Using residual H1* for voice quality research[a)]

Mykel Loren Brinkerhoff[b)] (iD) and Grant McGuire

Department of Linguistics, University of California, Santa Cruz, California 95064, USA

mlbrinke@ucsc.edu, gmcguir1@ucsc.edu

Abstract: One of the most common ways of assessing voice quality is through the spectral slope measure $H1^*$–$H2^*$. However, Chai and Garellek [J. Acoust. Soc. Am. **152**(3), 1856–1870 (2022)] propose and demonstrate the robustness of a new measure—residual $H1^*$. In this study, we test the reliability of residual $H1^*$ in capturing voice quality contrasts in Santiago Laxopa Zapotec. We find that residual $H1^*$ is a more robust measure of voice quality than $H1^*$–$H2^*$ and can capture the differences in voice quality in Santiago Laxopa Zapotec, which supports the use of this measure in voice quality research. © 2025 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Languages use voice quality distinctions to convey phonemic distinctions (Garellek, 2019) and to convey paralinguistic information by "indexing the biological, psychological, and social characteristics of the speaker" (Laver, 1968; Podesva, 2016). Voice quality contrasts have been studied extensively by examining their correlates in the acoustic signal (e.g., Esposito and Khan, 2020), resulting in a large and complex literature on acoustic correlates of phonation differences (see Garellek, 2019).

One measure that is by far the most commonly applied is $H1^*$–$H2^*$, which measures the relative amplitudes of the first harmonic and second harmonic. As established by Fischer-Jørgensen (1968), the relative strength of the fundamental is a correlated measure of breathy voice in contrast with modal voice). In order to normalize the amplitude of the fundamental frequency ($f0$) and counteract some of the effects of high-pass filtering and differences in sound pressure in the signal, the author proposed that you could subtract the amplitude of a higher harmonic, in this case, the second harmonic (H2), from the amplitude of the fundamental frequency (H1). Since its introduction, $H1^*$–$H2^*$ has been used in many studies to measure not only breathy voice but other voice quality contrasts as well (Chai and Garellek, 2022; Garellek, 2019).

Despite the large amount of evidence in support of $H1^*$–$H2^*$, it is not without its problems. At a fundamental level, it is not clear that $H1^*$–$H2^*$ adequately measures the strength of the fundamental frequency. Sundberg (2022) found that H1 and H2 are affected differently by subglottal pressure, compromising some of the original reasoning behind the use of $H1^*$–$H2^*$ from Fischer-Jørgensen (1968). Similarly, in a comprehensive overview of the main concerns with $H1^*$–$H2^*$ as a phonation type measure, Chai and Garellek (2022) found that in addition to the issues mentioned above, errors in measuring $H1^*$–$H2^*$ are uncomfortably high. This is mainly due to the need to precisely measure two different harmonic amplitudes when there are errors in calculating H1; this, in turn, leads to errors in calculating H2 (Arras, 1998). An example of this type of error propagation is that errors in measuring the fundamental frequency, which is especially common with non-modal phonation, are introduced into measuring harmonics because they are based on the fundamental frequency. Despite algorithms correcting for vowel height, a common error that occurs is when a high fundamental frequency co-occurs with a low first formant (Chai and Garellek, 2022). This situation causes errors in tracking the fundamental frequency and the first formant. A final issue that can occur when measuring the harmonics is in contexts where the vowel is nasalized. Simpson (2012) shows that in these nasalized contexts, the first nasal pole ($P0$) can increase the amplitude of H2 and, when the fundamental frequency is high, H1 increases instead.

This collection of errors leads Chai and Garellek (2022) to propose a new measure—residual $H1^*$. This measure is calculated by first regressing $H1^*$ on energy and then subtracting the product of energy and the energy factor from H1. Chai and Garellek (2022) argue that this measure better reflects the initial purpose of using $H1^*$–$H2^*$. Furthermore, they find that residual $H1^*$: (i) provides better differentiation between phonation types in !Xóõ; (ii) was more robust for measuring creak in Mandarin with respect to different utterance positions; and (iii) has a stronger relationship to the open quotient than $H1^*H2^*$ based on a comparison using electroglottographic data.

[a)]Portions of this work were presented in "On Residual $H1^*$ as a measure of voice quality" at the 19th Conference on Laboratory Phonology, Seoul, Korea, June 2024.

[b)]Author to whom correspondence should be addressed.

The contributions Chai and Garellek (2022) make are very intriguing and have the potential to alter how spectral analyses of voice quality are performed. However, they are not convincing on their own. If residual H1* is to be widely adopted in linguistics, speech pathology, and other speech sciences, then it requires considerable evidence of its effectiveness in acoustic studies. This paper offers additional evidence for the effectiveness of residual H1*.

Given the promising nature of this measure, we tested residual H1* with data from Santiago Laxopa Zapotec. Although Chai and Garellek (2022) evaluated residual H1* for !Xóõ and Mandarin, both of which contain tone and voice quality; they do not factor tone into their analysis.[1] This is where testing the effectiveness of residual H1* in Zapotec languages is beneficial. Zapotec languages are often described as laryngeally complex (Ariza-García, 2018; Silverman, 1997). Laryngeal complexity is defined as a language that allows contrastive tone and contrastive voice quality that are unrestricted in their interactions. This laryngeal complexity between tone and phonation types presents a unique challenge for acoustic analysis and testing of the residual H1* measure.

Other Zapotec languages have also been studied for voice quality research. For example, Esposito (2010) found that in Santa Ana del Valle Zapotec, there was a biological sex difference in which acoustic measures best capture the voice quality contrasts similar to observations from Klatt and Klatt (1990). Arellanes Arellanes (2010) found that the realization of laryngealization in a variety of San Pablo Güilá Zapotec is highly variable and depends on the position of laryngealization in the phrase. This was also found to be the case in Betaza Zapotec (Crowhurst et al., 2016). Barzilai and Riestenberg (2021) found that tone and phrasal position also played a role in how voice quality is realized in San Pablo Macuiltianguis Zapotec. These studies show that Zapotec languages have much to contribute to our understanding of voice quality.

We find that residual H1* can adequately capture differences in voice quality and is a more robust measure of voice quality than H1*–H2*, adding credence to the use of this measure instead of H1*–H2* in voice quality research. The remainder of this paper is organized as follows: Sec. 2 provides a brief overview of the Santiago Laxopa Zapotec language. Section 3 describes the methods used in the data collection, data processing, and statistical modeling used in this study. Section 4 presents the results of the study. Section 5 concludes the paper.

2. Santiago Laxopa Zapotec

Santiago Laxopa Zapotec (SLZ) is a Northern Zapotec language of the Oto-Manguean language family (Adler et al., 2018; Adler and Morimoto, 2016; Brinkerhoff et al., 2021, 2022; Foley et al., 2018; Foley and Toosarvandani, 2022; Sichel and Toosarvandani, 2020a,b). It is spoken by 981 people in the municipality of Santiago Laxopa, Ixtlán, Oaxaca, Mexico and a small number of other speakers from the diaspora in Mexico and the United States.

Santiago Laxopa Zapotec exhibits a four-vowel inventory (that is, /i/, /e/, /a/, and /u/), which is further distinguished by a four-way contrast in voice quality. This variety is unique because it is a Northern Zapotec that has developed a breathy voice (/V̤/), in addition to the two types of laryngealization that characterize the rest of the Zapotec languages, namely checked and articulated. Checked vowels are defined as a modal vowel that ends with a period of creaky voice or a glottal closure (/VV̰/ or /VʔV/). Rearticulated vowels are also defined as a modal vowel that also has a period of creakiness or glottal closure but aligned to the middle of a vowel (/V̰V̰V/ or /VʔV/). This makes an otherwise modal vowel appear as if it is interrupted by this laryngealization.[2] This difference in laryngeal timing is one of the key differences between these phonations (see Ariza-García, 2018, for a detailed typological study of voice quality distinctions in Zapotec languages).

Santiago Laxopa Zapotec is also tonal with three level tones (H, M, and L) and two contours (MH and HL) appearing in nominals (Brinkerhoff et al., 2022).[3] The language has a complex interaction between tone and phonation types. Every tone can appear with every phonation type, with two exceptions being that breathy voice cannot appear with the high tone and checked voice cannot appear with the rising contour tone. It is unclear whether these are accidental gaps or have a phonetic underpinning.

These interactions between voice quality and tone present a rich environment for testing the reliability of voice quality measures in laryngeally complex languages.

3. Methods

3.1 Elicitation

Ten native speakers of SLZ (five female; five male) participated in a wordlist elicitation. Elicitation was performed in the pueblo of Santiago Laxopa, Ixtlán, Oaxaca, Mexico during the summer of 2022 using the built-in microphone of a Zoom H4n handheld recorder (16 bit, 44.5 kHz).

The wordlist consisted of 72 items repeated three times in isolation and the carrier sentence *Shnia' X chonhe lhas* [ʃnːiaʔ X tʃone ɾas] "I say X three times." This sentence was chosen to minimize any effect of the phrasal position [see Crowhurst et al. (2016), for a study about phrasal effects on voice quality in Zapotec]. Additionally, there are no co-articulatory effects on voice quality found with this frame and the items in the wordlist. Between these 72 words, there were 11 words with breathy voice, nine with rearticulated voice, ten with checked voice, and 42 with modal. Thirteen of the 72 words were disyllabic and eight of the 13 contained the same voice quality in each syllable. Of those 13 disyllabic words, only five contained different voice qualities in each of the syllables. This resulted in 85 different syllables.[4]

Table 1. Distribution of tone and voice quality in the wordlist.

	High	Mid	Low	Rising	Falling
Modal	14	9	15	2	10
Breathy	—	—	11	—	2
Checked	1	—	9	—	—
Rearticulated	1	—	4	—	4

The token selection for the wordlist was done in consultation with a native speaker. Similarly to Barzilai and Riestenberg (2021), we did not balance the tones with the different voice qualities. The frequency in which certain tones co-occur with certain voice qualities is uneven in the language, making it difficult to control for. The distribution of tones and voice quality across the 85 syllables used in this study are presented in Table 1. This imbalance was taken into account by including tone as a fixed effect in our statistical models in Sec. 3.3.

3.2 Data processing

Each vowel of the target words in the carrier sentence condition was labeled following Garellek (2020) for where the vowel began and ended. Each vowel in the word list was annotated for speaker, word, vowel, tone, voice quality, and utterance number. This labeling was conducted for each vowel in the target word from the elicitation list of the carrier sentences.

These vowels were then extracted and fed into VoiceSauce for acoustic measurement (Shue *et al.*, 2011). The formants were measured using Snack (Sjölander, 2004), while the fundamental frequency (*f0*) was measured using the STRAIGHT algorithm (Kawahara *et al.*, 1998). Spectral slope measures were corrected for formants and bandwidths (Hanson, 1997; Iseli *et al.*, 2007). Each vowel was measured with ten equal time intervals, resulting in 22 890 data points in total.

The data were cleaned of outliers following the same steps as taken by Chai and Garellek (2022) in their study. The H1*, H1*–H2*, and *f0* values were *z*-scored by speaker to reduce the variation between the speakers and provide a way to directly compare the different measures on the same scale. Data points with an absolute *z*-score value greater than three were considered outliers and excluded from the analyses. Within each vowel category, we calculated the Mahalanobis distance in the F1–F2 panel. Each data point with a Mahalanobis distance greater than six was considered an outlier and excluded from the analysis. This is comparable to what was done in Chai and Ye (2022), Garellek and Esposito (2023), and Seyfarth and Garellek (2018).

Data points whose *f0*, F1, or F2 that were marked as outliers were excluded from the analyses. Energy was excluded if it had a zero value and then logarithmically transformed to normalize its right-skewed distribution. Afterward, the resulting logarithmically transformed data were *z*-scored, and any data point with a *z*-score greater than three was excluded. This outlier removal resulted in 1918 data points being removed.

After removing the outliers, we calculated residual H1* for the remaining data points following Chai and Garellek (2022). First, a linear mixed-effects model was generated with the *z*-scored H1* as the response variable and the *z*-scored energy as the fixed effect. The uncorrelated interaction of the *z*-scored energy by speaker was treated as random. The energy factor resulting from this linear mixed-effects model was extracted. Finally, the *z*-scored H1* had the product of the *z*-scored energy and the energy factor subtracted from it, giving us the residual H1* measure.

The measures were then orthogonally coded according to their position in the vowel (first, middle, and third) and for tone (H, M, L, R, F) for statistical modeling.
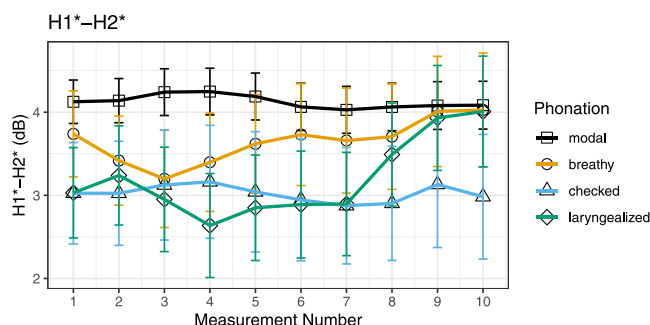

Fig. 1. H1*–H2* across the duration of the vowel. Points represent the mean of each measure across the ten intervals. The bars around each point represent a 95% confidence interval. A line was plotted over each to show how the acoustic measure functions across the ten intervals.

Fig. 2. Residual H1* across the duration of the vowel. Points represent the mean of each measure across the ten intervals. The bars around each point represent a 95% confidence interval. A line was plotted over each to show how the acoustic measure functions across the ten intervals.
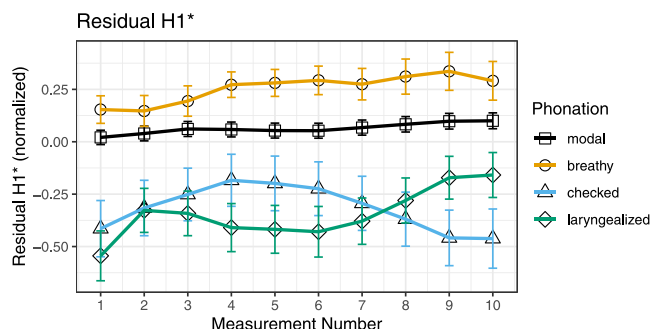
3.3 Statistical modeling

Two linear mixed-effects regression models were fitted: one for the z-scored H1*–H2* and one for the residual H1*. Each model had the tone and the interaction between voice quality and position in the vowel as fixed effects. Vowel and interaction between the speaker, the word, and the repetition were treated as random intercepts:

$$Measure \sim Phonation * Position + Tone + (1|Speaker : Word : Repetition) + (1|Vowel). \tag{1}$$

The tone and the interaction between voice quality and position in the vowel were selected as fixed effects for several reasons. The first is that five unique tones appeared in the data and it is well established that tone interacts with voice quality in different ways (see Esposito and Khan, 2020, and Garellek, 2019, for discussion). By treating tone as a fixed effect in our model, we can account for these interactions. The interaction between voice quality and position in the vowel as a fixed effect was included to account for the temporal differences between the two different laryngealizations—checked and rearticulated vowels. Checked vowels in Zapotec languages have a glottal occlusion or a short period of creaky voice located at the right edge of the vowel. This is in contrast to rearticulated vowels, where there is a glottal occlusion or creaky voice in the middle of the vowel. Because this difference between checked and rearticulated vowels is temporal in nature, we can account for this difference through the interaction of voice quality and position in the vowel.[5]

The interaction between speaker, word, and repetition was treated as a random intercept because this allows us to consider that each speaker said each word on the elicitation list three times. This intercept accounts for not only the intra-speaker variability, but also the inter-speaker variability during each time the word was uttered. Treating a vowel as a random intercept allows us to capture the fact that each voice quality occurred with different vowels during elicitation.

4. Results

In interpreting the results, there are certain expectations about how these measures should capture the voice quality contrasts. As discussed in Garellek (2019), it is expected that breathy vowels should have a higher spectral slope than modal vowels across the two measures in keeping with observations about breathy vowels in other languages. Because both checked and rearticulated vowels make use of creaky voice, they should have a lower spectral slope than modal vowels. Additionally, because there is a temporal distinction between checked and rearticulated with where the creakiness appears, this should also be captured in the two measures.

4.1 H1*–H2*

Figure 1 shows the mean H1*–H2* values for each voice quality at each of the ten vowel intervals. We see that the breathy, checked, and rearticulated values are lower than the modal values at each of the first nine intervals. In the final

Table 2. Model comparison between H1*-H2* and residual H1* in distinguishing Santiago Laxopa Zapotec voice quality.

Voice quality contrast	Model	β	Std. error	t-value	p-value	
Breathy vs Modal	H1*–H2*	0.046 31	0.038 06	1.216 80	0.223 72	—
	Res. H1*	0.236 25	0.028 66	8.241 77	<0.001	***
Checked vs Modal	H1*–H2*	−0.118 80	0.034 76	−3.417 93	<0.001	***
	Res. H1*	−0.400 99	0.026 21	−15.300 98	<0.001	***
Rearticulated vs Modal	H1*–H2	−0.091 75	0.045 88	−1.999 68	0.045 60	*
	Res. H1*	−0.441 62	0.034 50	−12.800 27	<0.001	***

Table 3. AIC for the H1*–H2* and residual H1* models.

Model	AIC	Δ AIC
H1*–H2* model	43 443.33	11 182.76
Residual H1* model	32 260.57	0

intervals, breathy and rearticulated are essentially equal to the modal value. In contrast, checked's value remains lower than the modal's value throughout the vowel. This measure does not match the expectations discussed above.

*4.2 Residual H1**

Figure 2 shows the mean residual H1* values for each voice quality at each of the ten vowel intervals. In contrast to Fig. 1, we see that breathy has a higher residual H1* measure than modal throughout the duration of the vowel, which is consistent with other observations for breathy voice (Fischer-Jørgensen, 1968). Checked and rearticulated both have lower values than the modal at each of the ten intervals. In addition, it shows that the checked voice has a lower residual H1* value than the rearticulated voice at intervals eight through ten. The rearticulated voice has a lower residual H1* value than the checked voice at intervals one through seven, showing the temporal distinction between these two voice qualities. This measure complies with the expectations discussed above.

4.3 Model comparison

To assess the robustness of the models, we compared the residual H1* linear mixed-effects model to the H1*–H2* linear mixed-effects model. This was done using two methods: direct comparison of the outputs of the two models in the same way as Chai and Garellek (2022) and the Akaike Information Criterion (AIC).

Table 2 compares the linear mixed-effects models for H1* -H2* and residual H1*. In comparing these models, we find that the residual H1* model performed better than the H1*–H2* model in distinguishing voice quality contrasts in Santiago Laxopa Zapotec. This is supported by the larger absolute value of the coefficient estimate, the lower standard error, and the higher *t*-value of the residual H1* to distinguish breathy, checked, and rearticulated vowels from modal vowels.

Table 3 shows the results of the AIC comparison between the H1*–H2* and residual H1* models. The residual H1* model had a lower AIC than the H1*–H2* model, indicating that the residual H1* model is a better fit for the data than the H1*–H2* model. Although AIC comparison is usually performed on nested models, it is still a useful tool for comparing non-nested models (Burnham and Anderson, 2004a,b; Burnham *et al.*, 2011).

5. Conclusion

In conclusion, we find that residual H1* is a more robust measure of voice quality than H1*–H2* in Santiago Laxopa Zapotec. This is supported by the results of the linear mixed-effects models, which show that residual H1* is better at distinguishing breathy, checked, and rearticulated vowels from modal vowels. This is further supported by the AIC comparison, which shows that the residual H1* model is a better fit for the data than the H1*–H2* model. These results lend credence to the claims of Chai and Garellek (2022) and support using residual H1* instead of H1*–H2* in voice quality research, especially in laryngeally complex languages.

However, the results neither suggest nor support the notion that residual H1* should be the sole measure used to determine phonation quality. Instead, we suggest that residual H1* should be included as one of the measures that researchers should consult during their acoustic studies.

Author Declarations
Conflict of Interest

The authors have no conflicts to disclose.

Ethics Approval

This research was carried out under the Internal Review Board Protocol #3289 at the University of California, Santa Cruz. Informed consent was obtained from all participants.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

[1] !Xóõ is most like Zapotec in that it has three phonologized phonation categories. It also has tone, which to our understanding is not restricted by phonation and which is not analyzed in Chai and Garellek (2022). The Mandarin case is less relevant as its phonation is prosodic and tonally linked.

[2] This is different from how rearticulated vowels are defined in other languages, where they are a modal vowel followed by an "echo vowel" which may or may not have a glottal closure intervening between the modal and "echo" vowel (see Baird, 2011). A fuller description of how these vowels are realized in one variety of Zapotec and the terms used are found in?.

[3] The tonal system of Santiago Laxopa Zapotec for verbs and other lexical categories is still being evaluated.

[4] There is ongoing debate about the status of strong syllables in Zapotec. The majority of evidence for the presence of stress comes from non-Northern Zapotec varieties (e.g., Chávez-Peón, 2008). At this time, evidence for stress in Northern Zapotec remains to be seen.

[5] Tone and voice quality are closely linked. By including only the positional interaction with voice quality, we can avoid collinear interactions that appear when we try to include tone in the interaction.

Adler, J., Foley, S., Pizarro-Guevara, J., Sasaki, K., and Toosarvandani, M. (**2018**). "The derivation of verb initiality in Santiago laxopa zapotec," in *A Reasonable Way to Proceed: Essays in Honor of Jim McCloskey*, edited by J. Merchant, L. Mikkelsen, D. Rudin, and K. Sasaki (University of California, Santa Cruz, Berkeley, CA), pp. 31–49.

Adler, J., and Morimoto, M. (**2016**). "Acoustics of phonation types and tones in Santiago Laxopa Zapotec," J. Acoust. Soc. Am. **140**(4), 3109.

Arellanes Arellanes, F. (**2010**). "Dos 'grados' De laringización con pertinencia fonológica en el zapoteco de San Pablo Güilá" ("Two 'degrees' of laryngization with phonological relevance in the Zapotec of San Pablo Guila"), in *Entre Cuerdas y Velo Estudios Fonológicos de Lenguas Otomangues* (*Between Strings and Veil Phonological Studies of Otomanguese Languages*), 1st ed., edited by E. H. Zendejas (El Colegio de México, México D. F.), pp. 85–122.

Ariza-García, A. (**2018**). "Phonation types and tones in Zapotec languages: A synchronic comparison," Acta Linguist. Petrop. **XIV**(2), 485–516.

Arras, K. O. (**1998**). An introduction to error propagation: Derivation, meaning and examples of equation CY= FX CX FXT, Technical Report EPFL-ASL-TR-98-01 R3 [Swiss Federal Institute of Technology Lausanne (EPFL), Lausanne, Switzerland].

Baird, B. (**2011**). "Phonetic and phonological realizations of 'broken glottal' vowels in k'ichee," in *Proceedings of FAMLi 1: Formal Approaches to Mayan Linguistics*, MIT Working Papers in Linguistics No. 63, edited by K. Shklovsky, P. Mateo Pedro, and J. Coon (MIT Press, Cambridge, MA).

Barzilai, M. L., and Riestenberg, K. J. (**2021**). "Context-dependent phonetic enhancement of a phonation contrast in San Pablo Macuiltianguis Zapotec," Glossa **6**(1), 18.

Brinkerhoff, M. L., Duff, J., and Wax Cavallaro, M. (**2021**). "Downstep in Santiago Laxopa Zapotec and the prosodic typology of VSO languages," in *Manchester Phonology Meeting*, Manchester, England.

Brinkerhoff, M. L., Duff, J., and Wax Cavallaro, M. (**2022**). "Tonal patterns and their restrictions in Santago Laxopa Zapotec," in *2022 Annual Meeting of the Society for the Study of Indigenous Languages of the Americas (SSILA)*, Virtual.

Burnham, K. P., and Anderson, D. R. (**2004a**). *Model Selection and Multimodel Inference* (Springer, New York).

Burnham, K. P., and Anderson, D. R. (**2004b**). "Multimodel inference: Understanding AIC and BIC in model selection," Sociol. Methods Res. **33**(2), 261–304.

Burnham, K. P., Anderson, D. R., and Huyvaert, K. P. (**2011**). "AIC model selection and multimodel inference in behavioral ecology: Some background, observations, and comparisons," Behav. Ecol. Sociobiol. **65**(1), 23–35.

Chai, Y., and Garellek, M. (**2022**). "On H1–H2 as an acoustic measure of linguistic phonation type," J. Acoust. Soc. Am. **152**(3), 1856–1870.

Chai, Y., and Ye, S. (**2022**). "Checked syllables, checked tones, and tone sandhi in Xiapu Min," Languages **7**(1), 47.

Chávez-Peón, M. E. (**2008**). "Phonetic cues to stress in a tonal language: Prosodic prominence in San Lucas Quiaviní Zapotec," in *Proceedings of the 2008 Annual Conference of the Canadian Linguistic Association*, edited by S. Jones (University of British Columbia, Vancouver, Canada).

Crowhurst, M. J., Kelly, N. E., and Teodocio, A. (**2016**). "The influence of vowel laryngealisation and duration on the rhythmic grouping preferences of Zapotec speakers," J. Phon. **58**, 48–70.

Esposito, C. M. (**2010**). "Variation in contrastive phonation in Santa Ana Del Valle Zapotec," J. Int. Phonetic Assoc. **40**(2), 181–198.

Esposito, C. M., and Khan, S. ud D. (**2020**). "The cross-linguistic patterns of phonation types," Lang. Linguist. Compass **14**(12) e12392.

Fischer-Jørgensen, E. (**1968**). "Phonetic analysis of breathy (murmured) vowels in Gujarati," Annu. Rep. Inst. Phonetics Univ. Copenhagen **2**, 35–85.

Foley, S., Kalivoda, N., and Toosarvandani, M. (**2018**). "Forbidden clitic clusters in zapotec," in *Proceedings of the Fifty-Third Annual Meeting of the Chicago Linguistic Society*, edited by D. Edmiston, M. Ermolaeva, E. Hakgüder, J. Lai, K. Montemurro, B. Rhodes, A. Sankhagowit, and M. Tabatowski (Chicago Linguistic Society, Chicago. IL), pp. 87–102.

Foley, S., and Toosarvandani, M. (**2022**). "Extending the person-case constraint to gender: Agreement, locality, and the syntax of pronouns," Ling. Inq. **53**, 1–40.

Garellek, M. (**2019**). "The phonetics of voice," in *The Routledge Handbook of Phonetics*, edited by W. F. Katz and P. F. Assmann (Routledge, Abingdon, Oxon), pp. 75–106.

Garellek, M. (**2020**). "Acoustic discriminability of the complex phonation system in !Xóõ," Phonetica **77**(2), 131–160.

Garellek, M., and Esposito, C. M. (**2023**). "Phonetics of White Hmong vowel and tonal contrasts," J. Int. Phonetic Assoc. **53**, 213–232.

Hanson, H. M. (**1997**). "Glottal characteristics of female speakers: Acoustic correlates," J. Acoust. Soc. Am. **101**(1), 466–481.

Iseli, M., Shue, Y.-L., and Alwan, A. (**2007**). "Age, sex, and vowel dependencies of acoustic measures related to the voice source," J. Acoust. Soc. Am. **121**(4), 2283–2295.

Kawahara, H., Cheveigne, A. D., and Patterson, R. D. (**1998**). "An instantaneous-frequency-based pitch extraction method for high-quality speech transformation: Revised TEMPO in the STRAIGHT-suite," in *5th International Conference on Spoken Language Processing (ICSLP 1998)* (ISCA), paper 0659–0.

Klatt, D. H., and Klatt, L. C. (**1990**). "Analysis, synthesis, and perception of voice quality variations among female and male talkers," J. Acoust. Soc. Am. **87**(2), 820–857.

Laver, J. D. M. (**1968**). "Voice quality and indexical information," Intl. J. Lang. Commun. Disord. **3**(1), 43–54.

Podesva, R. J. (**2016**). "Stance as a window into the language-race connection: Evidence from African American and white speakers in Washington, DC," in *Raciolinguistics: How Language Shapes Our Ideas About Race*, edited by H. S. Alim, J. R. Rickford, and A. F. Ball (Oxford University Press, Oxford, UK), pp. 203–219.

Seyfarth, S., and Garellek, M. (**2018**). "Plosive voicing acoustics and voice quality in Yerevan Armenian," J. Phon. **71**, 425–450.

Shue, Y.-L., Keating, P., Vicenik, C., and Yu, K. (**2011**). "VoiceSauce: A program for voice analysis," in *Proceedings of the 17th International Congress of Phonetic Sciences* (ICPhS XVII), Hong Kong), pp. 1846–1849.

Sichel, I., and Toosarvandani, M. (**2020a**). "The featural life of nominals."

Sichel, I., and Toosarvandani, M. (**2020b**). "Pronouns and attraction in sierra zapotec," in *Syntax and Semantics at Santa Cruz*, edited by A. Hedding and M. Hoeks (Linguistics Research Center, Santa Cruz, CA), Vol. IV.

Silverman, D. (**1997**). "Laryngeal complexity in Otomanguean vowels," Phonology **14**(2), 235–261.

Simpson, A. P. (**2012**). "The first and second harmonics should not be used to measure breathiness in male and female voices," J. Phon. **40**(3), 477–490.

Sjölander, K. (**2004**). "Snack sound toolkit" KTH. Stockholm, Sweden, available at https://www.speech.kth.se/snack/.

Sundberg, J. (**2022**). "Objective characterization of phonation type using amplitude of flow glottogram pulse and of voice source fundamental," J. Voice **36**(1), 4–14.