Probabilistic reduction in Spanish-English bilingual speech

Khia A. Johnson

University of British Columbia

khia.johnson@ubc.ca

International Congress of Phonetic Sciences 2019

Probabilistic reduction

- The more predictable a unit is, the more reduced (e.g. shorter duration) it is in production [Hall et al., 2018]
- Characterized by frequency, local predictability, and informativity [Cohen Priva, 2015]
- Reflects language-specific experience

Bilingualism

- A bilingual is someone with experience in two (or more) languages
- Probabilistic reduction has not yet been studied in bilingual speech
- Bilinguals' languages are simultaneously active, and **mutual influence** is unavoidable [Kroll et al., 2015]

Research question

Broad

How does probabilistic reduction operate in bilingual speech?

Narrow

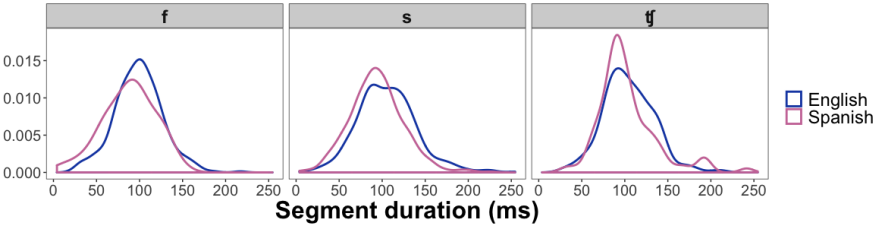
Are consonant reduction patterns in Spanish-English bilingual speech better accounted for by **separate-lexicon** or **pooled-lexicon** probabilistic measures?

Database & target segments

- Bangor Miami corpus of Spanish-English bilingual speech [Deuchar et al., 2014]
- Word-medial, intervocalic [f], [s], and [tʃ]
  - Occur in both languages
  - Identified from orthographic transcriptions, using pronunciation dictionaries
  - Exclusionary criteria & random sample ensuring equal numbers for each language ($n = 7896 \rightarrow n = 2052$)
  - Duration measured by forced alignment with hand-correction

Linear regression analysis

- Comparison of linear mixed effects models *with the same structure* using *lmer* in *R* [Bates et al., 2015]:
  1: **Separate-lexicon model**
  2: **Pooled-lexicon model**
- Dependent variable: Segment duration



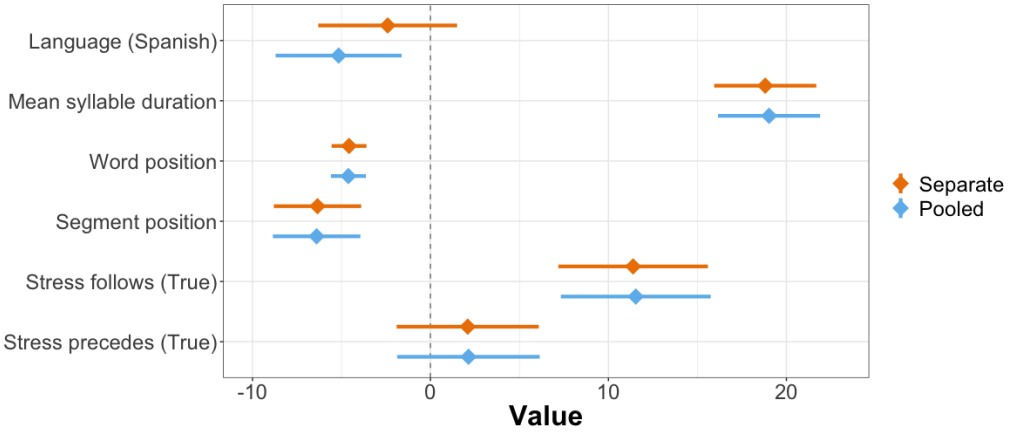
Fixed and random effects

- Control fixed effects (speech rate, stress, etc.)
- Probabilistic fixed effects:
  - Word frequency
  - **Segment frequency**
  - **Segment local predictability**
  - **Segment informativity**
- Random speaker intercepts, by-speaker random slopes for segment frequency, predictability, and informativity

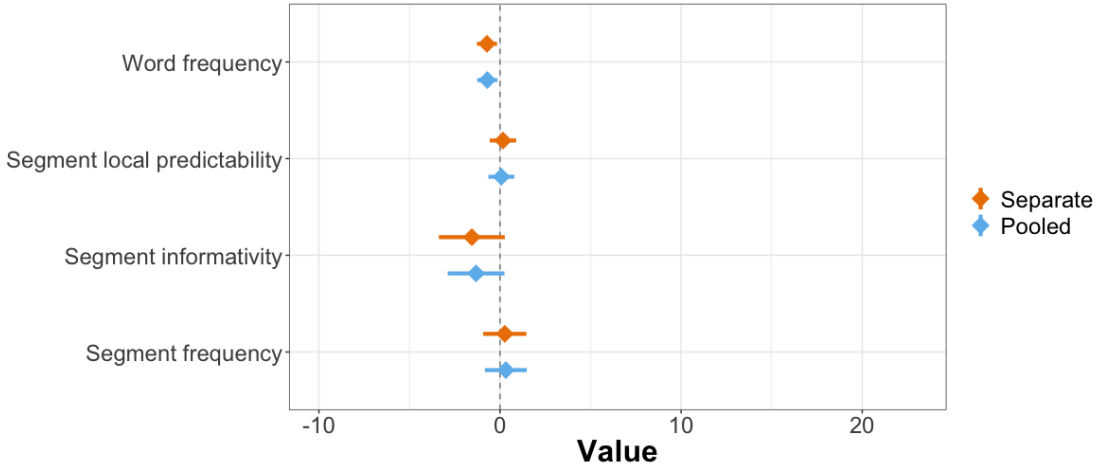
Model comparison results

- Compare AIC values (Akaike's Information Criterion)
  - Lower values $\rightarrow$ better fit
- **Difference in fit is not meaningful** ($\Delta = 2.6$) between *pooled-lexicon* ($AIC = 19509.5$) and *separate-lexicon* ($AIC = 19512.1$)

Control fixed effects pattern as expected



Probabilistic fixed effects have little to no effect



A puzzling null result

- Segmental probabilistic reduction is missing entirely in both models
- Lack of effect for segment frequency and informativity likely due to small number of target segments—*not enough coverage of the distribution*
- Local predictability result suggests fundamental difference in bilingual speech

Take home points

- Findings from monolingual speech do not transparently apply to bilingual speech
- Unclear if/how probabilistic reduction operates in bilingual speech
- I'm currently pursuing this in dissertation research

Thank you!

Contact

Email: **khia.johnson@ubc.ca**
Twitter: **@khia_johnson**

THE UNIVERSITY OF BRITISH COLUMBIA

Acknowledgements

Thank you to Dr. Margaret Deuchar for sharing the Bangor Miami corpus files, and to Dr. Molly Babel and Dr. Kathleen Currie Hall, for their advice during this project.

Selected references

- Bates, D., Mchler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Usinglme4. Journal of Statistical Software, 67(1).
- Cohen Priva, U. (2015). Informativity affects consonant duration and deletion rates. Laboratory Phonology, 6(2), 243278.
- Deuchar, M., Davies, P., Herring, J. R., Couto, M. C. P., & Carter, D. (2014). 5. Building Bilingual Corpora. Advances in the Study of Bilingualism, 93110.
- Hall, K. C., Hume, E., Jaeger, T. F., & Wedel, A. (2018). The role of predictability in shaping phonological patterns. Linguistics Vanguard, 4(s2).
- Kroll, J. F., Dussias, P. E., Bice, K., & Perrotti, L. (2015). Bilingualism, Mind, and Brain. Annual Review of Linguistics, 1(1), 377394.