⟨’⟩ in Tsimane’: a Preliminary Investigation

William N Havard, Yaya Sy, Camila Scaff, Loann Peurey, Alejandrina Cristia

INTERSPEECH 2023
20-24 August 2023, Dublin, Ireland

⟨'⟩ in Tsimane': a Preliminary Investigation

William N. Havard [1,2], *Yaya Sy* [1], *Camila Scaff* [3,1], *Loann Peurey* [1], *Alejandrina Cristia* [1]

[1] Language Acquisition Across Cultures Team, Laboratoire de Sciences Cognitives et de Psycholinguistique, Département d'Études Cognitives, ENS, EHESS, CNRS, PSL University, France
[2] Cognitive Machine Learning Team, INRIA, Paris, France
[3] Human Ecology Group, Institute of Evolutionary Medicine, University of Zurich, Switzerland

william.havard@gmail.com, yayasysco@gmail.com, camila.scaff@cri-paris.org, loannpeurey@gmail.com, alecristia@gmail.com

Abstract

Tsimane' is a language spoken in Bolivia by several thousand people. Yet, it has not been described in detail. We aim to take a step towards a better description by focusing on an aspect of language: the sound represented in spelling with ⟨'⟩, informally described as a glottal stop. We recorded two adult speakers of Tsimane' producing (near-)minimal pairs involving this sound. Perceptual analyses suggested ⟨'⟩ is very rarely realised as a full glottal stop, and is more often cued by creaky-voiced vowels and nasals. Despite the variability in implementation, presentation of syllabic minimal pairs to these two informants and two other adult Tsimane' listeners revealed evidence that they could easily perceive when ⟨'⟩ was intended. Together, these data suffice to rule out the hypothesis that ⟨'⟩ is systematically realised as a full stop, and suggests instead a more complex set of perceptual cues may be at speakers' and listeners' disposal.
Index Terms: phonology, perception, production, adapted lab experiments

1. Introduction

Tsimane' is a language spoken in Bolivia by several thousand people and yet the phonology of Tsimane' has not been described in detail. There are few linguistic descriptions of Tsimane' [1, 2, 3, 4, 5], and none covers the phonology of Tsimane' very well. In this paper, we describe a small-scale research project aimed at studying the sound marked in the orthography as ⟨'⟩.

⟨'⟩ is clearly distinctive in Tsimane', for instance it is a morpheme that allows speakers to distinguish feminine from masculine (e.g. bǫ́riɾǫ'/female donkey; and bǫ́riɾǫ/male donkey) [6]. ⟨'⟩ has been described to us as a glottal stop (Ritchie, undocumented personal communication, 2018), which has been well documented across languages. However, its distribution in Tsimane' is unusual: since glottal stops are hard to hear, languages only allow glottals in codas when they also allow them in onsets [7]. In contrast, in Tsimane', ⟨'⟩ occurs *only* in syllable offsets. So how is ⟨'⟩ cued, and how well is it perceived by native listeners? These are the questions we sought to answer.

Although some speech corpora exist in Tsimane', none have the high audio quality required for our analyses nor for eliciting judgments by native Tsimane' informants. We therefore designed small-scale production and perception studies to be carried out during a 6-week visit to San Borja, a city close to traditional Tsimane' territory. Although the amount of data finally analysed here is small, it is based on an extraordinary effort.[1] Additionally, as an effort to foster reproducibility and further analysis of this sound, and in order to create speech resources for this acutely under-resourced language, all the data used in this present article (recordings, annotations, scripts, etc.) is made available.

[1] We estimated the total number of person hour directly invested in this project to 596; this is without counting travel time or annex activities, like holding meetings with community members, etc.

Table 1: *Frequencies of oral vowels (top), nasal vowels (middle) and consonants (bottom) in general (baseline) and before ⟨'⟩.*

Grapheme	i	a	u	o	e	ä
Baseline freq.	11.30	8.44	4.77	1.80	7.62	1.21
Before ⟨'⟩ freq.	21.67	16.25	16.37	3.80	18.71	2.14

Grapheme	i̧	a̧	u̧	o̧
Baseline freq.	0.41	0.57	1.01	0.91
Before ⟨'⟩ freq.	1.53	1.28	1.27	4.53

Grapheme	m	n	r	v
Baseline freq.	6.32	5.90	1.84	1.05
Before ⟨'⟩ freq.	4.27	8.04	0.12	0.02

Before starting, it is relevant to describe the distribution of ⟨'⟩. In general terms, ⟨'⟩ can occur after a vowel, a nasal consonant, or an approximant, thus, only after sonorants. Table 1 provides the frequency with which ⟨'⟩ is observed by context based on analyses of the largest text available in Tsimane': the Bible. Not shown here is the fact that in Tsimane' phonology, complex onsets and codas are not allowed, and thus the fact that ⟨'⟩ can follow /m,n,r,v/ is in itself unusual. That said, ⟨'⟩ appears after vowels (87.6%) a great deal more than after consonants (12.4%), somewhat beyond what would be predicted by these sounds' baseline frequency (vowels represented 36.04% of segments in the Bible, and /m,n,r,v/ jointly represented 12.45%). Although at first inspection, ⟨'⟩ seems to be biased towards *oral* vowels specifically, since it follows them more frequently than nasal ones (78.94% v. 8.61%), consideration of the baseline frequency reveals that, if anything, nasal vowels are relatively over-represented before ⟨'⟩, where they occur up to four times as frequently compared to their baseline frequency. Overall, ⟨'⟩ most frequently follows /i/, with a fifth of ⟨'⟩ tokens following this vowel. In contrast, it follows /r,v/ almost anecdotally, and much less frequently than would be expected given their baseline frequencies. With these considerations in mind, the minimal pairs employed for the production study contain ⟨'⟩ following vowels and nasal consonants /m,n/; and the perception study focuses on the most common context /i/.

2. Participants

This study was approved by CER U-Paris ethics committee under the reference 2022-84-CRISTIA. All participants are native Tsimane' speakers, aged between 30 and 45 years old (3 males, 1 female), all come from different communities, and

10.21437/Interspeech.2023-431

Table 2: *List of the 18 minimal pairs used in our production study. The perception study only focused on the -qui'/-qui pairs.*

#	Word 1	Word 2	#	Word 1	Word 2
1	ä'ä**m'**	ä'ä**m**	10	fó'jeya**qui'**	fó'jeya**qui**
2	á'nii'**tyi'**	á'nii'**tyi**	11	fú**rqui'**	fú**rqui**
3	aj**á'**	á**ja**	12	**jä'**mij	**jä**míj
4	án'dye**m'**	án'dye**m**	13	**já'**na'	**ja**ná'
5	bóri**cọ'**	bóri**cọ**	14	**jí'**ĉun	**ji**ĉún
6	bubá**qui'**	bubá**qui**	15	jí'**jun'**taqui	jí'**jun**taqui
7	**ĉó'**chaqui	**ĉo**cháqui	16	jí'junta**qui'**	jí'junta**qui**
8	ĉó'cha**qui'**	ĉó'cha**qui**	17	jí'jun'ta**qui'**	jí'jun'ta**qui**
9	**có'**co'	**co**có'	18	**jí'**jun'**tye'**	**jí**jun'**te**

Table 3: *Tsimane' carrier sentences and their translation.*

Begin	TARGET-WORD mo' nash peyacdye' yu yi
	TARGET-WORD is the word I am saying
Middle	yu ra' yi TARGET-WORD jeñej peyacdye'
	I will say TARGET-WORD as a word
End	yu ra' yi mo' peyacdye' TARGET-WORD
	I will say the word TARGET-WORD

have great experience assisting research projects. Two male participants recorded the entirety of the production study, while all four participated in the perception study. Participants for both the production as well as perception task were also informants on other research tasks. While they were not specifically compensated for participating in these studies, they received a compensation for their work-day.

3. Production study

3.1. Methods

Word List. We used a Tsimane'-English dictionary [3] to find potential minimal or quasi-minimal pairs featuring the distinction between ⟨'⟩ and no ⟨'⟩. The tentative word list was then curated by two literate native male Tsimane' speakers to ensure its validity. The data was collected in three pseudo-random batches, totalising 97 minimal pairs and 7 quasi-minimal pairs. This study reports an analysis of the first batch, which consists of 17 minimal pairs and 1 quasi-minimal pair (Table 2); the other batches are left for future work. The appearance of the two items of each pair was randomised within the batch during the recording sessions. Each item was recorded in 5 contexts: isolation, a natural phrase (elicited from our informants) and three carrier sentences (Table 3), where the position of the target word varied, so as to mitigate the effects of prosodic variations.

Recordings. We intended to use LIG-Aikuma [8] to collect this data, using a simple smartphone. However, due to low audio quality (16kHz) as well as random recording failures, we wrote a computer-based version that re-implements the text-elicitation and respeaking modes and allows to record at a 44.1kHz sampling rate.[2] In a quiet room, each informant read aloud the sentences into JBL Quantum 300 headphones, equipped with a foam windscreen, and connected via a USB audio cable to a Dell Precision 3561 computer running Ubuntu 20.04.5 LTS. This first batch consists of 380 recorded sentences across two speakers, totalising 20min of speech.[3]

[2] `https://github.com/William-N-Havard/williaikuma`

[3] `https://gin.g-node.org/William-N-Havard/tsimane-glottal-public` contains the full data set (02h13 of recorded speech, 20min of which fully segmented and annotated).

3.2. Annotation

We call a "clip" each audio recording of one sentence (i.e., one item in one context: isolation, natural sentence, or carrier phrase). We used Praat to annotate each clip through a three-step process. Figure 1 provides some example segmentations. We first sought to develop a forced aligner using the Montreal Forced Aligner [9] and recordings from the Bible (cf. supra) in order to use it to force-align our recordings and their transcriptions.[4] However, the forced-alignments on our data revealed too brittle to be used. Moreover, even though the alignments were good at word and phone level on the Bible data, they were particularly inaccurate for the sound ⟨'⟩ we are interested in. Thus, the target word of each clip was *manually* segmented from the sentence it was embedded into (e.g. a'ja), and an interval tier was used to indicate the segmentation into syllables (e.g. a'.ja). We then randomised these clips, so as to annotate the acoustic characteristics without potentially biasing knowledge of the source. Second, we segmented ⟨'⟩ jointly with the preceding sound and separate from a following sound on an other tier.

Finally, we attempted to identify phonetic realisations based on spectrogram inspection of ⟨'⟩ and its context. We separated: (i) vowel or nasal consonant sections with modal from those with amodal voicing, (ii) sections of silence as closure, and (iii) sections that were not silent but did not seem to be clear resonances above 1kHz as ʔ (regardless of whether a voicing bar was present, which may mean these are rather glottal approximants). This resulted in a fine-grained phonetic transcription (e.g. /a'/ may have been realised [a̰aa̰] in some cases, but only [a̰] in others).

3.3. Analysis

Our perceptual annotations reveal that there is a large variation in the way ⟨'⟩ are realised. Figure 1a and 1b show two occurrences of the same syllable of the same word in two different sentences uttered by the same speaker. While a clear creak is visible for the first syllable to signal ⟨'⟩ where /i'/ is realised [ḭi], no such creak is visible for the second occurrence. Instead, /i'/ is realised [iiʔi] with a slight glottal constriction which is neatly visible in the spectrum and visible as a slighter brighter part in the spectrogram. In view of this variation, we describe the acoustic correlates of ⟨'⟩ through a series of questions.

1. Is ⟨'⟩ systematically realised with ʔ? Collapsing across all observations (18 pairs, 5 realisations, 2 speakers, for a total of 380[5] observations), we found that only 7% of observations with a ⟨'⟩ contained a ʔ in the transcription. We therefore turned to other potential cues. We started by the most obvious, the presence of a closure. We then turn to amodal voicing, which is commonly observed across languages in the context of glottal and glottalised consonants. Finally, we inspect segment duration.

2. Are ⟨'⟩ realised with a closure? For this analysis, we only consider the 7 minimal pairs for which ⟨'⟩ appears word internally, so that silence can be unambiguously attributed to closure (and not to a boundary pause). In these items, ⟨'⟩ was followed either by a nasal or fricative consonant (/m, n/ or /h/), or a voiceless stop (/t, p, .../), which are analysed separately.

In the pre-nasal/fricative context (3 pairs in 5 different contexts for 2 speakers, totalling 60 observations), only 4 out of 30 observations where a nasal was followed by ⟨'⟩ showed the presence of a closure. One closure was observed for the 30

[4] `https://github.com/LAAC-LSCP/TsimaneForcedAligner`

[5] The quasi-minimal pair **jí'**jun'**tye'/jí**jun'**te** has two target syllables.

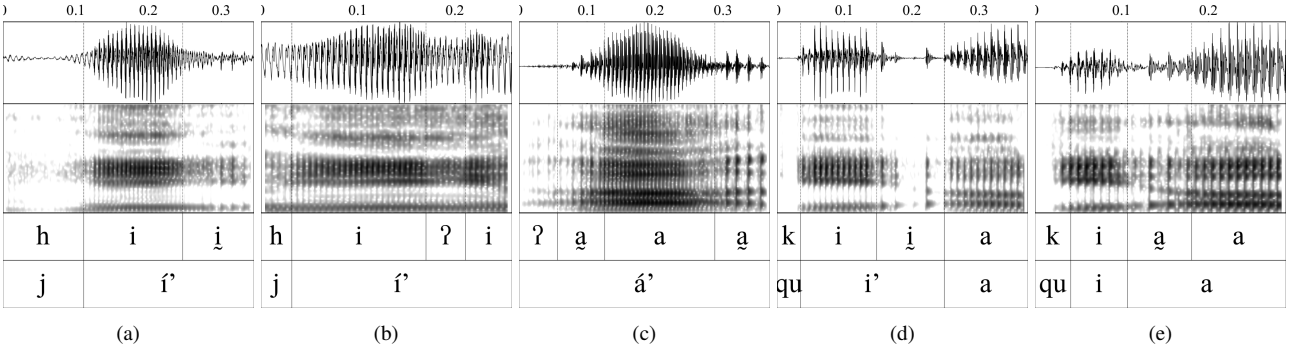

Figure 1: *(a) shows the realisation of* **ji’**jun’tye’ (v. to give work to s.o.) *with a clear creak (labelled ḭ) while (b) shows a token of the same syllable by the same speaker in the same word where only a slight constriction is visible (labelled ʔ). (c) shows the complex realisation of* **á’**nii’tyi’ (adj. good, appealing) *where the first vowel may be decomposed into 4 distinct states [ʔa̰aa̰] featuring the same cues that signl a ⟨’⟩. (e) Show the realisation of* fo’jeyaqu**i’ a**rosh (**she** throws the rice away) *and (d)* fo’jeyaqu**i a**rosh (**he** throws the rice away) *where in the first case /i’/ is realised with a creaky voice while /a/ is realised with a modal voice, while in the latter case /i/ is realised with a modal voice while /a/ is realised as a creaky voice, presumably to avoid a hiatus.*

paired observations when glottal was absent. Given the low incidence of closures in this context, we did not perform additional analyses.

In the pre-voiceless stop context (4 pairs in 5 contexts, totalling 80 observations), a closure could belong to ⟨’⟩ or to the following consonant. We therefore do not report on the presence of a closure (since it is 100% regardless of whether the item is supposed to contain ⟨’⟩), but analyse instead the closure *duration*. A paired t-test revealed that the duration of the closure is significantly linked to the presence of a ⟨’⟩ (p-value < 0.01): the duration of a closure preceding the voiceless occlusive is on average longer in the presence of ⟨’⟩, than in its absence (means 0.127 versus 0.103, respectively).

3. Is ⟨’⟩ cued by non-modal voicing? Previous literature [10] suggests that glottal stops are often cued by creaky and other forms of non-modal voicing. We analysed vowels versus nasal consonant contexts separately, as they present different physiological conditions. In the 160 observations where ⟨’⟩ followed vowels, 125 contained at least one section judged to have a non-modal phonation, whereas only 12 out of 160 paired vowels without ⟨’⟩ did so. This was significantly different according to a Chi-Squared test (p-value < 0.001). In contrast, although more nasals had some amount of non-modal voicing when followed by ⟨’⟩ (6/30), the incidence of non-modal voicing was almost as high when ⟨’⟩ was absent (3/30).

4. Is ⟨’⟩ cued by duration? We also inspected whether there may be any differences in segment length as a function of the presence of ⟨’⟩. Whereas in previous analyses we were relying on the top tier of Figure 1, containing the fine-grained segmentation, in this one we focus in the middle tier, to look at overall duration of the section attributed to ⟨’⟩ and the preceding vowel or nasal stop. Duration is not significant in the 160 paired observations with preceding vowels (mean with = 0.170, mean without = 0.159; paired t-test p-value > 0.05). In contrast, we observe a significant effect of the presence of ⟨’⟩ on the duration of preceding nasals according to an analysis of the 30 paired observations (mean with = 0.070, mean without = 0.099; paired t-test p-value < 0.05). Notice that the means for vowels and nasals go in opposite directions, with longer vowels in the presence v. absence of ⟨’⟩, but shorter nasals in the presence v. absence of ⟨’⟩.

Table 4: *Percent of correct responses, separated by participant (rows), stimulus speaker and target (columns). Participants 1 and 2 correspond to our first and second informants of the Production study described above.*

	Speaker 1		**Speaker 2**	
Participant	/ki/	/ki’/	/ki/	/ki’/
1 (M)	85	100	95	93
2 (M)	95	97	100	83
3 (M)	82	100	86	72
4 (F)	60	72	70	69

3.4. Discussion

Our results suggest that ⟨’⟩ in Tsimane’ are primarily realised via non-modal phonation (i.e. non-modal voice, observed in 69% of observations intended to include ⟨’⟩) and not by a stop (observed in only 7% of observations intended to include ⟨’⟩). For pairs in which ⟨’⟩ is followed by voiceless stops, the presence of ⟨’⟩ may be cued through increased closure duration. Additionally, our current results for segment duration are ambiguous, and additional cues may be inspected in the future.

Although we started this study with an open mind regarding whether ⟨’⟩ was indeed a glottal stop, these results suggest that the behaviour of ⟨’⟩ resembles that of glottal stops as described in previous literature: “In the great majority of languages [...], *glottal stops are apt to fall short of complete closure* [...]. In place of a true stop, a very *compressed form of creaky voice* or some less extreme form of *stiff phonation* may be superimposed on the vocalic stream.” [10, cited by [11]]. For example, [12] and [13] (both cited by [14]) report that glottal stops are realised as stops 25% and 7% of the time in Arapaho and Hawaiian respectively, and either omitted (13% and 20% of the time respectively) or realised with a creaky voice otherwise (i.e. 62 and 73% of the time respectively). Our results, although preliminary as done on productions of two speakers, point in the same direction, as we observe that ⟨’⟩ is cued by some amount of non-modal phonation 69% of the time.

Finally, previous literature on Mosetén and Tsimane’ reports that ⟨’⟩ only occurs in syllable-final positions in Tsimane’, and indeed there are no minimal pairs based on ⟨’⟩ presence v. absence outside of this position. This is not to say that cues like

non-modal voicing occur unambiguously in the presence of ⟨'⟩. Our perceptual annotations reveal non-modal voicing can occur in other positions (corresponding to the 7.8% of non-modal voicing found for paired items that were not intended to include ⟨'⟩), such as in word-initial positions (see Figure 1c), and to avoid vowel hiatus (see Figure 1d).

4. Perception study

To complement our perceptual annotation made based on spectrogram inspection, we also collected some information about the overall recoverability of ⟨'⟩, by presenting our two speakers from the production study and two other informants with extracts, focusing on the most common syllable in our stimuli. The annotations above were made after the 6 weeks were over, and thus our perception study could not benefit from insights gained through those annotations.

4.1. Methods

Equipment and procedure. Praat was used to present the stimuli and collect judgments. Participants were tested one at a time in a quiet room. Participants sat in front of the computer and wore JBL Quantum 300 headsets. There were three experiments presented in succession. The first was simply to acquaint participants with the procedure: two maximally different syllables (/ko'/ and /yi/) were extracted from the speech of our first informant. These two clips were presented a total of three times in a random order. On the screen, participants could see the syllables "co'" and "yi", as well as a button that allowed them to listen to the same stimulus up to three times before making a decision. We explained the procedure and stayed with them while they did this. We then answered any questions that they could have, and provided them feedback on the procedure if needed. The next two experiments were self-paced. In one, they heard all the stimuli for our first informant, and in the other they heard all the stimuli for the second informant.

Stimuli. Given time constraints in the field, we could not submit all 380 observations to a perceptual study, but had to rely instead on a subset of data which had been segmented at the time. We extracted all /ki/ and /kiʔ/ syllables, except for two syllables of our second speaker that contained a click. Due to an error, some syllables were extracted several times with slightly different segmentations, resulting in 138 stimuli to be presented, 72 from our first speaker and 66 from the second.

4.2. Results & Discussion

Results. Table 4 shows the percent of correct responses (i.e., listener reported hearing a /kiʔ/ when the target was /kiʔ/ and vice versa) separated by participant, speaker, and whether the speaker intended /ki/ or /kiʔ/. It is obvious that all four listeners were overwhelmingly capable of retrieving the intended syllable. This is remarkable given that there was wide variation in implementation, as described above.

Discussion. Despite variation in the input, participants were well above chance, suggesting they probably used a mix of cues, some beyond vowel glottalisation and closure. We know that in other contrasts, a multitude of acoustic cues can correspond to several alternate gestures. Here, listeners could be relying on differences in formant structure, formant transitions, and even pitch levels, cues that were not studied in our production study above, which focused on perceptual annotation rather than acoustic analyses.

5. General discussion

Our goal was to study how ⟨'⟩ was produced and perceived by native Tsimane' speakers. Our results reveal that the production of ⟨'⟩ resembles that of glottal stops as described in the previous literature. Our annotation reveals that this sound is realised in a wide variety of manners, ranging from (rare) full closures and slight glottis constriction to strong glottis constriction with creakiness, the latter being the most frequent. In future work, we wish to have a fine-grained annotation of the different kinds of creaky voice, such as proposed by [15] (e.g. distinction between prototypical creaky voice, from vocal fry, aperiodic voice, etc.) to understand if these realisations occur freely or on the contrary, are context-dependant. The perceptual study showed that both speakers were signaling the contrast clearly, since all four listeners (two who had participated in the production study,[6] and two who had not collaborated with us on this project) were well above chance level for both speakers and both items of the pair. In the future, analyses using random forests on a variety of acoustic cues could help us understand the acoustic indices listeners were employing. In addition, analyses could inspect whether relevant acoustic cues vary as a function of the position in which the target syllable occurred.

The present study has several limitations. We were only able to study production in two native speakers, who also have extensive experience with Spanish, and it is unclear to what extent their use of this second language may affect their pronunciation. That said, there is no particular reason to suppose that Spanish could affect it, since glottal stops and vowel glottalisation do not occur in any salient way in the local variety of Spanish. Another limitation was the use of elicited phrases, and the focus on one specific syllable, /ki/, as context for the perception study. We also hope to utilise other extant corpora (such as the Bible, mentioned in the Introduction) to study this contrast in a wider variety of materials, including some from spontaneous conversations and extant child-centered corpora [16]. Our perceptual study is also affected by a relatively small sample size, although results are quite clear in that they suggest the contrast is easily recovered by listeners. Nonetheless, sample size will be a limitation when attempting to use these data to infer which acoustic cues are most useful to listeners. Instrumental studies may be necessary (e.g. electroglottograph), so as to recover the precise gestures that are made, although transporting such equipment to field conditions will be challenging. Finally, taking into account the goal of establishing what is the phonological category underlying this contrast (e.g. is V+⟨'⟩ or C+⟨'⟩ represented as one or two units?), further studies are necessary, and should aim to include non-literate speakers and listeners, as literacy might have an effect on the cognitive representation of ⟨'⟩ [17]. While much remains to be done, we hope the present study, as well as the accompanying open data and code, will facilitate additional exploration of the still-mysterious ⟨'⟩.

6. Acknowledgements

We thank the J. S. McDonnell Foundation Understanding Human Cognition Scholar Award; European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ExELang, Grant No. 101001095).

[6] It was not the case that the two participants who provided production data performed differently when classifying occurrences of ⟨'⟩ in their own speech (127 successes) than when classifying occurrences pronounced by the other speaker (131 successes; 138 trials, z-test $p > 0.05$).

7. References

[1] J. Sakel, *A grammar of Mosetén*. Walter de Gruyter, 2011, vol. 33.

[2] W. Gill, "A pedagogical grammar of the Chimane (Tsimane') language," *San Borja, Bolivia: New Tribes Mission*, 1999. [Online]. Available: https://www.pueblos-originarios.ucb.edu.bo/Record/106001027

[3] W. Gill and R. Gill, "Chimane-English Dictionary," *San Borja, Bolivia: New Tribes Mission*, 1999.

[4] S. Ritchie, "Agreement with the internal possessor in chimane*: A mediated locality approach," *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"*, vol. 41, no. 3, pp. 660–716, 2017. [Online]. Available: https://www.jbe-platform.com/content/journals/10.1075/sl.41.3.05rit

[5] ——, "Posesión y Relactiones Gramaticales en Chimane," *Lenguas Indígenas de Bolivia: Teoría y Práctica*, p. 7, 2017.

[6] S. Ritchie and J. Sakel, *7 Chimane-Mosetén*. Berlin, Boston: De Gruyter Mouton, 2023, pp. 301–370. [Online]. Available: https://doi.org/10.1515/9783110419405-007

[7] M. Garellek, "Production and perception of glottal stops," Ph.D. dissertation, UCLA, 2013.

[8] E. Gauthier, D. Blachon, L. Besacier, G.-N. Kouarata, M. Adda-Decker, A. Rialland, G. Adda, and G. Bachman, "LIG-AIKUMA: A mobile app to collect parallel speech for under-resourced language studies," in *Interspeech 2016 (short demo paper)*, 2016.

[9] M. McAuliffe, M. Socolof, S. Mihuc, M. Wagner, and M. Sonderegger, "Montreal forced aligner: Trainable text-speech alignment using kaldi," in *Interspeech 2017*. ISCA, Aug. 2017. [Online]. Available: https://doi.org/10.21437/interspeech.2017-1386

[10] P. Ladefoged and I. Maddieson, *The sounds of the world's languages*, ser. Phonological Theory. London, England: Blackwell, Dec. 1995.

[11] F. Rose, "Mojeño trinitario," *Journal of the International Phonetic Association*, vol. 52, no. 3, p. 562–580, 2022.

[12] D. H. Whalen, C. DiCanio, C. Geissler, and H. King, "Acoustic realization of a distinctive, frequent glottal stop: The arapaho example," *The Journal of the Acoustical Society of America*, vol. 139, no. 4, pp. 2212–2213, Apr. 2016. [Online]. Available: https://doi.org/10.1121/1.4950615

[13] L. Davidson, "Effects of word position and flanking vowel on the implementation of glottal stop: Evidence from hawaiian," *Journal of Phonetics*, vol. 88, p. 101075, Sep. 2021. [Online]. Available: https://doi.org/10.1016/j.wocn.2021.101075

[14] C. T. DiCanio, "Phonetic variation in the production of glottal stops and glottalization," 12 2021.

[15] P. A. Keating, M. Garellek, J. Kreiman, and Y. Chai, "Acoustic properties of subtypes of creaky voice," *The Journal of the Acoustical Society of America*, vol. 153, no. 3 supplement, pp. A297–A297, 03 2023. [Online]. Available: https://doi.org/10.1121/10.0018918

[16] C. Scaff, M. Casillas, J. Stieglitz, and A. Cristia, "Characterization of children's verbal input in a forager-farmer population using long-form audio recordings and diverse input definitions," *PsyArxiv*, 2022.

[17] R. Kolinsky, A. L. Navas, F. V. de Paula, N. R. de Brito, L. de Medeiros Botecchia, S. Bouton, and W. Serniclaes, "The impact of alphabetic literacy on the perception of speech sounds," *Cognition*, vol. 213, p. 104687, 2021.