**Word frequency drives message choices**

Arella E. Gussow[1,2], Jennifer S. Trueblood[1] & Nazbanou Nozari[1]
[1]University of Richmond, [2]Indiana University Bloomington

Language production research has typically focused on turning a definite message into linguistic representations [1]. But is the choice of the message itself affected by linguistic properties? We investigated how word frequency and length affect participants' choices in a picture naming task, when faced with equally probable messages. Interactive theories of production [2] predict that linguistic properties should affect naming choices, as feedback within the language system allows for lower-level production stages (i.e., lexical retrieval) to affect higher-level message choices. Conversely, serial accounts [3] would predict no effect of linguistic properties, as lower-level processes should only proceed after the message was determined. While current research generally supports interactivity in production [4], to date there is no experimental evidence of feedback all the way up to the earliest stage of message choice. Moreover, our study included a control motor task, allowing to contrast domain-general versus language-specific decision-making processes: while some aspects of message choice are expected to be domain-general, feedback effects of frequency and length should occur within the language system only.

**Procedure.** 48 native English speakers were recruited on Prolific. On every trial, two images appeared side-by-side, and participants had to produce one of them. Two identical cartoon characters were used as cues; see Fig. 1. In a Forced-choice condition, both characters immediately appeared above the target picture that participants were required to produce. In a Free-choice condition, one character appeared above each picture, indicating that participants could choose either target. In language trials, participants named the target aloud and pressed the spacebar to end the trial, or timed out after 4 seconds. In motor trials, participants used a standard keyboard and pressed the 'g' key for a left-side choice, or 'h' for a right-side choice. Motor trials ended when participants pressed a key, or timed out after 2.5 seconds.

**Materials.** 120 color photographs were divided into four categories according to word length (Short, Long) and word frequency (Low, High). Short words were monosyllabic with up to four phonemes, long words were di- or tri-syllabic with five to nine phonemes. Care was taken to orthogonally vary length and frequency [5] across categories; there was no correlation between number of phonemes and word frequency ($r$ = -0.12, $p$ = .18). Next, we created 300 items of two pictures each. Pictures were paired according to all category combinations (HighShort-HighLong, HighShort-HighShort, etc.), with 30 unique pairs per combination. We then created eight lists to counterbalance within-subject condition, target side, and which picture within the pair was the target. Each participant completed two trial lists per task, for a total of 600 trials per task. Each list was presented in a trial block, with two language and two motor blocks. Order of trials within a block was randomized per participant. Language and motor blocks alternated, with initial domain block randomly assigned per participant.

**Results.** Overall, participants showed left-side facilitation that correlated across motor and language tasks, suggesting a domain-general left-side bias (Fig. 2). Similarly, there was a significant correlation between participants' RT savings for forced-choice compared to free-choice trials in language and motor tasks, indicating a domain-general decision-making component (Fig. 3). Mixed-effects regression models showed that for language trials, RTs in the Forced-choice condition were shorter when the target was of higher frequency, but there was no effect of length (Fig. 4). Critically, speakers were more likely to choose higher-frequency words in the Free-choice condition (Fig. 5), suggesting lexical frequency drove message choices. As expected, frequency and length did not have reliable effects on responses in the motor task.

**Discussion.** While some aspects of message choice appear to be domain-general, there are language-specific effects resulting from feedback within the production system: the same lexical indices that affect lexical retrieval also affect message selection. Results thus support interactive theories of language production, while adding a critical finding – even high-level message choices are influenced by feedback within the language system.

**References**
[1] Levelt, W. J., & Speaking, M. (1989). From intention to articulation. *Cambridge, MA: The MITPress*.
[2] Dell G.S. (1986). A Spreading activation theory of retrieval and sentence production. Psychological Review. 93:283-321
[3] Bock, K., and Levelt, W.J.M. (1994). Language production. Grammatical encoding. IN M.A. Gernsbacher (Ed.). Handbook of psycholinguistics (pp.741-779). New York: Academic Press
[4] Pinet, S., & Nozari, N. (2018). "Twisting fingers": The case for interactivity in typed language production. *Psychonomic Bulletin & Review*, *25*, 1449-1457.
[5] Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior research methods*, *41*(4), 977-990.

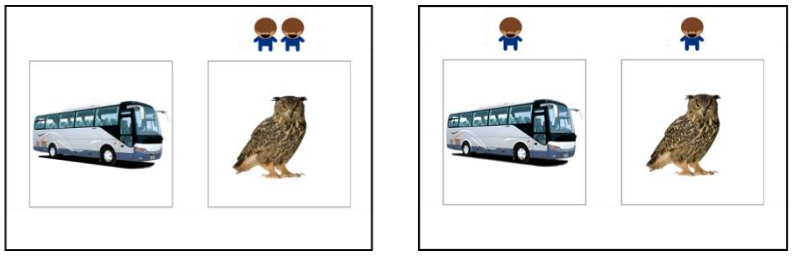

*Figure 1.* Example of trial displays in the Forced-choice (left panel) and Free-choice (right panel) conditions.

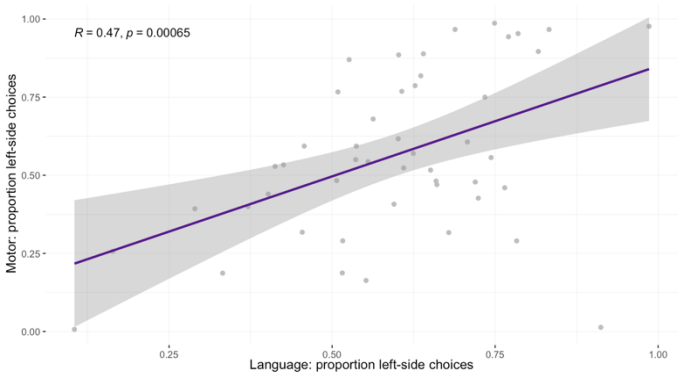


*Figure 2.* Correlation between the proportion of left-side choices in language and motor Forced- choice trials.

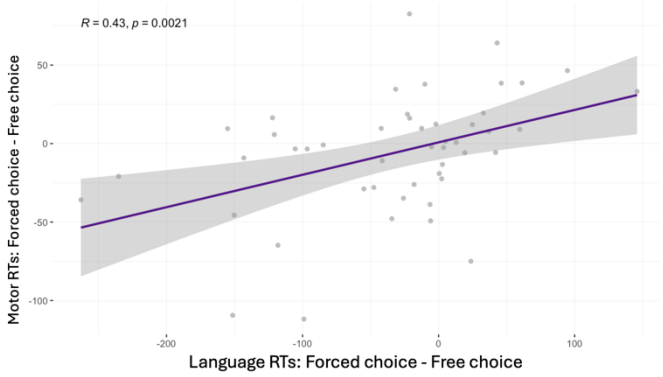


*Figure 3.* Correlation between RT differences (in ms) for responses in Forced-choice compared to Free-choice trials in the language and motor tasks.

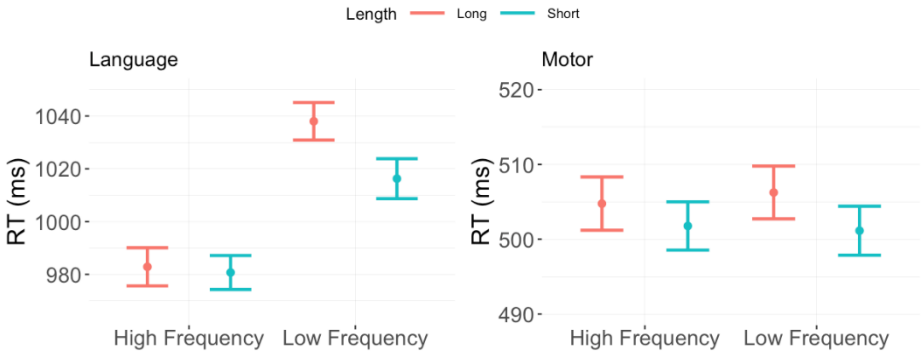


*Figure 4.* Mean response times in Forced-choice trials as a function of word frequency and length, in the Language (left panel) and Motor (right *panel) domains.*

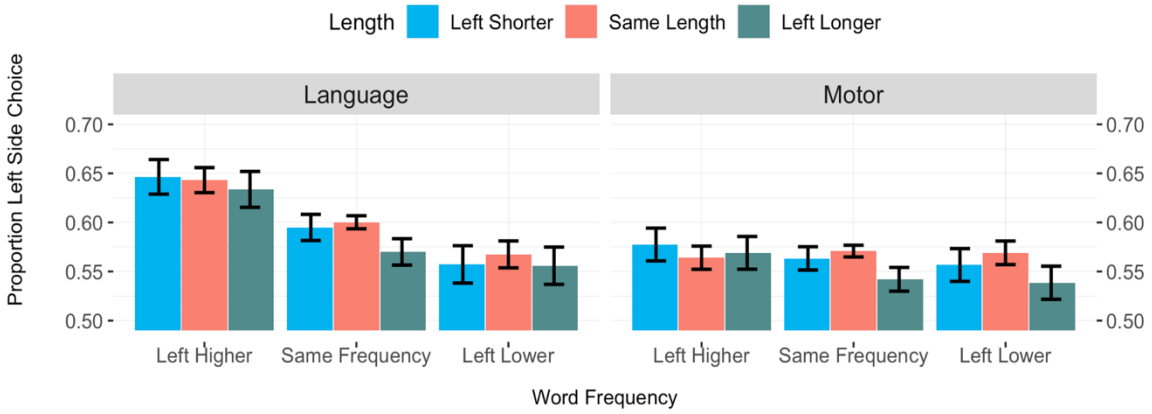


*Figure 5.* Proportion of left-side picture choices in the Free-choice trials of the language (left panel) and motor (right panel) tasks, as a function of relative frequency and length between the two pictures.