

Prediction illusion due to retrieval interference: a computational model

Umesh Patil (University of Cologne) & Sol Lago (Goethe University Frankfurt)

umesh.patil@gmail.com

While illusions have been mostly studied in the computation of retrospective dependencies, it has been recently shown that they can also occur in predictive processing [1]. [1] conducted a visual-world study with German possessive pronouns, which have a bi-directional pattern of gender agreement: their stem encodes agreement with a previously mentioned antecedent/possessor (as in English *his/her*) but their suffix encodes agreement with a following noun/possessum. Crucially, while only the gender of the suffix is relevant for the prediction of the possessum, participants' predictions in [1] were faster when the possessum and possessor matched in gender (match condition) than when they mismatched (mismatch condition). Since the gender of the possessor doesn't provide any information about the upcoming possessum, the prediction advantage in the match condition is an illusion. We propose that this illusion is due to interference during antecedent retrieval: the gender cue used in the retrieval of the possessor in the match condition boosts the activation of the target picture due to overlap in the gender feature (a similarity-based interference), but the gender cue in the mismatch condition boosts the activation of the competitor picture. This difference in the activation of the target and competitor pictures leads to a faster prediction in the match condition. We present a computational model of the prediction illusion. The model extends the cue-based retrieval architecture proposed in [2] (henceforth CBR).

Data. In a German visual-world study, participants heard an instruction with a possessive pronoun, e.g. "Klicke auf seinen/ihren blauen Knopf" (*Click on his/her blue button*), while seeing a target object and a competitor with different gender on the screen (e.g., a blue bottle). The findings were: (i) predictive looks to the target object before its mention, during the adjective time window; (ii) an earlier onset of predictions when the possessor and target object matched in gender, as estimated by a bootstrapping approach [3].

Modeling. We combined the CBR model of pronoun resolution [4] and the model of eye fixations in a sentence-picture matching task [5] in ACT-R [6] such that the antecedent retrieval process was carried out as per the standard CBR architecture, while the possessum prediction was carried out as a retrieval of the target picture's memory representation (see *Model assumptions* for more details).

Results & discussion. The model captures the two key effects in the data: (i) the prediction of the target object before hearing its name, and (ii) the earlier onset of prediction in the match than the mismatch condition (see *Figure*). The model captures (i) by using the gender feature of the suffix of the possessive and color feature of the adjective to retrieve the target object (e.g. "masculine" and "blue"). The model captures (ii) through an interaction between retrieval and predictive processes. Specifically, the gender cue in the antecedent retrieval in the match condition boosts the activation of objects in memory that match this gender cue, including the possessum, i.e. the target picture (a similarity-based interference effect). On the other hand, the gender cue in the antecedent retrieval in the mismatch condition boosts the activation of the competitor picture but not of the target picture. This difference in activation at the possessive leads to a prediction advantage in the match condition. Thus, our model proposes that retrieval interference in the resolution of the pronoun-antecedent dependency influences the prediction of the following possessum, leading to the prediction illusion.

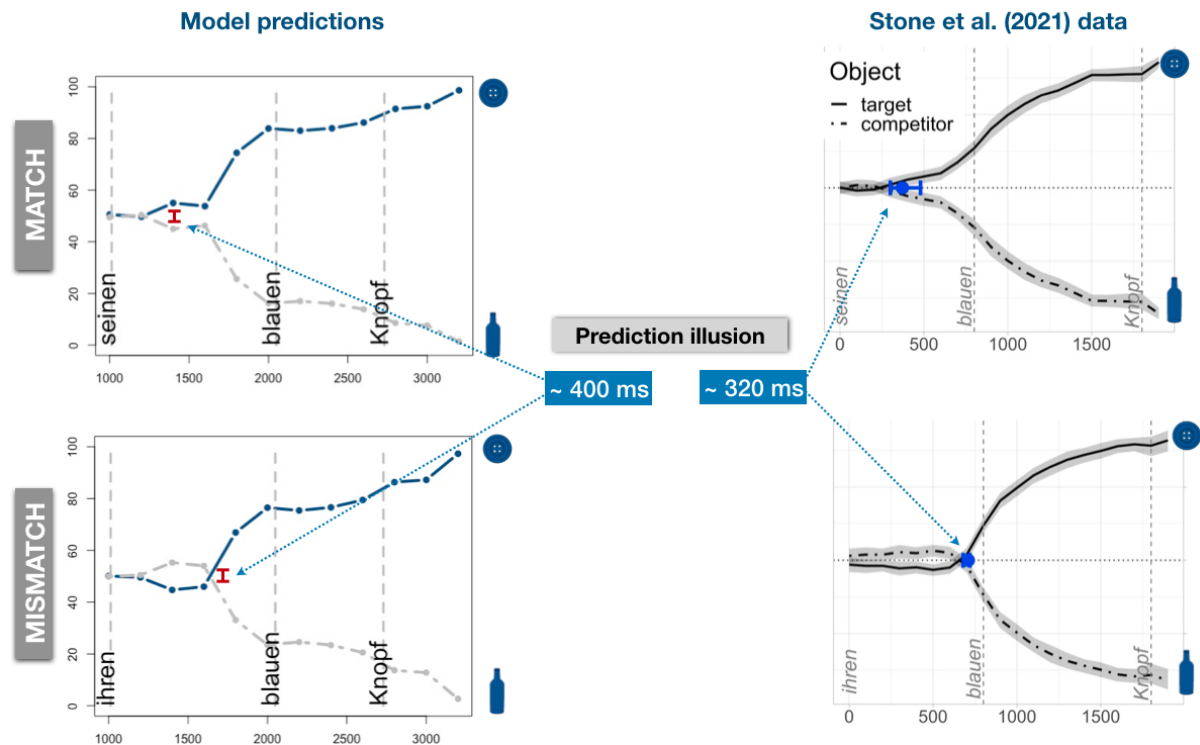


Figure. Right-hand side panels show the visual-world data in match and mismatch conditions. Left-hand side panels show corresponding predictions of the model. Blue circles in the data plots denote the earliest point when target and competitor diverged together with bootstrapped 95% credible intervals, and red bars in the model plots denote the corresponding divergence points predicted by the model. The divergence onset was delayed on average by 320 [200-400] ms in the mismatch condition in the data and by 400 ms in the model predictions (model predictions are generated with 2000 simulations per condition).

Model assumptions. In order to model the task reported in [1], the following assumptions were added to the CBR model of antecedent retrieval.

1. At each input word, the model tries to predict the target picture (the possessum) based on the information in the sentence encountered up to this point in time.
2. The target picture prediction is implemented as the cue-based retrieval of the memory representation of this picture. Additionally, the cues used for retrieval were weighted, such that color cues were weighted more highly than the linguistic cues. This was done to reflect the importance of visual features over linguistic features due to the nature of the visual world task.
3. At the onset of the possessive pronoun, first the antecedent is retrieved and then the target picture is predicted.
4. The activation level of the memory representations of a picture corresponds to its fixation probability.

References

[1] Stone, Oltrogge, Vasishth & Lago (2020) *CUNY 2020*. [2] Lewis & Vasishth (2005) *Cognitive Science*. [3] Stone, Lago & Schad (2021) *Bilingualism: Language and Cognition*. [4] Patil, Vasishth, & Lewis (2016) *Frontiers in Psychology*. [5] Patil, Hanne, Buchert, De Bleser & Vasishth (2016) *Cognitive Science*. [6] Anderson et al., 2004, *Psychological Review*.