VOT ANALYSIS OF L1 AND L2 SPEAKERS OF ITZA'

Jack Mahlmann [*1]

[1]Department of Linguistics, University of Toronto, Toronto, Ontario, Canada

1 Introduction

This study explores Voice Onset Time of plain and ejective stops (VOT) in Itza' (ISO 639-3: itz), a critically endangered Mayan language spoken in northern Guatemala. VOT is the amount of time (usually measured in milliseconds) elapsed between the release burst of a stop sound such as /p/ or /t/ and the onset of voicing associated with the following vowel; this measurement can vary for the same phoneme across languages, or within a language according to social or other factors. In addition to the plain stops (/p/, /t/, /k/), Itza' has a series of ejective stops (/p'/, /t'/, /k'/) which are produced with the glottalic ingressive airstream mechanism: upward movement of the larynx compresses the air between the glottal folds and an oral constriction (at the velum, for example, in the case of /k'/), creating the characteristic 'popping' sound of an ejective [1].

Ejectives can be seen as a mark of identity for speakers of languages that have them, especially in the case of endangered languages. Hyper-articulation of particular sounds that are unlike those found in the dominant language (often for the benefit of second-language learners) can result in increased awareness of their connections to the speaker's linguistic identity. In the case of SENĆOŦEN (an endangered Salishan language spoken in the Pacific Northwest), Bird [2] found that L2 teachers produced a "stronger" ejective /t'/ than L1 elders, with a longer VOT. A similar effect may occur in Itza', where there are two main populations of speakers: L2 learners and L1 elders. I hypothesize that if L1 speakers of Itza' have a weaker ejective with a shorter VOT, then L2 speakers will produce a stronger one with a longer VOT. However, if L1 speakers have a strong ejective, then I would not expect the production of L2 speakers to differ because stronger ejectives are more salient and less likely to weaken over time [2].

2 Method

2.1 Participants

Elicitation sessions for this study were carried out in San José, Petén, Guatemala. Participants were recruited through the Itza' Linguistic Community (a community organization associated with the *Academia de Lenguas Mayas de Guatemala*). A total of 8 participants were included in the analysis, 4 L1 (mean age of 72) and 4 L2 speakers (mean age of 45). Each participant was compensated with 50 Guatemalan Quetzales.

2.2 Stimuli

All participants were presented with 12 Itza' words with plain/ejective stops in initial position — 6 initial consonants

each followed by /u/ and /a/ — written in the standard orthography. Each word also had an accompanying picture conveying the word intended for elicitation. Words were selected from Hofling & Tesucún's dictionary [3] and approved by a member of the Itza' community. Table 1 shows the full wordlist used.

Table 1: List of Itza' words elicited, with English translations.

Itza'	English
p'ak	*tomato*
p'ul	*smoke*
t'a'	*forehead*
t'ut'	*parrot*
k'ab'	*arm*
k'u'	*nest*
pach	*back*
put	*papaya*
tat	*father*
tup	*earring*
kal	*neck*
kum	*pot*

2.3 Procedure

After signing a consent form, participants self-reported their status as an L1 or L2 speaker of Itza'. They were then instructed to say the name of each item they see using the carrier sentence "Kinwa'lik tech __ ti tech" ("I say __ to you"). During the recorded elicitation session, participants were presented with each word one-by-one, in orthography and with the matching picture at the same time. Once the participant said the word within the provided carrier sentence three times, the stimuli switched to the next word and picture until all 12 words were completed.

2.4 Data analysis

Audio files were analyzed using Praat [4]. VOT for plain and ejective stops in the target words was manually measured from initial release burst of the stop to the onset of voicing for the following vowel. Some tokens were excluded for reasons of audio quality.

The extracted data was analyzed using R [5] to acquire averages and create a mixed-effects model [6] with place of

* jack.mahlmann@mail.utoronto.ca

articulation, following vowel, airstream mechanism, and L1/L2 status as fixed effects, and speaker ID as the grouping variable. 216 tokens were included in the analysis: 99 tokens of plain stops and 117 tokens of ejective stops.

3 Results

Figure 1 below shows the distribution of VOT in milliseconds for plain and ejective stops in Itza', separated by L1/L2 status of the speaker and place of articulation.
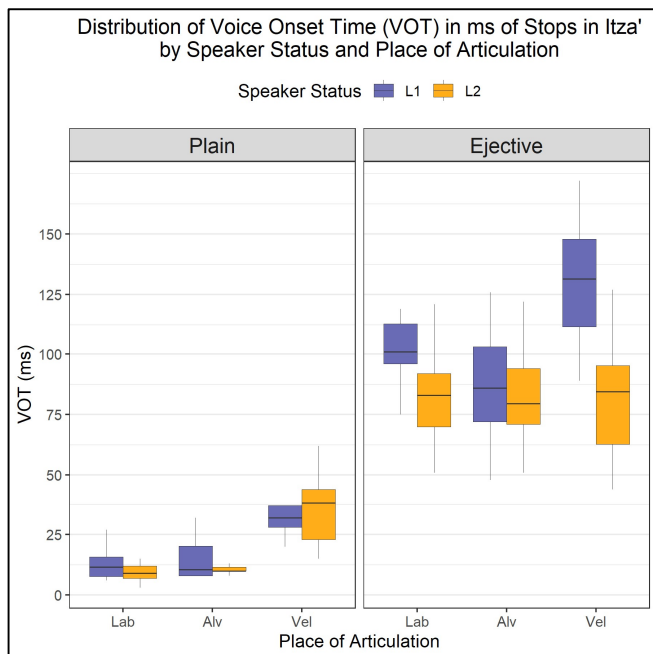

Figure 1: Distribution of VOT (ms) of plain and ejective stops produced by L1 and L2 speakers of Itza' by speaker status and place of articulation.

A higher median value and inter-quartile range can be seen for L1 speakers as opposed to the L2 speakers for the ejectives, particularly at the labial and velar places of articulation. Differences between L1 and L2 speakers in production of plain stops are less prominent.

The mixed-effects model found that L1/L2 status was a significant factor at the $p < 0.01$ level, with the isolated effect being that L1 speakers have a +16 ms higher VOT for stops than L2 speakers. The difference between labial and alveolar places of articulation was not significant, however there was a significant effect ($p < 0.001$) for the velar place of articulation, which is to be expected. Airstream mechanism also had a significant effect at the $p < 0.001$ level; the isolated effect of airstream was that ejectives had a VOT of +72 ms when compared to plain stops. Additionally, the following vowel was also significant at the $p < 0.05$ level. Stops with a following /a/ were found to have a VOT of +5 ms when compared to stops with a following /u/.

4 Discussion

The results of the statistical analysis stand in contrast with the findings of Bird's SENĆOŦEN study [2]. An apparent time interpretation (coupled with the understanding that L1 speakers are few in number and all elderly, and generational transmission will be restarted by L2s) suggests that Itza' ejectives are moving towards the VOT criteria more associated with weaker ejectives. This information could be used to provide Itza' language instructors with advice for L2 students for whom it is a goal to sound as native-like as possible.

5 Conclusion

This study has found that VOT for ejective stops in Itza' is higher in the L1 speakers analyzed than the L2 speakers, while differences in plain stops are less prominent. If one of the goals of the language revitalization efforts is to preserve the way of speaking of the community elders, then Itza' language educators have an opportunity to modify this aspect of their approach to ensure that ejectives are retained in the language in the long-term such as by intentionally hyper-articulating these sounds to increase their salience. While this type of change — the potential loss of a contrasting feature — is natural and doesn't represent a corruption or bastardization of the language in the eyes of linguists, it is important to consider the needs of the language community being served, as determined by members of the community itself.

Acknowledgments

References

[1] Fallon, P. D. (1998). The Synchronic and Diachronic Phonology of Ejectives [PhD Thesis]. The Ohio State University.

[2] Bird, S. (2020). Pronunciation among adult Indigenous language learners. Journal of Second Language Pronunciation, 6(2), 148–179.

[3] Hofling, C. A., & Tesucún, F. F. (1997). Itzaj Maya-Spanish-English Dictionary. University of Utah Press.

[4] Boersma, P., & Weenink, D. (2022). Praat: Doing phonetics by computer (6.3.02).

[5] R Core Team. (2022). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. https://www.R-project.org/

[6] Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1–48.