UNIVERSITY OF ALBERTA
DEPARTMENT OF LINGUISTICS

Alberta Phonetics Laboratory

Recognition of spoken pseudowords

Matthew C. Kelley & Benjamin V. Tucker
WECOL 2017
October 20, 2017

Introduction: Spoken Word Recognition

- Spoken word recognition studied in phonetic and psycholinguistic research
- Tells us things about the lexicon
    - E.g., more lexically frequent → faster to recognize (Dahan, Magnuson, & Tanenhause, 2001; Dupoux & Mehler, 1990; Ernestus & Cutler, 2015)
        - Usually explained as resting levels for activation or different connection strengths (Dahan et al., 2001)
- Studies generally get at mental processes ongoing during word recognition

Introduction: Pseudowords

- Most word recognition studies use pseudowords (usually phonotactically legal)
  - Ensures linguistic processing in experimental tasks
- Responses to pseudowords often **thrown out**, or else examined to understand real word processing
- Restricted research in this area points to lack of knowledge
  - E.g., what happens when heard in an experiment? (represents 50% of stimuli)

Introduction: Present Study

- Seeks to describe some of the processes involved in pseudoword recognition
  - Bears some relation to a number of linguistic phenomena
    - Hearing a word a listener hasn’t encountered before
    - Detecting what’s been heard is not a real word (and possibly recovering from that)
- Effects of several lexical predictors analyzed with linear mixed-effects modeling
- Trends from fit models examined and framed in greater speech processing context

Analysis: Data set

- Comes from Massive Auditory Lexical Decision data set (Tucker et al., 2017)
  - Responses to auditory lexical decision task
  - 232 monolingual western Canadian English speakers
  - 26,800 real words, 9,600 pseudowords recorded to be phonotactically legal
    - Recorded by 28 year-old western Canadian English speaker trained in phonetics
    - Mean of 11.88 responses per pseudoword
    - Pseudowords on average 132 ms longer than words
  - Pseudowords generated using Wuggy (Keeulers & Brysbaert, 2010), set to substitute a third of the sub-syllabic units in real words with other sub-syllabic units (e.g., onset cluster, phoneme, etc.)

Analysis: Lexical Predictors

- Phonotactic probability
  - How often certain phones, phone combinations, or transitions occur
  - Positive correlation to pseudoword "goodness" (Vitevitch, Luce, Charles-Luce, & Kemmerer, 1997; Bailey and Hahn, 2001)
  - High values facilitative to auditory lexical decision but overshadowed by effect of lexical status (Vitevitch & Luce, 1998)
- Calculated here as product of diphone co-occurrence probabilities, using Google Unigram corpus (Michel et al., 2011) and augmented copy of CMU Pronouncing Dictionary 0.6 (Weide, 2005)
- Hypothesis: positive correlation to difficulty in recognizing pseudoword
  - Higher values should suggest that an item is less remarkable, and more competitors to decide between

Analysis: Lexical Predictors

- Phonological neighborhood density
  - Measure of how many phonologically similar items there are to an item in question
  - Usually, for a given item, the count of entries in lexicon with an edit-distance of 1 from said item
  - Inhibitory effect for high values in auditory lexical decision with pseudowords (Luce, 1986; Luce & Pisoni, 1998)
  - Inhibitory effect for high values on accuracy in primed naming tasks (Goldinger, Luce, & Pisoni, 1989)
- Hypothesis: positive correlation to difficulty in recognizing pseudoword
  - Higher values should suggest more competitors to decide between

Analysis: Lexical Predictors

- Uniqueness point
  - Phoneme where sequence can be uniquely identified from among other items in the lexicon
  - Found to be more important than phonological neighborhood density in audio-primed visual lexical decision (Marslen-Wilson & Zwitserlood, 1989)
  - Effect size found comparable to lexical frequency (Balling & Baayen, 2012)
- Hypothesis: positive correlation to difficulty in recognizing pseudoword
  - Higher values should suggest more time needed to determine the item being heard

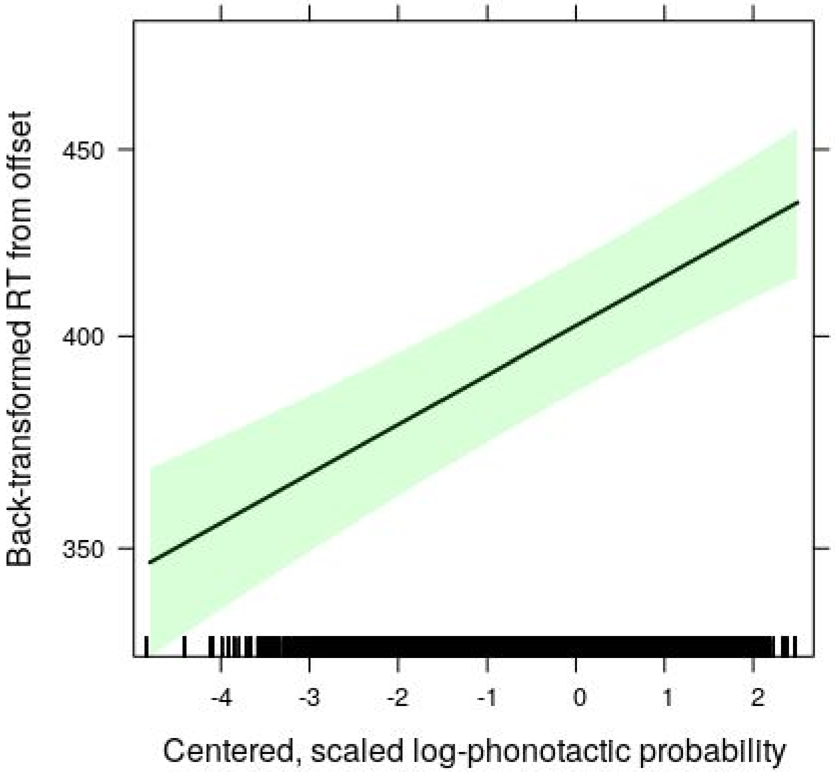
Analysis: Data Subsetting and Transforming

- Correctly identified pseudowords (n=96,049)
- Responses less than 500 ms from onset, before the word offset, or to items with phonotactic probability calculated to be 0 were dropped
  - 94,199 responses remained to analyze (98.07%)
- Reaction time (from offset), phonotactic probability, phonological neighborhood density+1, and uniqueness point were all logged for model fitting
  - Normal distribution of residuals
- All continuous variables were centered and scaled in the model fitting to bring the predictors to similar scales and help the models to converge

Results: Model

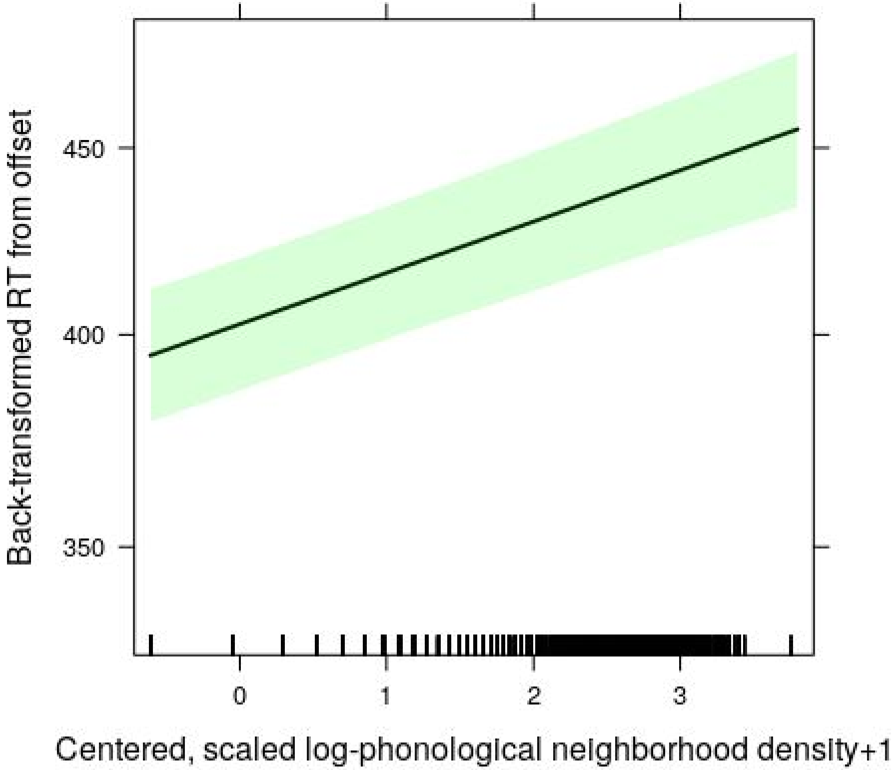
- Predictors of interest: log phonotactic probability, log phonological neighborhood density+1, log uniqueness point
- Controls: pseudoword duration, trial number
  - Dropped during fitting: age, sex, booth number, all two-way interactions between predictors of interest
- Random effects: random intercept for subject with a random slope for trial, random intercept for item with a random slope for trial

Results: Phonotactic Probability



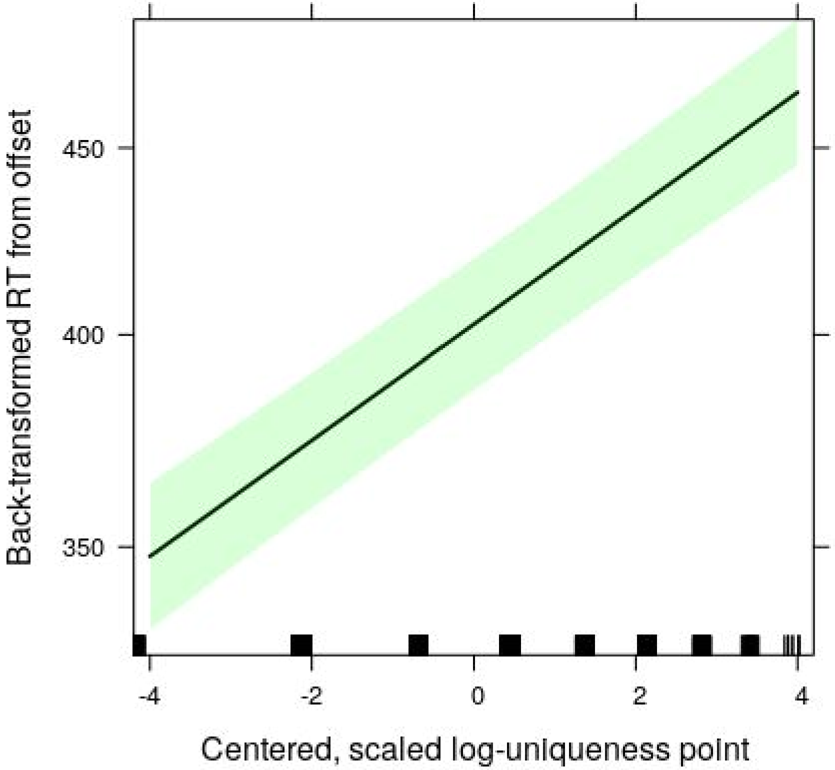
- Rare sequences should be easier to identify, and common sequences harder
    - Like distinctive vs. common writing styles
- Agrees with Vitevitch & Luce (1998)
    - Their data set is smaller and restricted to CVC items
    - Our results show effect's robustness across possible pseudowords

Results: Phonological Neighborhood Density



- More possible candidates to compare, so more difficult to decide
- Matches previous trends (Luce, 1986; Luce & Pisoni, 1998)
- Effect size is approximately the same sa phonotactic probability
    - Suggests its role may be smaller than has been described in previous studies (see above)

Results: Uniqueness Point



- Further in → need to wait longer for enough evidence to decide
- Probably segment that contains most information, as in Balling & Baayen's account of surprisal and uniqueness point (2012)
  - Effect size larger than other predictors of interest

Discussion: Info Used in Pseudoword Recognition

- Significance of each trend suggests multiple pieces of lexical information are used in pseudoword recognition
  - Likely that same mechanisms used in real word recognition are used in pseudoword recognition
  - No “magic bullet” predictor
- Task responses as the product of multiple characteristics of an item
  - Uniqueness point does has largest effect
  - Effects of phonotactic probability and phonological neighborhood density are similar
    - Suggests similar importance?

Discussion: The Lexicon

- There must exist some mechanism to decide if pseudoword/nonword is being heard
  - "If all else fails... Nonword!" accounts are less than satisfying...
  - Nonword identification itself could help determine when a perception error has occurred, as perhaps in Shortlist B (Norris & McQueen, 2008)
- Based on significance of all trends, unlikely to be organized around one particular characteristic (e.g., phonological neighborhoods)
  - If it were, we would expect one characteristic to explain a large amount of variation

Discussion: The Experimental Tasks

- Speech processing is going on during pseudoword trials (as we would hope)
  - Phonological priming and semantic priming could inadvertently occur
- If characteristics of the pseudowords skew too far from wordlikeness (e.g., consistently low phonotactic probability or phonological neighborhood density) confounds could arise
  - Lexical decision: are listeners really only deciding lexical status at that point?

Conclusions and Future Directions

- We should be paying attention to our pseudowords
  - Responses should not be neglected in data analysis
  - Processing is still ongoing when a pseudoword is heard in experiments
  - There is some order to be found in the responses to them
- Future directions:
  - Effect of morphological complexity?
  - Timing of uniqueness point (as opposed to position)?
  - Acoustic similarity vs. phonological similarity
  - Effects of wordlikeness?

References

Bailey, T. M., & Hahn, U. (2001). Determinants of Wordlikeness: Phonotactics or Lexical Neighborhoods? *Journal of Memory and Language*, 44(4), 568–591.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.

Balling, L. W., & Baayen, R. H. (2012). Probability and surprisal in auditory comprehension of morphologically complex words. *Cognition*, 125(1), 80–106.

Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology*, 42(4), 317–367.

Dupoux, E., & Mehler, J. (1990). Monitoring the lexicon with normal and compressed speech: Frequency effects and the prelexical code. *Journal of Memory and Language*, 29(3), 316–335.

Ernestus, M., & Cutler, A. (2015). BALDEY: A database of auditory lexical decisions. *The Quarterly Journal of Experimental Psychology*, 68(8), 1469–1488.

References

Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language*, 28(5), 501–518.

Luce, P. A. (1986). *Neighborhoods of words in the mental lexicon. research on speech perception. technical report no. 6*. Bloomington, IN: Department of Psychology, Indiana University.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 19(1), 1.

Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance*, 15(3), 576.

Michel, J.-B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Pickett, J. P., ... others. (2011). Quantitative analysis of culture using millions of digitized books. *Science*, 331(6014), 176–182.

Norris, D., & McQueen, J. M. (2008). Shortlist B: a Bayesian model of continuous speech recognition. *Psychological review*, 115(2), 357.

References

R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Tucker, B. V., Brenner, D., Danielson, D. K., Kelley, M. C., Nenadić, F., & Sims, M. (2017). Massive auditory lexical decision: Toward reliable, generalizable speech research. Manuscript submitted.

Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science* (0956-7976), 9(4), 325–329.

Vitevitch, M. S., Luce, P. A., Charles-Luce, J., & Kemmerer, D. (1997). Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. *Language and Speech*, 40(1), 47–62.

Weide, R. (2005). The Carnegie mellon pronouncing dictionary [cmudict. 0.6].

Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A., & Smith, G. M. (2009). *Mixed effects models and extensions in ecology with R*. New York: Springer.

Appendix: Model fitting process

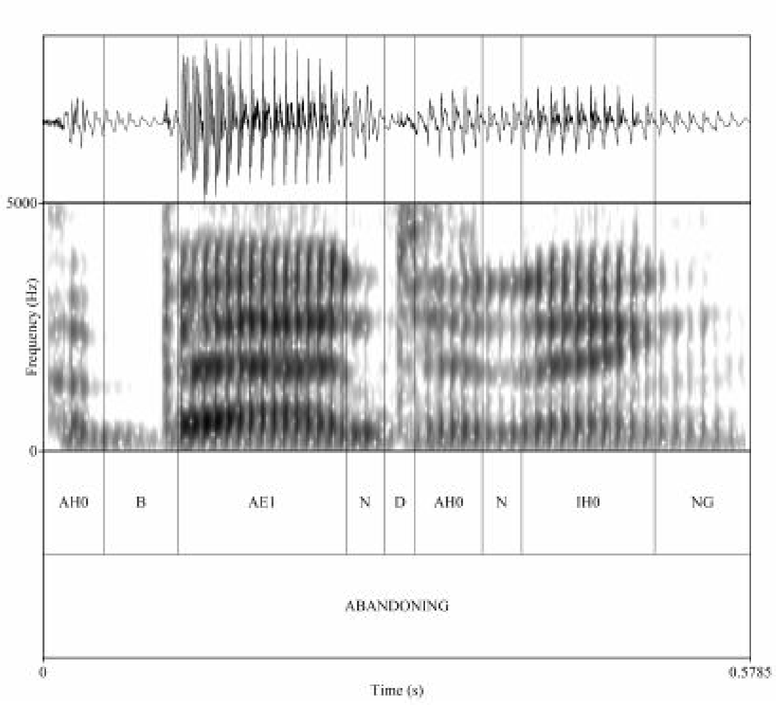
- Linear mixed-effects regression using lme4 package (Bates et al., 2015) in R (R Core Team, 2015)
- In fitting model, nested models were compared via maximum likelihood and restricted maximum likelihood, as in Zuur (2009)
    - Random structure forward-fit
        - Complexity added to random-effect structure if maximum likelihood indicated it was warranted
    - Fixed structure backward-fit
        - Complexity removed from fixed-effect structure if restricted maximum likelihood indicated it was not warranted

Appendix: Table of Coefficients

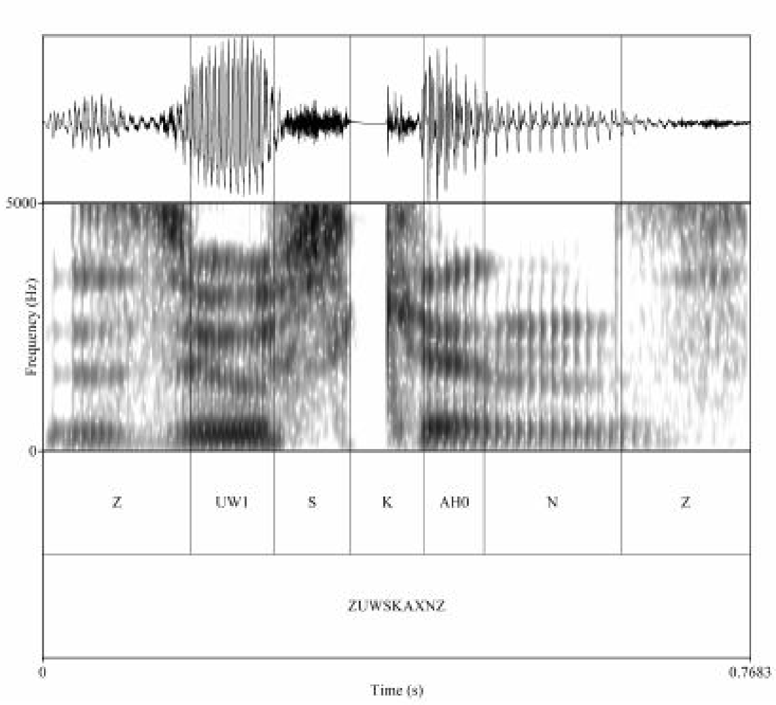
Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	5.997586	0.021122	283.95
scalelogpp	0.031039	0.004530	6.85
scalelognd	0.032334	0.003567	9.06
scalelogup	0.036521	0.002596	14.07
scaletrial	-0.110018	0.005464	-20.13
scaledur	-0.200223	0.004004	-50.01

Appendix: Sample Spectrograms



abandoning



zuwskaxnz