TITLE

Mykel Loren Brinkerhoff

Contents

1 Introduction

- Recent work from Keating et al. (2023) has shown that voice quality across languages map very similarly.
- This handout takes the work from Keating et al. (2023) and applies it to Santiago Laxopa Zapotec.
- The goal is to see how well SLZ compares to the results of Keating et al. (2023) and how residual H1 functions in these analyses.
- This analysis uses the same NMDS analysis as Keating et al. (2023) and the a more robust method of doing classification and regression tress known as bootstrap aggregating or bagging (Breiman 1996, Hastie, Tibshirani & Friedman 2009).

2 Preparing the data for analysis

- The data came from fieldwork conducted in Santiago Laxopa, Oaxaca, Mexico in 2022.
- The data was identical to the data used in Brinkerhoff & McGuire (2024).
- The data was preprocessed in the same way as Keating et al. (2023).
  - This consisted of extracting the middle portion of each vowel from the data.
  - This was done to ensure that the data was as close to the steady state portion of the vowel as possible.
  - After extracting the middle portion of the vowel, the associated acoustic measures were z-scored to bring everything into the same frame of reference.

  - Because the data was z-scored, some values were negative which produces a problem for the NMDS analysis.
  - To counteract this, the lowest value was added to all the data to make all the values positive while still maintaining the same relative distance between the values.

3 Multidimensional Scaling Analysis Results

•

4 Bagging

5 Discussion

References

Breiman, Leo. 1996. Bagging predictors. *Machine Learning* 24(2). 123–140. https://doi.org/10.1007/BF00058655.

Brinkerhoff, Mykel Loren & Grant McGuire. 2024. On residual H1* as a measure of voice quality. In Taehong Cho, Sahyang Kim, Jeff Holiday & Sang-Im Lee-Kim (eds.), *Proceedings of the 19th Conference on Laboratory Phonology*. Seoul, Korea: Hanyang Institute for Phonetics and Cognitive Sciences of Language5.

Hastie, Trevor, Robert Tibshirani & Jerome Friedman. 2009. *The Elements of Statistical Learning* (Springer Series in Statistics). New York, NY: Springer. https://doi.org/10.1007/978-0-387-84858-7.

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.