WORD LENGTH AFFECTS LANGUAGE PRODUCTION IN (NON)PREDICTIVE CONTEXTS BUT NOT LANGUAGE COMPREHENSION

Max S. Dunn & Zhenguang G. Cai
Department of Linguistics and Modern Languages, The Chinese University of Hong Kong
1155131488@link.cuhk.edu.hk

Language is an optimized communicative system in that longer words are used to convey more information than shorter words, as evidenced in corpus analysis (Piantadosi et al., 2011). Mahowald et al. (2013) provides further support for this, showing that people prefer a shorter expression over a longer one (e.g., *bike* vs. *bicycle*) to complete a sentence preamble when the expression is predictable from the preamble (hence carries less information; e.g., *For commuting to work, John got a 10-speed…*) than when it is not (hence carries more information; e.g., *Last week John finally bought himself a new…*). In this study, we investigated if the effect of predictability (or informativeness) on word choice (e.g., bike vs bicycle) generalizes from a force-choice task to language production and language comprehension.

Experiment 1 first aimed to replicate the finding in Mahowald et al. (2013). Participants chose between two expressions of nearly identical meanings but with different lengths (e.g., *bike* vs. *bicycle*) to complete a sentence preamble that was either predictive or nonpredictive of the expressions. Participants more often chose short expressions when the preamble was predictive than nonpredictive ($p=.001$; logit linear mixed effect model). Experiment 2 tested for this predictivity effect on word length in language production. Participants read a sentence with the last word consisting of two letters (e.g., *For commuting to work, John got a 10-speed bi__*) and supplied the last word. Critically, the partial word could be continued into the short or long expression (e.g., bike or bicycle). Participants more often supplied the short expression in the predictive than nonpredictive sentence ($p=.006$; logit linear mixed effect model), replicating the predictivity effect on word length in a production task. Experiment 3 tested for the effect in comprehension, using a self-paced reading task. We manipulated the sentence to be predictive or nonpredictive of the target expression and the target expression to be short or long (e.g., predictive: *Susan was very bad at algebra, so she hated* ***math/mathematics*** *more than science*; nonpredictive: *Susan introduced herself to me as someone who loved* ***math/mathematics*** *more than science*). We did not find a significant interaction between predictivity and length on reading time of the target expression ($p=.740$; linear mixed effect model), failing to replicate the predictivity effect on word length in comprehension.

Overall, the results indicate that people follow the principle of assigning shorter and longer words less and more information respectively in word choice and language production (Experiments 1 and 2). However, in language comprehension people do not use information of likely contexts shorter and longer words are found in to aid language processing. Thus, word length plays a role in language production in predictive and nonpredictive contexts but not in language comprehension. This dichotomy may come about from people being able to control the information rate of their productions according to word length in order to increase comprehensibility for their interlocutor, while not needing to do this and not expecting their interlocutor to do this during comprehension.