Voice Quality and Laryngeal Complexity in Santiago Laxopa Zapotec

Mykel Loren Brinkerhoff

University of California, Santa Cruz

6 June 2025

Outline

Outline

Research Overview

Overview:

This presentation explores what characterizes voice quality in a Santiago Laxopa Zapotec (SLZ), how is it is structured acoustically, and how this structure helps explain SLZ's laryngeal complexity.

Research Overview

Research Questions:

- How is phonation's acoustic space structured in a single language?
- Which measures are important for capturing phonation contrasts?
- How do these measures help explain SLZ's laryngeal complexity?

Research Overview

Answers to Research Questions:

- The acoustic space is three-dimensional and correlated with glottal-airflow (D1/D3) and nonmodal-to-modal (D2) continua.
- Only a handful of measures are needed for capturing contrasts.
- SLZ's laryngeal complexity is weakly articulated and shows phasing.

Outline

What is Voice Quality?

- Broad sense = long-term characteristics of an individual's voice (Abercrombie 1967, Laver 1980).
- Narrow sense = how larynx affects phonetic characteristics of speech sounds (e.g., Esling et al. 2019).
- Sometimes used interchangeably with phonation
  - Barzilai & Riestenberg 2021 use *voice quality* for phonetic description and *phonation* for phonological contrasts.

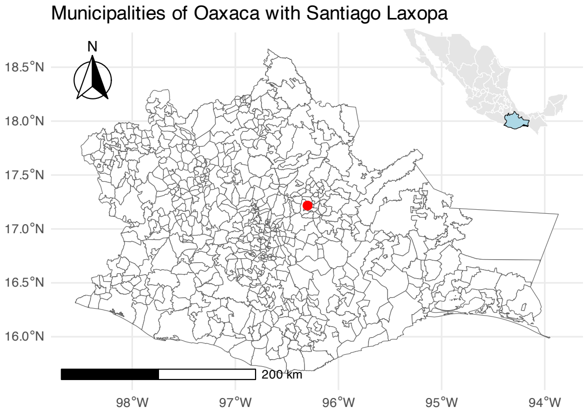
How is voice quality used?

- Paralinguistic information by "indexing the biological, psychological, and social characteristics of the speaker" (e.g., Laver 1968, Podesva 2016)
- Phonological contrasts (e.g., Esposito & Khan 2020).
  - Gujarati's breathy and modal vowels (Fischer-Jørgensen 1968).
  - Mazatec's breathy, creaky, and modal vowels (Silverman 1997a).

Outline

What is Santiago Laxopa Zapotec

- Santiago Laxopa Zapotec (SLZ; *Dille'xhunh Laxup*) is a Sierra Norte variety of Zapotec (Oto-Manguean).
- Spoken by c. 1,000 speakers in Santiago Laxopa and in diaspora.



Why study SLZ?

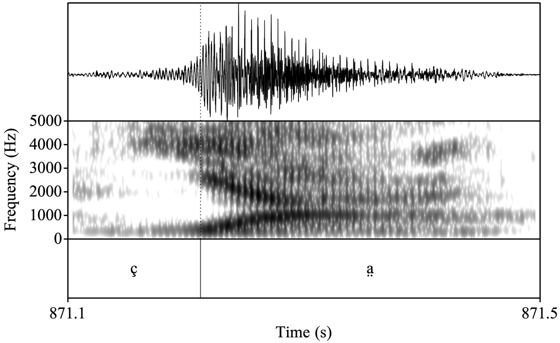
- SLZ is a relatively understudied language
  - Contributes to our understanding of language diversity.
  - Test claims about language.
- SLZ has a complex laryngeal system that has not been fully documented.
  - SLZ has a four-way phonation contrast.
  - SLZ has a five-way tonal contrast.
  - These contrasts are orthogonal, meaning that they can occur independently of each other.

Phonation in SLZ

- SLZ has a four-way phonation contrast:
  - Modal ([a])
  - Breathy ([a̤])
  - Checked ([aʔ] or [aa̰])
  - Rearticulated ([aʔa], [aa̰a], or [a̰])

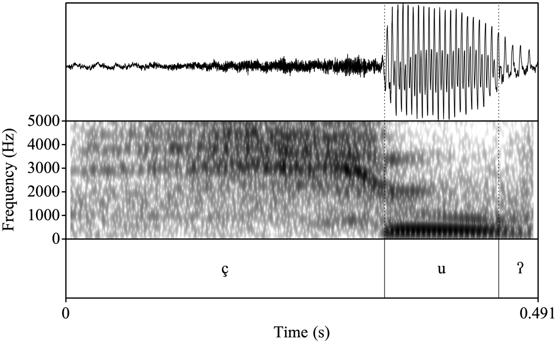
Breathy Phonation in SLZ

- Characterized by aperiodicity and noise in the harmonic structure.
- Longer duration than modal voice
- Lower fundamental frequency ($f0$)



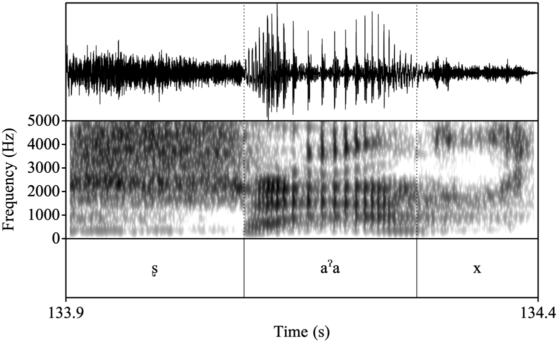
Checked Phonation in SLZ

- Glottal occlusion at the end of vowel or creaky voice in second half vowel.
- Decrease in voicing amplitude during second half.



Rearticulated Phonation in SLZ

- Glottal occlusion or creaky voice in the middle of the vowel.
- Longer duration than modal voice
- Decrease in voicing amplitude during first half.



Tone in SLZ

- SLZ shows a five-way tonal contrast on monosyllabic nouns (Brinkerhoff, Duff & Wax Cavallaro 2022).

Tone	Example	Transcription	Gloss
High	*xha*	[ʐa˥]	‘clothing.poss’
Mid	*lhill*	[ɾiʒ˧]	‘house.poss’
Low	*yu’*	[ç̦uˀ˨]	‘earth’
Rising	*yu’u*	[ç̦uˀu˧˥]	‘lime (Sp. *cal*)’
Falling	*yu’u*	[ç̦uˀu˦˨]	‘house’

Interaction of Tone and Phonation in SLZ

- Tone and phonation are orthogonal.

	Modal	**Breathy**	**Checked**	**Rearticulated**
High	✓	–	✓	✓
Mid	✓	✓	✓	✓
Low	✓	✓	✓	✓
Rising	✓	✓	–	✓
Falling	✓	✓	✓	✓

Outline

Measuring voice quality

- Long been established that phonation has correlates in the acoustic signal (e.g., Fischer-Jørgensen 1968, Klatt & Klatt 1990).
- Fischer-Jørgensen (1968) found that a strengthened fundamental (i.e., H1) correlated with breathy vowels in Gujarati.
    - Proposed H1—H2 amplitude as a measure to normalize for overall intensity differences
- Subsequent research shows H1—H2 is useful for classifying voice quality contrasts.
- However, it is not without its problems (see Chai & Garellek 2022 for an overview).

Measuring voice quality

- H2 and other normalizations (A1, A3) are really attempts to understand the relative strength of the fundamental (H1) to higher harmonic energy
- H1 and H2 are affected by subglottal pressure differently (Sundberg 2022)
  - This undermines the original reasoning behind H1—H2
- Researchers sometimes find that H1—H2 does not distinguish phonation types (e.g., Esposito 2010, 2012, Garellek & Esposito 2023).
- Error in measuring H1—H2 is uncomfortably high (e.g., Chai & Garellek 2022).

Measuring voice quality

- Many measures have been proposed to capture voice quality.
- Measures can be broadly categorized into several categories (Gordon & Ladefoged 2001):
  - Periodicity (e.g., HNR or CPP)
  - Energy
  - Spectral tilt (e.g., H1*−H2*, residual H1*)
  - Pitch (e.g., $f0$)
  - Duration
- Linguists have used combinations of these measures to model phonation (e.g., Blankenship 2002, Brunelle & Kirby 2016, Esposito 2012).

New measures of voice quality

- New measures are being developed all the time.
- Chai & Garellek's (2022) residual H1* is a new measure of spectral tilt that is more robust than traditional spectral tilt measures.
  - Does this by removing the the effects of Energy from the H1* measure.
  - This was the original goal of Fischer-Jørgensen (1968)
- Brinkerhoff & McGuire (2025) found that residual H1* is a better measure of spectral tilt than H1*−H2* for SLZ.

Too many measures

- There are many measures of voice quality, but it is not clear which ones are the most useful.
- It also makes it difficult to compare results across studies.
- VoiceSauce makes this a common issue (Shue et al. 2011).
  - Computes almost every acoustic measure related to voice quality.
  - Many of these measures are potentially redundant (e.g., three different algorithms for calculating $f0$, formants, and bandwidths).
  - Hard to know which measures to focus on.

Too many measures

- Kreiman et al. (2014, 2021) tackle this by proposing a psychoacoustic model of voice quality.
- Found certain measures are perceptually more important than others.

Model Component	**Parameters**
Harmonic source spectral shape	H1–H2 H2–H4 H4–2 kHz 2 kHz–5 kHz
Inharmonic source excitation	Spectrally-shaped noise-to-harmonics ratio
Time-varying source characteristics	$f0$ mean and standard deviation (or $f0$ track) Amplitude mean and standard deviation (or amplitude track)
Vocal tract transfer function	Formant frequencies/bandwidths Spectral zeroes/bandwidths

Outline

Modeling voice quality

- Early models proposed that voice quality is one dimensional and represents glottal airflow (Ladefoged 1971, Ladefoged & Maddieson 1996).

Open Glottis ⟷ Closed Glottis

[h] Breathy Modal Creaky [ʔ]

Voice quality's multidimensionality

- More recent work has shown that voice quality is not one-dimensional, but minimally five-dimensional (e.g., Garellek et al. 2016, Kreiman et al. 2021).
  - Especially in the case of individual speaker differences.
- Garellek et al. (2013) has argued that dimensionality might not be as complex for capturing phonation contrasts.

Keating et al. (2023)

- Explored phonation's cross-linguistic acoustic space.
- Found a two-dimensional space for phonation across 11 languages.
    1. First dimension = nonmodal-to-modal continuum.
    2. Second dimension = glottal-airflow continuum.
- Found languages with more contrasts used more of the acoustic space than languages with fewer contrasts.
- Found correlations between dimensions and acoustic measures.
    1. First dimension = periodicity and energy measures.
    2. Second dimension = spectral tilt and periodicity measures.

Keating et al. (2023)

- Colors = phonation.
- Labels = language and phonation.

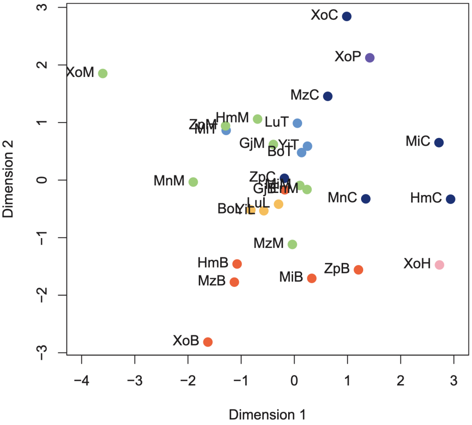


Figure: Figure 4 from Keating et al. (2023).

Outline

Data

- Data comes from SLZ fieldwork from Summer 2022.
- Production data was collected from 10 speakers (5 male/5 female).
- Recordings were made in a quiet room using a Zoom H4n Pro handheld recorder (16-bit, 44.1 kHz).

Methods

- Participants produced 72 words in a carrier phrase 3 times each.
  - *Shnia' X chonhe lhas* [ʃnːiaˀ X tʃone ɾas] "I say X three times."
- Words covered the four-way phonation contrast.
- Tone was not balanced due to frequency constraints.

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

Methods

- Vowels were segmented in Praat from the second glottal pulse to the last glottal pulse before the drop in amplitude (Garellek 2020)
- Acoustic measures were extracted using VoiceSauce (Shue et al. 2011).

Data cleaning

- All acoustic measures were normalized by speaker.
- Outliers were removed using the following criteria:
  - If the z-scored $f0$ was greater than 3 or less than -3.
  - If Energy was equal to 0.
  - If the Mahalanobis distance was greater than 6 in the F1-F2 panel for each vowel category.
- Calculated residual H1* (Chai & Garellek 2022).

Outline

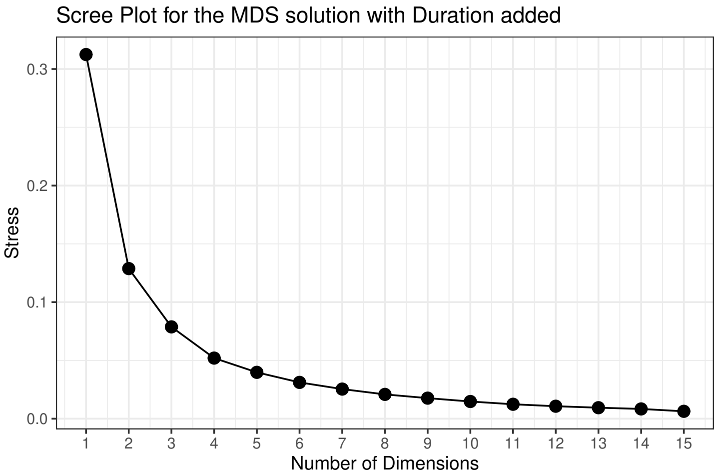
MDS analysis

- Multidimensional scaling (MDS; Kruskal & Wish 1978) is a dimensionality reduction technique.
- Used to visualize the structure of a high-dimensional space in a lower-dimensional space.
- Especially effective when many variables contribute to the structure of the space.
- More robust than principal component analysis (PCA) or linear discriminant analysis (LDA) because it does not assume linearity or normality of the data.

Predictions for the MDS analysis

- Small number of dimensions needed to capture the structure of the space.
- Phonation types will occupy distinct regions of the space.
- Dimensions should be comparable to those found in Keating et al. (2023).
- Dimensions should show correlations with acoustic measures.

Number of Dimensions needed



Explanation of the plots

- Each point = unique speaker x phonation combination.
- Colors = phonation type.
  - Modal = black
  - Breathy = orange
  - Checked = blue
  - Rearticulated = green
- Axes = dimensions of the acoustic space.

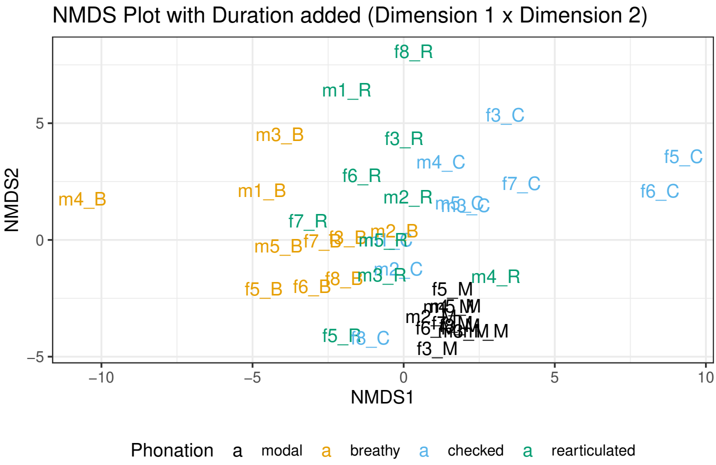
Dimensionality in SLZ

- Scan the QR code to see the three-dimensional space.
- Or use the link: https://bit.ly/3St9S6g

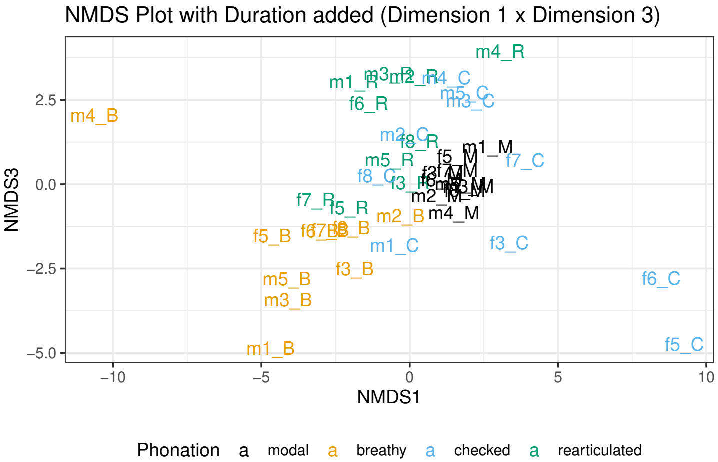
Takeaways from 3D plot

- Occupies a three-dimensional space.
- Clear groupings of phonation types.
- Very little overlap between phonation types.

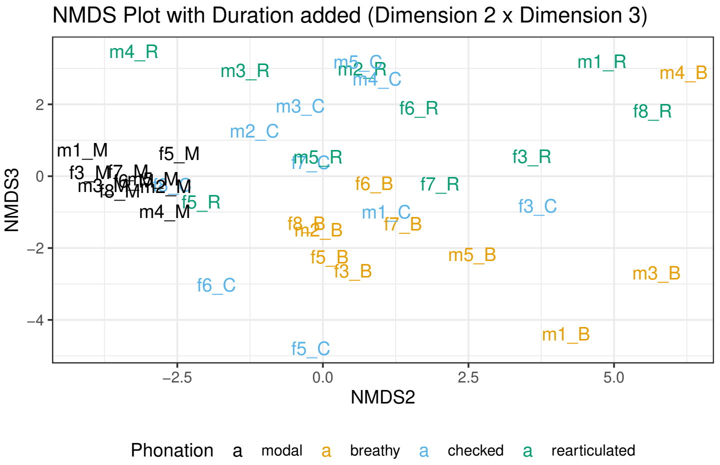
Dimensionality in SLZ



Dimensionality in SLZ



Dimensionality in SLZ



Summary of Dimensions

- Dimension 1 (D1) gives a rough continuum from breathy to creaky.
- Dimension 2 (D2) gives a rough continuum from modal to nonmodal.
- Dimension 3 (D3) gives a rough continuum from breathy to creaky.

Correlation to Acoustic Measures

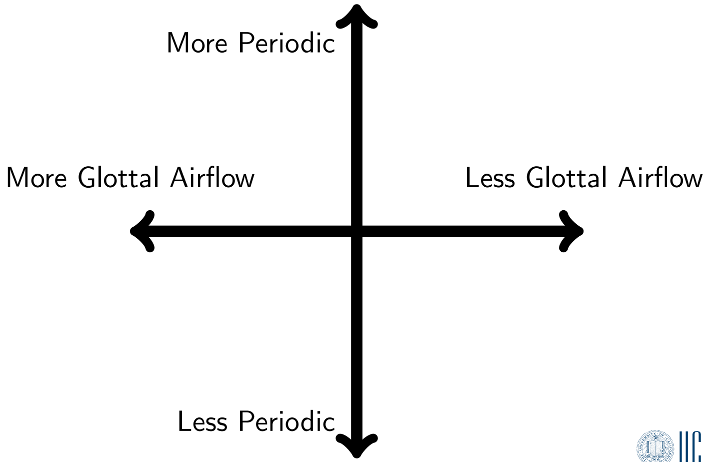
- D1 correlated with spectral tilt measures:
  - H1*–A1* ($r^2 = -0.83$)
  - H1*–A2* ($r^2 = -0.86$)
  - H1*–A3* ($r^2 = -0.81$)
- D2 correlated with periodicity and energy:
  - HNR<500 Hz ($r^2 = -0.79$)
  - HNR<1500 Hz ($r^2 = -0.80$)
  - Energy ($r^2 = -0.79$)
- D3 correlated with different spectral tilts:
  - residual H1* ($r^2 = -0.72$)
  - H2*–H4* ($r^2 = -0.69$)
  - H2* ($r^2 = -0.68$)

Summary of Acoustic Landscape

- SLZ's phonation occupies a three-dimensional space.
- Dimensions are correlated with glottal-airflow continuum (D1/D3) and nonmodal-to-modal continuum (D2).
- Dimensions are similar to those found in Keating et al. (2023).

Summary of Acoustic Landscape

- Acoustic space can be reduced to two dimensions.
- More dimensions add information about these two dimensions.



Outline

What are decision trees?

- Decision trees are a type of statistical modeling (Breiman et al. 1986).
- Used for classification and regression tasks.
- Splits the data into subsets based on variable values.
- Splits and rules are graphically represented as a tree structure.

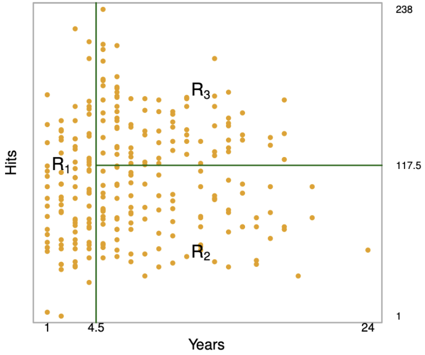


Figure: Figure 8.2 from James et al. (2021)

What are decision trees?

- Splits and their rules are graphically represented as a tree structure.

Figure: Figure 8.1 from James et al. (2021)

Use in linguistics

- Decision trees are not commonly used in linguistics.
- Used by Tagliamonte & Baayen (2012) for sociolinguistic variation.
- Used by Keating et al. (2023) to analyze phonation contrasts.
- One issue with decision trees is *high variance*.

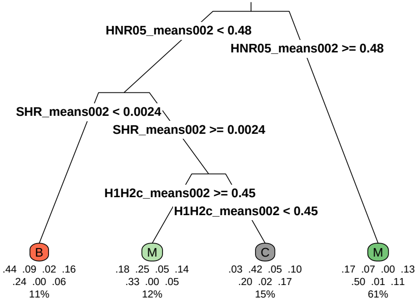


Figure: Figure 8 from Keating et al. (2023).

What are Random Forests?

- Random forests are an extension of decision trees (Breiman 2001).
- Instead a single decision tree, multiple decision trees (i.e., a forest) are created.
- Trees trained on a random subset of the data and a random subset of the parameters.
- The final prediction is made by averaging the predictions of all trees.
- This reduces overfitting and improves accuracy.

Tuning Random Forests

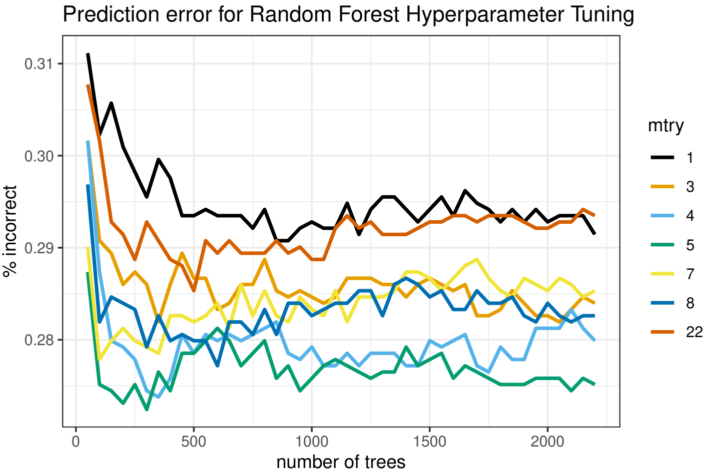
- Requires tuning to find the best parameters (e.g., Boehmke & Greenwell 2019).
- Two important parameters need to be tuned:
  - How many trees to create in the forest.
  - How many variables to consider when splitting the data.

Tuning Random Forests

- Tune with hyperparameter grid search (Boehmke & Greenwell 2019).
- Number of trees is sequence of trees up to 2000.
- Number of variables (i.e., $mtry$) = 5%, 15%, 25%, 40%, 100%, of the total number of variables (p).
  - Also include $mtry = \sqrt{p}$ (default for classification tasks) and $mtry = \frac{p}{3}$ (for regression tasks) as options.

Parameters for SLZ Analysis

- $mtry = 5$
- 300 trees.



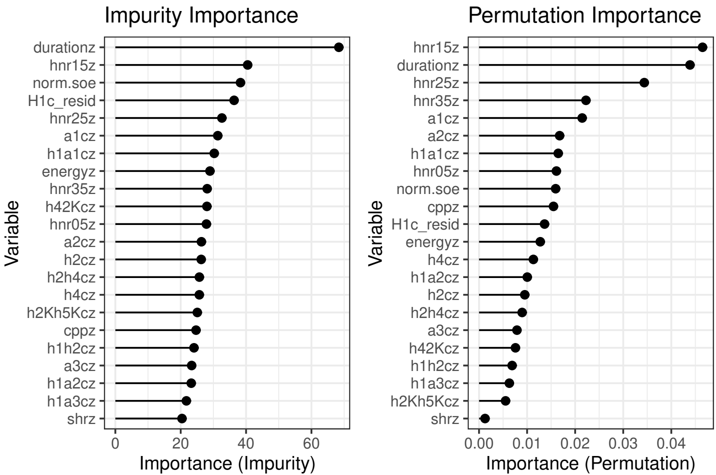
Interpreting Random Forests

- Improvement in accuracy comes at the cost of interpretability.
- No single tree can be interpreted, but the forest as a whole.
- Variable importance can be calculated to determine which variables are most important for the model.
- Two common methods for calculating variable importance:
  - How much they reduce impurity at the splits.
  - How much the model's accuracy decreases when a variable is permuted.

Predictions for Random Forests

- Certain measures expected to be more important than others.
- Should be similar to the measures found in the MDS analysis.
  - Spectral tilt measures (e.g., H1*−A1*, residual H1*).
  - Periodicity measures (e.g., HNR, CPP).
  - Energy measures (e.g., Energy, SoE).

Variable Importance in SLZ



Variable Importance

- Variable importance shows that only a handful of measures are important for capturing phonation contrasts.
  1. Duration
  2. A1*
  3. H1*−A1*
  4. Residual H1*
  5. HNR $<$ 1500 Hz
  6. Strength of Excitation (SoE)

Summary of Random Forest Results

- Only a handful of measures are important for capturing phonation contrasts.
- Duration, A1*, H1*−A1*, residual H1*, HNR $<$ 1500 Hz, and SoE are the most important measures.
- Overlap between MDS' correlated measures and the important measures from the Random Forests.
- Provides us with a subset of measures to use for quantifying laryngeal complexity.
  - Strength of Excitation (SoE)
  - HNR $<$ 1500 Hz

Outline

What is Laryngeal Complexity

- *Laryngeal complexity* describes languages that use both contrastive tone and phonation (Silverman 1997a,b).
- Very common in the Oto-Manguean language family, but is also found in other languages.

Phonation's phasing

- One of the key features of laryngeal complexity is *phasing*.
- To maximize production and perception constraints, tone and phonation are not produced simultaneously, but rather in sequence.
- This produces a complex segment that has a modal portion for tone and a nonmodal portion for phonation.
- There are three types of phasing:
  - *Prevocalic* phasing = phonation is produced before modal portion.
  - *Postvocalic* phasing = phonation is produced after modal portion.
  - *Interrupted* phasing = phonation is produced in the middle of the modal portion.

Implicational hierarchy of patterns

- Silverman (1997a) proposed a hierarchy of laryngeal complexity phasing pattern.

Language	Prevocalic	Postvocalic	Interrupted
Jalapa Mazatec	hV˥, ʔV˥	—	—
Comaltepec Chinantec	hV˥, ʔV˥	Vh˥, Vʔ˥	—
Copala Trique	hV˥, ʔV˥	Vh˥, Vʔ˥	VhV˥, VʔV˥

Previous research on laryngeal complexity

- *Descriptive studies* focus on the patterns of tone and phonation (e.g., Ariza-García 2018, Frazier 2013, Pickett, Villalobos & Marlett 2010).
  - Often use impressionistic data or small datasets.
  - Do not provide a detailed analysis of the acoustic properties of tone and phonation.
- *Instrumental studies* focus the acoustic properties of laryngeal complexity (e.g., DiCanio 2012, Garellek & Keating 2011, Kelterer & Schuppler 2020).
  - $f0$ perturbations (e.g., DiCanio 2012).
  - Strength of Excitation (SoE) also used (Weller et al. 2023a,b, 2024).

Previous research on Zapotec laryngeal complexity

- Herrera Zendejas (2000) studies laryngeal complexity in Amuzgo and Zapotec
- Herrera Zendejas (2000) claims that Zapotec languages:
    - Lack phasing altogether.
    - Tone and laryngealization "have reached a degree of equilibrium, weakening enough to be present simultaneously (558)".

Outline

Quanitfying Laryngeal Complexity

- Three measures were used to quantify laryngeal complexity:
  1. $f0$
  2. HNR $<$ 1500 Hz
  3. Strength of Excitation (SoE)
- These measures were chosen based on:
  - Previous research on laryngeal complexity (e.g., DiCanio 2012).
  - Correlations with the MDS dimensions.
  - Importance in the Random Forests analysis.
- Assessed these measures using Generalized Additive Mixed Models (GAMM; Hastie & Tibshirani 1986, Wood 2017).

What are Generalized Additive Mixed Models (GAMMs)

- Similar to GLMs, but without linearity.
- Allows modeling of nonlinear relationships between predictors and response variables.
- Allows smoothing functions of continuous predictors.
- Excellent for modeling complex non-linear relationships in linguistic data (e.g., Coetzee et al. 2022, Wieling 2018).

Evaluating GAMMs

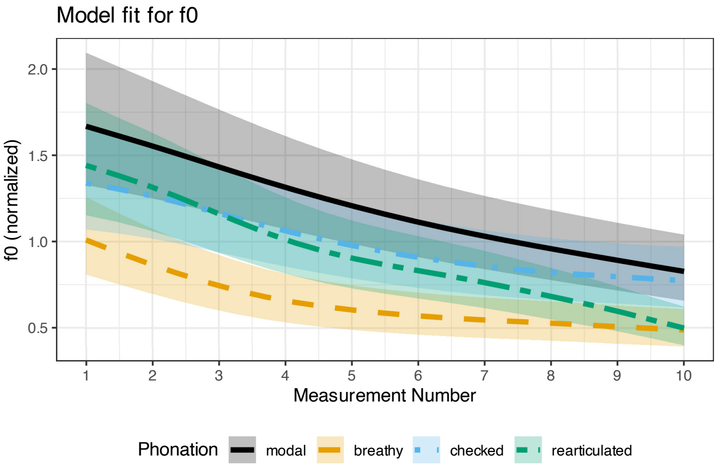
- GAMMs can be difficult to evaluate (Sóskuthy 2017, 2021).
- General approach is to compare model fits and difference plots visually.
- Model fits show how well the model captures the data.
- Difference plots show how different the fixed effects are from each other.

- GAMMs were fitted for each measure using the bam function from mgcv package in R (Wood 2017).
- Each model included:
  - Phonation as a fixed effect.
  - Smoothing term for time (i.e., measurement number)
  - Smoothing term for time to vary by phonation type.
  - Speaker as a random smooth.
  - Interaction between speaker and phonation as a random smooth.
  - Tensor product interaction between time and repetition as a random smooth.

GAMMs for laryngeal complexity

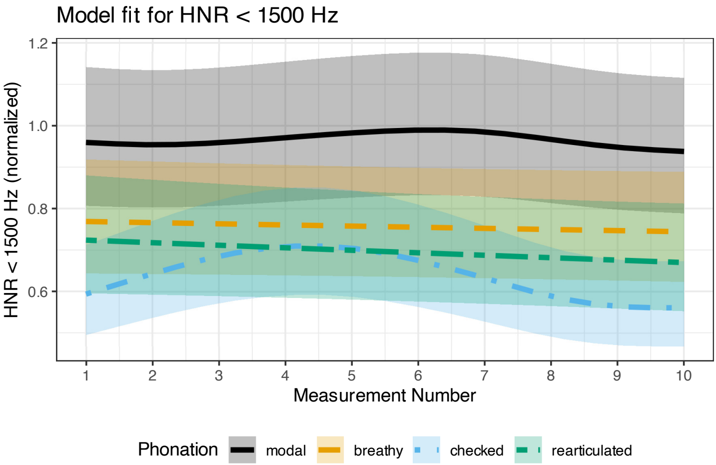
- Models capture four things:
  1. Main effect of phonation on the acoustic measures.
  2. Nonlinear relationships in time both overall and specific for each phonation type.
  3. Speaker variability and its interaction with phonation.
  4. Variability in the data due to the different repetitions from the speakers.
- More accurate modeling of the data.
- Better understanding of how phonation affects the acoustic measures.

Model fit for $f0$



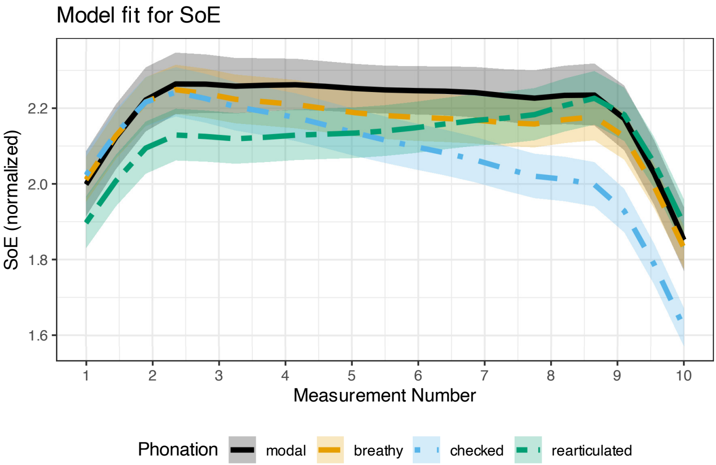
Difference plots for $f0$

Model fit for HNR

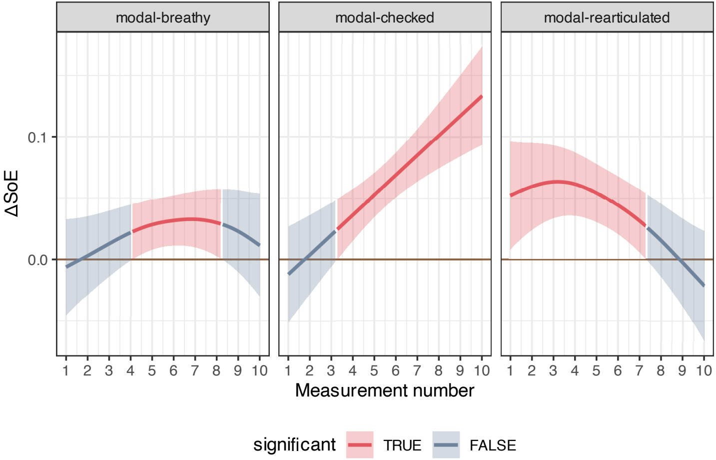


Difference plots for HNR

Model fit for SoE



Difference plots for SoE



Summary of Laryngeal Complexity

- SLZ shows phasing.
  - Breathy shows postvocalic phasing.
  - Checked shows postvocalic phasing.
  - Rearticulated shows prevocalic phasing.
  - No evidence for interrupted phasing.
- Implicational hierarchy of laryngeal complexity holds.
- Laryngealization is weakly articulated in SLZ.

Summary of Research Questions

Research Questions:

- How is phonation's acoustic space structured in a single language?
- Which measures are important for capturing phonation contrasts?
- How do these measures help explain SLZ's laryngeal complexity?

Summary of Answers

Answers to Research Questions:

- The acoustic space is three-dimensional and correlated with glottal-airflow (D1/D3) and nonmodal-to-modal (D2) continua.
- Only a handful of measures are needed for capturing contrasts.
- SLZ's laryngeal complexity is weakly articulated and shows phasing.

Implications and big picture

- Showcases Keating et al.'s (2023) methodology for narrowing down which acoustic measures to use.
- Analysis is congruent with Keating et al.'s (2023) findings.
- SLZ's phonation is consistent with Silverman's (1997a) claims about laryngeal complexity.
- Shows that HNR and SoE are useful measures for quantifying laryngeal complexity.

Next steps

- What are the perceptual cues that SLZ speakers use to distinguish phonation types?
- How do the dimensions of the acoustic space relate to phonological features?
- Do we actually observe interrupted phasing in other Zapotec languages?
- How do we analyze laryngeal complexity's phasing patterns in the phonology?

Duxhklhenhu' lhe' (Thank you all!)

Acknowledgements

- Thank you to Maestra Fe Silva-Robles for introducing me and teaching me about *Dille'xhunh*
- Thank you to the other speakers of *Dille'xhunh* for sharing their time and language expertise.
- Thank you to Grant McGuire, Jaye Padgett, Marc Garellek, Ryan Bennett, Jack Duff, Maya Wax Cavallaro, and many others for their help and discussions during all stages of this project.

Acknowledgements

This work is supported by funding from:

- The National Science Foundation under Grant No. 2019804
- The Humanities Institute at UC Santa Cruz
- The Jacobs Research Funds

References I

Abercrombie, David. 1967. *Elements of general phonetics*. Edinburgh: Edinburgh University Press.

Ariza-García, Andrea. 2018. Phonation types and tones in Zapotec languages: A synchronic comparison. *Acta Linguistica Petropolitana* XIV(2). 485–516. https://doi.org/10.30842/alp2306573714220.

Barzilai, Maya L. & Katherine J. Riestenberg. 2021. Context-dependent phonetic enhancement of a phonation contrast in San Pablo Macuiltianguis Zapotec. *Glossa: a journal of general linguistics* 6(1). 1–36. https://doi.org/10.5334/gjgl.959.

Blankenship, Barbara. 2002. The timing of nonmodal phonation in vowels. *Journal of Phonetics* 30(2). 163–191. https://doi.org/10.1006/jpho.2001.0155.

Boehmke, Brad & Brandon M. Greenwell. 2019. *Hands-On Machine Learning with R*. New York: Chapman and Hall/CRC. 484 pp. https://doi.org/10.1201/9780367816377.

References II

Breiman, Leo. 2001. Random Forests. *Machine Learning* 45(1). 5–32. https://doi.org/10.1023/A:1010933404324.

Breiman, Leo, Jerome H. Friedman, Richard A. Olshen & Charles J. Stone. 1986. *Classification and regression trees*. Boca Raton, Fla.: Taylor and Francis. 358 pp.

Brinkerhoff, Mykel Loren, John Duff & Maya Wax Cavallaro. 2022. Tonal patterns and their restrictions in Santago Laxopa Zapotec.

Brinkerhoff, Mykel Loren & Grant McGuire. 2025. Using residual H1* for voice quality research. *JASA Express Letters* 5(2). 025501. https://doi.org/10.1121/10.0035881.

Brunelle, Marc & James Kirby. 2016. Tone and Phonation in Southeast Asian Languages. *Language and Linguistics Compass* 10(4). 191–207. https://doi.org/10.1111/lnc3.12182.

References III

Chai, Yuan & Marc Garellek. 2022. On H1–H2 as an acoustic measure of linguistic phonation type. *The Journal of the Acoustical Society of America* 152(3). 1856–1870. https://doi.org/10.1121/10.0014175.

Coetzee, Andries W., Patrice Speeter Beddor, Will Styler, Stephen Tobin, Ian Bekker & Daan Wissing. 2022. Producing and perceiving socially structured coarticulation: Coarticulatory nasalization in Afrikaans. *Laboratory Phonology* 13(1). https://doi.org/10.16995/labphon.6450.

DiCanio, Christian. 2012. Coarticulation between tone and glottal consonants in Itunyoso Trique. *Journal of Phonetics* 40(1). 162–176. https://doi.org/10.1016/j.wocn.2011.10.006.

Esling, John H., Scott R. Moisik, Allison Benner & Lise Crevier-Buchman. 2019. *Voice Quality: The Laryngeal Articulator Model*. 1st edn. (Cambridge Studies in Linguistics 162). Cambridge University Press. https://doi.org/10.1017/9781108696555.

References IV

Esposito, Christina M. 2010. Variation in contrastive phonation in Santa Ana Del Valle Zapotec. *Journal of the International Phonetic Association* 40(2). 181–198. https://doi.org/10.1017/S0025100310000046.

Esposito, Christina M. 2012. An acoustic and electroglottographic study of White Hmong tone and phonation. *Journal of Phonetics* 40(3). 466–476. https://doi.org/10.1016/j.wocn.2012.02.007.

Esposito, Christina M. & Sameer ud Dowla Khan. 2020. The cross-linguistic patterns of phonation types. *Language and Linguistics Compass* 14(12). https://doi.org/10.1111/lnc3.12392.

Fischer-Jørgensen, Eli. 1968. Phonetic Analysis of Breathy (Murmured) Vowels in Gujarati. *Annual Report of the Institute of Phonetics University of Copenhagen* 2. 35–85. https://doi.org/10.7146/aripuc.v2i.130674.

References V

Frazier, Melissa. 2013. The phonetics of Yucatec Maya and the typology of laryngeal complexity. *STUF - Language Typology and Universals* 66(1). 7–21. https://doi.org/10.1524/stuf.2013.0002.

Garellek, Marc. 2020. Acoustic Discriminability of the Complex Phonation System in !Xóõ. *Phonetica* 77(2). 131–160. https://doi.org/10.1159/000494301.

Garellek, Marc & Christina M. Esposito. 2023. Phonetics of White Hmong vowel and tonal contrasts. *Journal of the International Phonetic Association* 53(1). 213–232. https://doi.org/10.1017/S0025100321000104.

Garellek, Marc & Patricia Keating. 2011. The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. *Journal of the International Phonetic Association* 41(2). 185–205. https://doi.org/10.1017/S002510031100019З.

References VI

Garellek, Marc, Patricia Keating, Christina M. Esposito & Jody Kreiman. 2013. Voice quality and tone identification in White Hmong. *The Journal of the Acoustical Society of America* 133(2). 1078–1089. https://doi.org/10.1121/1.4773259.

Garellek, Marc, Robin Samlan, Bruce R. Gerratt & Jody Kreiman. 2016. Modeling the voice source in terms of spectral slopes. *The Journal of the Acoustical Society of America* 139(3). 1404–1410. https://doi.org/10.1121/1.4944474.

Gordon, Matthew & Peter Ladefoged. 2001. Phonation types: a cross-linguistic overview. *Journal of Phonetics* 29(4). 383–406. https://doi.org/10.1006/jpho.2001.0147.

Hastie, Trevor & Robert Tibshirani. 1986. Generalized Additive Models. *Statistical Science* 1(3). 297–310. https://doi.org/10.1214/ss/1177013604.

References VII

Herrera Zendejas, Esther. 2000. Amuzgo and Zapotec: Two More Cases of Laryngeally Complex Languages. *Anthropological Linguistics* 42(4). 545–563. http://www.jstor.org/stable/30028566.

James, Gareth, Daniela Witten, Trevor Hastie & Robert Tibshirani. 2021. *An introduction to statistical learning: with applications in R*. Second edition (Springer Texts in Statistics). New York: Springer. 1 p. https://doi.org/10.1007/978-1-0716-1418-1.

Keating, Patricia, Jianjing Kuang, Marc Garellek, Christina M. Esposito & Sameer ud Dowla Khan. 2023. A cross-language acoustic space for vocalic phonation distinctions. *Language* 99(2). 351–389. https://doi.org/10.1353/lan.2023.a900090.

Kelterer, Anneliese & Barbara Schuppler. 2020. Phonation type contrasts and tone in Chichimec. *The Journal of the Acoustical Society of America* 147(4). 3043–3059. https://doi.org/10.1121/10.0001015.

References VIII

Klatt, Dennis H. & Laura C. Klatt. 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America* 87(2). 820–857. https://doi.org/10.1121/1.398894.

Kreiman, Jody, Bruce R. Gerratt, Marc Garellek, Robin Samlan & Zhaoyan Zhang. 2014. Toward a unified theory of voice production and perception. *Loquens* 1(1). e009. https://doi.org/10.3989/loquens.2014.009.

Kreiman, Jody, Yoonjeong Lee, Marc Garellek, Robin Samlan & Bruce R. Gerratt. 2021. Validating a psychoacoustic model of voice quality. *The Journal of the Acoustical Society of America* 149(1). 457–465. https://doi.org/10.1121/10.0003331.

Kruskal, Joseph & Myron Wish. 1978. *Multidimensional Scaling*. SAGE Publications, Inc. https://doi.org/10.4135/9781412985130.

References IX

Ladefoged, Peter. 1971. *Preliminaries to linguistic phonetics*. Chicago: University of Chicago.

Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. (Phonological Theory). Oxford, OX, UK ; Cambridge, Mass., USA: Blackwell Publishers. 425 pp.

Laver, John D. M. 1968. Voice Quality and Indexical Information. *British Journal of Disorders of Communication* 3(1). 43–54. https://doi.org/10.3109/13682826809011440.

Laver, John D. M. 1980. The Phonetic Description of Voice Quality. *Cambridge Studies in Linguistics London* 31. 1–186.

Pickett, Velma B., María Villalobos Villalobos & Stephen A. Marlett. 2010. Isthmus (Juchitán) Zapotec. *Journal of the International Phonetic Association* 40(3). 365–372. https://doi.org/10.1017/S0025100310000174.

References X

Podesva, Robert J. 2016. Stance as a Window into the Language-Race Connection: Evidence from African American and White Speakers in Washington, DC. In H. Samy Alim, John R. Rickford & Arnetha F. Ball (eds.), *Raciolinguistics: How Language Shapes Our Ideas About Race*, 203–219. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190625696.003.0012.

Shue, Yen-Liang, Patricia Keating, Chad Vicenik & Kristine Yu. 2011. VoiceSauce: A program for voice analysis. In *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII)*, 1846–1849. Hong Kong.

Silverman, Daniel. 1997a. Laryngeal complexity in Otomanguean vowels. *Phonology* 14(2). 235–261. https://doi.org/10.1017/S0952675797003412.

Silverman, Daniel. 1997b. *Phasing and recoverability*. (Outstanding Dissertations in Linguistics). New York: Garland Pub. 242 pp.

References XI

Sóskuthy, Márton. 2017. *Generalised Additive Mixed Models for Dynamic Analysis in Linguistics: A Practical Introduction*. arXiv: 1703.05339 [stat]. Pre-published.

Sóskuthy, Márton. 2021. Evaluating generalised additive mixed modelling strategies for dynamic speech analysis. *Journal of Phonetics* 84. 101017. https://doi.org/10.1016/j.wocn.2020.101017.

Sundberg, Johan. 2022. Objective Characterization of Phonation Type Using Amplitude of Flow Glottogram Pulse and of Voice Source Fundamental. *Journal of Voice* 36(1). 4–14. https://doi.org/10.1016/j.jvoice.2020.03.018.

Tagliamonte, Sali A. & R. Harald Baayen. 2012. Models, forests, and trees of York English: *Was/were* variation as a case study for statistical practice. *Language Variation and Change* 24(2). 135–178. https://doi.org/10.1017/S0954394512000129.

References XII

Weller, Jae, Jeremy Steffman, Félix Cortés & Iara Mantenuto. 2023a. Interactions of tone, glottalization, and word shape in San Sebastián Del Monte Mixtec.

Weller, Jae, Jeremy Steffman, Félix Cortés & Iara Mantenuto. 2023b. Lexical tone and vowel duration in San Sebastián del Monte Mixtec. *The Journal of the Acoustical Society of America* 153. A293. https://doi.org/10.1121/10.0018900.

Weller, Jae, Jeremy Steffman, Félix Cortés & Iara Mantenuto. 2024. Voice Quality and Tone in San Sebastián del Monte Rearticulated and Modal Vowels.

Wieling, Martijn. 2018. Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics* 70. 86–116. https://doi.org/10.1016/j.wocn.2018.03.002.

References XIII

Wood, Simon N. 2017. *Generalized Additive Models: An Introduction with R*. 2nd edn. (Texts in Statistical Science Series). Boca Raton, FL: Chapman and Hall/CRC. https://doi.org/10.1201/9781315370279.

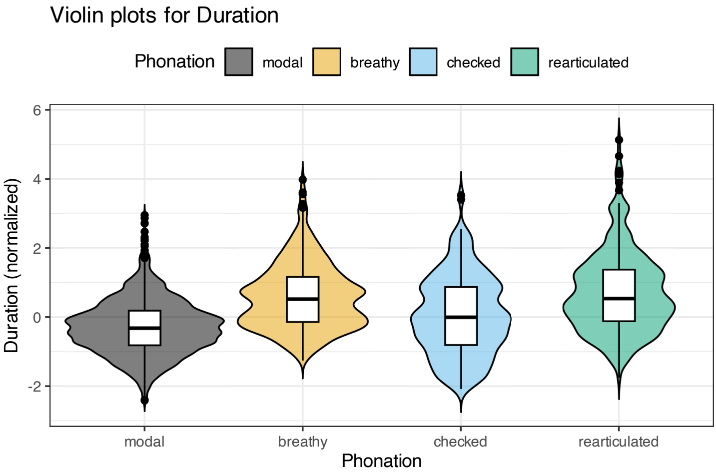
Correlations of Acoustic Measures to Dimensions I

Acoustic Measure	NMDS1	NMDS2	NMDS3	NMDS4
H1*—H2*	-0.221	-0.339	0.031	0.314
H2*—H4	-0.437	0.239	-**0.689**	-**0.364**
H1*—A1*	-**0.828**	0.048	-**0.459**	0.044
H1*—A2*	-**0.855**	-0.067	-0.343	0.114
H1*—A3*	-**0.809**	-0.218	-0.297	0.126
H4*—H2k*	-0.452	-0.598	0.294	**0.366**
H2k*—H5k*	0.152	0.023	0.101	0.057
residual H1*	-0.290	-0.443	-**0.722**	0.084
H2*	-0.157	-0.555	-**0.679**	0.114
H4*	0.295	-**0.778**	0.078	**0.479**
A1*	0.756	-0.549	0.092	0.124
A2*	**0.779**	-0.476	-0.103	0.086
A3*	0.735	-0.416	-0.211	0.093

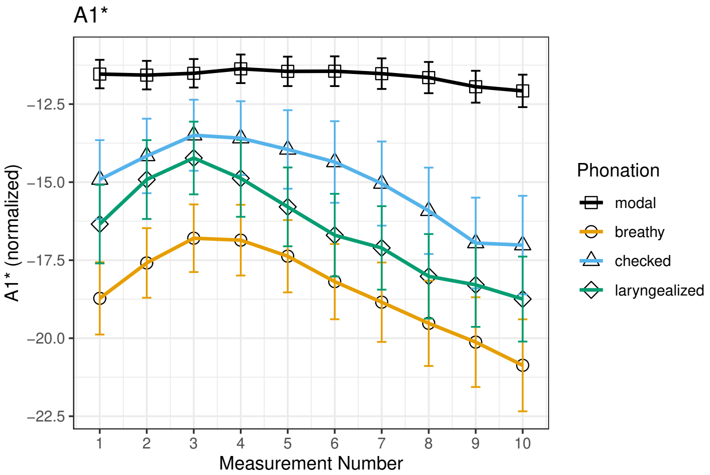
Correlations of Acoustic Measures to Dimensions II

Acoustic Measure	**NMDS1**	**NMDS2**	**NMDS3**	**NMDS4**
CPP	-0.590	-0.606	0.209	-0.179
HNR $<$ 500 Hz	-0.513	**-0.792**	0.152	-0.202
HNR $<$ 1500 Hz	-0.275	**-0.799**	0.323	-0.290
HNR $<$ 2500 Hz	-0.327	-0.714	0.391	-0.348
HNR $<$ 3500 Hz	-0.446	-0.644	0.393	-0.356
Strength of Excitation	-0.013	-0.741	-0.238	0.145
SHR	0.144	-0.176	0.122	**-0.597**
Energy	-0.080	**-0.793**	-0.015	0.341
Duration	-0.622	0.539	0.257	0.030

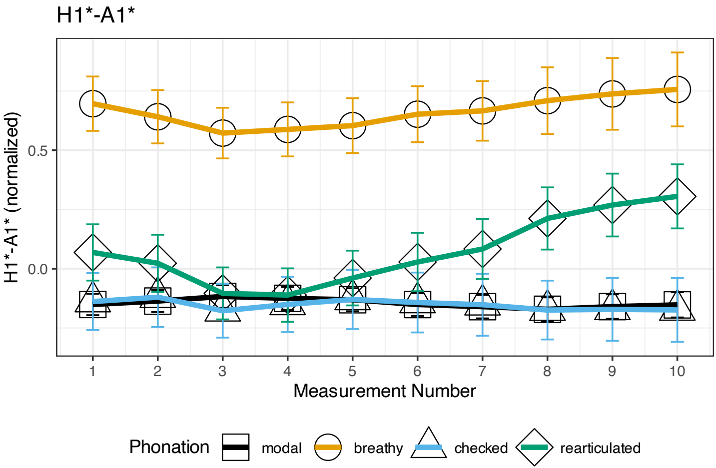
Duration


Violin plots for Duration
Phonation
modal
breathy
checked
rearticulated
6
4
2
0
-2
Duration (normalized)
modal
breathy
checked
rearticulated
Phonation

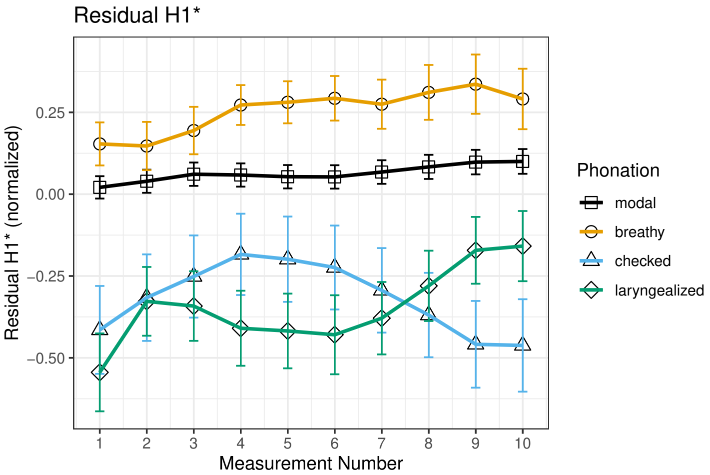
A1*


A1*
A1* (normalized)
−12.5
−15.0
−17.5
−20.0
−22.5
1
2
3
4
5
6
7
8
9
10
Measurement Number
Phonation
modal
breathy
checked
laryngealized

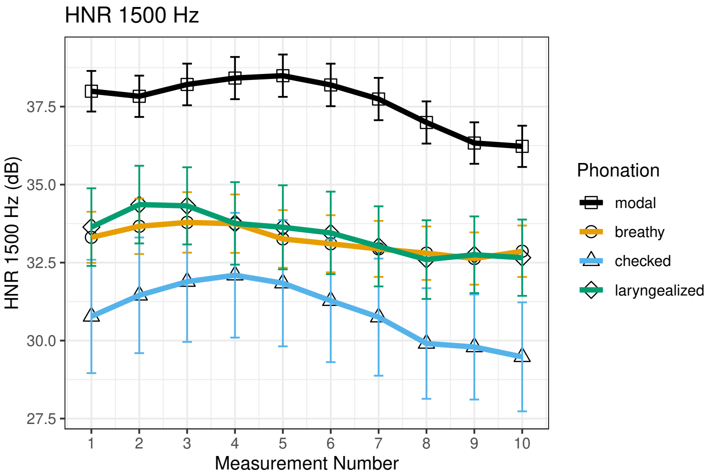
H1*—A1*


H1*-A1*
H1*-A1* (normalized)
0.5
0.0
1
2
3
4
5
6
7
8
9
10
Measurement Number
Phonation
modal
breathy
checked
rearticulated

Residual H1*


Residual H1*
Residual H1* (normalized)
0.25
0.00
−0.25
−0.50
1
2
3
4
5
6
7
8
9
10
Measurement Number
Phonation
modal
breathy
checked
laryngealized

HNR < 1500 Hz



HNR 1500 Hz
HNR 1500 Hz (dB)
37.5
35.0
32.5
30.0
27.5
1
2
3
4
5
6
7
8
9
10
Measurement Number
Phonation
modal
breathy
checked
laryngealized

Strength of Excitation (SoE)