Perception of word boundaries in spontaneous speech in English

ASA Victoria
Nov 5, 2018

Jiaer Tao, Francisco Torreira, Meghan Clayards
Contact: meghan.clayards@mcgill.ca

McGill CRBLM

Research Question

Does spontaneous speech contain acoustic information that listeners can use to segment words?

Background

In READ SPEECH

- Evidence for acoustic cues to word/syllable boundaries (e.g. Lehiste, 1960; Turk & Shattuck-Hufnagel, 2000)
- Evidence that listeners use acoustic cues (e.g. Mattys & Melhorn, 2007) to segment ambiguous phrases, especially when top-down information is weak (e.g. Cutler, 2012).

In SPONTANEOUS SPEECH

- context may be more readily available (especially in conversation).
- Acoustic cues may be less available (Lindblom, 1990).

Production Methods

Participants

- Native speakers of North American English
- 3 females & 3 males in their 20s

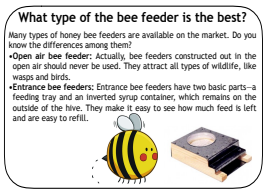
Materials

- 27 pairs of near-homophonous phrases
- each word the topic of an 'article'

Procedure

- **Spontaneous Speech** Participants read the article silently first, then explained its content to the confederate.
- **One-week gap** Half the articles were presented in an initial session, and their counterparts one week later.
- **Read Speech** At the end of the experiment participants read all 27 pairs.

Article example: beef eater vs. bee feeder



Perception Methods

Participants

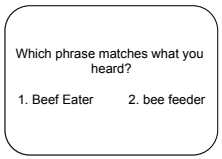
- Native speakers of North American English
- 30 new participants

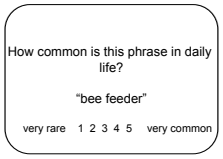
Materials

- ambiguous phrases extracted from recordings
- 316 tokens spontaneous speech
- 146 tokens read speech
- randomized together

Procedure

- **2 AFC**
  474 test trials in 3 blocks
  3 'catch' trials in each block (no ambiguity)*
- **Rating task**
  all 30 phrases (15 pairs)





* one participant was excluded for failing 5/9 catch trials

Models vs. Humans

Boundary Classification Accuracy

- using duration variables, consonant type and prosodic phrasing information
- Logistic regression model
- Random Forest model [Strobl et al., 2009] (ntree = 1000, mtry = 2) performance on unseen data

	Total	Read	Spont
Logistic Regression	74%	84%	80%
Random Forrest	63%	64%	66%
Participants	76%	81%	73%

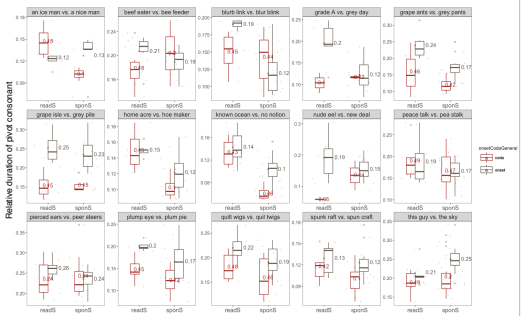
Conclusion

Spontaneous speech contains cues that listeners can use to segment.

Some consonants are more reliable than others.

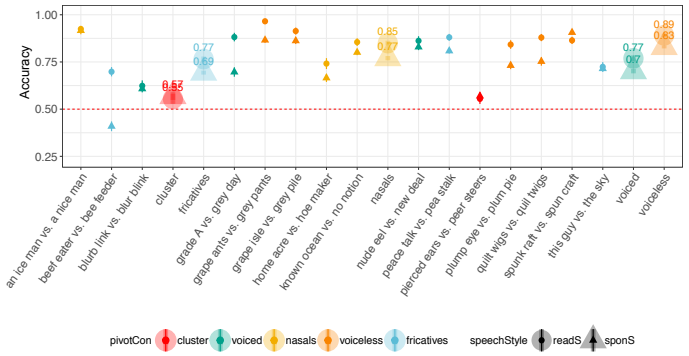
Production Results

Pivot consonant duration relative to phrase duration



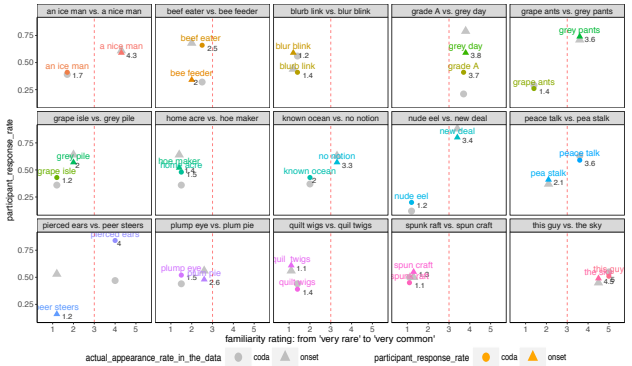
- Longer pivot consonants in onset position
- Differences greater in read vs. spontaneous speech
- biggest differences for voiceless stops and nasals (but consistent for fricatives and voiced stops as well)
- See Tao, Torreira & Clayards, 2018 for details

Perception Accuracy



- Overall accuracy above chance for both spontaneous (73%) and read speech (81%)
- Highest accuracy for voiceless stops and nasals
- lowest accuracy for /st/ clusters

Familiarity Ratings



- Familiarity drives responses for items with limited acoustic cues (e.g. *pierced ears vs. peer steers)*
- In other cases, duration cues are clear and responses are accurate (e.g. *grape ants vs. gray pants)*
- In some cases accuracy is high despite inconsistent durational differences (e.g. *a nice man vs. an ice man)*

References: A. Cutler, "Native listening: Language experience and the recognition of spoken words". Mit Press. 2012. I. Lehiste, "An Acoustic – Phonetic Study of Internal Open Juncture," Phonetica, vol. 5, no. 1, pp. 5–54, 1960. B. Lindblom, (1990). "Explaining phonetic variation: A sketch of the H&H theory". In Speech production and speech modelling (pp. 403-439). Springer, Dordrecht, 1990. S. L. Mattys & J.F. Melhorn, "Sentential, lexical, and acoustic effects on the perception of word boundaries". The Journal of the Acoustical Society of America, 122(1), 554-567, 2007. C. Strobl, J. Malley and G. Tutz, "An Introduction to Recursive Partitioning: Rationale, Application, and Characteristics of Classification and Regression Trees, Bagging, and Random Forests.," Psychological Methods, vol. 14, no. 4, pp. 323, 2009. J. Tao, F.Torreira & M. Clayards, "Durational cues to word boundaries in spontaneous speech". In Proc. 9th International Conference on Speech Prosody (pp. 240-244), 2018. A. E. Turk and S. Shattuck-Hufnagel, "Word-boundary-related duration patterns in English," Journal of Phonetics, vol. 28, no. 4, pp. 397–440, 2000.

This work was supported by SSHRC grant 435-2014-1504 to Meghan Clayards.
Social Sciences and Humanities Research Council of Canada
Conseil de recherches en sciences humaines du Canada
Canada