Agreement attraction effects in the comprehension of grammatical sentences

Anna Laurinavichyute (Higher School of Economics, Moscow; University of Potsdam),
Titus von der Malsburg (University of Potsdam; MIT)
anna.laurinavichyute@uni-potsdam.de

Agreement attraction errors are known to ease the processing of ungrammatical sentences, such as *The key to the cabinets are rusty*[1]. However, there is only scarce evidence suggesting that agreement attraction can also have the reverse effect in grammatical sentences such as *The key to the cabinets is rusty*, i.e. cause a slowdown at the verb due to erroneous representation of subject number effectively leading to an ungrammaticality illusion. The majority of studies haven't found any evidence for such effects, and most of those which did had design confounds. As a result, agreement attraction effects in comprehension have been attributed to faulty repair processes triggered only in ungrammatical sentences.[2] Here we investigate the alternative that attraction effects are present in grammatical sentences as (1b), but masked by an interference effect in the control condition (1a). According to cue-based parsing theories[3] and the encoding interference account[4], inhibitory interference in (1a) arises due to the attractor matching the head noun's number feature[5] which leads to a slowdown at the verb. Therefore, if the interference effect and the attraction effect are similar in size, they might cancel out. We conducted three high-power preregistered online experiments that aimed to attenuate the potentially confounding interference effect so that agreement attraction effects could surface. In each experiment, we selected attractor nouns that were either semantically compatible (a,b) or incompatible (c,d) with the verb hence manipulating the amount of interference.

In three experiments, 16 items were tested in a 2x2 design manipulating semantic and number match/mismatch between the verb and the attractor noun. Participants each read a single experimental sentence (thus preventing adaptation to the stimuli and strategic effects) in an online self-paced reading paradigm. Bayesian LMMs were used for the analysis. Spill-over from the attractor noun was statistically controlled in Experiments 1 and 3 by adding the reading time on the pre-critical word as a covariate in the model for reading times on the critical word[6].

Experiment 1 (N=4296) showed a main effect of attraction on the critical verb (CrI: [0.3 ms, 54 ms], $P(\beta > 0)$ = 97.6%) but no evidence for an effect of semantic interference or an interaction. Experiment 2 (N=3920) introduced long parentheticals with 4 to 6 words (see 2) between the attractor noun and the verb. We found a reliable interaction in the region following the verb: There was an inhibitory effect of attraction in semantic match conditions (1b vs. 1a, CrI: [4.1 ms, 62.6 ms], $P(\beta > 0)$ = 98.6%), but not in semantic mismatch conditions. Experiment 3 (N=3600) probed for grammatical attraction in object relative clauses with 3 words intervening between the attractor noun and the verb (see 3). Again we observed a inhibitory main effect of attraction on the critical verb (CrI:[7.8 ms, 70.6 ms], $P(\beta > 0)$ = 99%).

None of the experiments supports the idea that agreement attraction in grammatical sentences is masked by inhibitory interference in the control condition - in fact, no interference effects were observed in any of the experiments. Instead, all three experiments provide evidence for agreement attraction (ungrammaticality illusion) in configurations where it has not been convincingly demonstrated in previous studies. We suggest that the single-trial design was critical for observing these effects: Participants were not exposed to ungrammatical sentences and did not adapt over the course of the experiment. Taken together, these results suggest that attraction effects do cause ungrammaticality illusion in normal sentence processing and therefore can not be attributed to repair mechanisms alone.

(1)   a. The admirer of the singer   supposedly **thinks** that ...
             b. The admirer of the singers   supposedly **thinks** that ...
             c. The admirer of the play   supposedly **thinks** that ...
             d. The admirer of the plays   supposedly **thinks** that ...
...the show was a big success.

(2)   a. The admirer of the singer, according to the Daily Mail,   supposedly **thinks** that ...
             b. The admirer of the singers, according to the Daily Mail,   supposedly **thinks** that ...
             c. The admirer of the play,   according to the Daily Mail,   supposedly **thinks** that ...
             d. The admirer of the plays,   according to the Daily Mail,   supposedly **thinks** that ...
...the show was a big success.

(3)   a. The singer   that the actor openly **admires** apparently …
             b. The singers that the actor openly **admires** apparently …
             c. The play   that the actor openly **admires** apparently …
             d. The plays   that the actor openly **admires** apparently ...
...received some harsh criticism.
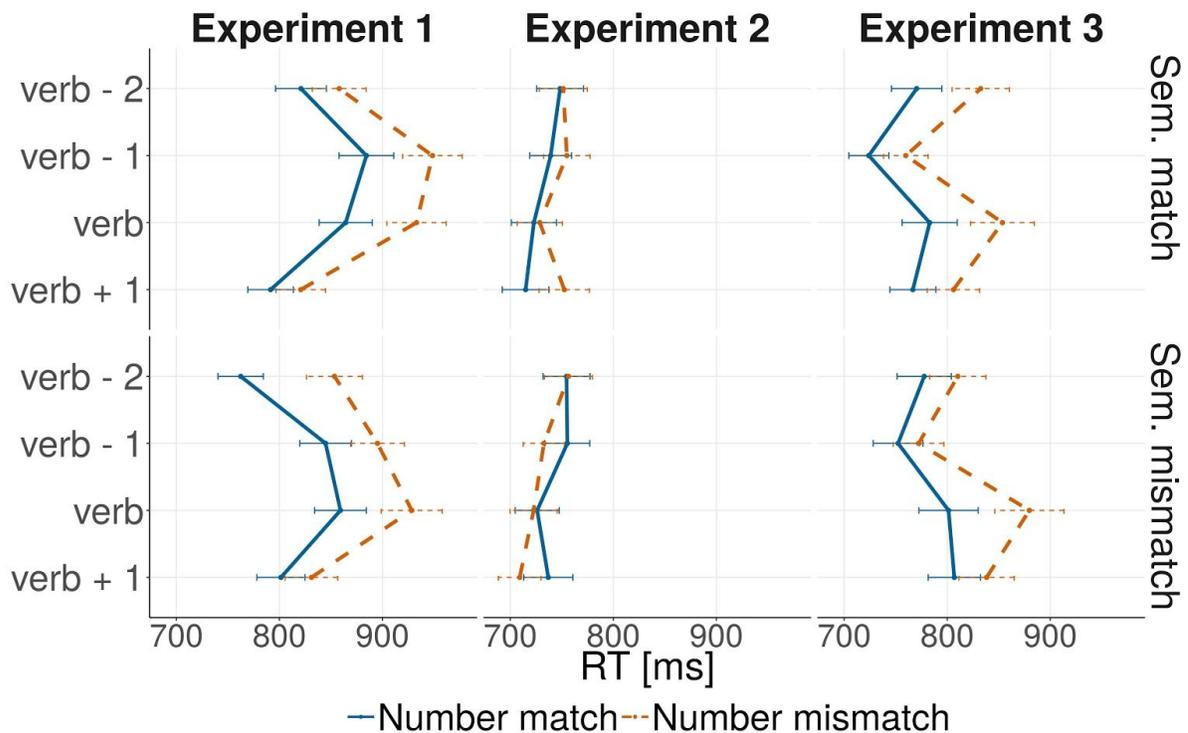
Mean reading times and 95% confidence intervals across conditions.

[1] Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *JML*, *41*(3), 427-456. [2]Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. JML, 61(2), 206–237. [3]Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *CogSci*, 29(3), 375-419. [4]Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *JML*, 55(4), 601-626. [5]Nicenboim, B., Vasishth, S., Engelmann, F., and Suckow, K. 2018. "Exploratory and Confirmatory Analyses in Sentence Processing: A Case Study of Number Interference in German." *CogSci* 42: 1075–1100. [6]Vasishth, S. (2006). On the proper treatment of spillover in real-time reading studies: Consequences for psycholinguistic theories. In *Proceedings of the International Conference on Linguistic Evidence*.