Segmenting Words in Two Languages: Cue Weighting of Prosodic vs. Statistical Information in English and Cantonese

Helen Shiyang Lu*, Janet F. Werker, Alexis K. Black
University of British Columbia

Background

- Language learners can use both statistical cues (e.g., syllable transition probabilities) and prosodic cues (e.g., stress patterns) to segment speech [1-4]
- As learners gain experience with a language, they adjust their reliance on different segmentation strategies [5]
- Learners of languages with predominant stress patterns in words (e.g., English and German) tend to prefer stress-based prosodic cues when these conflict with statistical cues [6-8]
- Some languages (e.g., Cantonese) do not have a predominant stress pattern in multisyllabic words, thus making this type of prosodic cues less informative for word segmentation
- **Bilinguals** exposed to two typologically distinct languages must navigate **competing segmentation cues**

Current Study

- Compared **English monolinguals** and **Cantonese-English bilinguals** in word segmentation tasks conflicting statistical and prosodic cues – one in *English* and one in *Cantonese* context
- In addition to an explicit recognition task, we also used pupillometry measures
  - **Larger pupil dilation** at test shows greater surprisal in response to unexpected or unfamiliar words
  - **Pupil entrainment** in training reveals alignment with statistical vs. prosodic cues
- Entrainment in training has been shown to predict test performance [8]

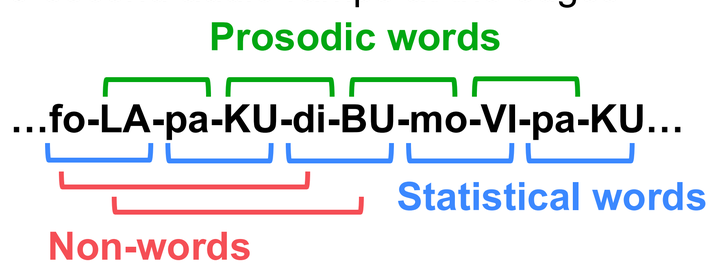
Methods

Stimuli

- For each context, four disyllabic words were created from
  - *English* syllables : *vi, pa, ku, mo, fo, la, di, bu*
  - *Cantonese* syllables: *caa2, ge6, je2, ngo3, wu5, zi4, zo1, zyu5*
- Stressed syllables were 6 dB louder than unstressed syllables

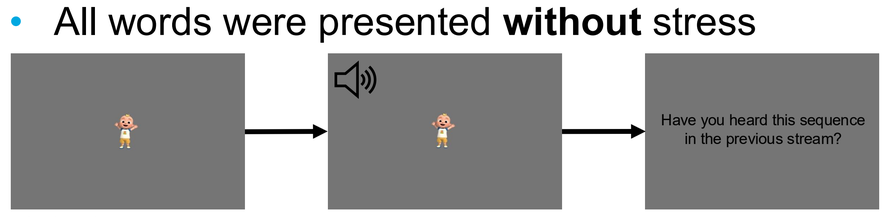
Familiarization (3 minutes)

- Participants watched an aquarium video while listening to a continuous speech stream, with 3-second audio ramps at the edges



Test Phase (3 * 12 trials)

- Half of the statistical and prosodic words were matched in frequency [9]
- All words were presented **without** stress



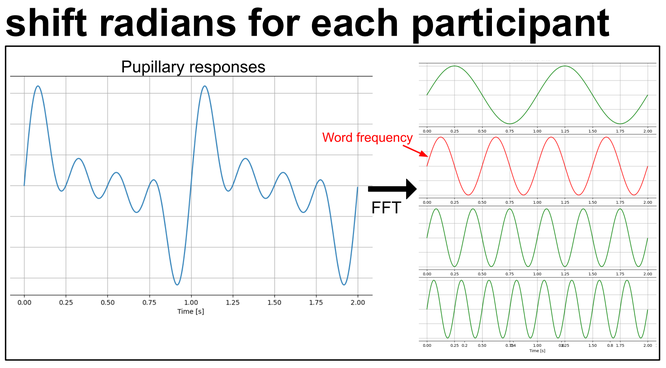
A visual target occurs; Participant accumulates 1-s baseline looking | *A test word plays for 650 ms, and the visual target remains for another 2.5 s* | *Participant answers with a button box*

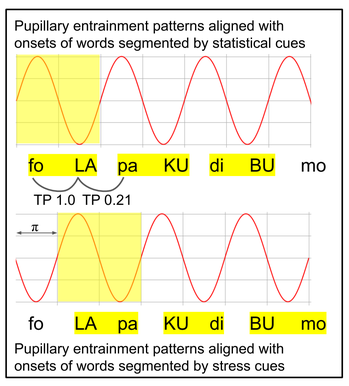
Pre-processing: Pupillary data from both phases were pre-processed with methods adapted from prior research [8,10]

Preliminary Results

Familiarization: Pupil Entrainment

Transforming pupillary data to phase shift radians for each participant

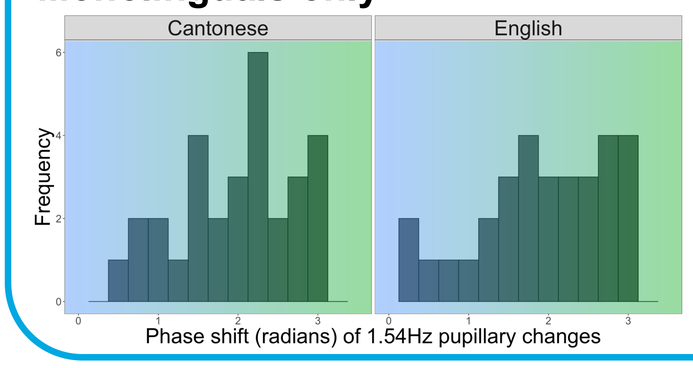




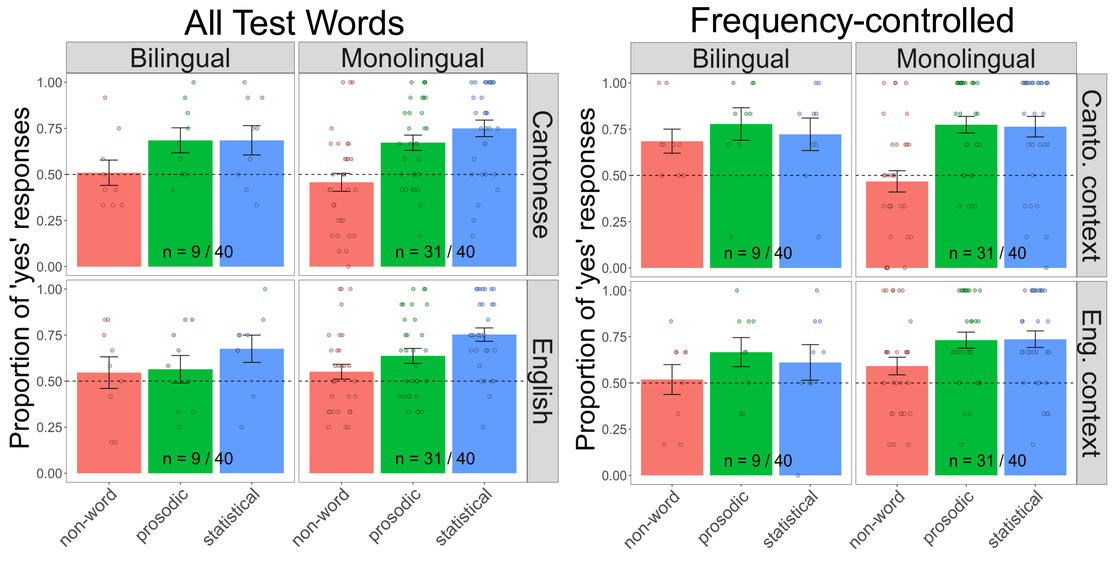
Familiarization always started with a statistical word:

- Phase shift = 0 → statistical
- Phase shift = π → prosodic

Monolinguals only



Test Phase: Recognition

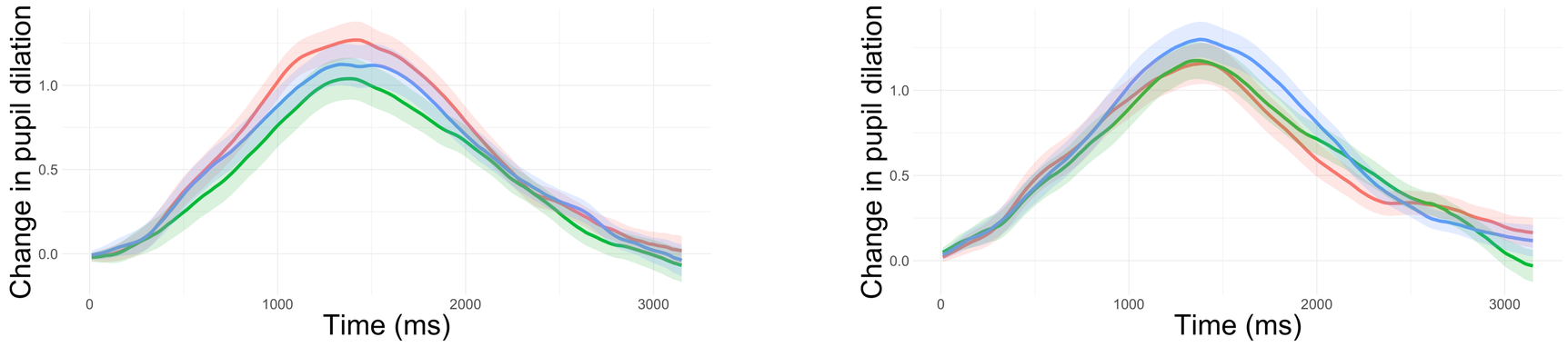


Monolinguals only

- Across all words, statistical > prosodic > non-words ($p < .001$ for both)
- When frequency was controlled, no differences were found between statistical and prosodic words ($p = .919$) but prosodic > non-words ($p < .001$)
- Differences between prosodic and non-words were **larger** in the *Cantonese* than in the *English* context (all words: $p = .007$; freq-ctrl words: $p = .01$)

Test Phase: Pupil Dilation

Monolinguals - Frequency-controlled words only – Non-word – Prosodic – Statistical



- Pupil dilation for non-words > prosodic between 770 to 1900 ms after word onset
- No differences were found in pupil dilation across word types

Summary: English monolinguals showed greater familiarity with prosodic words than non-words, especially in the *Cantonese* context, suggesting successful segmentation of the stream. However, data do not demonstrate a clear preference for either prosodic or statistical segmentation strategies. Ongoing analyses will explore whether cue reliance shifts over the course of familiarization.

References

1) Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language, 35*(4), 606–621.
2) Choi, D., Batterink, L. J., Black, A. K., Paller, K. A., & Werker, J. F. (2020). Preverbal infants discover statistical word patterns at similar rates as adults: Evidence from neural entrainment. *Psychological Science*, *31*(9), 1161-1173.
3) Matzinger, T., Ritt, N., & Fitch, W. T. (2021). The influence of different prosodic cues on word segmentation. *Frontiers in Psychology*, *12*, 622042.
4) Curtin, S., Mintz, T. H., & Christiansen, M. H. (2005). Stress changes the representational landscape: Evidence from word segmentation. *Cognition*, *96*(3), 233-262.
5) Thiessen, E. D., & Saffran, J. R. (2007). Learning to learn: Infants' acquisition of stress-based strategies for word segmentation*. Language Learning and Development, 3*(1), 73–100.
6) Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants*. Developmental Psychology, 39*(4), 706.
7) Marimon, M., Langus, A., & Höhle, B. (2024). Prosody out-weighs statistics in 6-month-old German-learning infants' speech segmentation*. Infancy, 29*(5), 750–770.
8) Marimon, M., Höhle, B., & Langus, A. (2022). Pupillary entrainment reveals individual differences in cue weighting in 9-month-old German-learning infants. *Cognition*, *224*, 105054.
9) Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 9*(4), 321–324.
10) Mathôt, S., & Vilotijevíc, A. (2023). Methods in cognitive pupillometry: Design, preprocessing, and statistical analysis. *Behavior Research Methods, 55*(6), 3055–3077.