What Drives Interference? A Mega-Analytic Approach to Predicting Reaction Times from Word Properties

Louis Schiekiera[1,2], Vincent Gruber[1], Rasha Abdel Rahman[1], Kirsten Stark[1], Antje Lorenz[3], Fritz Günther[1]

[1] *Humboldt Universität Berlin, Germany*

[2] *Freie Universität Berlin, Germany*

[3] *Universität Bielefeld, Germany*

Background: Picture–word interference (PWI) tasks probe spoken word production by asking participants to name pictures while ignoring simultaneously presented distractor words [1, 2]. Despite decades of work, reported effect sizes and even effect directions vary across studies, likely because naming latencies are shaped by multiple interacting design factors as well as lexical and semantic properties of both targets and distractors [3]. To move beyond single-predictor explanations, we assembled a large trial-level dataset and adopted a prediction-oriented, mega-analytic approach to quantify what is reliably predictable in PWI naming latencies from a broad set of word properties.

Method: We compiled and harmonized raw trial-level data from 86 experiments (42 studies), comprising 3,353 participants and 657,605 trials in English and German (see Figure 1). Reaction times were log-transformed after standard cleaning. We extracted 39 predictors spanning (i) design characteristics (e.g., trial number, language, stimulus onset asynchrony, modality), (ii) target–distractor similarity measures (distributional, orthographic, and phonological distances; graded semantic similarity), and (iii) lexical features of targets and distractors (see Figure 2 for the full predictor set and Figure 3 for predictor collinearity). We first ran a simulation study to compare variable-selection methods (Figure 4). The simulation indicated that stepwise backward elimination combined with stability selection performs best. We then applied this procedure on the training partition to obtain a robust and interpretable fixed-effect set [4, 5]. Final models were fit as linear mixed-effects models [6]. Generalization was assessed on three disjoint held-out test sets targeting increasingly demanding transfer: unseen trials from known participants/experiments (Test A), unseen participants from known experiments (Test B), and unseen experiments (Test C).

Results: Variable selection retained a compact subset of predictors drawn from all major predictor families, indicating that naming latencies in PWI reflect a combination of design factors (e.g., practice/fatigue and timing), lexical properties of distractors and targets, and target–distractor similarity (see Figure 5 for all retained predictors and their stability scores). Surprisingly, predictive evaluation on the held-out sets showed a systematic increase in performance as the generalization target shifted from new participants to new experimental contexts (see Figure 6 for prediction results on the held-out sets).

Discussion: This mega-analytic, prediction-first framework provides (i) an integrated resource for studying interference across PWI designs and (ii) a transparent benchmark for how much variance in naming latencies can be accounted for by a broad set of word and design properties [2]. Beyond adjudicating isolated effects, the approach supports theory building by identifying which predictors are robust and which effects fail to generalize when experimental context changes [4, 5].

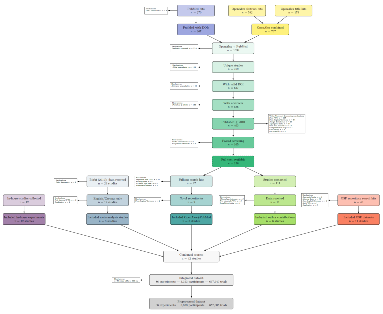

Figure 1: Data collection.

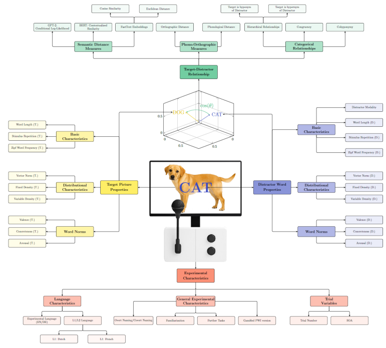

Figure 2: Predictor overview.

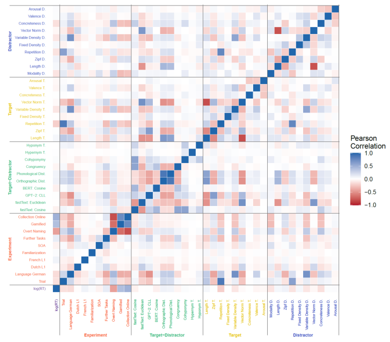

Figure 3: Correlation matrix of predictors.

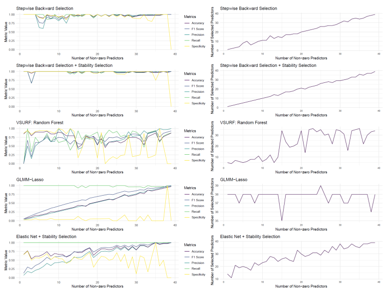

Figure 4: Preregistered simulation results.

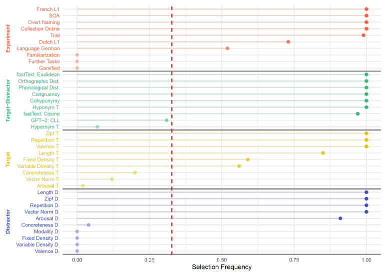

Figure 5: Variable selection.

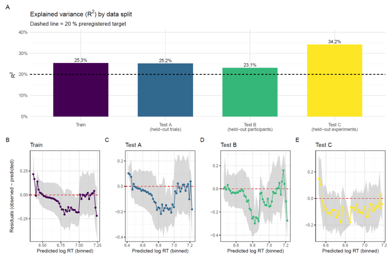

Figure 6: Held-out prediction.

References

[1] Rahman, R. A., & Aristei, S. (2010). Now you see it . . . and now again: Semantic interference reflects lexical competition in speech production with and without articulation. *Psychonomic Bulletin & Review*, *17*(5), 657–661. https://doi.org/10.3758/PBR.17.5.657

[2] Bürki, A., Elbuy, S., Madec, S., & Vasishth, S. (2020). What did we learn from forty years of research on semantic interference? A Bayesian meta-analysis. *Journal of Memory and Language*, *114*, 104125. https://doi.org/10.1016/j.jml.2020.104125

[3] Bürki, A., & Madec, S. (2022). Picture-word interference in language production studies: Exploring the roles of attention and processing times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48*(7), 1019.

[4] Meinshausen, N., & Bühlmann, P. (2010). Stability Selection. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *72*(4), 417–473. https://doi.org/10.1111/j.1467-9868.2010.00740.x

[5] Hyde, R., O'grady, L., & Green, M. (2022). Stability selection for mixed effect models with large numbers of predictor variables: A simulation study. *Preventive Veterinary Medicine*, *206*, 105714. https://doi.org/10.1016/j.prevetmed.2022.105714

[6] Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01