Phonetic effects of syntactic probability

Harry Tily
Stanford University

Speakers are sensitive to a diversity of information sources when making a choice between syntactic constructions (e.g. Jaeger, 2006; Bresnan et al., 2007; Wasow et al., 2007). When multiple cues coincide to favour one construction over another, that outcome is more probable than when cues conflict, and much more probable than when cues converge to favour an alternative construction. Previous results have shown that the production of high-probability words and syllables tends to be faster or otherwise reduced, as measured by frequency, local cooccurence probability, or topicality (e.g. Bybee, 2000; Jurafsky et al., 2001; Bell et al., 2003; Aylett and Turk, 2004; Pluymaekers et al., 2005). It has also been shown that articulation reflects verb bias, a single-cue estimate of syntactic probability (Gahl and Garnsey, 2004).

Using word durations measured from spoken, naturalistic data, I show that syntactic probabilities incorporating diverse information sources are reflected in articulation: certain words tend to be relatively shorter in more probable constructions. This result confirms that speakers' processing or representation of language includes knowledge about syntactic probabilities, and suggests that probabilistic knowledge is relevant at all stages of language production.

This work was completed jointly with Inbal Arnon, Joan Bresnan, Anubha Kothari and Neal Snider, and was presented as Tily et al. (2007).

Aylett, M. and Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence and duration in spontaneous speech. *Language and Speech*, 47:31–56.

Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., Gregory, M., and Gildea, D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in english conversation. *Journal of the Acoustical Society of America*, 113:1001–1024.

Bresnan, J., Cueni, A., Nikitina, T., and Baayen, R. H. (2007). Predicting the dative alternation. In Boume, G., Kraemer, I., and Zwarts, J., editors, *Cognitive Foundations of Interpretation*, pages 69–94. Royal Netherlands Academy of Science, Amsterdam.

Bybee, J. L. (2000). The phonology of the lexicon: Evidence from lexical diffusion. In Barlow, M. and Kemmer, S., editors, *Usage-based Models of Language*, pages 65–85. CSLI Press, Stanford.

Gahl, S. and Garnsey, S. (2004). Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunctiation variation. *Language*, 80:748–775.

Jaeger, T. F. (2006). *Probabilistic Syntactic Production: Expectedness and Syntactic Reduction in Spontaneous Speech*. PhD thesis, Stanford University.

Jurafsky, D., Bell, A., Gregory, M., and Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In Bybee, J. and Hopper, P., editors, *Frequency and the emergence of linguistic structure*, pages 229–254. John Benjamins, Amsterdam.

Pluymaekers, M., Ernestus, M., and Baayen, R. H. (2005). Articulatory planning is continuous and sensitive to informational redundancy. *Phonetica*, 62:146–159.

Tily, H., Arnon, I., Bresnan, J., Kothari, A., and Snider, N. (2007). What makes a construction predictable? Using semantic and contextual cues to better model phonetic reduction. Paper presented at the 20th CUNY conference on Human Sentence Processing, La Jolla.

Wasow, T., Jaeger, T. F., and Orr, D. (2007). Lexical variation in relativizer frequency. In *Workshop on Expecting the unexpected: Exceptions in Grammar, 27th Annual Meeting of the German Linguistic Association*, Germany. University of Cologne.