

Trading off precision in conveyed meaning for robust communication

We address an assumption implicit in the last four decades of psycholinguistic work on alternations. Alternations are environments in which speakers can select between several seemingly meaning-equivalent variants to encode their intended message. Psycholinguistic research has investigated a large number of environments under this assumption (e.g. ditransitive alternation; active/passive; left-dislocation; *that*-omission; e.g. Bock & Warren, 1985; Branigan et al, 2008; Ferreira & Dell, 2000), using the findings to make inferences about the production system. Linguistic work has, however, often assumed that two different forms never carry the same meaning and that speakers' choices in alternations are meaning-driven, rather than reflecting processing biases. We hypothesize that this view is too simplistic: even when two variants do not have exactly the same meaning, processing and communicative pressures can affect speakers' preferences.

Specifically, we test the hypothesis that speakers trade off precision in intended meaning with communicative pressures that ensure robust communication. The latter is assessed via the principle of Uniform Information Density (UID, Jaeger 2006, Levy & Jaeger 2007), which refers to speakers' preference to distribute information uniformly across the linguistic signal, thereby maximizing information transmission. Support for UID comes from research on alternations: speakers prefer the form with more linguistic signal when the encoded meaning is contextually unexpected (e.g. different word pronunciations, auxiliary-contraction, optional *that*-mentioning, Bell et al. 2003; Frank & Jaeger 2008; Jaeger 2006).

We investigate the choice of simple versus partitive "some" (1a,b respectively). These variants are generally assumed to differ in meaning (Ladusaw 1982), although they also arguably overlap in meaning. We assess this shared information, I(SOME), as the logarithm-transformed summed contextual probability of simple and partitive "some" (Shannon 1948). If speakers are willing to tolerate deviation from subtler aspects of the intended meaning in order to safely communicate the core meaning SOME, UID predicts that speakers prefer partitive over simple "some", the higher I(SOME).

We extracted 1362 cases of "some"-NPs from spontaneous speech data. We used linear mixed models to assess the effects of three independent estimates of the information of SOME: I(SOME) conditioned on (a) the previous word ("ate" above), (b) the NP head ("apples" above), and (c) the NP's grammatical function (direct object above). For all three measures, speakers preferred the longer partitive form, the higher the information of SOME. These effects are additive in that they hold independently of each other and while simultaneously controlling for other measures known to affect speakers' preferences, such as animacy, givenness, frequency, and bigram predictability of the head noun.

This result is surprising given that simple and partitive "some" are assumed to at least partially differ in the message they encode. The data suggest that even when two forms do not encode the same (but a similar enough) message, speakers may sacrifice precision in meaning for increased processing efficiency. We discuss these findings in light of recent game-theoretic approaches to pragmatics and expression choice (Franke, 2009).

- Bell, A., D. Jurafsky, E. Fosler-Lussier, C. Girand, M. Gregory, and D. Gildea (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustical Society of America*, 113(2), 1001 – 1024.
- Bock, J. K. and R. K. Warren (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, 21(1), 47 – 67.
- Branigan, H.P., Pickering, M.J, & Tanaka, M. (2008). Contributions of animacy to grammatical function assignment and word order during production. *Lingua*, 118, 172-189.
- Ferreira, V. S. and G. S. Dell (2000). The effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology*, 40, 296 – 340.
- Frank, A. and T. F. Jaeger (2008). Speaking rationally: Uniform information density as an optimal strategy for language production. In *The 30th annual meeting of the Cognitive Science Society (CogSci08)* (pp. 939 - 944), Washington, DC.
- Franke, M. (2009). Signal to act: Game Theory in Pragmatics. Ph.D. thesis, Universiteit van Amsterdam.
- Jaeger, T. F. (2006). Redundancy and syntactic reduction in spontaneous speech. Ph.D. thesis, Stanford University, Stanford, CA.
- Ladusaw, W. (1982). Semantic constraints on the English partitive constructions. In *Proceedings of WCCFL 1*, ed. By D. Flickinger, M. Macken and N. Wiegand, 231 – 242. Stanford: Stanford Linguistics Association.
- Levy, R. and T. F. Jaeger (2007). Speakers optimize information density through syntactic reduction. In B. Schölkopf, J. Platt, and T. Hoffman (Eds.). *Advances in neural information processing systems (NIPS)* (Vol. 19, pp. 849 - 856). Cambridge, MA: MIT Press.
- Shannon, C. (1948). A mathematical theory of communications. *Bell Systems Technical Journal*, 27(4), 623 – 656.

Examples

(1a) Alex ate some cashews.

(1b) Alex ate some of the cashews.