Phonetics of glottalized phonations in Yateé Zapotec

Yuan Chai[1], Adrián Fernández, Briseida Mendez

[1]Department of Linguistics, UCLA, Los Angeles CA USA 90095
yuanchai@g.ucla.edu; Dyadee03@gmail.com; briseida.mendezfer@gmail.com

ABSTRACT

This study presents a language, Yateé Zapotec, that contrasts modal, rearticulated (VˀV), and checked (Vˀ) phonations independently of tone. Rearticulated and checked phonations both involve glottalization but differ in the phasing of glottalization in vowels. We describe the three phonation types in Yateé Zapotec in terms of their f0, voice quality, amplitude, and duration based on the production of 147 words by a native speaker, and ask what acoustic parameter distinguishes them most effectively. Using multidimensional scaling analysis, we see that rearticulated vowels are best differentiated from modal and checked vowels by amplitude dipping. Checked vowels are best differentiated from modal and rearticulated vowels by their short duration. The results are in accordance with the findings of previous perception studies, such that listeners rely on amplitude dip to perceive vowels interrupted by glottal constriction, whereas they use short duration to identify a checked vowel.

Keywords: checked, rearticulated, phasing of glottalization, amplitude of voicing, duration

1. INTRODUCTION

Yateé Zapotec is a variety of Zapotec spoken in San Francisco Yateé, Oaxaca, Mexico, as well as by diaspora communities around Los Angeles, USA. There were around 480 people in the village of Yateé in 2017, according to a census conducted by the local clinic. Existing documentation of Yateé Zapotec is sparse [1, 2], and both studies focus on the consonants in the language. The current study focuses on the vowel phonation types in Yateé Zapotec and provides a quantitative analysis of their acoustic properties.

There are three vowel phonation types in Yateé Zapotec: modal (/V/), rearticulated (/VˀV/), and checked (/Vˀ/). Rearticulated phonation and checked phonation both involve glottalization but differ in the phasing of glottalization. Rearticulated phonation can be realized with either glottalization in the middle of vowels, or with glottalization spanning the whole duration of vowels. Checked phonation, in contrast, has glottalization at the end of vowels. The relative phasing of glottalization and vowel is in Figure 1. Examples of each phasing type in Yateé Zapotec are in Figure 2.

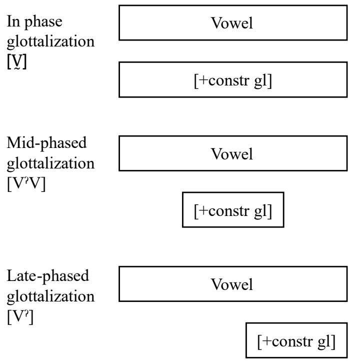

Figure 1: The phasing of in-phase, mid-phased, and late-phased glottalization relative to vowels.

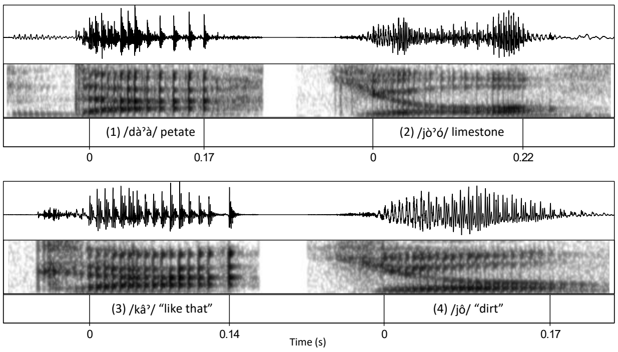

Figure 2: Sample tokens for in phase (1), mid-phased (2), late-phased (3) glottalization, and modal (4) phonation.

Yateé Zapotec has four tones: low, high, rising, and falling. The four tones are fully crossed with the three phonations. Examples of the interaction between phonation and tone are in Table 1.

	Modal	Rearticulated	Checked
Low	/ʂnà/ <xna> “red”	/ʂnːàˀà/ <xna’a> “my mother”	/nàˀ/ <na’> “there”
High	/lːá/ <la> “hot”	/zíˀí/ <zi’i> “heavy”	/ʐáˀ/ <xha’> “clothes”
Rising	/gǎ/ <ga> “nine”	/wèˀé/ <we’e> “get drunk”	/wěˀ/ <we’> “wound”
Falling	/jô/ <yo> “dirt”	/jóˀò/ <yo’o> “house”	/kâˀ/ <ka’> “like that”

Table 1: Examples of phonation and tone in Yateé Zapotec. Orthography is in angle brackets.

The research question of the current study is: how do the three phonations differ acoustically? We test the acoustics of the three phonations in five dimensions: f0, spectral tilt, noise, amplitude of voicing, and duration. The first four parameters are related to glottalization: glottalized vowels can have an extra-low f0 or high f0, lower spectral tilt, noisy voicing, and decreased voicing amplitude [3]. Checked vowels have been found to be shorter than modal and rearticulated vowels in several varieties of Zapotec (Yalálag: [4]; Betaza: [5]; Quiaviní: [6]). Further, we ask, among those acoustic parameters, which parameter(s) distinguish(es) the three phonations in Yateé Zapotec most effectively? To answer these questions, we conducted a production experiment and performed quantitative analysis.

2. METHOD

2.1. Participant

The participant is a male speaker of Yateé Zapotec at the age of 37. Yateé Zapotec is his native language acquired at home. He acquired Spanish at the age of 10 and acquired English at the age of 21. He uses Zapotec, Spanish, and English at home and uses English and Spanish at work.

2.2. Stimuli and procedure

The stimuli are 49 Yateé Zapotec words and were elicited in isolation. The distribution of the words in terms of tone and phonation is in Table 2. Table 1 is part of the word list. The full word list and the R script for data analysis are available at https://osf.io/a98sf/?view_only=dc40c474d5064dda a649892131d41394.

	Modal	Rearticulated	Checked
Low	4	6	4
High	5	5	7
Rising	7	3	2
Falling	5	1	3

Table 2: Distribution of the stimuli by tone and phonation.

During the experiment, each target word was elicited by a picture, its Spanish translation, and its Zapotec orthography. The stimuli were presented on a webpage three times in random order. The participant produced 147 words in total (49 words * 3 repetitions).

2.3. Acoustic measures and data cleaning

The recordings of the target words were processed using VoiceSauce [7]. We drew the f0 track of each token and visually inspected the f0 track. If we saw an f0 track that has an octave jump, we regarded it as an f0 tracking error and excluded it from the analysis of f0 and spectral tilt. We also drew the formant track of each vowel. If a token has a Mahalanobis distance larger than 6 from the mean F1 and F2 of the vowel, we regarded it as a formant tracking error and excluded the token from the analysis of spectral tilt [8]. After manually excluding the f0, F1, and F2 tracking errors, we converted parameter values to z-scores, then treated z-scores larger than 3 as outliers and excluded those tokens from the analysis.

3. RESULTS

In this section, we describe the results of six acoustic parameters for each phonation type: f0, H1*–H2* and residual H1* (spectral tilt), Harmonic-to-noise ratio (noise), Strength of Excitation (amplitude of voicing), and duration. Then we pass all these parameters to a multi-dimensional scaling analysis to illustrate what parameters best distinguish the three phonations.

3.1. F0

We fit f0 contours using a generalized additive model: *gam(f0~phonation*tone+s(time,k=9,by=interaction(phonation,tone)))*. *Phonation*, *tone*, and their interaction are the independent factors. *Time* is the smooth term (degrees of freedom = 9), interacting with the interaction of *phonation* and *tone* to allow an independent time smooth spline for each phonation and tone. The predicted f0 contour of each phonation and tone is in Figure 3. Low tone and falling tone both have a falling contour. Falling tone has a larger falling slope than low tone. High tone and rising tone both have a rising contour. Rising tone has a larger rising slope than high tone. Rearticulated vowels are characterized by a dipping point in the middle of vowels for the four tones, reflecting middle-phased glottalization in rearticulated vowels. High-toned checked and rearticulated vowels have higher F0 than modal vowels, indicating that they are likely to be characterized by tense voice.

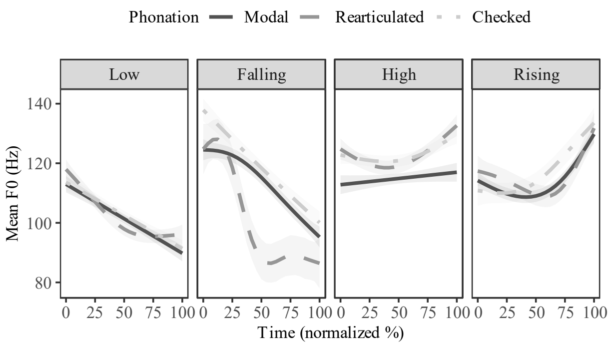

Figure 3: F0 of each tone and phonation.

3.2. Spectral tilt: H1*–H2* and residual H1*

Spectral tilt, such as H1*–H2*, is usually used to represent the degree of glottal constriction. Lower spectral tilt indicates a higher degree of glottal constriction [9, 10]. Residual H1* is a measure proposed in [11], which controls the amplitude of H1 by normalizing the energy of the signal. We fit H1*–H2* contours and residual H1* contours using a generalized additive model with *phonation* as the independent variable and *time* as the smooth term (degrees of freedom = 9), interacting with *phonation*. The same gam model was used for all the following time contours in this paper. Residual H1* (Figure 4b) differentiates the three phonations better than H1*–H2* (Figure 4a). The differences in residual H1* between every two phonations are plotted in Figure 4c. The blue portion of lines indicates that the difference in residual H1* is significant during that period, whereas the red portion means that the difference is not significant. As shown in Figures 4b and 4c, modal vowels have higher residual H1* than the other two phonation types almost all the time, indicating that modal vowels have the least glottal constriction among the three phonations. Rearticulated vowels have the lowest residual H1* in the middle of the vowel, whereas checked vowels have the lowest residual H1* at the end of the vowels. This verifies our prediction that rearticulated vowels have the most glottal constriction in the middle of vowels whereas checked vowels have the most glottal constriction at the end of the vowels.

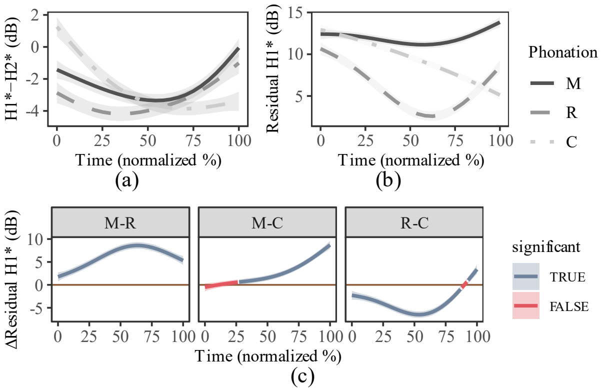

Figure 4: H1*–H2* (a) and residual H1* (b) contours and residual H1* differences (ΔResidual H1*) between phonations (c). M: modal; R: rearticulated; C: checked.

3.3. Noise level: Harmonic-to-noise ratio (HNR)

HNR measures the level of noise in the signal. A lower HNR value represents a higher level of relative spectral noise [12]. As shown in Figure 5, rearticulated vowels have lower HNR values than modal vowels after the first fourth of vowels. Checked vowels have lower HNR than modal vowels in the latter half. This indicates that rearticulated vowels have a noisy quality after the beginning of the vowel. Checked vowels are noisier at the end of the vowel. Modal vowels are the least noisy among the three phonation types.

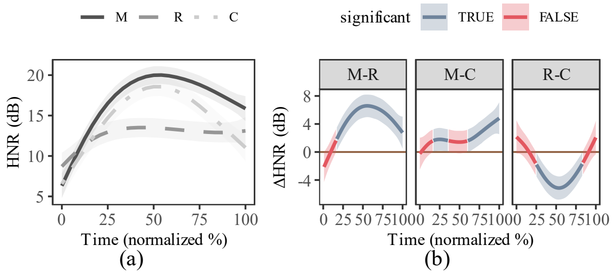

Figure 5: HNR contour (a) and HNR differences (b).

3.4. Amplitude of voicing

Strength of Excitation (SoE) measures the amplitude of voicing [7]. A larger SoE means that the voicing has a larger amplitude. As shown in Figure 8, modal vowels have a near-flat SoE contour. Among the three phonation types, checked vowels have the highest SoE at the beginning of the vowel and the lowest SoE at the end of the vowel. Rearticulated vowels exhibit a dipping SoE contour. The middle point has the lowest SoE. The amplitude of voicing is another parameter that reflects glottalization. Glottalization usually elicits damping in the amplitude of the voicing [3, 13]. The SoE contours in Figure 6 reflect that rearticulated vowels are characterized by amplitude damping in the middle of vowels, whereas checked vowels have amplitude damping at the end of vowels. The amplitude damping corresponds to the middle-phased and late-phased glottalization in rearticulated and checked vowels, respectively.

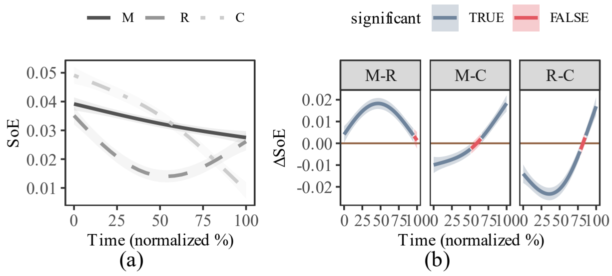

Figure 6: SoE contours (a) and SoE differences (b).

3.5. Duration

Among the three phonation types, checked vowels have the shortest duration (mean = 177 ms), followed by modal vowels (mean = 210 ms) and rearticulated vowels (mean = 216 ms) (Figure 7). A linear regression test shows that checked vowels have a significantly shorter duration than modal and rearticulated vowels ($p < 0.001$). The duration of modal and rearticulated vowels do not differ significantly ($p = 0.5$).

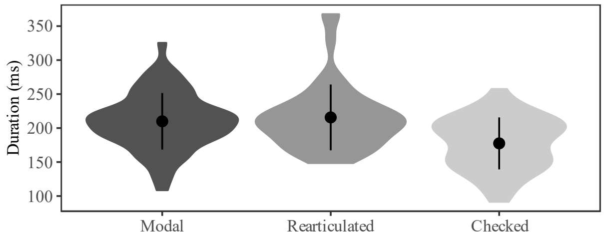

Figure 7: Duration of each phonation type.

3.6. Dimensionality reduction

Among the various parameters described above, which parameter(s) can most effectively distinguish the three phonation types? We performed a non-metric multi-dimensional scaling (MDS) analysis on the data. MDS can reduce the dimensions of the data, then project the data points to the reduced dimensions, and measure the distance among the data points [14]. MDS analysis has been used to analyze the acoustic cues associated with perceptual similarity [15, 16, 17] and the acoustic cues that differentiate phonation types across languages [18].

For each acoustic parameter described in Section 3 (except for duration), we divided the vowel into three equal-timed intervals and measured the mean value of each interval. We measured the Euclidean distances among the data points on a 3-dimensional MDS solution. The stress of the MDS model is 0.115. We plotted the data points on Dimension 1 (MDS1)-MDS2 and MDS1-MDS3 planes (Figure 8). MDS1 separates rearticulated vowels from modal and checked vowels. MDS2 separates checked vowels from modal and rearticulated vowels (Figure 8a). MDS3 separates modal vowels from checked and rearticulated vowels (Figure 8b). The top 3 parameters that have the biggest weight for each dimension are listed in Table 3. MDS1 is represented by the amplitude of voicing (SoE). MDS2 is primarily represented by duration, followed by the amplitude of voicing. MDS3 is represented by the final-third f0 and the last two-thirds of residual H1*.

	Top 1	Top 2	Top 3
MDS1	SoE_3 (5.132)	SoE_2 (4.852)	SoE_1 (2.577)
MDS2	Duration (2.326)	SoE_1 (2.181)	SoE_2 (2.065)
MDS3	F0_3 (2.899)	Residual H1*_2 (1.953)	Residual H1*_3 (1.486)

Table 3: Acoustic parameters with the highest weight on each MDS dimensions. Weight is in parenthesis. Numbers 1, 2, and 3 followed by the parameter name represent the initial, medial, and final third of the vowel, respectively.

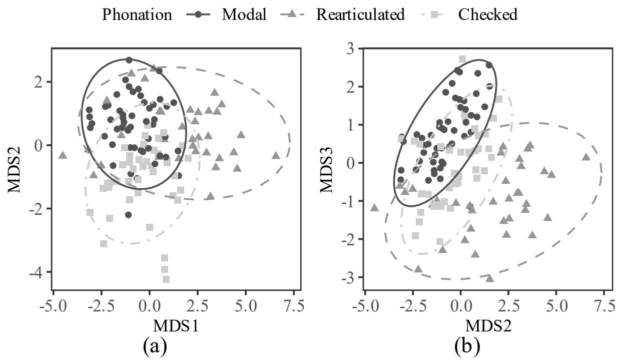

Figure 8: MDS distribution (confidence interval of the ellipses is 95%)

4. DISCUSSION AND CONCLUSION

The descriptive data of the acoustic parameters and the MDS analysis show us that rearticulated vowels are differentiated from modal and checked vowels primarily by their amplitude of voicing. The convex-shaped SoE contour of rearticulated vowels is distinguished from the falling contour of checked vowels and the flat contour of modal vowels. Checked vowels are distinguished from modal vowels primarily by their shorter duration. In addition, modal vowels are further distinguished from checked and rearticulated vowels by having less glottal constriction in the last two-thirds of vowels. Although f0 also emerges as a parameter in MDS3, we do not regard it as a defining feature of phonation because the tonal distribution in the stimuli is unevenly distributed among phonation types. There are more high and rising tones in modal vowels than in rearticulated and checked vowels in the stimuli.

The results agree with some previous findings on the perceptual cues of rearticulated and checked vowels. For rearticulated vowels, listeners of English [19] and Coatzospan Mixtec [20] use amplitude and f0 dips to identify VʔV or V̰. For checked vowels, short duration and falling f0 contour are more important cues than late-phased glottalization (White Hmon [21], Sgaw Karen [22], Xiapu Min [23]). The cues of amplitude and duration that listeners use to perceive rearticulated vowels and checked vowels are also the important cues for distinguishing rearticulated and checked vowels from modal vowels in the production of Yateé Zapotec. F0 not emerging as the primary cue in Yateé Zapotec is probably due to that phonation is fully crossed with all four tones in the language. A constant f0 is required to maintain the tonal contrast. In future studies, we will conduct perception tests with resynthesized stimuli and determine whether listeners of Yateé Zapotec attend to amplitude in voicing, duration, and spectral tilt when distinguishing checked, rearticulated, and modal vowels.

5. REFERENCES

[1] Jaeger, J. J. 1983. The fortis/lenis question: evidence from Zapotec and Jawoñ. *Journal of Phonetics*, 11, 177–189.

[2] Jaeger, J. J., Van Valin, R. D. 1982. Initial Consonant Clusters in Yateé Zapotec. *International Journal of American Linguistics*, 48, 125–138.

[3] Keating P., Garellek, M., Kreiman, J. 2015. Acoustic properties of different kinds of creaky voice. in *Proc. 18th ICPhS* Glasgow.

[4] Avelino, H. 2004. Topics in Yalálag Zapotec, with particular reference to its phonetic structures. Dissertation, University of California, Los Angeles, Los Angeles, CA.

[5] Teodocio Olivares, A. 2009. Betaza Zapotec Phonology: Segmental and Suprasegmental Features. Master's thesis, University of Texas at Austin, Austin, TX.

[6] Chávez-Peón, M. E. 2010. "The interaction of metrical structure, tone, and phonation types in Quiaviní Zapotec,"Dissertation, University of British Columbia, Vancouver, Canada.

[7] Shue, Y.-L., Keating, P., Vicenik, C., Yu, K. M., 2011. VoiceSauce: A program for voice analysis. in *Proc. 17th ICPhS* Hong Kong, 1846–1849.

[8] Chai, Y., Ye, S. 2022. Checked Syllables, checked tones, and tone sandhi in Xiapu Min. *Languages*, 7, 47.

[9] Klatt, D. H., Klatt, L. C. 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *J. Acoust. Soc. Am*. 87, 820–857.

[10] Zhang, Z. 2016. Cause-effect relationship between vocal fold physiology and voice production in a three-dimensional phonation model. *J. Acoust. Soc. Am*. 139, 1493–1507.

[11] Chai, Y., Garellek, M. 2022. On H1–H2 as an acoustic measure of linguistic phonation type. *J. Acoust. Soc. Am*. 152, 1856–1870.

[12] Garellek, M. 2019. The phonetics of voice. in *Routledge Handbook of Phonetics*, Katz, W., Assmann, P. (eds), Oxford: Routledge, 75–106.

[13] Pan, H.-h. 2017. Glottalization of Taiwan Min checked tones. *Journal of the International Phonetic Association*, 47, 37–63.

[14] Cox, T. F., Cox, M. A. A. 2001. *Multidimensional scaling*, Boca Raton: Chapman & Hall/CRC.

[15] Bradlow, A., Clopper, C., Smiljanic, R., Walter, M. A. 2010. A perceptual phonetic similarity space for languages: Evidence from five native language listener groups," Speech Communication, 52, 930–942.

[16] Fox, R. A., Flege, J. E., Munro, M. J. 1995. The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis. *J. Acoust. Soc. Am*. 97, 2540–2551.

[17] Gandour, J. T., Harshman, R. A. 1978. Crosslanguage differences in tone perception: a multidimensional scaling investigation. *Language and Speech*, 21, pp. 1–33.

[18] Keating, P., Kuang, J., Garellek, M., Esposito, C., Khan, S. Submitted. A cross-language acoustic space for vocalic phonation distinctions. *Language*.

[19] Hillenbrand, J. M., Houde, R. A. 1996. Role of F0 and amplitude in the perception of intervocalic glottal stops. *Journal of Speech, Language, and Hearing Research*, 39, 1182–1190.

[20] Gerfen, C., Baker, K. 2005. The production and perception of laryngealized vowels in Coatzospan Mixtec. *Journal of Phonetics*, 33, 311–334.

[21] Garellek, M., Keating, P., Esposito, C. M., Kreiman, J. Voice quality and tone identification in White Hmong," *J. Acoust. Soc. Am*., 133, 1078–1089.

[22] Brunelle, M., Finkeldey, J. 2011. Tone perception in Sgaw Karen. in *Proc. 17th ICPhS* Hong Kong.

[23] Chai, Y. 2022. *Phonetics and phonology of checked phonation, syllables, and tones*. PhD diss., University of California San Diego, San Diego, CA.