Listeners’ predictions of sentence lengths are categorical, not gradient

Amanda Rysling1, Charles Clifton, Jr.2, Nick Van Handel1, Ria Mai Geguera2, Hanna Choi2, & Anne Cutler3

1University of California Santa Cruz, 2University of Massachusetts Amherst, 3The MARCS Institute, Western Sydney University

Predicting sentence length
Do listeners predict the length of a sentence given its beginning, and if so, how?

Previous evidence of prediction
Grosjean (1983): native listeners of American English (AmE) gradiently predicted how many words followed a potential last word (PLW) of a sentence.

- +0: Earlier my sister took a dip.
- +3: Earlier my sister took a dip in the pool.
- +6: Earlier my sister took a dip in the pool at the club.
- +9: Earlier my sister took a dip in the pool at the club on the hill.

Sample item - Potential last word (PLW) underlined

• Native AmE speaker spoke sentences
  - +0: PLW was last word produced
  - +3, +6, +9: productions continued for three, six, or nine words past the PLW
• Each produced entirely in one breath group
• Listeners always and only heard words up to PLW

Means from Grosjean’s (1983) study

• Listeners reliably predicted more upcoming words for +6 than +3 than +0 conditions.
• No difference between +9 and +6 [4,5].

Acoustic analyses implicated F0 declination cue; only F0 range (start – PLW) changed with length.

Is sentence length prediction gradient or categorical?

+0 vs. +3 vs. +6 only partially replicable
• All later studies found difference in predicted upcoming words for +0 vs. other conditions.
• But failed to find +3 vs. +6 difference [5,6,7].
  - Native European French listeners
  - Native and L2 Australian English listeners
  - Native and L2 German listeners
• Problematic for both F0 declination explanation and gradient length prediction claim.

Gradient vs. categorical prediction
• Gradient – listeners form prediction of how many words follow a potential last word
• Categorical – listeners form prediction only of whether a sentence continues

Hypothesis: Listeners form only categorical predictions of upcoming sentence length.
• Prediction: AmE listeners with new AmE stimuli will not display +3 vs. +6 difference.

Predicted words after potential last word

Two registered replications:
• native AmE listener undergraduates
• UC Santa Cruz and UMass Amherst
  - both n=32: same as previous work
  - 33 items, just +0, +3, +6 conditions
• predicted no +3 vs. +6 difference

Materials and procedure:
• native male AmE speaker recorded stimuli
• listeners heard all three versions of each sentence
  - on each trial, listeners heard one version and saw the full +0, +3, +6 sentences in one item
• they selected which the production came from

Significant Differences

- +0 vs. +3 vs. +6 differences: reliably predicted more upcoming words in conditions in which productions continued past the PLW
- no +3 vs. +6 differences: no prediction of even more upcoming words when productions continued for longer after PLW
- F0 (Hz) declination differed significantly in +0 vs. +3, +0 vs. +6, and +3 vs. +6

Summary of Results
Listeners exhibited same pattern of responses as in studies that failed to fully replicate original findings

- +0 vs. (+3 and +6) differences: reliably predicted more upcoming words in conditions in which productions continued past the PLW
- no +3 vs. +6 differences: no prediction of even more upcoming words when productions continued for longer after PLW
- F0 (Hz) declination differed significantly in +0 vs. +3, +0 vs. +6, and +3 vs. +6

Conclusions

• Listeners predict whether sentences continue.
  - do not predict different continuation lengths
  - do not predict directly from declination
• Explicable if listeners only generate prosodic predictions from cues of high validity.
  - speakers modulate F0s quickly, mid-sentence
  - rate of pitch change relatively unreliable
• Failure to form gradient predictions when evidence is weaker parallels previous work:
  - native English listeners do not leverage regularities in stress for word segmentation
  - but Dutch and Spanish native listeners do make use of comparable stress regularities, with higher cue validity due to frequency of occurrence [1, 2, 3, 8]

References

Acknowledgements

Many thanks to John Duff, Lyn Frazier, Janeen Jeske, Shayne Sloggett, UCSC s/lab, and the ARC Centre of Excellence for the Dynamics of Language.

Contact: rysling@ucsc.edu