The Acoustic Landscape of Voice Quality

Mykel Loren Brinkerhoff

University of California, Santa Cruz

BayPhon 2025

What is Voice Quality?

- Describes how the vocal folds vibrate.
- Used for both paralinguistic (Laver 1968, Podesva 2016) and phonological contrasts (Esposito & Khan 2020).
- Long been established that phonation contrasts have correlates in the acoustic signal (e.g., Fischer-Jørgensen 1968).

UC SANTA CRUZ

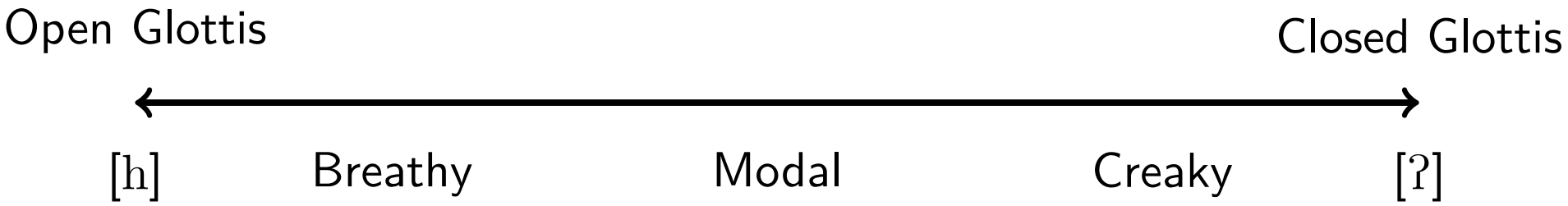
Modeling voice quality

- Long been established that phonation contrasts have correlates in the acoustic signal (e.g., Fischer-Jørgensen 1968, Klatt & Klatt 1990).
- Gordon & Ladefoged (2001) list several types of measures types that can be used:
  - Periodicity
  - Energy
  - Spectral tilt
  - Pitch
  - Duration
- Linguists have used these measure, or combinations of them, to model voice quality in a variety of languages (e.g., Blankenship 2002, Brunelle & Kirby 2016, Esposito 2012).

UC SANTA CRUZ

Voice quality's dimensionality

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



Voice quality's dimensionality

- 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.
  1. First dimension = nonmodal-to-modal continuum.
  2. Second dimension = glottal-airflow continuum.
- Found that languages with more contrasts used more of the acoustic space than languages with fewer contrasts.
- Found correlations between dimensions and acoustic measures.
  - First dimension = periodicity and energy.
  - Second dimension = spectral tilt and periodicity.

UC SANTA CRUZ

Research Question

Question:

- Can this same dimensionality reduction be found for a single language?
- If so, what do the dimensions correlate with?

UC SANTA CRUZ

Research Question

Answer:

- Yes; we find a three-dimensional space.
- Dimensions in SLZ are correlated with:
  1. First/third dimension = glottal-airflow continuum.
  2. Second dimension = nonmodal-to-modal continuum.

UC SANTA CRUZ

Data

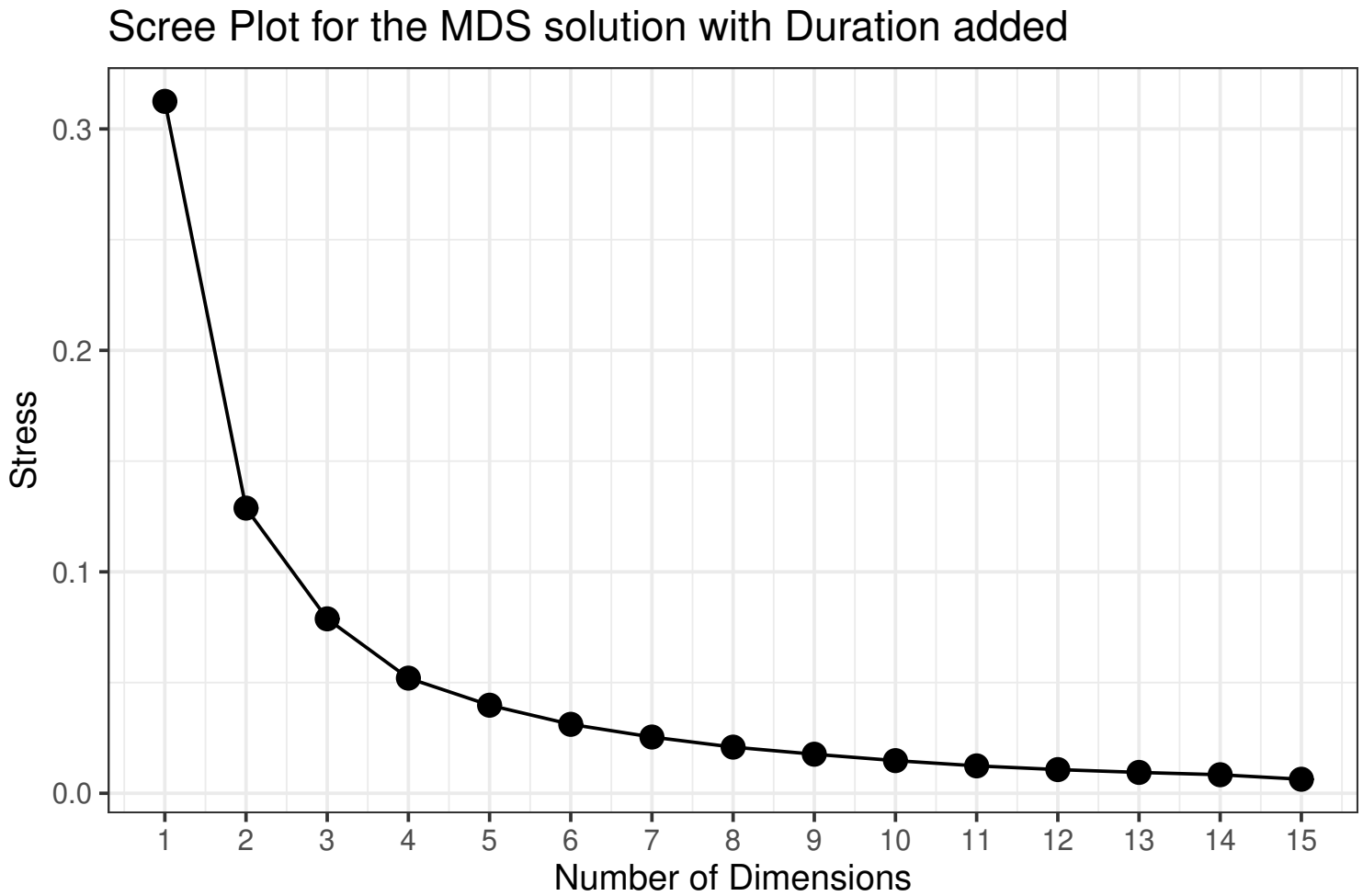
- Data comes from fieldwork on Santiago Laxopa Zapotec (SLZ).
- SLZ is a good candidate because of its four-way phonation contrast.
  - Modal ([a])
  - Breathy ([a̤])
  - Checked ([a͡ʔ] or [a͡a̰])
  - Rearticulated ([a͡ʔa] or [a͡a̰a])
- Data was collected from 10 speakers (5 male/5 female).
- Acoustic measures were calculated using VoiceSauce (Shue et al. 2011).

UC SANTA CRUZ

MDS analysis

- Multidimensional scaling (MDS; Kruskal & Wish 1978) was used to reduce the dimensionality of the data.
- Acoustic measures were used to define the acoustic space, following Keating et al. (2023).

Number of Dimensions



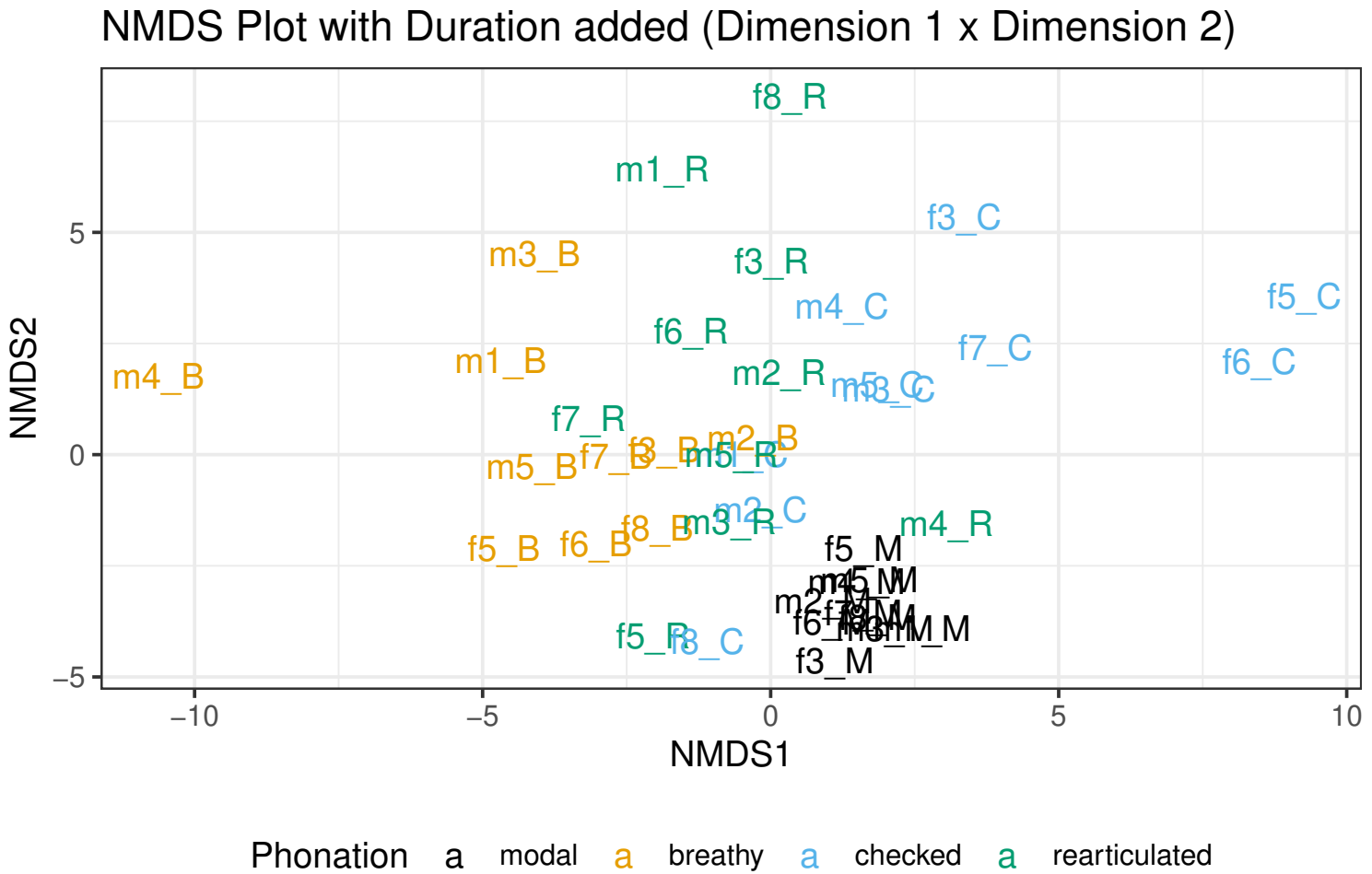
UC SANTA CRUZ

Dimensionality in SLZ

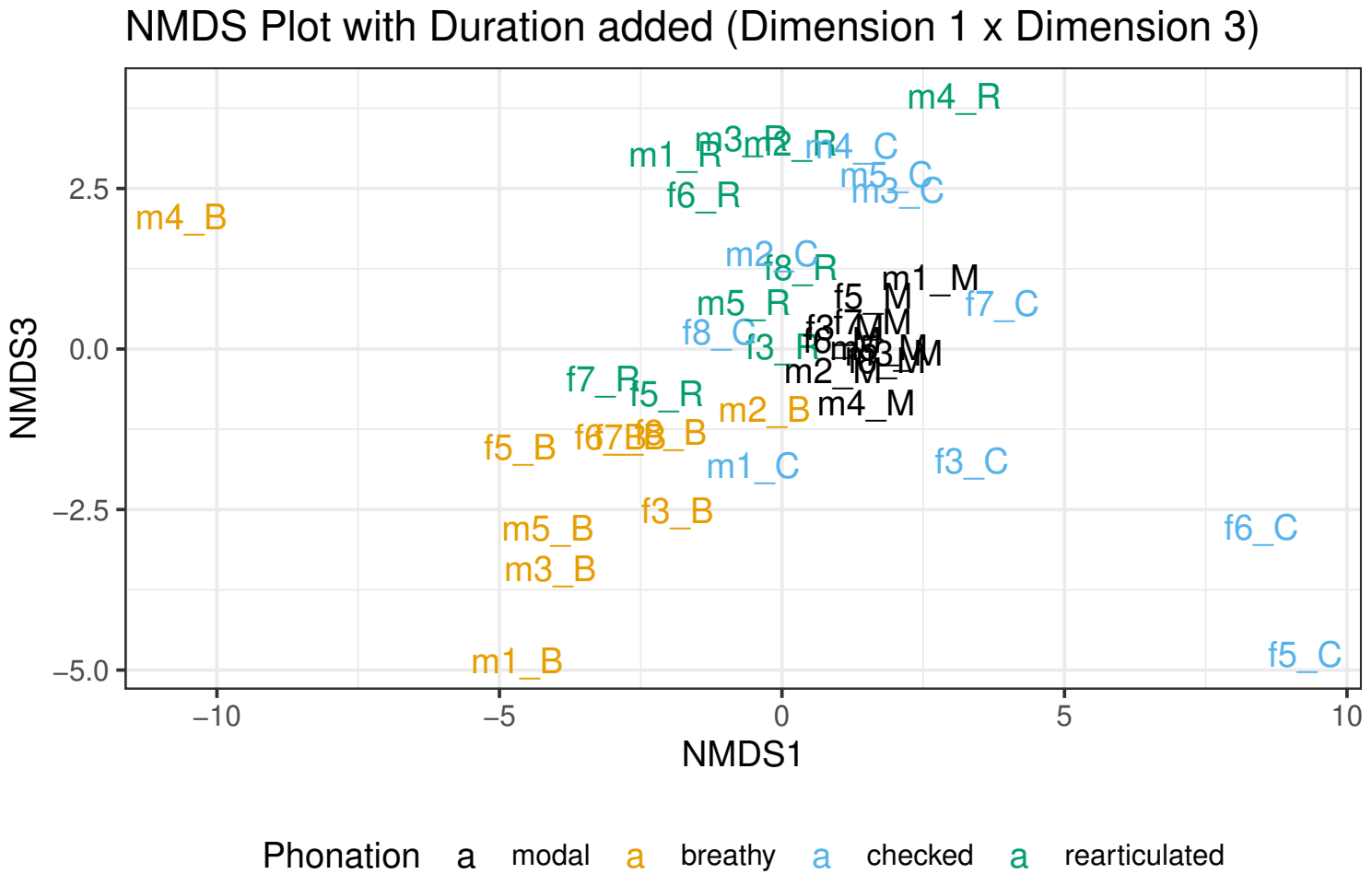
- Scan the QR code to see the three-dimensional space.

UC SANTA CRUZ

Dimensionality in SLZ

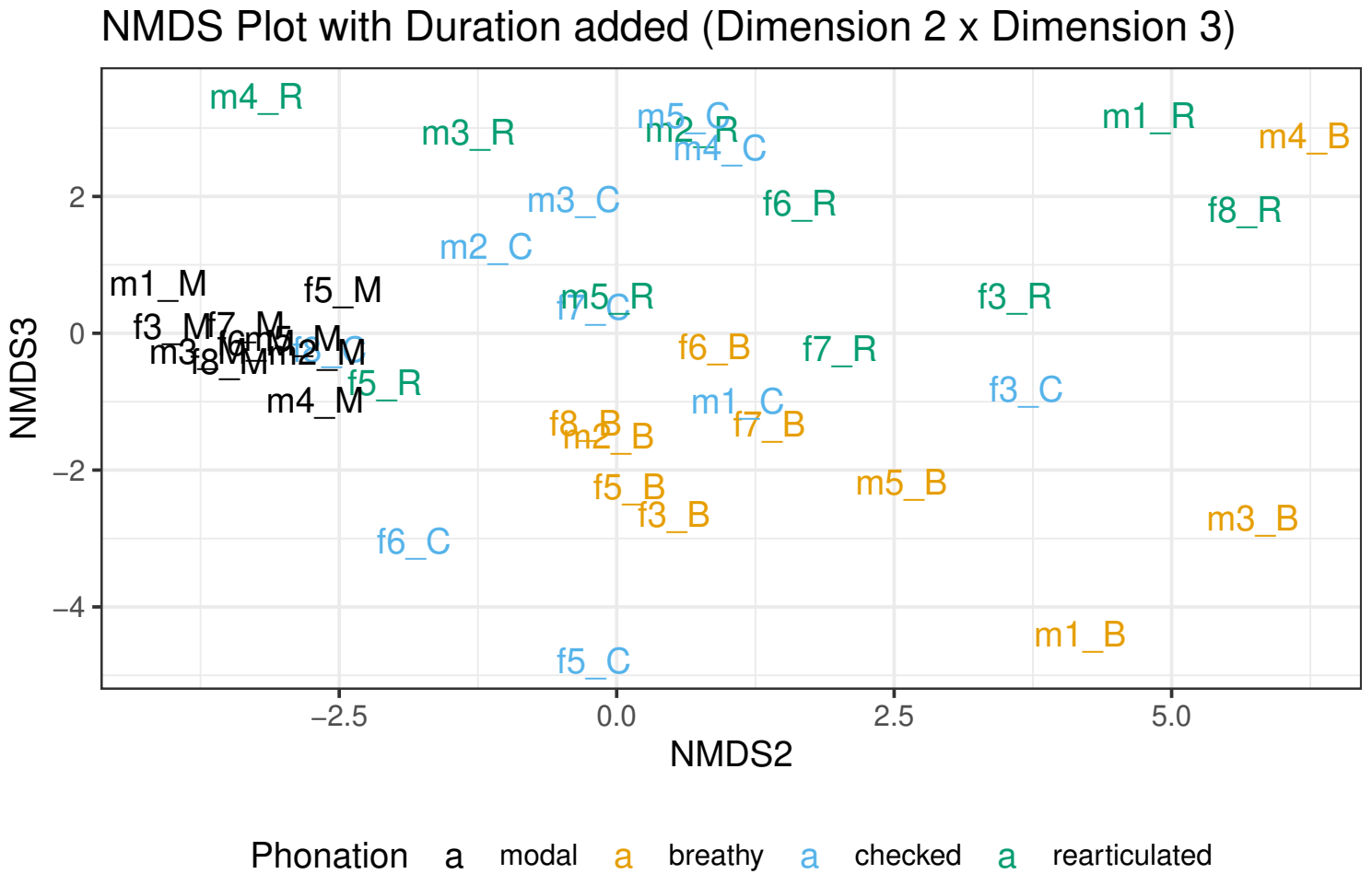


Dimensionality in SLZ



Phonation a modal a breathy a checked a rearticulated

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 spectral tilt:
  - residual H1* ($r^2 = -0.72$)
  - H2*−H4* ($r^2 = -0.69$)
  - H2* ($r^2 = -0.68$)

UC SANTA CRUZ

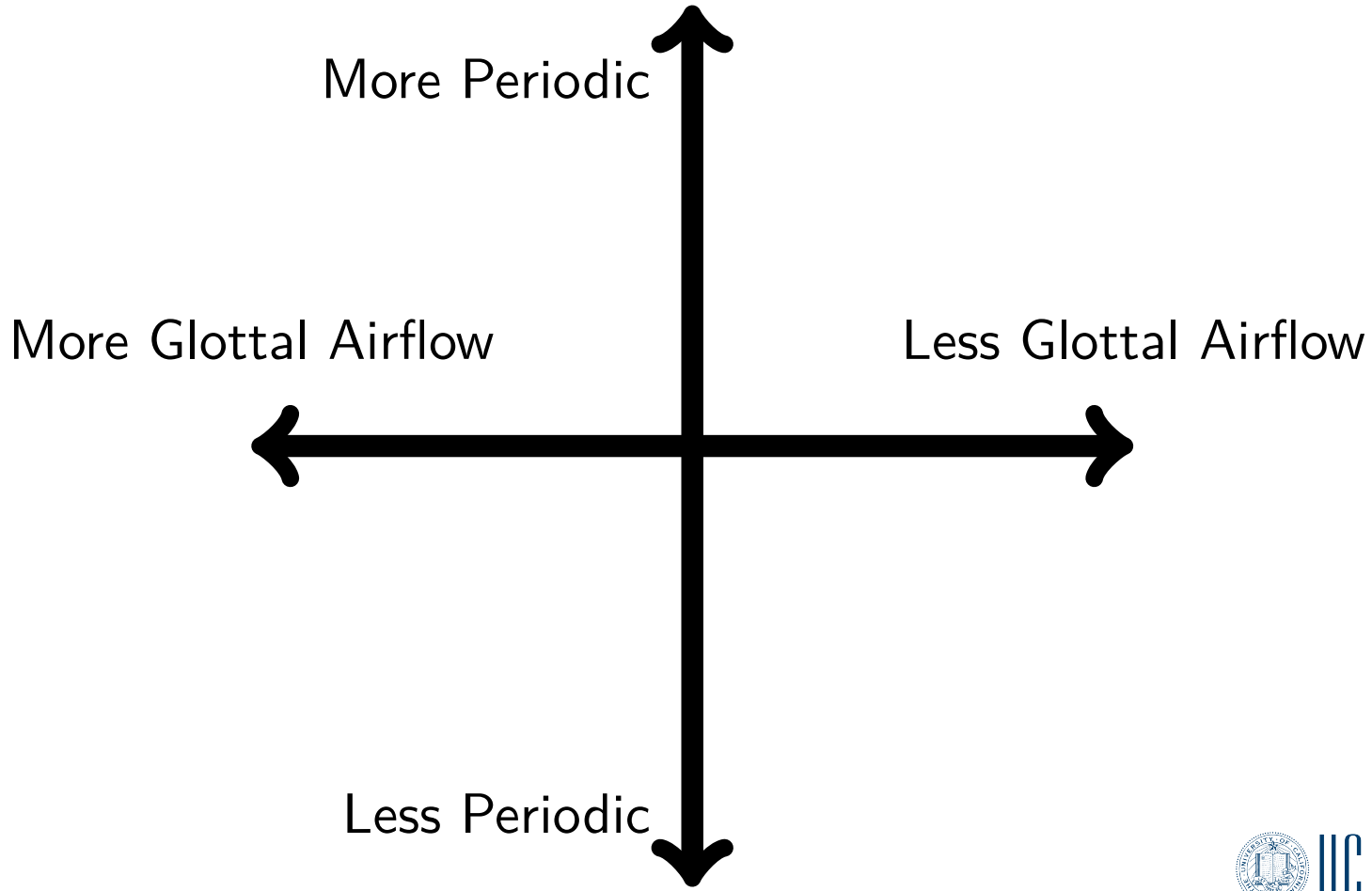
Summary of Results

- SLZ's phonation contrasts occupy a three-dimensional space.
- Dimensions are correlated with glottal-airflow (D1/D3) and nonmodal-to-modal (D2).
  - Collaborated with acoustic measure correlations.
- Dimensions are similar to those found in Keating et al. (D1 = nonmodal-to-modal; D2 = glottal-airflow).

UC SANTA CRUZ

Summary

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



Summary

- Dimensionality reduction also occurs in a single language.
- Dimensions correspond to glottal-airflow and periodicity within a language and cross-linguistically.
- If additional dimensions are added, they only add additional information about these two dimension.

- Outlook
  - What are the perceptual cues that speakers use to distinguish between phonation types?
  - How do these dimensions relate to the phonology?

UC SANTA CRUZ

Acknowledgements

- Thank you to the speakers in Santiago Laxopa 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.

UC SANTA CRUZ

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

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.

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.

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.

UC SANTA CRUZ

References II

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.

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.

References III

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.

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, 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.

UC SANTA CRUZ

References IV

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.

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.

UC SANTA CRUZ

References V

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.

Variable Importance