Background

- Speech categories are extremely multidimensional
  - Many cues are present to make same phonetic distinction
  - Not all cues are equally diagnostic of category
- How can we concisely describe distinctions with many contributing cues?
  - Massive tables delineating effects of various cues are difficult to interpret
  - These also may mask cues that are correlated or interact
  - Can we determine groups of cues that work together as a class?
- Do different talkers use groups of cues in different ways?
  - Examining classes of cues may make it easier to investigate individual differences

Approach

1) Standardized cue measurements (get on a similar scale)
   - Example: duration (time) and spectral mean (Hz)
   - Z-score transform → common scale

2) Cues vary in importance: must assign weight to each.
   - Example: Spectral mean (dominant cue) vs. F2 (secondary cue) for s vs /ʃ distinction
     - Logistic regression – find each cue’s discriminative capacity

Combined score = \( w_1 \cdot \text{dur} + w_2 \cdot \text{fmax} + w_3 \cdot \text{dur} + w_4 \cdot F2 \ldots \)

Fric. Category = logistic (combined score)
Choose weights (\( w_1, \ldots, w_n \)) to maximize correct categorization.

3) Select combinations of cues to answer theoretical questions.
   - Cues in different segments
     - e.g., Fricatives: cues in frication vs. cues in vocoid
   - Spectral cues vs. non-spectral cues

4) How to combine them?
   - Linear combination of weight x Z-score value

\[
w_1 \cdot \text{dur} + w_2 \cdot \text{fmax} \qquad \text{Frication cues}
\]
\[
w_3 \cdot \text{dur} + w_4 \cdot F2 \ldots \quad \text{Vocoid cues}
\]

Test Data

- 2880 CV tokens of the eight English fricatives (Jongman, Wayland, & Wong, 2000)
- 20 talkers, 6 vowel contexts, 3 repetitions
- Measured 24 cues to fricative identity (McMurray & Jongman, 2011)
- Weights calculated based on cue’s ability to perform some [binary] discrimination
  - based on the corpus (e.g. place of articulation, voicing, sibilance)
- Weights gauge discriminability of dimension based on each cue

Fricative vs. Vocoid cues

- Question: what is the relative contribution of cues in frication vs. cues in the vocoid?
- Discriminate place of articulation within voicing/sibilance (/ʃ/, /f/, /z/, /θ/)
- The combined cue values correspond to continuous representation of place of articulation (e.g., positive = more /ʃ/, negative = more /s/)

Conclusions

- We can combine cues into classes to reflect production patterns
  - Correlation of classes how well these cues work together
- Cues within frication and within vocoid are useful classes
  - Listeners or scientists may be able to collapse within these classes to form a “megacue”
    - For sibilants, equal contributions of both classes of cues
    - For non-sibilants, vocoid cues more effective for place
- Spectral vs. non-spectral: only spectral useful
  - Spectral information discriminates (most) place distinctions
  - Non-spectral not informative
  - Place of articulation carried by spectral cues (as expected)

- Examining cues as classes shows clear individual differences in fricative production
  - Some talkers show greater variance within categories
  - Some show greater overlap between categories
  - Some show asymmetric variance between categories