Durational cues in spontaneous speech

Jiaer Tao, Meghan Clayards, Francisco Torreira
Contact: jiaer.tao@mail.mcgill.ca

Research Questions
1. Are durational cues to word boundaries present in spontaneous speech as in lab speech?
2. Which of two possible mechanisms – word-initial lengthening or word-final lengthening – best explains our data?

Methods

Participants
• Native speaker of North American English
• 27 pairs of near-homophonous phrases e.g., beef eater vs. bee feeder; grade A vs. grey day [Lehiste, 1960]

Materials
• One article for each target phrase

Procedure
• Conversation 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.
• Annotation & analysis
  • 355 tokens
    • 133 coda tokens
    • 222 onset tokens
  • Exclusion unusual pronunciations & salient boundary cues
  • Prosodic-phrasing information
    • the presence/absence of an intonational boundary adjacent to the phrase (i.e., ip/Ip)

Results

Model Prediction

Predictors
• Pivot consonant duration
• Post-juncture section duration
• Pivot consonant type
• Prosodic-phrasing information

Response
• Boundary prediction (coda vs. onset)

Method
• Random Forest model [Strobl et al., 2009] (ntree = 1000, mtry = 2)

Results
• Out-of-bag classification accuracy of 64%
• Word-initial lengthening was an important cue assisting in word-boundary. Still, durational cues alone are not sufficient to segment word boundaries in many instances of spontaneous speech.

Follow-up Work

Perception study
• Can listeners correctly identify the boundary placement in spontaneous speech?
• Are word-initial lengthening effects identified in this study are perceptually relevant?
  • Phrases were played in isolation to 30 Native English listeners in a 2AF segmentation task
  • Accuracy was above chance (73.1%)
• Measures of word-initial lengthening were good predictors of performance
• Accuracy varied according to the pivot consonant type

References

Acoustic cues to word boundaries
• F0, amplitude contour, durational patterns, allomorphic variation, etc.

Boundary-related lengthening
• Word-initial lengthening
  Domain-initial strengthening (DIS) Strongest at the initial segment and becomes gradually weaker

Word-final lengthening
Preboundary lengthening e.g., Phrase-final lengthening Evidence at word-level is less clear [Cho & Keating, 2009; Turk & Shattuck-Hufnagel, 2000; Cho, 2016]

Speech style
• Laboratory read speech
  Hyper-articulation, careful speech
• Spontaneous speech
  reduction and lenition, casual speech

Voices
• Voiceless Voiced Nasal Fricative Cluster
  /p/ /t/ /k/ /b/ /d/ /g/ /v/ /m/ /n/ /s/ /z/ /s/ /z/ /s/ /s/ /s/
  67 26 21 42 33 30 17 41 56 22
  • Exclude unusual pronunciations & salient boundary cues

Pre-juncture Pivot C Post-juncture Relative duration proportions of phrase duration

1. Pivot consonant duration – Evidence for word-initial lengthening

Pivot consonant

Pivot consonants – longer in onset position (POSITION; β = -0.02, t = -4.23, p < .001)

2. Post-juncture section
– Evidence for word-initial lengthening

Post-juncture section

Post-juncture section duration – Longer when it involved an onset (POSITION; β = 0.12, t = 3.28, p < .01)

3. Pre-juncture section – No clear evidence for word-final lengthening

Pre-juncture duration e.g., [bi:] in bee vs. bee – Do not differ significantly