Untangling the relationship between acceptability, grammaticality and processing

Michelle Suijkerbuijk (michelle.suijkerbuijk@ru.nl), Stefan L. Frank
Centre for Language Studies, Radboud University

Acceptability, grammaticality, and sentence processing are all frequently linked concepts in the study of syntax. Although acceptability judgments were originally introduced to measure grammaticality [1], recent research has asked which other factors, such as sentence processing, affect them [2,3]. While studies have independently linked acceptability judgments to grammaticality and sentence processing (via reading times; RTs) [4,5], it remains unclear how the two jointly contribute: Does lower acceptability reflect ungrammaticality, increased processing difficulty, or both? This study addresses that question.

Two hundred Dutch sentence pairs, consisting of a grammatical and a minimally different ungrammatical sentence, where selected from the BLiMP-NL corpus [5]. An example pair is shown in (1) (the critical word, indicating where (1b) becomes ungrammatical, in bold font).

(1) a. Ik bekijk de foto van **mezelf** in de kamer.
b. *Wij bekijken de foto van **mezelf** in de kamer.
'I/*We look at the photo of myself in the room.'

These pairs cover 20 paradigms across 11 syntactic phenomena (e.g., number agreement paradigm of anaphor agreement phenomenon in (1)). Each sentence was presented to at least 30 participants in an eye-tracking study, who also rated the sentence for acceptability after reading it. Thus, grammaticality, acceptability and sentence processing variables were measured within-participants. As measures of sentence processing difficulty, we take the log-transformed first-pass critical word RT, and a binary variable indicating whether an eye-movement regression was made directly following critical word fixation.

We use path analysis to explore the joint contributions of Grammaticality and processing on Acceptability, and the possible directions of causality (Fig. 1). For RT, we find full mediation: Grammaticality affects RT only via Acceptability. For regression, there is no mediation: Ungrammaticality, not unacceptability, causes regressions. If we assume that the effect of Grammaticality on Acceptability may be mediated by processing difficulty (Fig. 1B), we find that such mediation has no noticeable effect, if any.

Our findings suggest that a sense of unacceptability (as opposed to ungrammaticality per se) leads to slower reading, whereas regressive eye-movements are directly due to ungrammaticality. While ease of processing may only have a very small effect (not fully accounted for by Grammaticality), Grammaticality seems to be the main factor driving a sentence's perceived acceptability.

References

[1] Chomsky, N. (1957). *Syntactic Structures*. Mouton de Gruyter.

[2] Sprouse, J. (2018). Acceptability judgments and grammaticality, prospects and challenges. In N. Hornstein, H. Lasnik, P. Patel-Grosz, and C. Yang (Eds.), *Syntactic Structures after 60 years* (pp. 195-224). Mouton de Gruyter. https://doi.org/10.1515/9781501506925-199

[3] Hofmeister, P., Jaeger, F.T., Arnon, I., Sag, I.A., & Snider, N. (2013). The source ambiguity problem: Distinguishing the effects of grammar and processing on acceptability judgments. *Language and Cognitive Processes*, *28*(1-2), 48-87. https://doi.org/10.1080/01690965.2011.572401

[4] Hildebrandt, T. (2023). *The Relationship between Online and Offline Measures of Gradient Sentence Acceptability*. [Doctoral dissertation, University of Michigan]. University of Michigan Library. https://hdl.handle.net/2027.42/176581

[5] Suijkerbuijk, M., Prins, Z., de Heer Kloots, M., Zuidema, W., & Frank, S.L. (2025). BLiMP-NL: A Corpus of Dutch Minimal Pairs and Acceptability Judgments for Language Model Evaluation. *Computational Linguistics*. https://doi.org/10.1162/coli_a_00559

Figures

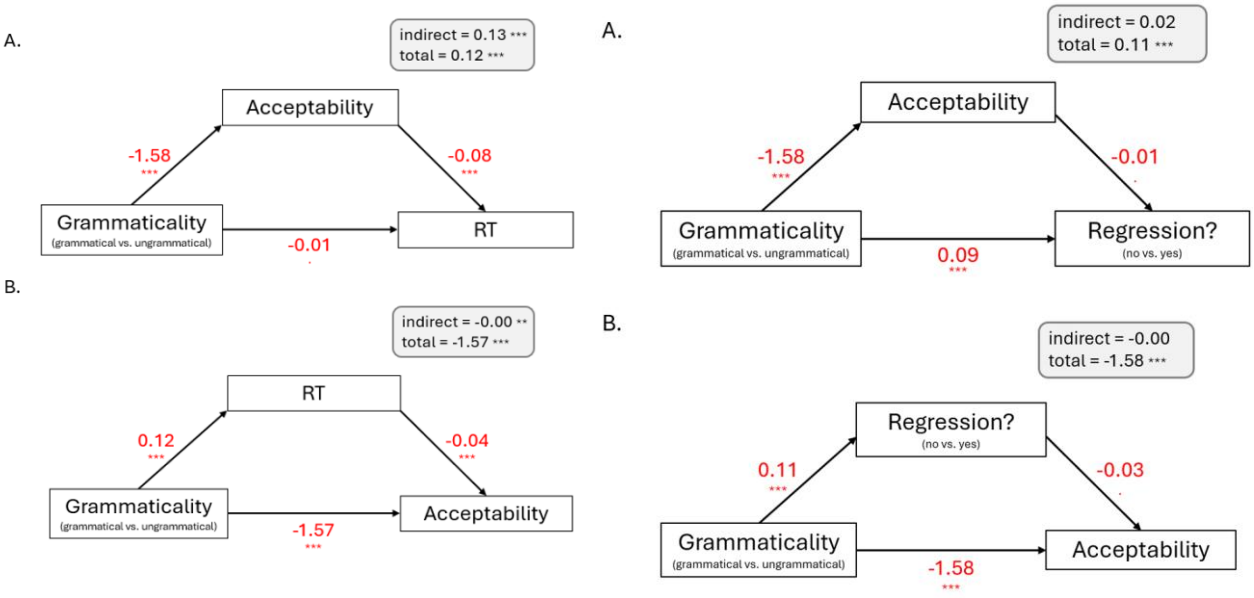

Figure 1: Path diagrams for the relationships between Grammaticality, Acceptability, and one measure of sentence processing (either log-RT on the critical word or eye-movement regression directly following fixation on the critical word). Numbers in red indicate direct effects of one variables on another. The indirect effect is the effect of Grammaticality on sentence processing mediated by Acceptability (A), or on Acceptability mediated by sentence processing (B). The total effect is the sum of the direct and indirect effects of Grammaticality on sentence processing (A) or Acceptability (B).