *b'* and *q'* to estimate an interaction (e.g. *b'* occurs in the context $V_{[\text{-str}]}$_ 35 times, and *q'* just 3 times). The same is true for additional interactions involving SPEAKER. This model was then simplified using the `stepAIC` function in the `MASS` package [26]. The final model (2b) has reasonably low collinearity for a model which includes interaction terms (VIF $\in [1.49, 2.92]$). To save space, we only report statistically significant levels of each predictor in the final model (Tab. 1).

(2) a. **Initial model**: PLOSIVE TYPE ~ C PLACE * SPEAKER + PRECEDING SEGMENT TYPE + FOLLOWING SEGMENT TYPE

b. **Final model**: PLOSIVE TYPE ~ C PLACE + SPEAKER + PRECEDING SEGMENT TYPE

	$\hat{\beta}$	Odds ratio ($e^{\hat{\beta}}$)	SE($\hat{\beta}$)	p-value
(Intercept)	0.95	2.60	0.25	$< .001$
LABIAL	1.30	3.68	0.19	$< .001$
SPEAKER 5	-1.51	0.22	0.46	$< .01$
SPEAKER 7	0.78	2.19	0.39	$< .05$
SPEAKER 9	-0.92	0.40	0.38	$< .05$
$V_{[\text{-str}]}$—	1.09	2.96	0.55	$< .05$
$V_{[\text{+str}]}$—	-1.54	0.21	0.29	$< .001$
SPEAKER 2 × LABIAL	1.25	1.39	3.49	$< .05$
SPEAKER 6 × LABIAL	-0.80	1.80	0.45	$< .05$
SPEAKER 7 × LABIAL	-0.80	2.19	0.45	$< .05$
SPEAKER 8 × LABIAL	-0.83	0.98	0.44	$< .05$

Table 1: Significant predictors in final logistic regression model. Odds ratios > 1 indicate higher likelihood of implosive, < 1 lower likelihood. $\hat{\beta}$ is relative to mean across all levels of the predictor.

Labial *b'* is realized as implosive more often than uvular *q'* is (Tab. 1). This finding parallels well-known typological tendencies, grounded in aerodynamic and articulatory factors [8, 22, 27]. In particular, labial stops have a larger oral cavity, and so it is more difficult to generate the increased intra-oral pressure needed to produce the high intensity 'pop' associated with ejective release. This may favor implosive realizations of *b'* across all contexts.

The segmental environment also plays a role: ejectives are more common after stressed vowels. This finding may actually indicate conditioning by word boundaries, which are highly correlated with stress in our data: most stops in $V_{[\text{+str}]}$_ contexts are word-final (82/110 = 75%), while no word-final stops occur in the environment $V_{[\text{-str}]}$_ (stress is often word-final [2, 4], and many target words were monosyllabic). In this dataset, *b'* is ejective in 42% of word-final tokens (26/62), and just 11% elsewhere (17/152); *q'* is ejective in 74% of word-final tokens (28/38), and 57% elsewhere (43/101) (see also Fig. 4). This is consistent with reports that word-final position favors ejective realizations of *b' q'* in K'ichean languages like Uspanteko [1, 2]. We speculate that word-final position favors ejectives because their releases are more salient, and so word-final stops at different places of articulation may be easier to distinguish when ejective [28]. (Word-final stops were also prosodically phrase-final in our study, due to our use of a frame sentence.)

Most of our significant effects concern inter-speaker variation. Compared to the overall sample of speakers, Speaker 2 is more likely to produce *b'* as implosive [ɓ̥]. Speakers 6 and 8 are instead more likely to produce *b'* as ejective [pˀ]. Speakers 5 and 9 favor ejectives for both *b'* and *q'*, when compared to the overall mean. Speaker 7 favors ejective realizations of *q'*, but not *b'*.

The source of this inter-speaker variation is not apparent. Some speakers who pattern together in their use of ejective vs. implosive variants of *b' q'* have different demographic characteristics (e.g. Speakers 6 and 8; Fig. 3). Notably, dialect variation is not a plausible explanation for these inter-speaker differences: our speakers were almost exclusively from one town, San Miguel Uspantán, and so geographically-based dialect variation cannot be the source of the observed variation [7, 14, 15]. We leave open the possibility that other sociolinguistic factors condition these patterns of phonetic variability.

5. CONCLUSION

While variation between implosive and ejective realizations of *b' q'* is sensitive to place of articulation and segmental context, a substantial component of this variation comes down to seemingly idiosyncratic differences between speakers. Our results are consistent with the view that glottalized consonants in Mayan languages are united by the presence of a laryngeal constriction of some kind, rather than by implosive or ejective articulations as such [1, 5]. Further, implosive *b'* was more common than implosive *q'*, in keeping with claims that the most canonical allophone of *b'* in K'ichean languages is implosive [1, 4]. If implosive and ejective productions of *b' q'* reflect qualitatively distinct articulatory maneuvers, variation between these sounds amounts to a variable allophonic process with categorical, rather than continuous, phonetic outcomes (cf. [6, 24]).

Coding glottalized stops as ejective or implosive on the basis of their acoustic properties can be challenging, particularly for *b'*. But the fact that it is sometimes difficult to distinguish ejective vs. implosive allophones of *b' q'* may help explain why we find so much intra- and inter-speaker variation in the production of these sounds in the first place: their acoustic similarity may faciliate their phonological equivalence.

6. REFERENCES

[1] R. Bennett, "Mayan phonology," *Language and Linguistics Compass*, vol. 10, no. 10, pp. 469–514, 2016.

[2] T. Can Pixabaj, *Gramática descriptiva Uspanteka*. Antigua, Guatemala: Oxlajuuj Keej Maya' Ajtz'iib' (OKMA), 2007.

[3] M. A. Vicente Méndez, *Diccionario bilingüe uspanteko-español*. Antigua, Guatemala: Oxlajuuj Keej Maya' Ajtz'iib' (OKMA), 2007.

[4] R. Bennett, R. Henderson, M. Harvey, and T. Méndez López, "The phonetics and phonology of Uspanteko (Mayan)," *Language and Linguistics Compass*, vol. e12467, 2022. [Online]. Available: https://dx.doi.org/10.1111/lnc3.12467

[5] S. Pinkerton, "Quichean (Mayan) glottalized and nonglottalized stops: a phonetic study with implications for phonological universals," in *Experimental phonology*, J. Ohala and J. Jaeger, Eds. Orlando: Academic Press, 1986, pp. 125–139.

[6] J. Kuang, "Contextual variation of glottalic stops in Q'anjob'al," in *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019*, S. Calhoun, P. Escudero, M. Tabain, and P. Warren, Eds. Canberra, Australia: Australasian Speech Science and Technology Association Inc., 2019, pp. 1114–1118. [Online]. Available: https://assta.org/proceedings/ICPhS2019/papers/ICPhS_1163.pdf

[7] S. Russell, "Some acoustic characteristics of word initial pulmonic and glottalic stops in Mam," Master's thesis, Simon Fraser University, 1997.

[8] P. Ladefoged and I. Maddieson, *The sounds of the world's languages*. Malden, MA: Blackwell, 1996.

[9] G. Clements and S. Osu, "Explosives, implosives, and nonexplosives: the linguistic function of air pressure differences in stops," in *Papers in Laboratory Phonology VII*, C. Gussenhoven and N. Warner, Eds. Berlin: Mouton de Gruyter, 2002, pp. 299–350.

[10] J. Kingston, "The phonetics and phonology of the timing of oral and glottal events," Ph.D. dissertation, University of California, Berkeley, 1984.

[11] M. Oh, D. Byrd, L. Goldstein, and S. S. Narayanan, "Enriching the understanding of glottalic consonant production: Vertical larynx movement in Hausa ejectives and implosives," *The Journal of the Acoustical Society of America*, vol. 144, no. 3, pp. 1940–1941, 2018.

[12] C. Ludlow, "Central nervous system control of the laryngeal muscles in humans," *Respiratory physiology & neurobiology*, vol. 147, no. 2-3, pp. 205–222, 2005.

[13] R. Shosted, "Towards a glottalic theory of Mayan," in *New Perspectives in Mayan Linguistics*, H. Avelino, J. Coon, and E. Norcliffe, Eds. Cambridge, MA: MIT Working Papers in Linguistics, 2009, vol. 59, pp. 80–113.

[14] L. Campbell, *Quichean linguistic prehistory*, ser. University of California Publications in Linguistics. Berkeley, CA: University of California Press, 1977.

[15] J. A. Us Maldonado, *Monografía Uspanteka*. San Miguel Uspantán: Comunidad Lingüística Uspanteka, no date.

[16] P. Boersma and D. Weenink, "Praat: doing phonetics by computer (version 6.1.05)," Computer program, 2020, retrieved from http://www.praat.org/.

[17] M. Lindau, "Phonetic differences in glottalic consonants," *Journal of Phonetics*, vol. 12, pp. 147–155, 1984.

[18] C. Henton, P. Ladefoged, and I. Maddieson, "Stops in the world's languages," *Phonetica*, vol. 49, no. 2, pp. 65–101, 1992.

[19] J. Kingston, "The phonetics of Athabaskan tonogenesis," in *Athabaskan Prosody*, S. Hargus and K. Rice, Eds. Amsterdam: John Benjamins, 2005, pp. 137–184.

[20] R. Wright, S. Hargus, and K. Davis, "On the categorization of ejectives: data from Witsuwit'en," *Journal of the International Phonetic Association*, vol. 32, no. 1, pp. 43–77, 2002.

[21] N. Warner, "Acoustic characteristics of ejectives in Ingush," in *4th International Conference on Spoken Language Processing*. New York: IEEE, 1996, pp. 1525–1528.

[22] H. Javkin, "Towards a phonetic explanation for universal preferences in implosives and ejectives," in *Berkeley Linguistic Society (BLS) 3*. Berkeley, CA: University of California, Berkeley Linguistics Society, 1977, pp. 559–565.

[23] R Development Core Team, *R*, Computer program, R Foundation for Statistical Computing, Vienna, Austria, 2020, version 4.0.2, retrieved from http://www.R-project.org/.

[24] R. Bennett, R. Henderson, and M. Harvey, "Vowel deletion as phonologically-controlled gestural overlap in Uspanteko," *Language*, to appear.

[25] D. Bates, M. Maechler, and B. Bolker, *lme4*, R package, 2020, version 1.1-23, retrieved from http://CRAN.R-project.org/package=lme4.

[26] W. N. Venables and B. D. Ripley, *Modern Applied Statistics with S*. New York: Springer, 2002.

[27] L. Campbell, "On glottalic consonants," *International Journal of American Linguistics*, vol. 39, no. 1, pp. 44–46, 1973.

[28] J. Ohala, "The phonetics and phonology of aspects of assimilation," in *Papers in Laboratory Phonology I*, J. Kingston and M. Beckman, Eds. Cambridge, UK: Cambridge University Press, 1990, pp. 258–275.

[1] Some Mayan languages have a *p' b'* contrast, realized phonetically as a contrast between ejective [pʔ] and preglottalized or implosive [ʔb ɓ]. Uvular consonants are limited to Eastern Mayan and Q'anjob'alan languages.