

Agreement attraction in grammatical sentences

Anna Laurinavichyute (Higher School of Economics, Moscow; University of Potsdam),

Titus von der Malsburg (University of Potsdam)

anna.laurinavichyute@uni-potsdam.de

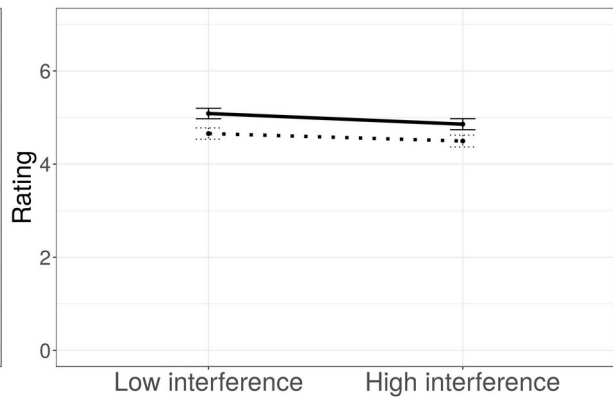
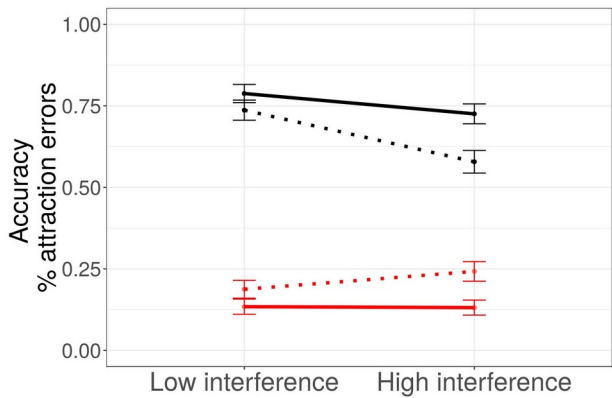
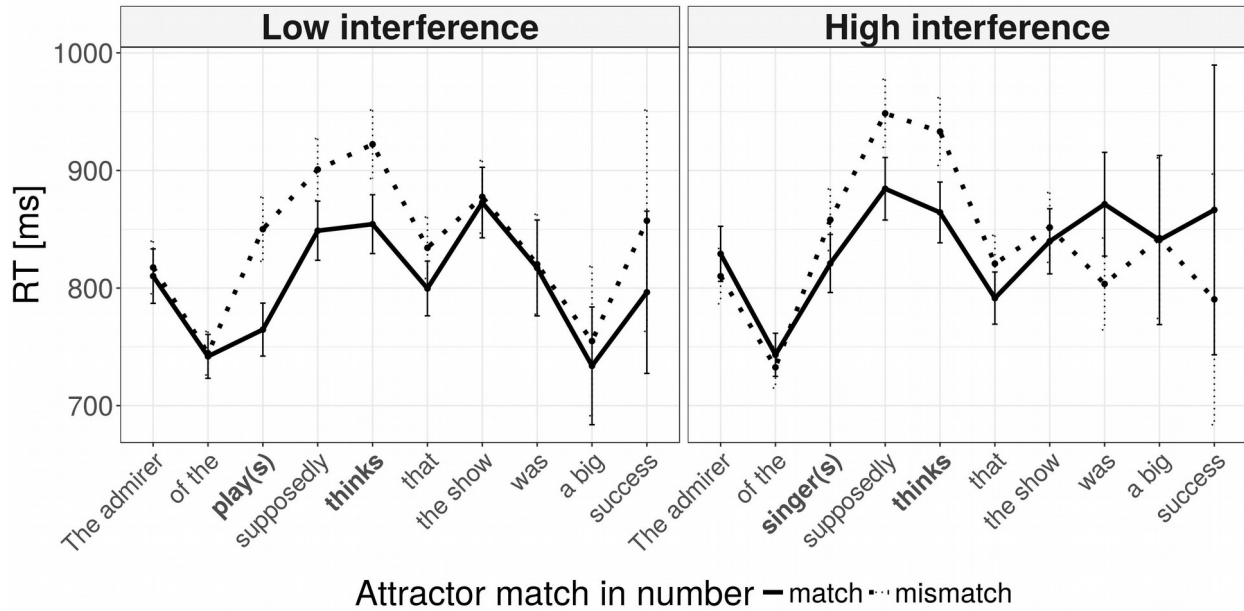
The processing of grammatical agreement is known to occasionally fail in sentence comprehension. For example the sentence “The key to the cabinets were on the table” is sometimes perceived as being grammatical even though the verb “were” does not match the subject in number (“key”). According to the feature percolation account this can happen when the plural feature of the attractor noun (“cabinets”) is accidentally projected onto the root node of the surrounding complex NP (“the key to ...”).^[1] Under this account, attraction effects are also expected to occur in grammatical sentences (“The key to the cabinets *is* on the table”) where they should result in an illusion of ungrammaticality. However experimental evidence for this idea is scarce which led Wagers et al. to argue that agreement attraction can only occur in ungrammatical sentences.^[2] They propose that, in ungrammatical sentences, the parser correctly detects the number-mismatch at the verb, which then triggers a reanalysis process that may occasionally misretrieve the number-matching attractor noun (“cabinets”). In grammatical sentences, reanalysis is not triggered and agreement attraction can therefore not strike.

We propose an alternative explanation for why agreement attraction has not been observed in grammatical sentences. The effects of interest are usually studied using sentences such as as (1a) and (1b). The feature percolation account predicts a slowdown in (1b) as compared to (1a) due to the perceived ungrammaticality. However, according to cue-based parsing theories^[3] and the encoding interference account^[4] we could also expect the opposite pattern, a slowdown in (1a) in comparison to (1b) due to the interference arising from the attractor matching the head's number feature. Effectively, attraction and interference effects might cancel out and this may explain why earlier studies did not find evidence for agreement attraction in grammatical sentences.

The present study aimed to attenuate the potentially confounding interference effect: In conditions (1c) and (1d), attractor nouns were inanimate rendering them incompatible with the verbs hence reducing interference. Thus we expected that, in (1d) vs (1c), agreement attraction effects would not cancel out the way they presumably do in (1b) vs (1a). 16 items were tested in a 2x2 design with the factors of semantic interference and number match/mismatch between the verb and the attractor. In a pre-registered online experiment, 4296 participants each saw a single experimental sentence (thus excluding adaptation to the stimuli and strategic effects). We collected reading times (SPR), sentence acceptability ratings, and question responses (*Who considered the show a success? – Admirer/Singer/Admirers/Singers/I'm not sure*).

In the verb region of conditions (1b) and (1d), we found a robust slowdown in reading times compatible with agreement attraction (est.=0.03, t=3.75). However, inspection of Fig. 1 suggests that this effect can be plausibly explained in terms of spill-over from the attractor noun (plural-complexity effect larger than in previous studies, e.g. [2], potentially due to higher power). Accuracy in the question-response task was lower in the attraction conditions (b/d vs a/c, est.= -0.24, p<0.0001) and the mistakes arose because participants tended to choose the subject noun but with incorrect number marking (e.g. *admirers*). In sum, even though the reading time data were inconclusive, question response accuracies suggest that processes postulated by the feature percolation account might be deployed at least at the post-interpretative stage.^[5] In contrast, interference accounts of sentence processing, which assume that only words from the input are available as candidate answers in the question response task, cannot explain the observed question response pattern.

- (1) a. The admirer of the singer supposedly thinks ...
 b. The admirer of the singers supposedly thinks ...
 c. The admirer of the play supposedly thinks ...
 d. The admirer of the plays supposedly thinks ...



(Accuracy is in black, attraction errors are in red)

[1] Nicol, J. L., Forster, K. I., & Veres, C. (1997). Subject-verb agreement processes in comprehension. *Journal of Memory and Language*, 36(4), 569-587. [2] Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, 61(2), 206-237. [3] Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive science*, 29(3), 375-419. [4] Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory and Language*, 55(4), 601-626. [5] Patson, N. D., & Husband, E. M. (2016). Misinterpretations in agreement and agreement attraction. *The Quarterly Journal of Experimental Psychology*, 69(5), 950-971.