



UNIVERSITY OF
TORONTO

Investigation of vowel length contrast in online learning materials: The case of Itza' revitalization

Jack Mahlmann, Laura Griffin, Radu Craioveanu, & Pedro Mateo Pedro
University of Toronto

Roadmap

Itza' language background & revitalization

Pedagogical grammar

Quantifying the diaspora

Data & Methodology

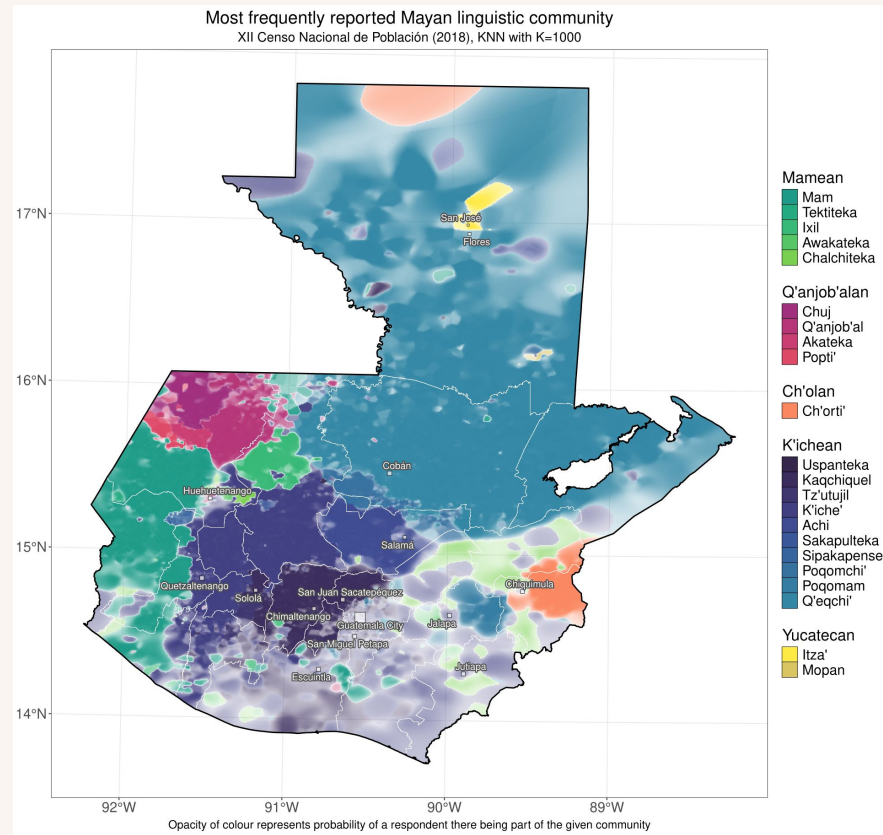
Results & Analysis

Discussion

References

Itza' language

- Itza' is an endangered Mayan language of the Yucatecan branch, spoken in Petén, northern Guatemala, around Lake Petén Itzá
 - (bright yellow region at right)
- L1 elders rarely speak Itza' among themselves, but there is a growing number of young adult L2 learners
- Children are not acquiring the language at home from parents or caregivers anymore



Itza' language

- Within the community, there are currently only 37 known speakers (Comunidad Lingüística Itza' 2021)
 - 16 L1 speakers (~70-80 years old)
- 21 proficient L2 speakers (as judged by Comunidad Lingüística Itza')
However, the 2018 Guatemalan census suggests more (INE, 2018)
 - 406 said Itza' was the language in which they first learned to speak
 - 2,926 self-identified as part of the Itza' linguistic community
 - But only 65 of these also said Itza' was their L1
 - We'll see some patterns in this data a little later

Revitalization efforts

- Heavily suppressed in both the colonial and post-colonial periods in Guatemala
 - In the 1930s there was a ban on speaking the language in the entire Petén department (Hofling, 1996)
- Revitalization movements began in the 1990s, such as the Itza' Maya Language Rescue Project led by local and non-local experts (Hofling, 1996, p. 112)
- The Academia de las Lenguas Mayas de Guatemala (ALMG) was involved because of the language's dire state, and also because the Itza' were the last independent Mayan group to fall to the Spanish (Jones, 1998, p. xix)

Revitalization efforts

- 2022: new “total immersion method” introduced in the community
 - All teachers are L2 speakers
 - Some L1 speaker involvement, but not ongoing/constant
- 2024: new pedagogical grammar published for use in schools, following a new immersion teaching method based on the Oxlajuj Aj immersion program (Maxwell 2012, Maxwell & Little 2006, Maxwell, et al. 2014)
 - L1 speakers were involved in proof-reading
 - online version of this grammar coming in 2025
 - funded by the University of Toronto’s Itza’ Revitalization Project (Connaught grant)

Itza' Phonology

		Labial	Alveolar	Palatal	Velar	Glottal
Stops	voiceless	p	t		k	ʔ
	glottalized	p'	t'		k'	
	voiced	b	(d)		(g)	
Affricates	voiceless		ts	tʃ		
	glottalized		ts'	tʃ'		
Fricatives	voiceless	(f)	s			h
Liquid			l			
Rhotic			r			
Nasals		m	n			
Semivowels		w		j		

(Hofling & Fernando Tesucún 2000: 3; Hofling, 1991, 2000; Mahlmann, 2023; Schumann Gálvez, 2000)

Itza' Phonology

	Front		Central		Back	
High	i	ii		i	u	uu
Mid	e	ee			o	oo
Low			a	aa		

(Hofling & Fernando Tesucún 2000: 3; Hofling, 1991, 2000; Mahlmann, 2023; Schumann Gálvez, 2000)

Itza' Phonology

- Ejectives/implosives, the lax vowel /i/, and vowel length all appear in Itza' but not in Spanish
 - Plain-ejective contrast is still robust in L2s but may be weakening (Mahlmann, 2023)
 - Virtually no recent research on vowels or their acquisition by L2 speakers

Honing in on vowel length...

What is the process of acquiring vowel length like for L2 learners?

L2 Acquisition of Vowel Length

- For speakers of non-quantity-sensitive languages, L2 acquisition of length contrasts is more difficult (McAllister et al., 2002)
- Vowel length is also susceptible to loss through contact-induced effects (cf. Romani in Matras, 2002)
- Vowel length is prominently marked in Itza' orthography

<i>b'ox</i>	<i>b'oox</i>	<i>chiich</i>	<i>chich</i>
/boʃ/	/boːʃ/	/tʃiːtʃ/	/tʃitʃ/
“black”	“tilapia”	“liquor”	“hard”

(Hofling & Tesucún, 1997)

“National Day of Indigenous Languages”



(Comunidad Lingüística Itza', 2025)

Pedagogical Grammar

- Grammar for pedagogical purposes written by Pedro Mateo Pedro and his team, published by the Guatemalan Ministry of Education (Mateo Pedro et al., 2024)
 - Modeled after Oxlajuj Aj method used in the Kaqchikel grammar (Brown et al., 2006)
- Used by the language revitalization schools, and will soon be released digitally with audio files
- We are presenting findings here based on the audio files from the first 5 chapters (vocabulary lists and dialogues), as well as supplementary data from a translation task which is unrelated to the grammar



Vocabulary

Uka' ka'ansajil: k'ab'a'

Itza'	Kaslan
1. inten	yo
2. intech	tú/usted
3. la'ayti'	él/ella
4. into'on	nosotros
5. inte'ex	ustedes
6. la'ayti'oo'	ellas/ellos
7. k'ab'a'	nombre
8. b'ix	cómo
9. b'a'ax	qué, cuál
10. max	quién
11. ajkänb'al	estudiante
12. ajkänb'aloo'	estudiantes
13. ajka'ansaj	maestro
14. ajka'ansajoo'	maestros
15. inwojel	yo sé
16. jaj	sí
17. ma'	no

k'ab'a'	nombre
ink'ab'a'	mi nombre
ak'ab'a'	tu nombre
uk'ab'a'	su nombre
kik'ab'a'	nuestro nombre (exclusivo)
kik'ab'a'e'ex	nuestro nombre (inclusivo)
ak'ab'a'e'ex	su nombre de ustedes
uk'ab'a'oo'	su nombre de ellas/ellos

Dialogue

Tzikb'al yok'ol ti ka'ansajil

I

- Aj Merex: Te' yoos.
 Ix Pet: Te' yoos.
 Aj Merex: ¿B'ix yanech?
 Ix Pet: Jach ma'lo'. ¿Intech b'ix ab'el?
 Aj Merex: Ki' inwool.
 Ix Pet: ¿B'ix ak'ab'a'?
 Aj Merex: Ink'ab'a' aj/ix ¿Intech b'ix ak'ab'a'?
 Ix Pet: Ink'ab'a' Ix Pet.
 Aj Merex: Inwojel ak'ab'a'.
 Ix Pet: Tak samal.
 Aj Merex: Tak samal.

II

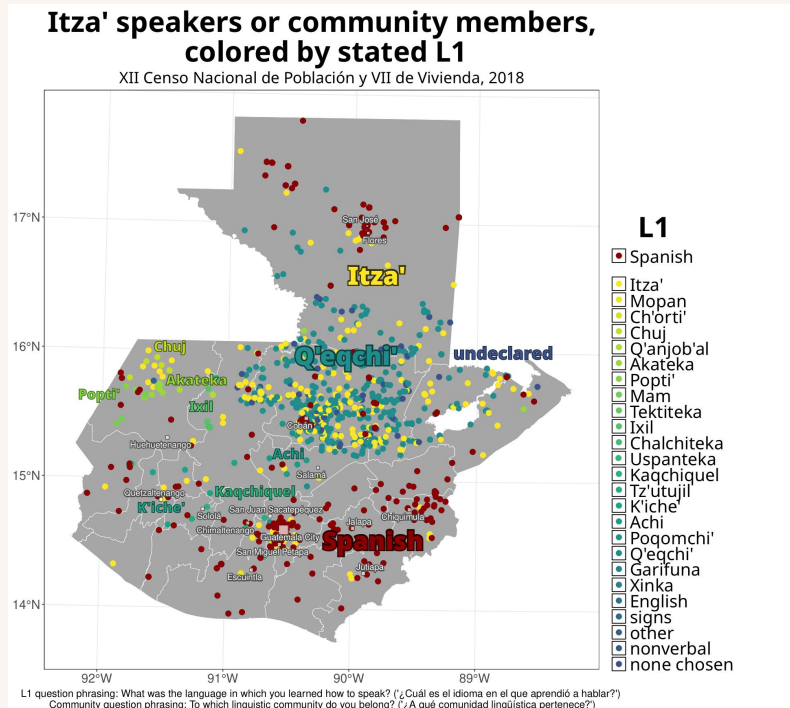
- Ix María: Te' yoos.
 Aj Pedro: Te' yoos.
 Ix María: ¿B'ix ab'el?
 Aj Pedro: toj inwool, ¿Intech b'ix ab'el?
 Ix María: inten xan
 Aj Pedro: ¿B'ix ak'ab'a'?
 Ix María: Ink'ab'a' aj/ix ¿Intech b'ix ak'ab'a'?
 Aj Pedro: Ink'ab'a' Aj Pedro
 Ix María: Jaj, inten Ix María.
 Aj Pedro: Tak samal
 Ix María: Tak samal

K'atalil

1. K'atej ti kántuul makoo' uk'ab'a'oo'.
2. Tz'iib'tej ka'p'eel sak'unoo' tu'ux uk'at a' k'ab'a'oo' ti makoo'.
3. "Peksäb'ajiloo' ti uya'lik".

Implications for the Diaspora

- Itza' diaspora speakers can be found in Guatemalan census data (INE, 2018)
- The largest number of people who are both Itza' L1 speakers and in the Comunidad Lingüística Itza' are in Petén
- The recorded dialogues and vocabulary lists included in the digital version are especially important for these learners



Guiding Questions for our Research

1. What type of input are L2 learners getting to help them acquire the Itza' vowel system?
 - How is the short/long vowel contrast represented in material directed to learners?
2. Who are the learners? What does the Itza' diaspora look like in Guatemala?
3. To what extent are L2 learners acquiring this contrast?
4. How can we improve input materials for learners and support language revitalization efforts through our phonetic analysis?

Predictions

- In the case of the learning materials, there may be hyperarticulation due to the educational context or a similar effect due to influence from the orthography
- Predictions depend on the context:
 - In vocabulary lists, hyperarticulation will exaggerate the duration difference
 - In dialogues, the contrast will be weaker because the words are not in isolation
 - In translation tasks, it is hard to say. Likely highly variable depending on many factors

Methodology

- Vowels segmented from 25 files total, in three contexts
- Dialogues
 - Two or more L2 teachers participate in a rehearsed conversation
 - 13 files, 7 minutes
- Vocabulary lists
 - One L2 speaker reads a list of vocabulary items for each chapter
 - 10 files, 18 minutes
- Translation task
 - One L2 speaker translates words from Spanish to Itza' when prompted by interviewer
 - Assortment of nouns & verbs, drawn from Itza' dictionary (Hofling & Tesucún 1997; word list in appendix)
 - 1 file, 25 minutes
 - NOT part of the digital grammar, but represents more naturalistic productions
- 1331 total vowel tokens, excluded lax /i/ and rearticulated /VʔV/



Methodology

- Length segmented using **Praat** (Boersma & Weenink, 2025)
- Coded for vowel quality and phonemic length category
- Data analyzed in **R**, figures use **ggplot2** (R Core Team, 2015; Wickham, 2016)

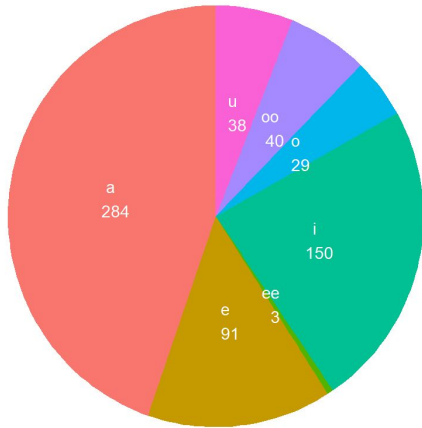
Token Counts

- Short vowels were much more common, long vowels comparatively scarce
- Dialogues do not include any /aa/, /ii/, or /uu/
- Vocabularies and Translation task have a more even distribution, but still skewed towards short vowels

Column Labels ▾		Dialog		Transl.		Vocab	
Row Labels ▾	Sum of n_V	Sum of n_VV	Sum of n_V	Sum of n_VV	Sum of n_V	Sum of n_VV	
a	284	0	58	25	122	7	
e	91	3	29	10	64	26	
i	150	0	59	7	77	21	
o	29	40	38	18	41	17	
u	38	0	34	12	30	1	
Grand Total	592	43	218	72	334	72	

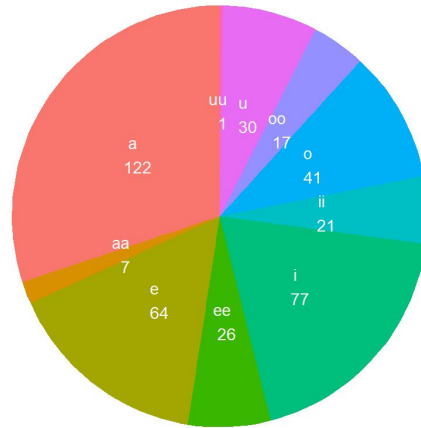
Summary of Distributions

Occurrence of Long and Short Vowels in Dialogues



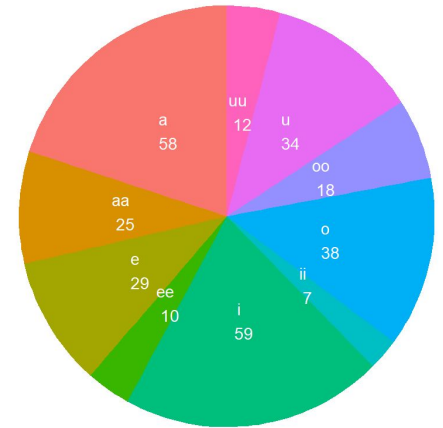
n = 635

Occurrence of Long and Short Vowels in Vocabularies



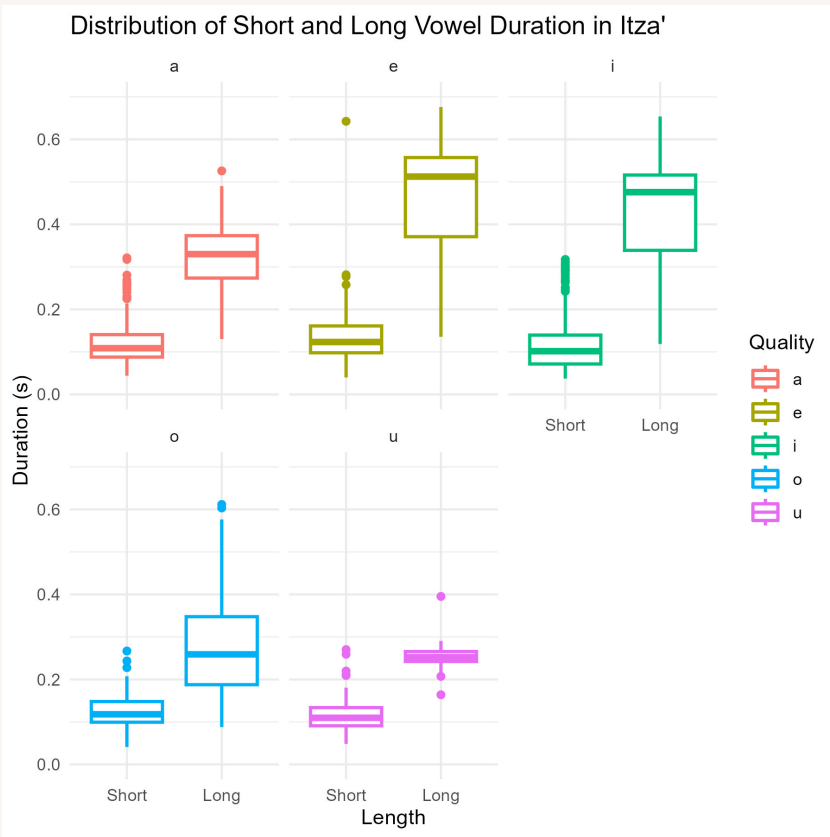
n = 406

Occurrence of Long and Short Vowels in Translation task



n = 290

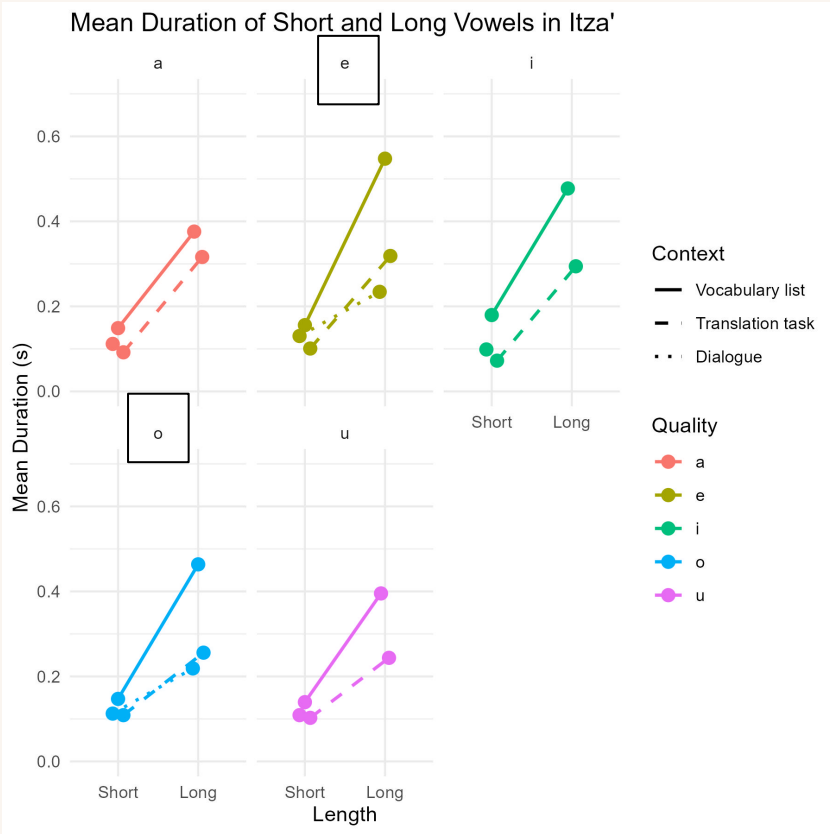
Results: Measurements

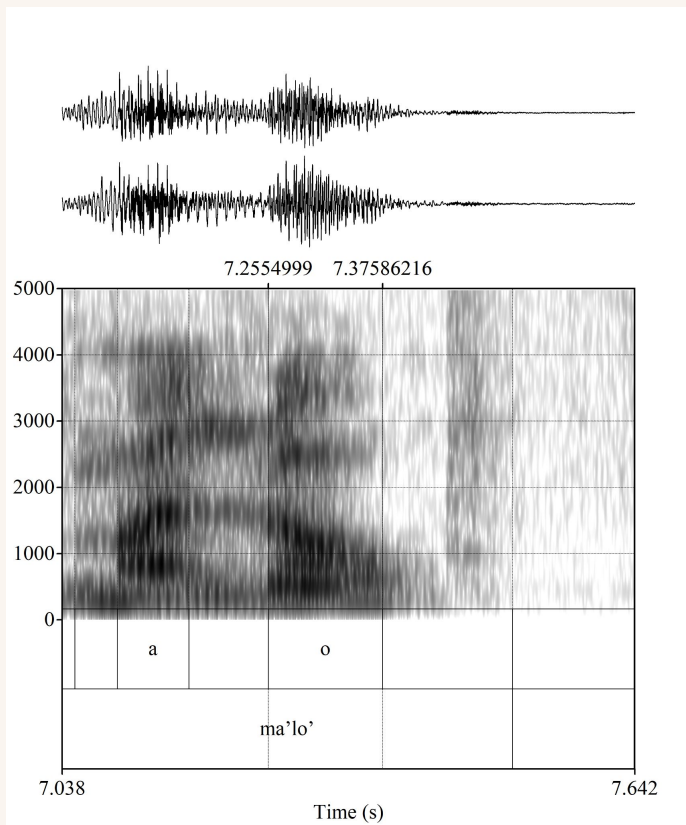


- Overall, vowels written as long were much longer than those written as short across all three tasks
- *Note:* these measurements did not account for stress, which does appear to lengthen vowels

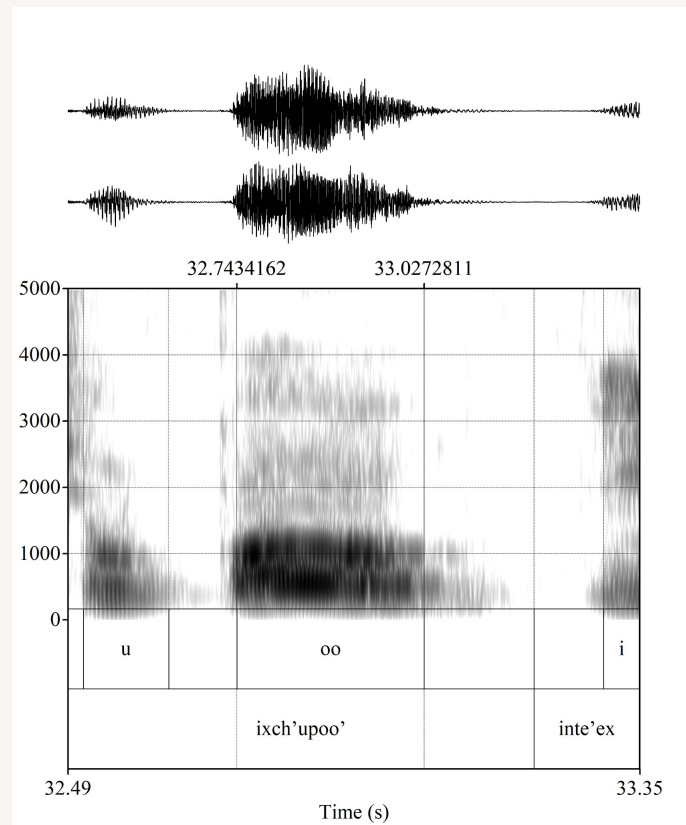
Measurements by Task

- The difference in duration was most apparent in the Vocabulary Lists
- Translation tasks and Dialogues showed generally similar patterns
- Long vowels range from 1.8x longer (/e/ in Dialogues) to 4.1x longer (/i/ in Translation task)





/o/ in *ma'lo'* with 120ms duration.



/o:/ in *ixch'upoo'* with 284ms duration.

Vowel	Context	n V	n VV	Short : Long Duration ratio	p
a	Dialogue	284	0	—	—
	Vocab list	122	7	2.5	*
	Transl task	58	25	3.4	**
e	Dialogue	91	3	1.8	
	Vocab list	64	26	3.5	**
	Transl task	29	10	3.2	**
i	Dialogue	150	0	—	—
	Vocab list	77	21	2.7	**
	Transl task	59	7	4.1	**
o	Dialogue	29	40	1.9	**
	Vocab list	41	17	3.2	**
	Transl task	38	18	2.3	**
u	Dialogue	38	0	—	—
	Vocab list	30	1	2.8	—
	Transl task	34	12	2.4	**

Results: Statistics

- T-tests reported that durations of short and long vowels were statistically significant in all but one subset (10/11)
- Comparing across contexts, the difference between Dialogue and Translation tasks were mostly non-significant
- Vocabulary list condition was significantly different compared to other two conditions

Discussion

- Based on these results, The digital grammar's first 5 chapters do convey to the learner that long vowels (written <aa>) have a significantly longer average duration than short vowels (written <a>)
- However, the distribution of long vowels and short vowels is quite unbalanced
 - Learners may not have enough tokens to make firm conclusions: e.g., only a single instance of long /uu/
 - Many of the long vowels are found within the same couple of common phrases (e.g. most /oo/ are in *te'yoos* "hello"), which does not provide a very diverse input

Discussion

- While rehearsed, the vowel lengths in the dialogues are comparable to the most naturalistic data we have (prompted spontaneous translation)
- We expected citation form (in the vocabularies) to have the largest difference in duration between short and long vowels, which is true

Future directions

- **In progress:** analyse L2 learners' vowel length, compared to fluent L2 and L1 speakers; investigate interactions between vowel length and other factors (such as stress, vowel quality)
- Propose improvements to available pedagogical grammar data to provide more diverse and balanced phonological input, or create additional modules focusing specifically on areas learners have difficulty with
- Find out more about Itza' diaspora members
 - Based on the census, these individuals largely identify with other Mayan language communities, and may know other Mayan languages
 - The most common of these is Q'eqchi', which also has contrastive vowel length and ejectives
 - Itza' learners who are speakers of Q'eqchi' (or are exposed to it) will likely show different acquisition patterns from learners in the San José area, who are mostly L1 Spanish speakers!

Conclusion

- The supplementary audio files included in the Itza' digital grammar do portray vowel length contrast as robust
- However, the distribution could be more balanced to help learners get a more holistic view of the vowel system
- Improving the content of the digital grammar is also important for those who may want to improve their language skills in diaspora
- Phonetic research can provide helpful information to language communities which they can use to work towards their own goals

D'yos b'o'tik!
(Thank you!)

Class Name



Student Name

References

- Bennett, R., Harvey, M., Henderson, R., & Méndez López, T. A. (2022). The phonetics and phonology of Uspanteko (Mayan). *Language and Linguistics Compass*, e12467.
- Boersma, P. & Weenink, D. 2025. Praat: Doing phonetics by computer. Version 6.4.27. URL <http://www.praat.org>.
- Brown, R. M., Maxwell, J. M., Little, W. E., & Bauer, A. (2006). *La üt awäch?: Introduction to Kaqchikel Maya Language*. University of Texas Press.
- Comunidad Lingüística Itza'. (2025). Comunidad Lingüística Itza'—ALMG on Facebook. Facebook.
- Hofling, C. A. (1991). *Itza' Maya Texts With a Grammatical Overview* (First Edition). University of Utah Press.
- Hofling, C. A. (1996). Indigenous Linguistic Revitalization and Outsider Interaction: The Itzaj Maya Case. *Human Organization*, 55(1), 108–116.
- Hofling, C. A. (2000). *Itzaj Maya Grammar*. University of Utah Press.
- Hofling, C. A., & Tesucún, F. F. (1997). *Itzaj Maya-Spanish-English Dictionary*. University of Utah Press.
- Instituto Nacional de Estadística. (2018). *XII Censo Nacional de Población y VII de Vivienda*. <https://censo2018.ine.gob.gt/>
- Jones, G. D. (1998). *The Conquest of the Last Maya Kingdom*. Stanford University Press.
- Mahlmann, J. (2023). VOT Analysis of L1 and L2 Speakers of Itza'. *Canadian Acoustics: Journal of the Canadian Acoustical Association*, 51(3), 214–215.
- Mateo Pedro, P., Mex Tesucún, J. F., Zhang, A., Im, S., Ramsey, D., Konyar, J., Lopez Loncar, D., López Sipac de Mateo, A., Chayax Tesucún, J. A., & Zacal Chayax, O. R. (2024). *Jelpalil Ka'ansajil Itza'/Gramática Pedagógica Itza'/Itza' Pedagogical Grammar*. Ministerio de Educación.

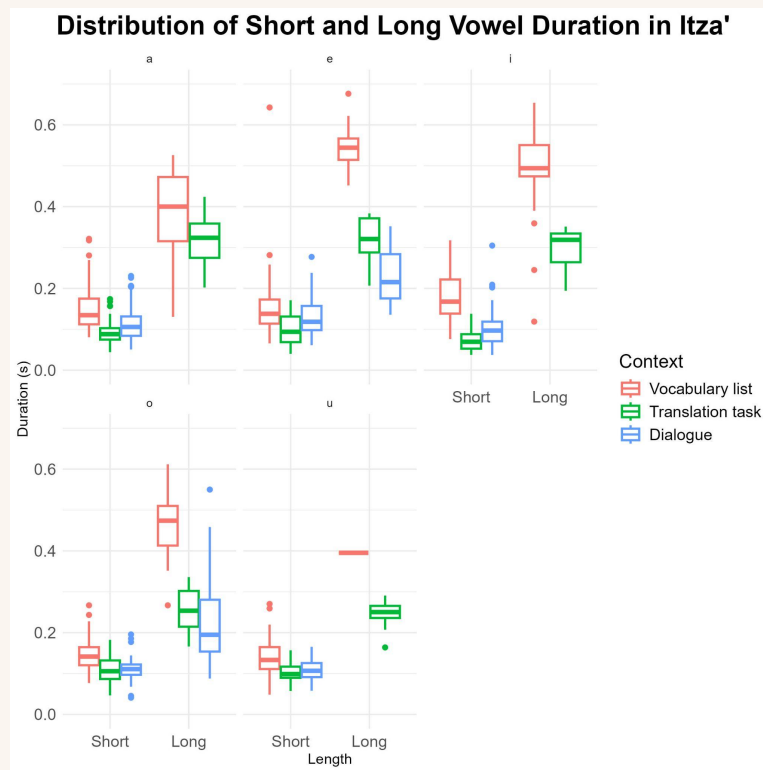
References

- Matras, Y. (2002). *Romani: A Linguistic Introduction*. Cambridge University Press.
- Maxwell, J. M. (2012). Pulsating galactic classrooms, immersion environments, individual vs. group language learning at home and abroad. Proceedings of *Intercultural Competence Conference 2*, 114–124.
- Maxwell, J. M., García Ixmatá, A. P., López Ixcoy, S. C., & Ajú Patal, C. A. (2014). *Manual para la enseñanza de un segundo idioma*. Guatemala: Universidad Rafael Landívar.
- Maxwell, I. J. M. & Little, W. W. E. (2006). *Tijonik Kaqchikel Oxlajuj Aj / Curso de Idioma y Cultura Maya Kaqchikel*. Guatemala: Editorial Junajpu.
- McAllister, R., Flege, J. E., & Piske, T. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics*, 30(2), 229–258. <https://doi.org/10.1006/jpho.2002.0174>
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL <https://www.R-project.org/>.
- Schumann Gálvez, O. (2000). *Introducción al Maya-Itzá*. Universidad Nacional Autónoma de México, Instituto de Investigaciones Antropológicas.
- Sobrino Gómez, M. (2015). Génesis de los tonos en el idioma maya yucateco. *Amerindia*, 37(2).
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>
- Zendejas, E. H. (2014). *Formas sonoras mapa fónico de las lenguas mexicanas; formas sonoras 1 y 2; (acompañado de CD-ROM)* (1. ed). El Colegio de México, Centro de Estudios Lingüísticos y Literarios.

Appendix A: Translation task word list

Yax ka'ansajil, Sak'unoo', Jiilt'an, Te', yoos, B'ix, ab'el, yanech, Intech, Toj, inwool, Ki', ma'lo'en, Inten, xan, ma'lo', Jach, B'aylo'ij, Ti, jatz'ka', ka'ka'at, ak'ä', Kiwilik, kib'aj, Känäntej, ab'aj, Uka', k'ab'a', max, ajkänb'al, ajka'ansaj, inwojel, jaj, ma', Uyox, kaj, nojkaj, tzukub', jajil, Nojkaj, Paxil, Job'onpich, lx, Tutz, Chäk'an, Itza', lxi'im, lu'umil, Waxaktun, Noj, P'eet, Tik'al, Mo'tul, Makanche', Säkluk', Xa'ayche', Ayimtun, Ukän, Makoo', mak, winik, tankelenpaal, ixch'up, ixnukna', ajchämach, ixch'upaal, ne', etel, junp'eel, sa', testen, Ujo', je'lo', a'lo', b'oon, k'atej, je'la', a'la', wa'ye', Uwak, B'a'axtak, unajilxok, nojb'ontz'iib', tuptz'iib', unu'ukultz'iib', ch'akche', k'anche', b'oniltz'iib', che'iltz'iib', jok'säni', ju'umilxok, waalju'um, ju'umiltz'iib', kuuchil ju'umilxok, k'aak', xok'al k'in, p'iisk'in, e'tes, e'tesa'an, ichil, kinajilxok, utz, tuka'ye', Uwuk, mentik, utz'iikchicht'an, wa'len, näklen, xi'manen, alka'nen, jok'en, oken, k'ayen, ok'otnen, xonlen, tiimes, siit'es, lub'es, emen, je'len, jeb'ix, Waxakp'eel, Pekesb'alil, b'aje'laj, wa'tal, näktal, ka'näl, je'lel, ok'ol, ichnil, ichkil, xot'ja', alka', jok'ol, okol, b'el, tal, siil, B'oniloo', chuntz'iib', ka'p'eel, oxp'eel, känp'eel, jo'p'eel, wakp'eel, wukp'eel, waxakp'eel, b'olonp'eel, lajp'eel, junlajp'eel, ka'lajp'eel, oxlajp'eel, känlajp'eel, jo'lajp'eel, waklajp'eel, waxaklajp'eel, b'olonlajp'eel, junk'älp'eel, ka'k'älp'eel, oxk'älp'eel, känk'älp'eel, jo'k'älp'eel, junb'ak'p'eel, tak, samal, k'äsmal'lo'en, ak'ab'a', ink'ab'a', aj, Mo', Sastun, uk'ab'a', akajal, inkajal, kajalej, Peet, Merex, Pet, säkluk', ti'ijoo', ek', ek' joche'en, b'ox, ch'aak, tz'am, chich, b'oj, yek'ja', ajchi'ik, paap, jol, tzaab', 'o'olki, kook, k'ook', kum, miis, 'oop, k'ochol, ke'entzil, ixk'ooch, p'uul, ajsaak', tu', malix, xak, xixi'ik, tat, xot', ch'ak, kokom, xaxak, ajme'ex, me'me'ex, kikil, kikilb'a, murux, mulix, b'ak, kab', k'ab', p'ak, b'aat, sum, k'am, paaychi', le', ixk'eech, b'ej, chem, pek', witz, k'ek'en, ajt'el, b'EEK, teek, chi', chip, ixkaax, kaax, ch'iich', xik', pix, ajmo', koj, pook, t'oot', much, kutz, sus, tunich, ixt'uut', b'uuk, kuk.

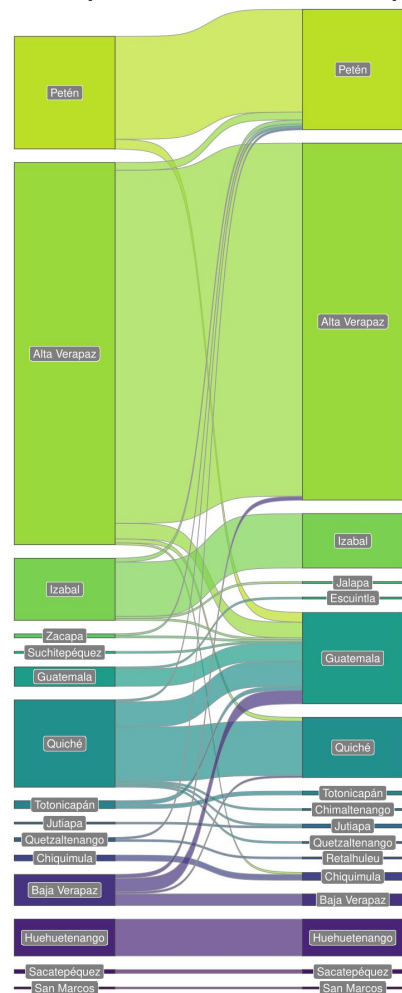
Appendix B: Boxplots of Itza' vowel durations



Appendix C: Are Itza' speakers moving?

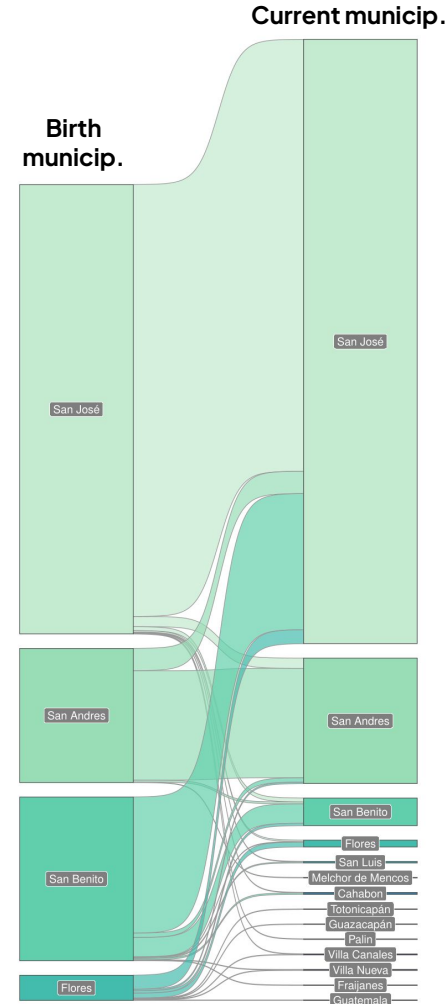
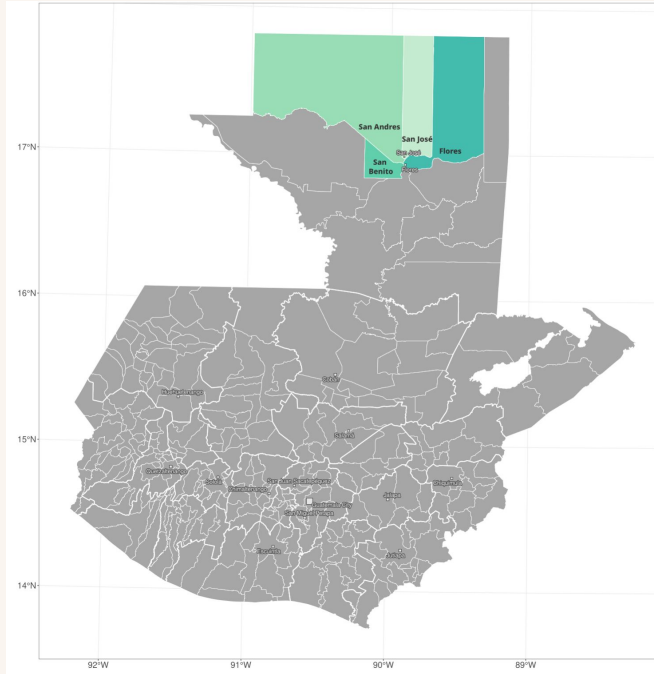
- Is the presence of a larger Itza'-adjacent population the result of people moving away from the Lake Petén Itzá region?
 - Not visibly! Migration away from Petén must have largely been in a previous generation
 - This suggests that the Itza' identity has been maintained in families that moved away
 - **There may be heritage speakers or people for whom Itza' is culturally significant who might be interested in language materials**

Birth dept. Current dept.



Are Itza' speakers moving?

- Even at the level of the municipality, it does not seem like the diaspora is the result of departure (in this generation) from Itza' heartland
 - If anything, people are moving back to San José from the more urban San Benito!



L1 / Community complementarity

- As the maps suggest, respondents who claim to be L1 Itza' overwhelmingly do not also say they belong to the Itza' linguistic community, while respondents who say they belong to the community largely do not choose Itza' as L1
- Perhaps ethnically Itza' people in the diaspora are using the L1 field to signal a kind of unofficial membership with the group?
- At right: speaker counts binned by age, divided into panes by department (rows) and L1 (columns)
- Itza' L1 respondents (1st col) are unlikely to also choose Itza' linguistic community
- Spanish & Q'eqchi' L1 respondents do select Itza' LC

