VOT ANALYSIS OF L1 AND L2 SPEAKERS OF ITZA': PRELIMINARY RESULTS

Jack Mahlmann
University of Toronto

1. Introduction

Itza' is a language in the Mayan family spoken in the Petén department of Guatemala's far north. Its speakers are mainly found surrounding Lake Petén, in the center of the region. Belonging to the Yucatecan branch, Itza' is closely related to Yucatec, Lacandon, and Mopan (Bricker 2019, Clemens 2019, Sobrino Gómez 2018). Ethnologue reports only 410 speakers of Itza' (Eberhard et al. 2023), and other sources report that the number of L1s may be as low as a few dozen (Gonzáles Toledo 2021, Hofling 1996, Mateo Pedro 2023); the language is critically endangered.

Of the 31 Mayan languages listed on Ethnologue, Itza' has the second lowest number of speakers for those which are not extinct. Other languages in the Yucatecan branch are not struggling in this way, as Yucatec itself has 775,000 speakers in Mexico (Eberhard et al. 2023). Yucatec, Lacandon, and Mopan all have an EGIDS rating of 5 ("Developing"), meaning that the language is in vigorous use but the written standardized form is not widely accessible. In contrast, Itza's rating is 8a ("Moribund"), conveying that generational transmission has ceased and therefore the future of the language is uncertain; despite this, revitalization efforts are ongoing. This study explores the production of one particular feature in the language – ejective stops – by L1 and L2 speakers to learn how pronunciation varies between the two groups.

1.1 History of the Itza' people and their language

Prior to its capture in 1697, the island in Lake Petén was the host of the last independent Mayan polity in the entire New World (Jones 1998). In the years since, colonial and post-colonial policies aimed to suppress the language in favor of Spanish; in the 1930s the Petén department banned the language entirely, initiating rapid language shift to Spanish (Hofling 1996, Schumann Gálvez 2000). "La Violencia" ('The Violence') in the early-mid 1980s was a phase of the Guatemalan Civil War described as the "most brutal expression to date" of violence against the country's Indigenous peoples (Zur 2019). With this history in mind, it is clear why the Itza' language is at risk of extinction.

Although Ethnologue describes Itza' as "moribund", revitalization efforts have been ongoing since the easing of tensions in the late 1980s and early 1990s. In October 2022, the latest development unfolded: a language immersion project run by the Comunidad Lingüística Itza' ('Itza' Linguistic Community'), supported by Dr. Pedro Mateo Pedro from the University of Toronto and funded by Connaught. This project is a part of the National Plan for the Decade of Indigenous Languages 2022-2032 organized by

Actes du congrès annuel de l'Association canadienne de linguistique 2023.
Proceedings of the 2023 annual conference of the Canadian Linguistic Association.

UNESCO (Dirección de Lingüística y Comunicación 2022). While there are very, very few L1 speakers left, there is a community of long-time L2 learners which this opportunity to teach the language appeals to; therefore, the demand for language learning cannot necessarily be met due to the supply of native speakers. For this reason, a relevant question to ask is whether certain sounds in the Itza' language are acquired more easily than others due to L1 influence, and if such effects make the language as it is currently being acquired by learners different from how it is spoken by the native speakers of the community.

1.2 Language endangerment and pronunciation

Minority languages being affected by a dominant language is well-documented, especially in the case of sounds which are similar between the two (Maclagan and King 2007, Nance and Stuart-Smith 2013). In the case of Itza' and the dominant language Spanish, ejectives are an entirely new sound category that L1 Spanish speakers must acquire. There is also evidence of marked features becoming more salient in L2 learners of a language, possibly as a result of certain instructional techniques; this can be especially true in the context of a language which is critically endangered (Wolfram 2008).

This is especially relevant for members of Indigenous communities, where speaking authentically (in a more similar manner to L2 elders) may be an additional goal alongside fluency (Bird 2020). Certain linguistic features may become something of an identity marker for speakers of languages that have them due to their distinctiveness when compared to the phonology of the dominant language. Glottalization often has this status, not because it is marked within the endangered language itself, but because it is marked in the context of the dominant language (Woolard 1989).

Like a typical Mayan language, Itza' has a series of ejective stops /p' t' k'/. Due to the absence of such sounds in Spanish and the current circumstances that the language is in, an examination of the acquisition of these sounds was worthwhile. This is most apparent because studies of other languages like SENĆOŦEN and Q'eqchi' have reported that there are measurable differences in how these sounds are produced by native and non-native speakers with respect to Voice Onset Time (Bird 2020, Wagner and Baker-Smemoe 2013). In the SENĆOŦEN study, Bird found higher average VOT for ejective /t'/ as produced by L2 teachers as opposed to L1 elders and interpreted this as the sound strengthening over time (2020). According to Kingston (1985), one characteristic associated with a strong rather than weak ejective (sometimes called stiff and slack) is a longer VOT; Bird (2020) suggests that the effect observed in SENĆOŦEN can be a result of hyperarticulation, emphasis, or exaggeration of these sounds in L2 teaching contexts. Due to Itza' being in a similar situation as SENĆOŦEN with very few L1 speakers but a notable community of L2 learners, it was plausible that such an effect may be happening in this language as well.

1.3 Research questions and predictions

The implications of this research are relevant on an academic level as a study of L2 language production, but they are also important due to their connection to the ongoing revitalization movement. The specific impact depends on the attitude of the community towards the language's ejective sounds, however. If the Itza' community members consider ejective sounds in the way they are produced by elders as worth preserving then identifying if the way L2 speakers produce them is different from L1 speakers would be assisting with a community-defined goal. To this end, comparing VOT measurements across L1 and L2 speakers will provide insight into the language's ejectives and provide typological information as well as results that are relevant to the Itza' community and their revitalization efforts. On the contrary, if Itza' speakers do not place any particular significance on the L1 realization of ejectives, then this study's results could still provide a look into the future of the language with regards to these sounds' place in the consonant inventory.

The research questions for this project are regarding whether L1 and L2 pronunciation of ejectives in Itza' differs along the dimension of VOT, and additionally whether such findings have implications on second-language learning in the context of the ongoing Itza' language revitalization movement. Based on Bird's results from SENĆOŦEN /t'/, we'd expect that if L1 speakers have a lower VOT for ejectives (like a weaker ejective), then VOT will be higher in L2s due to hyperarticulation, exaggeration, or emphasis in second-language learning contexts (2020). However, if L1 VOT is already high like in the closely-related Yucatec language (González Poot 2011), no change is expected due to the increased salience that is associated with stronger ejectives and their resistance to weakening. This project sought to answer both questions with a phonetic study of Itza' ejective and plain stop consonants, measuring their VOT.

2 Methodology

Elicitation sessions were carried out in Spanish by an associate of the University of Toronto's Revitalization Project in San José, Petén, Guatemala. Ten sessions were conducted with self-identified L1 and L2 speakers of Itza'; five L1s and five L2s were interviewed. The mean age for L1 speakers was 72, while the mean age for L2 speakers was 45. One L1 speaker was excluded for reasons of audio quality and one L2 speaker was excluded due to the fact that they had been learning the language for less than two years. Participants were compensated with 50 GTQ for their time.

Both groups were subject to the same stimuli: 12 Itza' words with both orthographic and visual representations (images), all of which had initial ejective or plain stops. Two words were chosen for each stop, one followed by /a/ and one followed by /u/. All words were sourced from Hofling and Tesucún's (1997) Itza' dictionary.

Table 1. List of words used in elicitation, with English translation and IPA transcription.

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

When prompted by the printed word and visual aid, the participant was instructed to say the word three times using the carrier sentence *Kinwa'lik tech* __ *ti tech* "I say __ to you". Each word was presented one-by-one until reaching the end of the list.

VOT was segmented using Praat (Boersma and Weenink 2023) by measuring from the release burst of the initial stop to the onset of voicing associated with the following vowel. Some tokens were excluded for reasons of audio quality; each speaker contributed an average of 14.6 ejective tokens and 12.4 plain stop tokens. In total, 216 VOT measurements were extracted and coded for the stop's place of articulation and airstream mechanism, the following vowel, and the Speaker Status (L1/L2) of the participant. The dataset was analyzed using R (R Core Team 2022) and two mixed-effects models (Bates et al. 2015) were created separating plain stops and ejective stops. Place of articulation, following vowel, and Speaker Status were the fixed effects, and speaker ID was a random effect. Graphs were created using ggplot2 (Wickham 2016).

3. Results

3.1 Overall findings

Figure 1 provides box plots for all data collected, separated by airstream mechanism, place of articulation, and Speaker Status. Across the ejective phonemes, median values for L1 speakers appear to be higher than those for L2 speakers, especially at the labial and velar places of articulation. For the labials and velars, the inter-quartile ranges (the solid sections above and below the medians) do not overlap at all. For all three ejective stops, the highest value recorded was produced by an L1 speaker and the lowest value recorded was produced by an L2 speaker. Interestingly, the distribution of the L2 speakers' VOT for ejective stops is very similar across all three places of articulation – hovering around 80ms. This is most surprising for the velars, for which VOT values are

usually higher across languages due to the vocal fold opening duration being fixed and the shorter closure time for velar stops (Maddieson 1997).

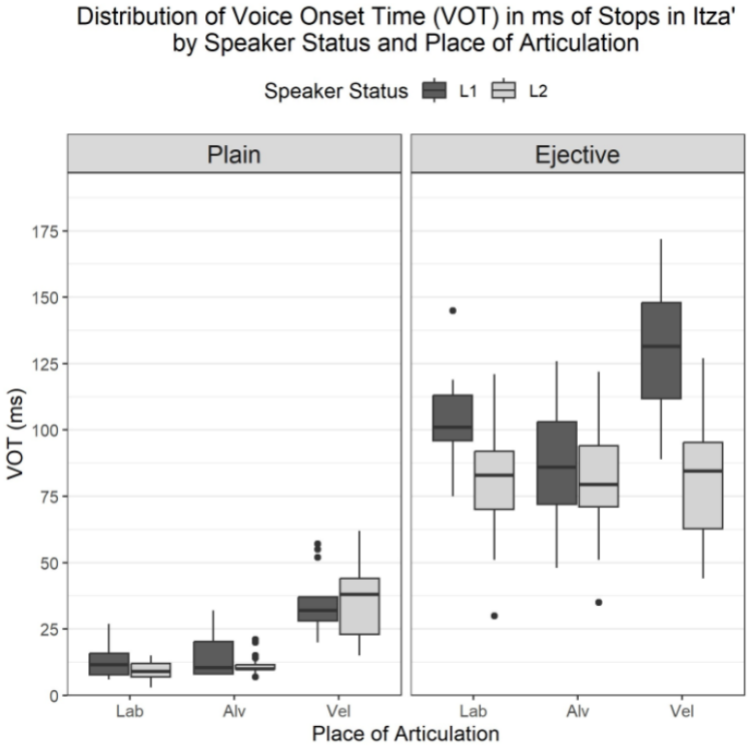

Figure 1. Distribution of VOT (ms) of plain and ejective stops in Itza', separated by Speaker Status.

Figure 2 and Figure 3 below show two tokens of the labial ejective in "p'ak" produced by a L1 speaker and L2 speaker, respectively. The L1 speaker's token has a VOT of 100ms, while the L2 speaker's token has a VOT of 78ms.

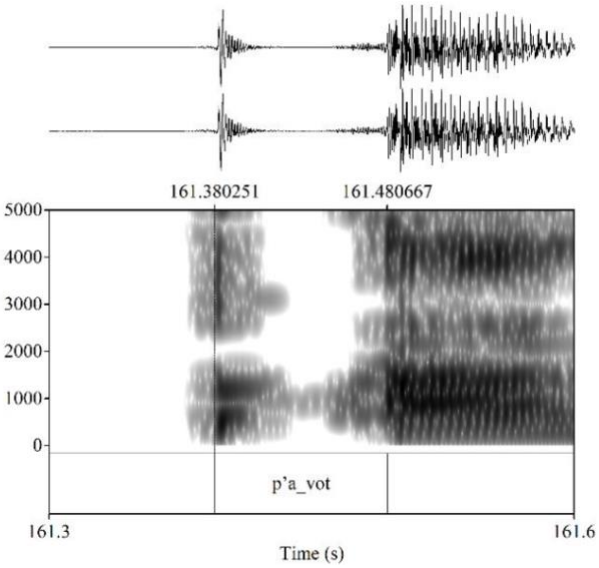

Figure 2. Spectrogram and waveform of a "p'ak" token from an L1 speaker showing a VOT of 100ms.

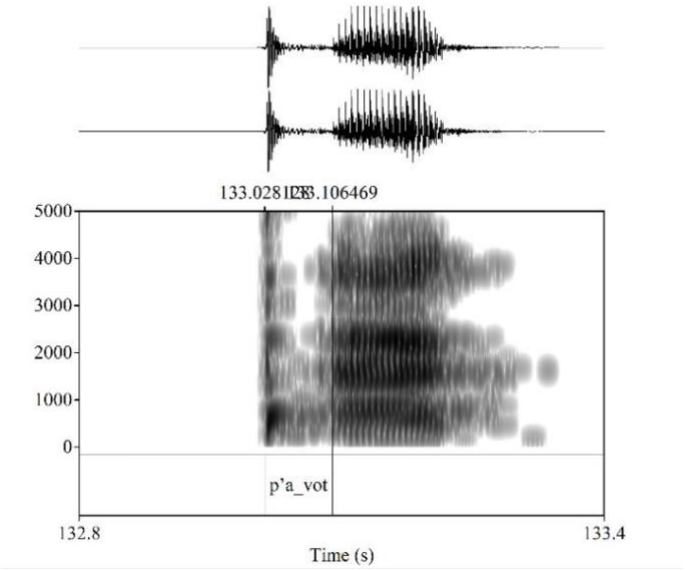

Figure 3. Spectrogram and waveform of a "p'ak" token from an L2 speaker showing a VOT of 78ms.

3.2 Mixed-effects models

This section provides the results from the two mixed-effects models used in the study. As explained above, each token was coded for place of articulation, airstream mechanism, following vowel, and Speaker Status. These four variables were the fixed effects in the mixed-effects models, while speaker ID was the lone random effect. If the predictions made in the Introduction section were to be true, the most obvious way for the relevant evidence to appear would be for the Speaker Status factor to have a significant effect in the model for ejective VOT but not in the plain stop model.

4.3 Ejective VOT model

The ejective VOT model included 117 tokens. Firstly, Speaker Status was statistically significant at the $p < 0.05$ level ($p = .0199$) with an estimated effect of -27ms; this indicates that, isolated from the effects of the other variables, L2 speakers have a VOT for ejectives that is 27ms shorter than L1 speakers. The quality of the following vowel was also found to be a significant factor, at the $p < 0.001$ level ($p = .00098$). This factor's estimated effect was -12ms, suggesting that ejective stops followed by vowel /a/ are 12ms higher than those followed by /u/. The velar place of articulation in comparison to alveolar was statistically significant at the $p < 0.001$ level ($p = .0009$), while there was no significant difference between the alveolar and labial places of articulation ($p = .0922$). Table 2 below shows the full table of results for the ejective stop model's fixed effects. The baseline factor levels for the model were vowel /a/, alveolar place of articulation, and L1 Speaker Status.

Table 2. Regression table for the results of the ejective stop model's fixed effects.

	Est.	**Std. Error**	**Deg. Fr.**	**T-value**	**P-value**	**Sig.**
Intercept	105.531	6.937	9.474	15.232	5.66e-08	***
Vowel	-12.201	3.601	107.383	-3.388	0.000984	***
Labial	7.324	4.310	106.571	1.699	0.092182	.
Velar	14.913	4.367	108.344	3.415	0.000899	***
Status	-26.986	8.565	5.981	-3.151	0.019880	*

3.4 Plain stop VOT model

Ninety-nine tokens of plain stops were included in this statistical model. This model diverges from the ejective model in that it does not mark Speaker Status as significant ($p = .3985$). Following vowel quality was significant in this model at the $p < 0.05$ level ($p = .0192$), however it had the opposite effect; plain stops followed by /u/ were 4ms higher than plain stops followed by /a/. The velar place of articulation is also significant at the $p < 0.001$ level in this model as well. Table 3 below shows the full table of results for the

plain stop model's fixed effects. The baseline factor levels for the model were vowel /a/, alveolar place of articulation, and L1 Speaker Status.

Table 3. Regression table for the results of the plain stop model's fixed effects.

	Est.	Std. Error	Deg. Fr.	T-value	P-value	Sig.
Intercept	13.4530	3.2953	7.4440	4.082	0.0041	**
Vowel	3.6864	1.5457	88.9424	2.385	0.0192	*
Labial	-0.6777	1.7941	87.9815	-0.378	0.07065	
Velar	23.3598	1.8523	88.5979	12.612	<2e-16	***
Status	-3.9231	4.2714	5.2659	-0.918	0.3985	

4. Discussion

Firstly, with regards to the aforementioned ejective typology, it appears that Itza' ejectives are more similar to stiff ejectives such as in Hupa, Navajo (Gordon et al. 2001), and Yucatec (González Poot 2011). In a previous study of Itza's ejective VOT (involving one Itza' speaker) by Burnett-Deas (2009), Itza' VOT was found to be higher than Yucatec VOT. This study corroborates this in finding relatively high VOT values; even higher than those reported by Burnett-Deas' study in the case of ejectives. Therefore, if Yucatec has stiff ejectives, then Itza' is likely to have stiff ejectives too – at least based on the dimension of VOT.

Another relevant comparison is to Bird's study of SENĆOŦEN. To review, Bird (2020) found that L2 teachers consistently produced /t'/ as a strong ejective (with longer VOT) while the L1 elders were more variable in their pronunciation. Previous descriptions of SENĆOŦEN referred to the ejectives as weak, so Bird explains this as a change over time induced by two things: firstly, the desire for retention of ejectives specifically due to their status as a distinguishing feature of the language, and secondly, hyperarticulation in second-language teaching used as a tool to make sure that learners can hear the sounds they are learning more clearly. Because this involves modifying the acoustic characteristics of the given sound, it is unsurprising that this could have long term effects on the way that learners acquire Itza'.

This study has found the opposite results when compared to SENĆOŦEN: based on VOT measurements, the ejectives produced by the L2 speakers involved in this study are more slack those of the L1 speakers. At all three places of articulation, the median VOT value for the L1 speakers is higher than that of the L2 speakers. Additionally, the mixed-effects models reported Speaker Status having a significant effect on ejective stop VOT ($p < 0.05$) but not on plain stop VOT. Therefore, the answer to the second research question is that ejective stops are measurably different when comparing L1 and L2 speakers; L1 speakers have longer VOT for initial ejectives than L2 speakers.

With respect to the second research question, an apparent time interpretation of these results would suggest that Itza' language as acquired by the elders featured stiff ejectives, but the language as it has been transmitted to the L2 speakers has slacker

ejectives despite the salience of stiffer ejectives as produced by the L1s in the community. Therefore, if the community decides that one of their goals is to preserve the language as it is spoken by the elders, then perhaps curriculum changes to target pronunciation of ejectives may help.

As mentioned in the methodology section, before the elicitation session the participants filled out an information sheet, and one of the questions was about where they learned Itza' (for the non-native speakers). Several of the speakers mentioned that they learned from the Itza' Linguistic Community, or the ALMG. One particular speaker said that they trained using the "total immersion method", which is the teaching style employed at the new language school in the Itza' community. Interestingly, this speaker had the lowest median and average VOT values across all three places of articulation. In other words, the only speaker which we know for certain was involved in the total immersion method had the most depressed VOT values. Because the main vehicle for Itza' transmission going forward will likely be second-language programs, it seems that the language has the potential to sound significantly different in the future as the community-leading roles are passed down to current L2 speakers.

5. Conclusion

In conclusion, this project's goal was to contribute meaningfully to the Itza' revitalization effort while also performing research that is relevant to the scholarship of the endangered languages and their phonetics, specifically with regards to ejectives. This class of sounds can have particular importance to speakers of Indigenous languages that have them; they can be relevant to identity formation with respect to their being distinctly non Spanish/non-English (Bird 2020). To these communities, it may also be a goal to learn the language as it is spoken by the elders, moving past just becoming intelligible to others. The realization of ejectives across languages is dynamic as highlighted by Kingston (1985); these cross-linguistic differences in VOT (as well as following vowel f0, burst intensity, etc.) have implications on the acquisition of ejective sounds in second-language contexts. Slack ejectives can be harder to hear, which may influence teachers to hyperarticulate their pronunciation to exaggerate the contrast between them and plain stops (Bird 2020).

A phonetic study was conducted to detail the differences (if any) between L1 and L2 pronunciation of the language's ejective stops, particularly along the dimension of VOT. Eight Itza'-speaking participants were included in the analysis, Four L1 speakers and four L2 speakers. The stimuli were 12 Itza' words which had initial ejective or plain stops, followed by either vowel /a/ or /u/. A carrier sentence was used and three repetitions were requested for each word from the participant. VOTs were segmented using Praat (Boersma and Weenink 2023) and mixed-effects models (Bates et al. 2015) were created using R (R Core Team 2022), to find that there was a difference between L1 and L2 VOT for ejectives: Speaker Status (L1/L2) was found to be significant for the ejective stops but not for the plain stops, with L1 VOT being higher overall than L2 VOT in ejectives.

Therefore, the answer to the first research question regarding L1 and L2 VOT of Itza' ejectives is that there are measurable differences across the two groups. As for the second research question, the phonetic implications (of ejectives becoming weaker) are meaningful only if the community decides that it is a goal for the language to be preserved as it is spoken by the elders. The obvious next step is to consult with Itza' speakers and community leaders to determine where to go from here; these results are not meant to recommend a course of action to the Itza' community, only to provide information that may be useful for their decision-making in the future.

References

Bates, Douglas, Martin Mächlet, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1): 1–48.

Bird, Sonya. 2020. Pronunciation among adult Indigenous language learners. *Journal of Second Language Pronunciation* 6(2): 148–179.

Boersma, Paul, and David Weenink. 2023. Praat: Doing phonetics by computer.

Bricker, Victoria R. 2019. *A Historical Grammar of the Maya Language of Yucatan: 1557–2000*. Salt Lake City: University of Utah Press.

Burnett-Deas, Andrea. 2009. *VOT of ejectives, implosives, and plain stops in Yukateko, Mopan, and Itzaj Maya*. Carbondale, Illinois: Southern Illinois University.

Clemens, Lauren. 2019. Mayan Languages. In *The Routledge Handbook of North American Languages*, ed. Daniel Siddiqi, Michael Barrie, Carrie Gillon, Jason Haugen, and Éric Mathieu, 365–396. New York: Routledge.

Dirección de Lingüística y Comunicación. 2022. *Tojmeyay Muk'yool Tikilu'umil ti Lajun ichiltanxelu'umil ti T'anoo Masewal ti Paxil 2022–2032*. Academia de Lenguas Mayas de Guatemala.

Eberhard, David M., Gary F. Simons, and Charles D. Fennig, eds. 2023. *Ethnologue: Languages of the World*. 26th ed. Dallas: SIL International.

Gonzáles Toledo, Juan Miguel Salvador. 2021. Palabras Inaugurales. Presented at the Hacia la revitalización del idioma Itza': contribuciones desde otros lugares de América Latina. https://www.funproeibandes.org/index.php/2021/07/19/seminario-internacional-hacia-la-revitalizacion-del-idioma-itza-contribuciones-desde-otros-contextos/.

González Poot, Antonio A. 2011. *Conflict resolution in the Spanish SLA of Yucatec ejectives: L1, L2 and universal constraints*. Doctoral dissertation, University of Calgary.

Gordon, Matthew, Brian Potter, John Dawson, William de Reuse, and Peter Ladefoged. 2001. Phonetic structures of Western Apache. *International Journal of American Linguistics* 67(4): 415–448.

Hofling, Charles A. 1996. Indigenous linguistic revitalization and outsider interaction: The Itzaj Maya case. *Human Organization* 55(1): 108–116.

Hofling, Charles A., and Félix Fernando Tesucún. 1997. *Itzaj Maya-Spanish-English Dictionary*. Salt Lake City: University of Utah Press.

Jones, Grant D. 1998. *The conquest of the last Maya kingdom*. Stanford, CA: Stanford University Press.

Kingston, John. 1985. *The phonetics and phonology of the timing of Oral and Glottal Events*. Doctoral dissertation, University of California, Berkeley.

Maclagan, Margaret, and Jeanette King. 2007. Aspiration of plosives in Māori: Change over time. *Australian Journal of Linguistics* 27: 81–96.

Maddieson, Ian. 1997. Phonetic universals. In *The handbook of phonetic sciences*, ed. John Laver and William J. Hardcastle. Oxford: Blackwell.

Mateo Pedro, Pedro. 2023. Revitalization of Itzaj: An Endangered Mayan language of Guatemala. Poster presented at the annual conference of the Canadian Linguistic Association, York University, Toronto.

Nance, Clare, and Jane Stuart-Smith. 2013. Pre-aspiration and post-aspiration in Scottish Gaelic stop consonants. *Journal of the International Phonetic Association* 43(2): 129–152.

R Core Team. 2022. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Schumann Gálvez, Otto. 2000. *Introducción al Maya-Itzá*. Mexico City: Universidad Nacional Autónoma de México, Instituto de Investigaciones Antropológicas.

Sobrino Gómez, Carlos Martin. 2018. *El sistema fonológico protoyucatecano: reconstrucción y desarrollo*. Doctoral dissertation, El Colegio de México.

Wagner, Karl Olaw Christian, and Wendy Baker-Smemoe. 2013. An investigation of the production of ejectives by native (L1) and second (L2) language speakers of Q'eqchi' Mayan. *Journal of Phonetics* 41: 453–467.

Wickham, Hadley. 2016. *ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag.

Wolfram, Walt. 2008. Language death and dying. In *The handbook of language variation and change*, ed. J. K. Chambers, Peter Trudgill, and Natalie Schilling-Estes, 764–787. Oxford: Blackwell.

Woolard, Kathryn A. 1989. Language convergence and language death as social processes. In *Investigating obsolescence: Studies in language contraction and death*, ed. Nancy C. Dorian, 355–368. Cambridge: Cambridge University Press.

Zur, Judith. 2019. *Violent memories: Mayan war widows in Guatemala*. New York: Routledge.