The effect of prominence and cue association in retrieval processes: A computational account

Felix Engelmann
University of Manchester, UK

Lena A. Jäger
University of Potsdam, Germany

Shravan Vasishth
University of Potsdam, Germany, and
CEREMADE, Université Paris Dauphine, France

December 4, 2016

Abstract

We present a model of cue-based retrieval in sentence processing that formalizes (i) memory accessibility (prominence) and (ii) a theory of associative cues as extensions to the ACT-R model of Lewis and Vasishth (2005). The extensions are independently motivated and, compared to the original model, enable more differentiated predictions with respect to the experimental design of individual experiments as well as differences between retrieval contexts. The predictions of the original and the extended model are compared with the results of a comprehensive Bayesian meta-analysis of published studies on retrieval interference in reflexive-/reciprocal-antecedent and subject-verb dependencies (Jäger, Engelmann, Vasishth, submitted). Quantitative simulations show that the extended model accounts for effects that are outside the scope of the original model. The results emphasize the importance of accounting for different aspects of memory accessibility, for individual study design, and context-based feature-selectivity in order to generate accurate predictions of a model of cue-based memory retrieval. The simulation results thus shed new light on the cognitive mechanisms underlying interference effects and should be considered in the interpretation of the available data and in the design of future experiments.

Keywords: Cue-based retrieval; Syntactic dependency processing; Retrieval interference; Computational modeling; Cue confusion

Introduction

Memory retrieval in human information processing has been argued to be a cue-based content-addressable mechanism (Anderson et al., 2004; Anderson & Lebiere, 1998; McElree, 2006; Ratcliff, 1978; Van Dyke, 2002; Watkins & Watkins, 1975). It has also been proposed that the formation of non-adjacent dependencies in sentence processing relies on the same process (Lewis, Vasishth, & Van Dyke, 2006; McElree, 2000; Van Dyke & Lewis, 2003; Van Dyke & McElree, 2011). Take, for example, what happens at the verb *was complaining* in Example 1. Our grammatical and lexical knowledge tells us that this verb is to be connected to a *noun phrase* that is *animate* and in *subject* position. Hence, these features are used as *retrieval cues* in order to seek out a linguistic element with the corresponding properties, i.e., *the resident*.

(1) The worker was surprised that **the resident** who was living near the dangerous neighbor **was complaining** about the investigation.

Cue-based retrieval predicts *interference* between memory items that are similar with respect to the retrieval cues even if one of them is in a syntactically illicit position.[1] In Example 1, there are two animate noun phrases, *the worker* and *the dangerous neighbor*, that could potentially distract attention from the grammatical subject, affecting retrieval accuracy or retrieval speed.[2] Numerous studies have found evidence for interference effects in subject-verb dependencies (Dillon, Mishler, Sloggett, & Phillips, 2013; King, Andrews, & Wagers, 2012; Lago, Shalom, Sigman, Lau, & Phillips, 2015; Pearlmutter, Garnsey, & Bock, 1999; Tucker, Idrissi, & Almeida, 2015; Van Dyke, 2007; Van Dyke & Lewis, 2003; Van Dyke & McElree, 2006, 2011; Wagers, Lau, & Phillips, 2009) as well as in reflexive-antecedent dependencies (Badecker & Straub, 2002; Chen, Jäger, & Vasishth, 2012; Cunnings & Felser, 2013; Felser, Sato, & Bertenshaw, 2009; Jäger, Engelmann, & Vasishth, 2015; Sturt, 2003).

Many researchers have explained their findings with respect to the computationally implemented cue-based retrieval model of Lewis and Vasishth (2005), henceforth LV05. The LV05 model is based on the general cognitive architecture ACT-R ("Adaptive Control of Thought-Rational", Anderson et al., 2004; Anderson & Lebiere, 1998).[3] The model provides quantitative predictions of retrieval speed and accuracy by using an incremental parser that

[1] This stands in contrast to structural search theories, which assume that memory access of the left part of a dependency is performed by a search through the syntactic representation of the sentence. A structural search is not subject to interference from items that are in syntactically inappropriate positions.

[2] Note that in an alternative model of cue-based retrieval proposed by McElree, Foraker, and Dyer (2003), interference is only reflected in a decreased retrieval probability of the target but not in retrieval speed. Effects observed in reading times are then explained by time-consuming repair processes which may be triggered by misretrievals (McElree, 1993).

[3] The source code of the original model is available from http://www.ling.uni-potsdam.de/vasishth/code/LewisVasishthModel05.tar.gz. The 2016 version developed by Engelmann (2016) includes eye movement control and is available from https://github.com/felixengelmann/act-r-sentence-parser-em.

Please send correspondence to felix.engelmann@manchester.ac.uk.

relies on rapid memory retrievals which are subject to activation decay and similarity-based interference.

In Jäger, Engelmann, and Vasishth (2016), we carried out a Bayesian meta-analysis of published eyetracking and self-paced reading studies on interference effects in dependency resolution and found that the empirical evidence is only partly consistent with the standard predictions of LV05. However, without doing computational simulations, the data can only be compared to generalized statements of the model's behavior, glossing over any variability or special cases that might lead to non-standard predictions. Ideally, differences between specific dependency environments and between individual experimental designs should be considered when deriving predictions in order to accurately compare the model with the data.

The current paper therefore complements the meta-analysis with a computational re-examination of the predictions of the LV05 cue-based retrieval model under variable conditions. Moreover, we extend the model to account for systematic differences in experimental design by introducing two independently motivated concepts that have not been implemented previously for studying sentence comprehension: *associative cues* and *distractor prominence*.

The idea of associative cues is that the associative strength between retrieval cues and the features of memory items is a result of gradual associative learning during the process of language acquisition. As a consequence, retrieval cues do not necessarily correspond to individual features in a categorical way but rather serve as abstract, experience-based heuristics that can be — strongly or weakly – associated with multiple features. We implement this idea in a cue-based retrieval model as an extension to LV05 in order to study its consequences for the model's predictions with respect to interference.

Distractor prominence accounts for the observation that accessibility of memory items affects similarity-based interference in sentence processing. The accessibility of a noun phrase is increased, for example, when it was encoded in subject position vs. object position, or when it is made salient in the discourse, e.g., by topicalization (Ariel, 1990; Chafe, 1976; Du Bois, 2003; Grosz, Weinstein, & Joshi, 1995). Throughout this article, we will refer to items in these positions as *prominent* items and assume that prominent items are more accessible than non-prominent items. The relevance of accessibility due to prominence for interference effects has been considered in Van Dyke and McElree (2011) for subject-verb dependencies and in Cunnings and Felser (2013) and Patil, Vasishth, and Lewis (2016) for reflexives. Regarding grammatical position, these studies found that a syntactically unlicensed noun-phrase (i.e., the *distractor*) in subject position as in Example 2 from Patil et al. (2016) causes stronger interference effects than when it is in object position as in Example 3 from Sturt (2003).

(2) The tough soldier that **Fred** treated in the military hospital introduced himself to all the nurses.

(3) The surgeon who treated **Jonathan** had pricked himself with a used syringe needle.

Using a design featuring a topicalized distractor, interference effects were found in Sturt (2003, Exp. 1) and Cunnings and Felser (2013, Exp. 2). Cunnings and Felser (2013) used sentences such as Example 4, where the distractor noun phrase was introduced in a context sentence and was referred to in the target sentence through the pronoun *he*. In their Discussion (pp. 212–213), Cunnings and Felser hypothesized that the distractor was more prominently encoded due to reactivation at the anaphora, and that this, in combination with the distractor occupying the subject position and interfering retroactively, may increase the probability of discovering an interference effect at the reflexive.

(4) **James** has worked at the army hospital for years.
The soldier that **he** treated on the ward wounded himself while on duty in the Far East.

These results suggest that grammatical position, discourse saliency, and interference type contribute to the *prominence* of a memory item relative to its competitors at the point of retrieval. Prominent distractors will thus cause stronger retrieval interference than non-prominent distractors.

In this paper, we address three questions: (i) what are the *quantitative* constraints on cue-based retrieval in the LV05 model, (ii) what consequences does a theory of associative cues have on the predictions of interference effects, and (iii) how can we can formalize distractor prominence in ACT-R. We compare the predictions of our simulations with the results of the meta-analysis in Jäger et al. (2016), showing that some of the previously unexplained effects can potentially be explained in terms of distractor prominence and associative cues. In the following section, we explain the retrieval process according to ACT-R and examine the precise predictions with respect to interference effects in the configurations investigated in the meta-analysis.

A re-examination of the Lewis & Vasishth (2005) predictions for retrieval interference

We discuss the predictions of the ACT-R LV05 model for retrieval interference in dependency resolution and contrast them with the results of the meta-analysis of Jäger et al. (2016). Performing a Bayesian meta-analysis is more informative than counting significant vs. non-significant results of individual studies because it quantitatively takes into account all available evidence (including the precision of the estimate) in order to derive an estimate of plausible values of the overall effect of interest. Jäger et al. (2016) included the data of 73 experimental comparisons from published eyetracking and self-paced reading studies in the analysis. A *random-effects* meta-analysis was carried out in order to account for between-study variance, and *covariates* were added for interference type (retro- vs. proactive interference) and distractor prominence in order to estimate the impact of these factors on the interference effect. See Jäger et al. (2016) for details on the procedure.

Target-match and target-mismatch configurations

Jäger et al. (2016) examined the results of studies on subject-verb dependencies, reflexive-antecedent, and reciprocal-antecedent dependencies. The syntactic configurations used in all included studies were similar to Example 5, which is taken from Sturt (2003). The critical position is the reflexive *himself* or *herself*, where a retrieval is assumed to be initiated in order to connect it with its antecedent. In all four configurations, the syntactically correct antecedent for the reflexive is the noun phrase *the surgeon* because it locally c-commands the reflexive, whereas the other noun phrase *Jennifer* or *Jonathan* is inside a relative clause and thus not in the binding domain of the reflexive (Principle A of Binding Theory, Chomsky, 1981). We therefore call the syntactically accessible antecedent the **target**, and the inaccessible noun phrase we call the **distractor**. In the case of subject-verb dependencies, the target is differentiated from the distractor on the basis of it being the *local subject* for the verb while the distractor could be a subject but not the one of the *local* phrase that contains the verb. Between the conditions in 5, the grammatical gender of both target and distractor is manipulated. From a cue-based retrieval perspective, the distractor is assumed to interfere with the retrieval process whenever its gender matches the gender of the reflexive. In 5, the relevant **retrieval cues** and corresponding **features** are shown next to the reflexive and the two noun phrases, respectively. The relevant cues used for retrieval of the antecedent are *c-command*[4] and the gender of the reflexive *masculine* and *feminine*. There are other cues that could be used for retrieval but usually only two cues are of experimental interest since: One cue is used to differentiate between target and distractor (e.g., c-command in the case of reflexives) and one cue is manipulated between conditions (gender in this case). A + or − in front of the features of target and distractor indicates whether there is a **match** or a **mismatch** with the respective retrieval cue.

(5) a. *Target-match; distractor-mismatch*
The surgeon$^{\texttt{+MASC}}_{\texttt{+CCOM}}$ who treated Jennifer$^{\texttt{-MASC}}_{\texttt{-CCOM}}$ had pricked himself$\left\{\begin{smallmatrix}\texttt{MASC}\\ \texttt{CCOM}\end{smallmatrix}\right\}$...

b. *Target-match; distractor-match*
The surgeon$^{\texttt{+MASC}}_{\texttt{+CCOM}}$ who treated Jonathan$^{\texttt{+MASC}}_{\texttt{-CCOM}}$ had pricked himself$\left\{\begin{smallmatrix}\texttt{MASC}\\ \texttt{CCOM}\end{smallmatrix}\right\}$...

c. *Target-mismatch; distractor-mismatch*
The surgeon$^{\texttt{-FEM}}_{\texttt{+CCOM}}$ who treated Jonathan$^{\texttt{-FEM}}_{\texttt{-CCOM}}$ had pricked herself$\left\{\begin{smallmatrix}\texttt{FEM}\\ \texttt{CCOM}\end{smallmatrix}\right\}$...

d. *Target-mismatch; distractor-match*
The surgeon$^{\texttt{-FEM}}_{\texttt{+CCOM}}$ who treated Jennifer$^{\texttt{+FEM}}_{\texttt{-CCOM}}$ had pricked herself$\left\{\begin{smallmatrix}\texttt{FEM}\\ \texttt{CCOM}\end{smallmatrix}\right\}$...

In 3a and 3b, the target matches both cues `CCOM` and `MASC`, i.e., it is a **full match** for the reflexive. We will call these sentences **target-match configurations**. In 3c and 3d, the target does not match the gender of the reflexive and is thus only a **partial match** for the reflexive. Examples 3c and 3d will therefore be referred to as **target-mismatch configurations**. Note that, in this example, the gender match/mismatch of *the surgeon* only refers to its prototypical gender, which is masculine. The meta-analysis included other designs of reflexive-antecedent as well as subject-verb dependencies where gender, case, or other features were manipulated unambiguously, such that the target-mismatch configurations were ungrammatical.

In the published papers that were included in the meta-analysis, interference effects were tested within target-match and target-mismatch configurations by manipulating the match of the distractor with the second cue, which is gender in the present example. In 3b, the distractor *Jonathan* is a partial match for the reflexive because it matches the masculine

[4] Mostly for reasons of simplicity, c-command is usually represented as a static feature similar to gender, case, etc., although it is actually a syntactic *relation* between two items. It is therefore debatable whether some sort of syntactic search mechanism is needed to determine a c-command relationship or whether it is approximated in some other way, e.g. by a *subject* and a *local-clause* feature. See Kush (2013) for an investigation of the computational complexity needed for keeping track of c-commanders.

cue. Under the content-addressable cue-based retrieval mechanism assumed in LV05, a partially cue-matching distractor is a potential retrieval candidate despite it being in a syntactically inaccessible position. Thus, the **distractor-match** condition 3b is considered to induce retrieval interference in comparison with the **distractor-mismatch** condition 3a, where the distractor does not match the gender cue. The same distractor manipulation is applied in the target-mismatch configurations 3c and 3d.

There are two different qualities of interference effects expected in reading time data for target-match and target-mismatch configurations. The presence of a partially matching distractor might either slow down or speed up reading at the critical region, i.e., at the reflexive, the reciprocal, or the verb. Slow-downs and speed-ups are interpreted as **inhibitory interference** and **facilitatory interference**, respectively, meaning that the presence of a distractor leads to an **inhibition** or a **facilitation** during the retrieval process. As we explain below, in the LV05 model, inhibitory effects are expected in target-match configurations, whereas facilitatory effects are expected in target-mismatch configurations.

Constraints on cue-based retrieval in the Lewis & Vasishth (2005) model

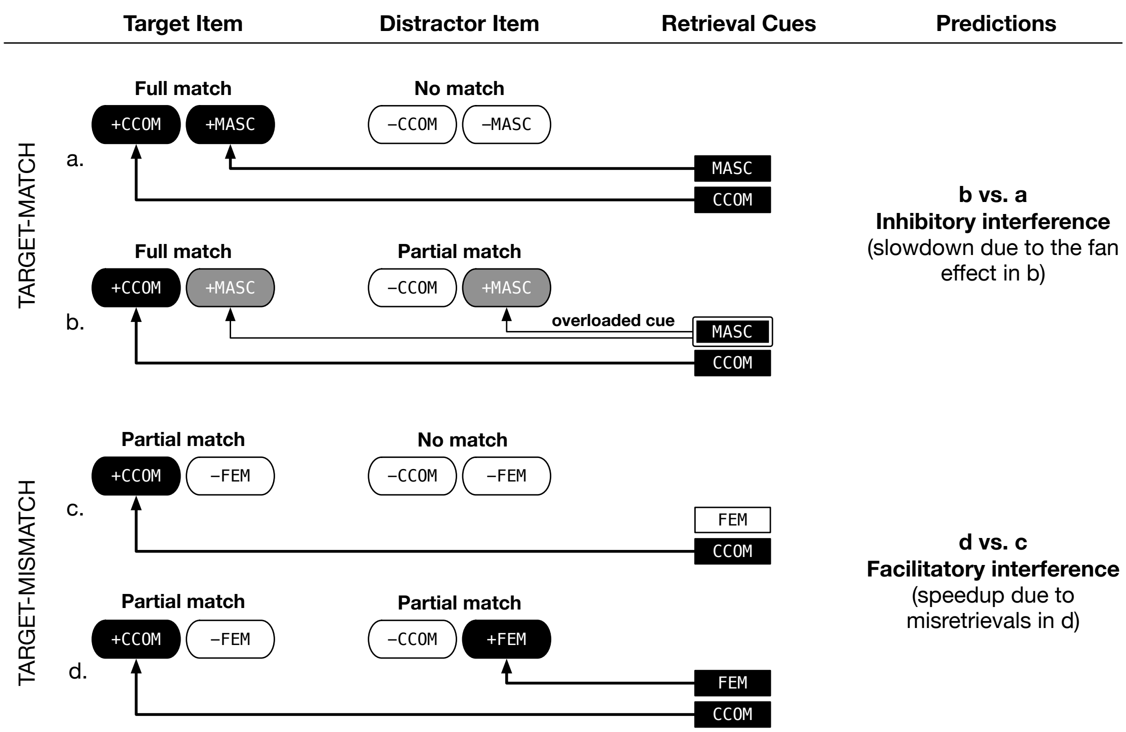


Figure 1. Predictions of ACT-R for the four conditions shown in Example 5. Line weights indicate the amount of spreading activation from a cue to an item. Black oval boxes represent a feature match. Gray oval boxes indicate features matching an ‘overloaded’ cue (`MASC` in b), and white boxes indicate a mismatch.

See Figure 1 for a graphical representation of the model predictions for Example 5. The oval boxes indicate matching (black or gray) or mismatching (white) features of an item with respect to the retrieval cues. The darker the boxes the better is the match of the item and the higher is its activation level. The relative activation levels of memory items in ACT-R determine their retrieval accuracy and retrieval speed. The item with the highest activation at the time of retrieval has the highest probability of being retrieved and the fastest retrieval time. Each item has a *base-level* activation that reflects past usage by accounting for all reactivations (i.e., access events at times t_j) and a time-based decay with rate d (this usually has the default value 0.5 in ACT-R):

$$B_i = \ln(\sum_{j=1}^{n} t_j^{-d}) + \beta_i \tag{1}$$

In addition to the base-level, *spreading activation* is added to every (partially) matching item at the time of retrieval. The spreading activation component is the core source of similarity-based interference effects in ACT-R. An item receives spreading activation from all matching cues j depending on the *associative strength* S_{ji} between cue j and item i and the cue's weight W_j:

$$S_i = \sum_j W_j S_{ji} \tag{2}$$

The arrows in Figure 1 show how activation from the retrieval cues is distributed to the target and the distractor based on their features. The weight of the arrowed lines indicates the amount of spreading activation that is added to an item due to that feature, assuming that each cue is weighted equally. In Figure 1a (cf. Example 5a), the target is a full match for the set of retrieval cues, `MASC` and `CCOM`. Both cues are also **unambiguous** because they are matched by the target only and not by the distractor. The target thus receives the maximal amount of spreading activation at retrieval. In the interference condition b in Figure 1 and Example 5, in contrast, the gender cue is matched by the distractor in addition to the target. Thus, the `MASC` cue is now **ambiguous** or "overloaded" (Watkins & Watkins, 1975), with the result that the activation from this cue is now split between the target and the distractor. This follows from Equations 3 and 4: The associative strength between each cue and the item is calculated by subtracting from a *maximal associative strength MAS* the logarithm of the number of competitor items (3), which is called the **fan** (4).

$$S_{ji} = MAS - \ln(fan_{ji}) \tag{3}$$

$$fan_{ji} = 1 + items_j \tag{4}$$

Each cue distributes the *limited* available activation equally between all matching items (with the maximally available amount being $W_j \times MAS$). The more competitor items are present that match a cue j, the weaker the association of this cue with an item i. Each competitor thus takes away some amount of spreading activation from the target item and thus makes it harder to distinguish from the other items. This is called the **fan effect**. In our example, the fan effect causes a reduction of the spreading activation received by the target in b in comparison with a, thus reducing the target's total activation, which is the sum of the base-level B_i and the spreading activation S_i plus random noise ϵ_i (cf.

Equation 5). A decrease in activation causes retrieval time RT_i to increase. As shown in Equation 6, retrieval time is a negative exponential function of the total activation at the time of retrieval, where F and f are two scaling parameters — the *latency factor* and the *latency exponent*, respectively.

$$A_i = B_i + S_i + \epsilon_i \tag{5}$$
$$RT_i = Fe^{-(f \times A_i)} \tag{6}$$

Hence, the similarity in gender between target and distractor in target-match configurations a vs. b predicts a slower retrieval latency due to the fan effect. We will refer to this slow-down as **inhibitory interference effect**. In addition, there is a higher probability in b compared to a that the distractor is erroneously retrieved instead of the target. This is because activation in ACT-R fluctuates due to the noise component in Equation 5. We refer to retrievals of the distractor as **misretrievals**.[5]

The predictions for retrieval time are different in target-mismatch configurations c and d of Figure 2 and Example 5. In c and d, the target is only a partial match as it does not exhibit the correct gender feature `+FEM`. When the distractor matches the gender in d, there is, however, no reduction in the target's activation. The reason is that both cues `FEM` and `CCOM` are only matched by one item each and are thus not ambiguous. Hence, no fan effect and no inhibitory interference is predicted. However, since target and distractor now both receive the same amount of spreading activation — each matches exactly one cue — their activation levels are relatively close to each other. As a consequence, both items enter into a *race process* where the retrieval of either item is almost equally probable. A race process has the effect that, on average, the retrieval latency is shorter than when there is a clear winner due to a bigger difference in activation as is the case in condition c (e.g., Van Gompel, Pickering, & Traxler, 2001; for simulations demonstrating a race process, see Logačev & Vasishth, 2016). Hence, the prediction for target-mismatch configurations in Figure 2d vs. 2c is a speed-up. We refer to this speed-up as a **facilitatory interference effect**.

Comparison of the LV05 predictions with the results of the meta-analysis

Figure 2A summarizes the predictions of ACT-R for interference effects for simulations of target-match and target-mismatch configurations.[6] The figure shows the possible ACT-R predictions for interference effects on retrieval latency over a range of values for the

[5] Notice that the cue-based retrieval model proposed by McElree et al. (2003) explains inhibitory interference in a different way than ACT-R. In the model proposed by McElree et al. (2003), interference is only reflected in a decreased retrieval probability of the target but not in the speed of the retrieval process. McElree's claim that interference affects only the retrieval probability of the target item and not the latency of the retrieval process is based on his observation that in speed-accuracy tradeoff (SAT) experiments, the intercept and the rate parameter (which represent the retrieval speed) are not affected by interference, whereas the asymptote (which represents the retrieval probability of the target) is sensitive to an interference manipulation (McElree, 2000, 2006; McElree et al., 2003). McElree explains the effects observed in reading times with self-paced reading or eyetracking as a by-product of changes in the retrieval probabilities. The idea here is that misretrievals may trigger a repair process that inflates reading times (McElree, 1993).

[6] Simulations were carried out in R (R Core Team, 2016). The code is available at https://github.com/felixengelmann/inter-act.

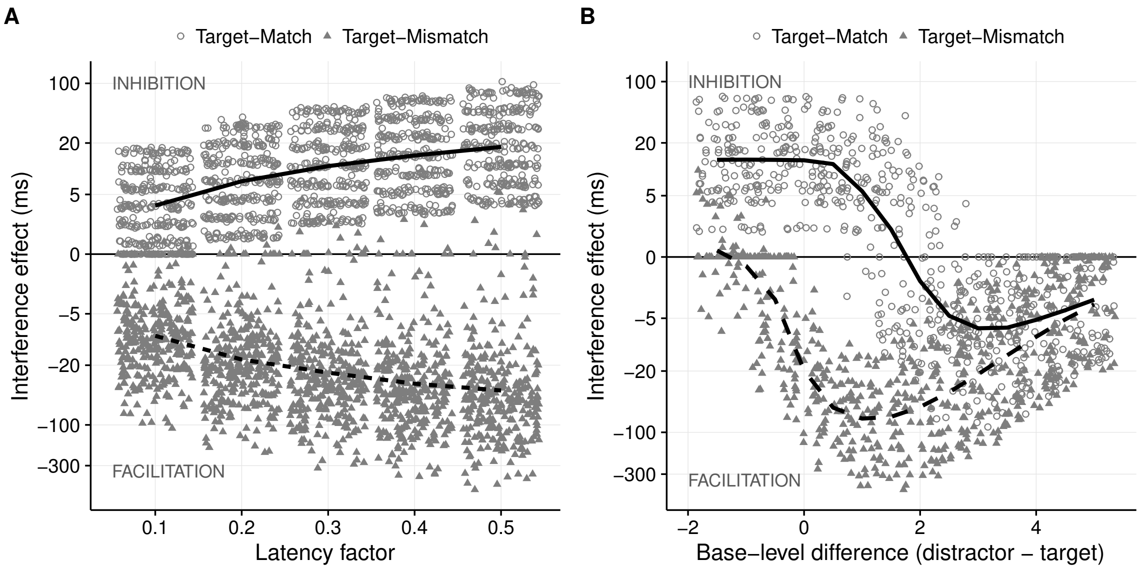


Figure 2. ACT-R predictions for an interference effect (distractor-match − distractor-mismatch) in target-match and target-mismatch configurations for a range of parameter values (768 combinations): Latency factor $0.1 \leq F \leq 0.5$, noise $0.1 \leq ANS \leq 0.3$, maximum associative strength $1 \leq MAS \leq 4$, mismatch penalty $0 \leq MP \leq 2$, retrieval threshold $-1 \leq \theta \leq 0$.
Lines represent the mean effect. Y-axis is log-scaled. A) Predictions as a function of the latency factor F (base-level activations B_i constant). B) Predictions as a function of the distractor base-level activation B_{dstr} while target activation is fixed at 0).

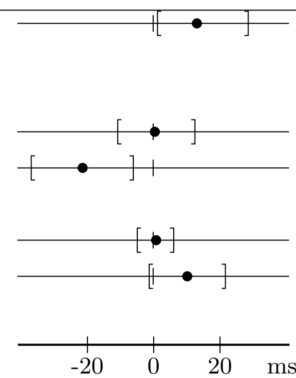
most relevant parameters in ACT-R. The interference effect is calculated as the latency difference between distractor-match and distractor-mismatch conditions (distractor-match − distractor-mismatch) *within* target-match and target-mismatch configurations so that values above zero indicate *inhibitory* interference (slow-down) and values below zero indicate a *facilitatory* effect (speed-up). Along the x-axis of Figure 2A, we plot increasing values of the latency factor F, which is usually the most freely varied parameter in ACT-R models and simply scales the retrieval latency. While there is variation in the mean interference effect along different parameter values, the figure clearly shows that the overall range of predictions is restricted to *inhibitory interference* in *target-match* configurations (caused by the fan effect) and *facilitatory interference* in *target-mismatch* configurations (caused by the race process between target and distractor).

Table 1 summarizes the results of the meta-analysis for the interference effect showing the mean posterior effect estimates and the posterior probability $P(b > 0)$ of the effect being greater than 0. Note that the posterior probability should not be confused with the frequentist p-value; the posterior probability gives the probability of the effect being positive, and therefore there is no concept of a conventional cut-off critical value such as 0.05, and a binary decision of "significant" or "non-significant" would be misleading.

Table 1

Results of the Jäger et al. (2016) Bayesian meta-analysis showing mean posterior effect estimates with Bayesian 95% credible intervals in the Evidence column and the posterior probability of the effect being greater than 0. The credible interval represents the range over which we can be 95% certain that the true value of the estimate lies, given the data (note that the posterior probability should not be confused with the frequentist p-value). A positive interference effect means inhibition, a negative one facilitation. Results are compared with the predictions of cue-based retrieval as implemented in the LV05 ACT-R model.

Dependency	Target	Evidence	$P(b > 0)$	ACT-R
Subject-verb non-agreement	Match		0.98	✓
Subject-verb agreement	Match		0.52	✗
	Mismatch		0	✓
Reflexives/ Reciprocals	Match		0.64	✗
	Mismatch		0.96	✗



In Table 1, we show the effects of interference in target-match and target-mismatch configurations for each of the dependency types separately. In Jäger et al. (2016), subject-verb dependencies were divided into *agreement* dependencies (e.g., Pearlmutter et al., 1999; Wagers et al., 2009) and *non-agreement* dependencies (e.g., Van Dyke, 2007; Van Dyke & McElree, 2011), because these constitute two distinct lines of research, usually showing different patterns. While agreement studies have focused on effects of number attraction, non-agreement studies investigated interference effects involving other semantic and syntactic cues. Reflexive-antecedent and reciprocal-antecedent dependencies were treated as one category in the meta-analysis because both follow Binding principle A and the data of only two publications on reflexives were available.

Clearly, the model cannot account for all the findings of the meta-analysis shown in Table 1. In *target-match* configurations, the predicted inhibitory effect was found only for non-agreement subject-verb dependencies. The other dependency types did not provide enough evidence for an effect in target-match configurations. Most problematic for the model predictions in target-match configurations are individual studies that found a facilitatory effect. For *target-mismatch* configurations, the prediction of a facilitatory effect is only supported by subject-verb agreement studies; reflexive-/reciprocal-antecedent dependencies show inhibition (for non-agreement subject-verb dependencies, no target-mismatch data was available). However, just like for target-match configurations, the individual results of different studies on target-mismatch configurations show a considerable range from facilitatory to inhibitory effects.

The next step is now to examine the effects of systematic differences that could follow from the individual experiments' designs. We do so by first modeling the effect of higher or lower distractor activation levels due to prominence. After that, we examine consequences of different associative strength levels between cues and features as a result of co-occurrence frequencies in different dependency types. We will show that some of the previously unexplained facilitatory effects in target-match and inhibitory effects in target-mismatch configurations can potentially be explained in terms of discourse prominence and associative cues, respectively.

Distractor activation

In ACT-R, an item's base-level activation (the general memory activation B_i irrespective of the item's match with the retrieval specification) is affected by recency due to a time-based activation decay. Consequently, in retroactive interference, where the distractor intervenes between target and retrieval site, the distractor base-level activation relative to the target activation is higher than in proactive interference, where the distractor is read before the target and is thus less recent. We further assume that the accessibility of an item in memory, i.e., its prominence, can also be represented in terms of its base-level activation, with a more highly activated item being more easily accessible. Thus, the question of ACT-R's predictions with respect to interference type and distractor prominence can be answered by inspecting the model behavior as a function of the distractor's base-level activation in relation to the base-level activation of the target. Figure 2B plots the possible range of predictions as a function of the distractor base-level activation in relation to a fixed target base-level.[7] We look first at the effect of distractor activation in target-match configurations.

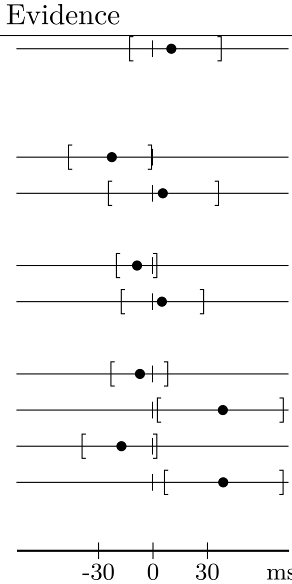
Target-match configurations. As shown in Figure 2A, ACT-R predicts inhibitory effects in target-match configurations. Equation 4 for the fan effect in ACT-R does not account for differences in activation between competing items. As a result, the predicted target-match interference effect would be unaffected by distractor activation as long as it remains lower than the target activation. This is shown in Figure 2B, which plots the range of possible predictions as a function of the distractor base-level activation (which is its activation B_{dstr} *before* adding spreading activation S_{dstr} from partially matching retrieval cues) in relation to a fixed target base-level. The left one-third of the solid line, showing the mean of possible target-match effects, is flat up to just above an activation difference of 0, i.e., when the distractor base-level starts exceeding the target base-level. Assuming that the prominence of an item is reflected by its base-level activation, this would mean that even distractors with a negligible activation level relative to the target would be predicted to cause the same amount of inhibitory interference in target-match configurations as distractors with an activation level equal to the target. Consequently, within this range, neither recency in terms of retro- vs. proactive interference nor prominence are predicted to affect similarity-based interference in target-match configurations in ACT-R.

[7] The target's base-level activation was at 0. Note that 0 does not mean *non-activated*. In ACT-R, activation levels can be positive or negative, with the default being 0.

However, from a certain point, when the distractor base-level activation exceeds the one of the target Figure 2B, the predicted target-match interference effect becomes smaller and finally turns into a facilitation. The explanation for this is a race process between the target and distractor in the same way as it happens in target-mismatch configurations depicted in Figure 1d vs. c. When the distractor activation increases to the point that its final activation at retrieval including base-level and cue match ($A_i = B_i + S_i$) is close to the target's final activation, a facilitatory effect emerges that counteracts the inhibitory fan effect. The decrease stops when the distractor activation is so high that the distractor is always retrieved instead of the target.

Table 2

Results of the Jäger et al. (2016) Bayesian meta-analysis showing mean posterior effect estimates $\bar{b}$ with Bayesian 95% credible intervals in the Evidence column and the posterior probability of the effect being greater than 0 (note that this is not a frequentist p-value). A positive effect of interference type means that retroactive interference led to a right-shift of the interference effect (i.e., an increase of inhibition or a decrease of facilitation) and proactive interference to a left-shift. The effect of distractor prominence is shown for sliding-contrast comparisons of the levels subject OR topic *vs.* other *and* subject AND topic *vs.* subject OR topic*. A positive effect of the distractor prominence comparisons means that a more prominent distractor led to a right-shift of the interference effect (i.e., an increase of inhibition or a decrease of facilitation).*

Dependency	Effect	Target	Evidence	$P(b > 0)$
Subject-verb non-agreement	Retro vs. pro	Match		0.82
Subject-verb agreement	Retro vs. pro	Match		0.02
		Mismatch		0.65
Reflexives/ Reciprocals	Retro vs. pro	Match		0.06
		Mismatch		0.68
All	Prominence OR	Match		0.19
		Mismatch		0.98
	Prominence AND	Match		0.04
		Mismatch		0.99

The meta-analysis estimated the effects of interference type and distractor prominence as shown in Table 2. The table shows the posterior estimates for the effect of retroactive interference (i.e., the distractor is read after the target) in comparison with proactive interference (i.e., the distractor is read before the target) for target-match and target-mismatch configurations for each of the dependency types separately. The prominence of the distractor was coded as three levels based on its grammatical position and topicalization status: *subject OR topic*, *subject AND topic*, and *other* when the distractor was neither a subject nor the topic of the sentence. Due to the sparsity of the data, the effect of prominence was not estimated within each dependency type but as an overall effect.

The analysis results in Table 2 show that increased distractor prominence caused a *left-shift* of the interference effect in *target-match* configurations, meaning that a more prominent distractor led to a decrease of an inhibitory effect or an increase of a facilitatory effect. However, the influence of prominence on the target-match interference effect was mainly evident when comparing *subject AND topic* vs. *subject OR topic* (the effect was below zero with a probability of 0.96), while the comparison of *subject OR topic* vs. *other* yielded weaker evidence for an effect (the effect was below zero with a probability of 0.81).

Interestingly, while in target-match configurations some individual studies found *inhibitory* interference with a distractor in subject position (Patil et al., 2016; Van Dyke & McElree, 2011), others found *facilitatory* interference with a distractor in subject *and* topic position (Cunnings & Felser, 2013; Sturt, 2003).

For interference type, the meta-analysis found an overall left-shift of the interference effect in target-match configurations due to a retroactive distractor (i.e., decreased inhibition or increased facilitation). However, as Table 2 shows, there were substantial differences between dependency types such that non-agreement subject-verb dependencies showed the opposite of other dependencies: While reflexive-/reciprocal-antecedent and subject-verb agreement dependencies showed a left-shift of the target-match interference effect due to retroactive interference, non-agreement subject-verb dependencies showed a tendency for a right-shift.

The finding of the meta-analysis that both distractor prominence and retroactive interference cause a left-shift of the target-match interference effect shown in Table 2 is consistent with the prediction of a decreasing inhibitory interference effect in target-match configurations for an increasing distractor activation in Figure 2B. As explained above, a decrease of an inhibitory effect or increase of a facilitatory effect in target-match configurations is only predicted when the distractor is more active in memory than the target, which corresponds to a base-level greater than zero in Figure 2B. A highly active distractor is a necessary assumption for predicting an influence of prominence or recency on interference in target-match configurations in ACT-R. In cases where the target remains at least as highly activated in memory as the distractor (due to the target occupying a prominent position such as the subject position in the sentence), no effect of distractor activation would be predicted in target-match configurations by ACT-R per se. It is thus not possible to make a meaningful general statement about the prediction of LV05 in the case of the effect of interference type without considering the specific sentence configurations of an experiment.

Judging from Figure 2B, the right-shift of the inhibitory interference effect in target-match configurations of non-agreement subject-verb dependencies due to a retroactive distractor is not predicted under any conditions. However, it is important to note that, when comparing interference effects in retroactive versus proactive designs, what matters is not only the position of the distractor relative to the target, but also the distance between the target and the retrieval site, irrespective of the distractor. In most retroactive interference designs, the target is further away from the retrieval site than in a proactive interference design and thus, due to decay, has a lower activation at the time of retrieval. Due to the negatively exponential relation between retrieval time and activation in Equation 6, a reduction in activation on an item with a lower activation causes a greater increase in retrieval time than the same reduction on an item with higher activation. This means that the inhibitory fan effect due to a partially matching distractor in target-match configurations is greater when the target's base-level activation is lower. Consequently, ACT-R predicts an increased inhibitory effect in target-match configurations for retroactive vs. proactive interference under the assumption that the target is more distant from the retrieval site in the retroactive case. Note that this prediction only depends on the spatial relation between target and retrieval site and not between target and distractor. As mentioned above, the relative activation of the distractor does not influence the fan effect directly. Because both the relation between target and distractor and the relation between target and retrieval site play a role with respect to interference type, the specific characteristics of the materials being tested need to be taken into account when deriving predictions of the LV05 model. We now turn to the effect of distractor activation in target-mismatch configurations.

Target-mismatch configurations. Figure 2B shows the prediction of a stronger facilitatory effect for more activated distractors in target-mismatch configurations, which becomes weaker in the extreme where the target activation becomes irrelevant relative to the very high distractor activation. At this point, the distractor activation exceeds the target activation so much that the distractor becomes practically the only retrieval candidate and, as a result, any interference effect disappears.

In the meta-analysis results shown in Table 2, an increased distractor prominence led to a right-shift of the interference effect in target-mismatch configurations for both *subject OR topic* and *subject AND topic*, i.e., inhibition was increased or facilitation decreased for more prominent distractors. A right-shift was also found for interference type in target-mismatch configurations across dependency types (i.e., increased inhibition or decreased facilitation for retroactive distractors).

As shown in Figure 2B, a right-shift of the interference effect in target-mismatch configurations due to prominence and retroactive interference is only predicted for an extremely high distractor activation that basically rules out the target as a retrieval candidate. That a reader would completely rule out the target for completing the dependency is arguably an unrealistic situation. One could, however, argue that, when the distractor is extremely salient, the reader would likely notice that the distractor is not suited for completing the dependency such that a facilitation due to undetected misretrievals would not occur. Rather, the reader would actively disregard the distractor in order to find the correct target, causing an additional slow-down. Thus, a right-shift of the interference effect in target-mismatch configurations due to distractor activation can be explained if we assume an extremely high saliency of the distractor compared to the target — this is the same assumption we had to make for explaining the left-shift in target-match configurations. Again, we have to emphasize the importance of accounting for the individual experimental design when deriving predictions.

In summary, by accounting for distractor prominence in the model, the variations in magnitude of the interference effect shown in Table 2 can potentially be explained. In addition, facilitatory interference effects in target-match configurations — a previously unexplained change of sign — are predicted for high distractor prominence.

However, inhibitory interference in target-mismatch configurations cannot be explained by distractor activation because no fan effect is present in target-mismatch configurations. The next section introduces the concept of associative cues, which predicts a fan effect for certain retrieval contexts even in target-mismatch configurations.

Associative Cues

The concept of associative cues states that (i) a cue-feature match is not a categorical *yes-or-no* decision but a continuously valued strength of association and (ii) a cue can be associated with *multiple* feature values at varying degrees, leading to *cue confusion*. This predicts that, under certain conditions, a fan effect can arise between two items in a target-mismatch configuration with the consequence of inhibitory interference — something which is not possible in standard ACT-R. We will explain this prediction shortly. First, we briefly introduce the principle of associative cues.

We assume that a language user's knowledge of the necessary properties or features to successfully identify the correct target is based on statistical knowledge derived from experience with a given dependency type. Based on this assumption, we propose that a cue-feature relation is not categorical but rather an association resulting from a usage-based process of learning to discriminate between relevant features. Expressed in terms of classical conditioning, two stimuli (features) are discriminated when they elicit different responses (Rescorla & Wagner, 1972). If two features frequently co-occur in constituting the correct target, they will be less discriminated. For each retrieval context there will be slightly different feature distinctions of relevance or, more broadly speaking, the content of our memory is categorized or clustered differently according to current needs.

Translated into ACT-R terms, discriminating features at retrieval means there are distinct cues activating each feature. If features are not discriminated at all, they are activated by the same cue, i.e., they are treated as one and the same feature by the retrieval process. Depending on the specific feature co-occurrence rate, there are gradations between these two extremes. Some features might be very similar but not identical in terms of relevance to the retrieval task. There would then be two cues that are to a certain degree activating both of the features. We will say that, in this case, the cues are *cross-associated*. The stronger the association between a cue and a feature, the more efficient the cue is for discriminating items with that feature in memory.[8]

As an example of cross-association, consider the difference between the retrieval contexts of reflexives and of reciprocals. As shown in Table 3, the correct antecedent for an English reflexive can exhibit different feature combinations depending on the specific form of the reflexive, i.e., *himself*, *herself*, *itself*, and *themselves*. A dissociation of c-command, number, and gender in the retrieval request is therefore necessary for identifying the correct target with respect to the individual form of the reflexive. Therefore, the retrieval cues in English reflexives are *highly selective*. In contrast, correct targets for a reciprocal invariably exhibit the features `+PLUR` and `+CCOM`. Because of their invariable co-occurrence, an effective

[8] In the context of the ACT-R architecture, an associative relation between cues and features, though not implemented, is a rather straightforward assumption. In ACT-R, any two memory items can be assigned a mutual similarity. This includes feature values and cue values as these are items in memory, too. Similarities are used, for example, in the equation for a component called *mismatch penalty*, which we do not discuss here. We propose that the same similarity values should be included in the computation of the fan effect.

retrieval cue specification for reciprocals does not require a strong discrimination between `+PLURAL` and `+CCOM`. Instead, it can be thought of as more efficient to also activate plural items with the `CCOM` cue and vice versa. This is a case of *low selectivity*. As a result, in the context of a reciprocal-antecedent dependency, the cues `CCOM` and `PLURAL` would both be associated to some degree with both the features `+CCOM` and `+PLUR`, i.e., they are *cross-associated*. A similar situation arises for the Chinese reflexive *ziji*, also shown in Table 3, that requires an animate c-commanding target. This, in the case of *ziji*, `CCOM` would be cross-associated with `ANIM`. For the situation when two cues are cross-associated, we will also use the term *cue confusion*.

Table 3
Possible feature combinations exhibited by correct antecedents of English reflexives, reciprocals, and Chinese ziji.

Context	Target features	Form
EN reflexive	{+MASC, +CCOM}	himself
	{+FEM, +CCOM}	herself
	{+NEUT, +CCOM}	itself
	{+PLUR, +CCOM}	themselves
EN reciprocal	{+PLUR, +CCOM}	each other
CN reflexive	{+ANIM, +CCOM}	ziji

We will say that two cues are cross-associated at a certain level C between 0 and 1, where $C = 0$ means that the corresponding features are maximally discriminated and $C = 1$ means that the corresponding features are treated as functionally identical, i.e., both cues activate the same features. In order to quantify the cross-association level for a specific context, such as reciprocal-antecedent vs. reflexive-antecedent dependencies, one could carry out a corpus analysis counting co-occurrence frequencies. For an illustration of the difference in cross-association levels between two dependency contexts, consider the following simplification: We define C_x as the relation between the number of possible combinations in a retrieval context x, where the correct target possesses both features k and l simultaneously and the number of all possible combinations in context x where at least one of the features is present:

$$C_x(k,l) = \frac{\sum_i [k \in Comb(x)_i \wedge l \in Comb(x)_i]}{\sum_i [k \in Comb(x)_i \vee l \in Comb(x)_i]}, \tag{7}$$

where $Comb(x)_i$ denotes feature combination i of all possible feature combinations $Comb(x)$ for the correct target of dependency context x, and the square brackets represent an Iverson bracket which denotes 1 if the enclosed condition is satisfied and 0 if not.

This way, we can say, e.g., that the confusion level between `CCOM` and `PLUR` in reciprocals is $C_{reci}(\texttt{CCOM}, \texttt{PLUR}) = 1/1 = 1.0$ compared to a level of $C_{refl}(\texttt{CCOM}, \texttt{MASC}) = 1/4 = 0.25$ between `CCOM` and `MASC` in English reflexives. This, of course, delivers a simplified picture which merely provides an estimate of the relative difference between two dependency contexts.

Predictions

Inhibitory interference in ACT-R is caused by the fan effect when a cue is *ambiguous*, i.e., it shares its spreading activation between target and distractor because they possess the same feature. As shown in Figure 1, this is only the case in target-match configurations such as Examples 5b vs. 5a. In target-mismatch configurations like 5c and d, on the other hand, target and distractor do not share the manipulated feature and, hence, no fan effect is predicted. This prediction rests on the simplified assumption of a categorical match between cues and features.

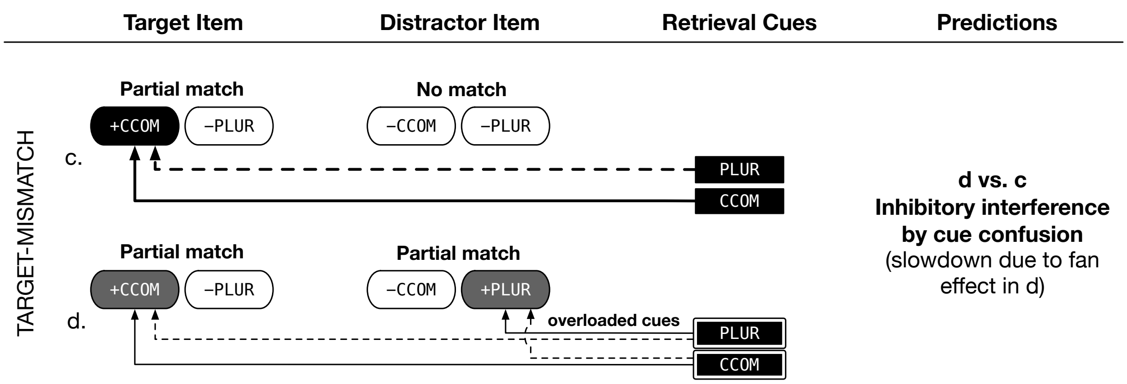


Figure 3. Predictions of the extended ACT-R model showing the consequences of cue confusion in target-mismatch configurations of, e.g., reciprocals. Line weight and box shading indicate the amount of spreading activation added to an item due to a feature match. Dashed lines represent spreading activation to a cross-associated feature as a result of cue confusion.

Under the theory of associative cues, a cue would in some situations share its spreading activation between categorically distinct features. In these situations, a fan effect would be predicted even in target-mismatch configurations (5d vs. d). Figure 3 illustrates how a hypothetical confusion of `CCOM` and `PLUR` for reciprocals would lead to inhibitory interference in target-mismatch configurations. In the target-mismatch/distractor-mismatch condition c, the target receives activation from the `CCOM` cue and the `PLUR` cue although it only carries the `+CCOM` feature. In the target-mismatch/distractor-match condition d, the distractor carries the `+PLUR` feature. Since there is a *cross-association* between cues and features, both cues are now *ambiguous* and share the available activation between the `+PLUR` and `+CCOM` features. As a consequence, both target and distractor are less activated in d than the target is in c. Hence, no matter which item is retrieved, it will be retrieved slower in d vs. c, predicting inhibitory interference for confused cues in target-mismatch conditions due to the fan effect.

The theory of associative cues predicts a higher confusion level for both reciprocals and the Chinese reflexive *ziji* compared to English reflexives. This could explain the result of Kush and Phillips (2014), who found inhibitory interference in target-mismatch conditions in Hindi reciprocals, as well as our finding of an inhibitory target-mismatch effect for Chinese *ziji* in Experiment 1 of Jäger et al. (2015).

Implementation of prominence and associative cues

We extend the standard ACT-R architecture with a mechanism for prominence and cue confusion by redefining the strength of the association S_{ji} between a cue j and an item i.[9] The association between cue j to item i reflects the probability of the item being needed given cue j (Schneider & Anderson, 2012, Eq. 2a):

$$S_{ji} = MAS + \ln[P(i|j)] \tag{8}$$

If all chunks associated with j are equally likely (i.e., *useful* in the context of cue j), $P(i|j) = 1/fan_j$, e.g., for two chunks that are equally useful, $P(i|j) = 0.5$. That is why, in ACT-R, the strength of association S_{ji} is usually simplified as:

$$MAS + \ln(1/fan_j) = MAS - \ln(fan_j) \tag{9}$$

However, in order to build a model that reflects differences in encoding strength between items (*prominence*) and selectivity between cues (*cue confusion*), we need to define how $P(i|j)$ is computed in the context of ACT-R. We compute $P(i|j)$ here as the **match quality** of item i with cue j in proportion to the match quality of all other active memory items v with j:

$$P(i|j) = \frac{Q_{ji}}{\sum_v Q_{jv}} \tag{10}$$

Associative cues. We assume that a cue can be of variable selectivity, i.e., it can be associated with multiple features to different degrees. The association between a cue j and a feature k is given by M_{jk}. The individual match quality Q_{ji} of cue j with item i then depends on the associative strength between j and all features K of i, weighted by the *general saliency* of the item:

$$Q_{ji} = \sum_{k \in K_i} M_{jk} \times saliency_i \tag{11}$$

Cue confusion predicts a fan effect also for items that do not share any of their features, as long as the same cue is associated with features from both items. For example, take an item that has feature f1 but not feature f2. Without cue confusion, the summed associations of cues c1 and c2 with item i would be $M_{c1,i} = 1$ and $M_{c2,i} = 0$. If we assume, however, a confusion level (or level of cross-association) of $C = 0.5$, cue c2 spreads 50% of its activation to feature f1 in addition to the activation from cue c1. Consequently, the association of the item with cue c2, $M_{c2,i}$, is 0.5 instead of 0, i.e., the item receives extra spreading activation from the cue that would not be matched categorically. Similarly, at a confusion level of $C = 0.25$, $M_{c2,i} = 0.25$, and at $C = 1$, $M_{c2,i} = 1$. This extra activation will, however, be shared when there is a distractor that matches this feature f2. Thus, because the cues have crossed associations, the item activation is reduced by the fan effect even though it does not overlap in features with the distractor. Figure 4A illustrates the effect of an increasing cue confusion level. In target-mismatch configurations, a higher cue confusion causes an inhibitory fan effect that eliminates the facilitatory effect.

[9] This implementation is inspired by a suggestion from the action editor, Klaus Oberauer.

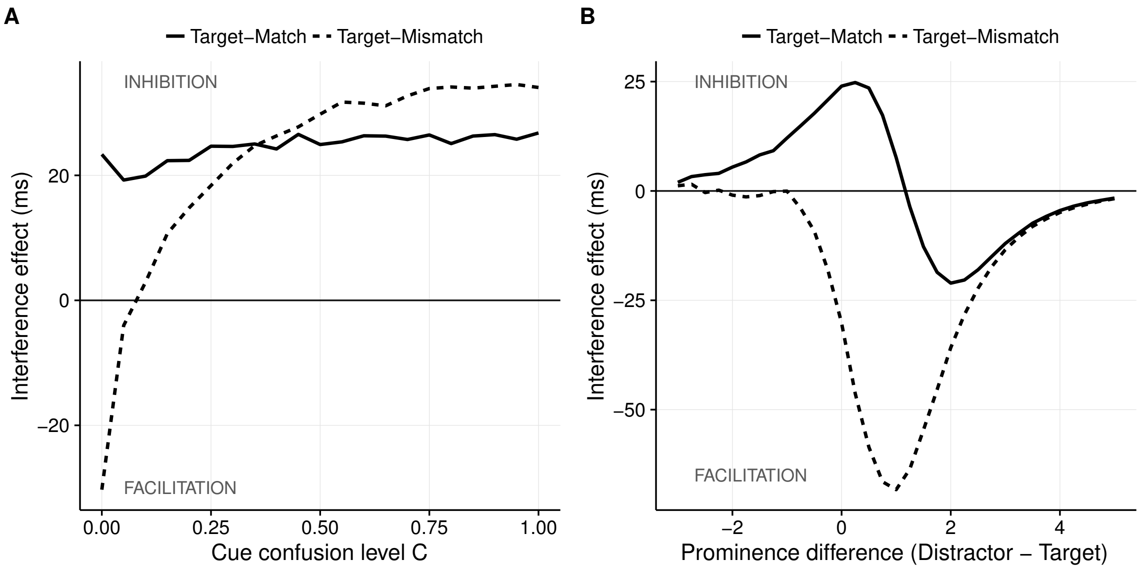


Figure 4. Predictions of the extended model. A) Predicted interference effect as a function of the cue confusion level C. B) Predicted interference effect as a function of distractor prominence p_{dstr} with fixed target prominence.

Prominence. In order to integrate distractor prominence into this account, we define the saliency of an item similarly to the equation of retrieval probability in Anderson et al. (2004, p. 1042), reflecting the current base-level activation B_i as well as a prominence component p_i:

$$saliency_i = \frac{1}{1 + qe^{-(B_i + p_i - \tau)}} \tag{12}$$

where τ is the retrieval threshold, and q is a scaling constant, the *quality correction factor QCF*. When $q = 0$, saliency or prominence is not reflected in $P(i|j)$. When $q = 0$ and all cues are maximally selective (i.e., exactly one feature matches one cue), $P(i|j) = 1/fan_j$, in which case the model behavior is identical to standard ACT-R. If, however, $q > 0$, the saliency of an item—and with it the item's prominence—affects the associative strength between the retrieval cues and the item.

Figure 4B illustrates the relationship between distractor prominence and the interference effect as predicted by the extended model, assuming that target prominence is fixed. In addition to the facilitatory effect of highly activated distractors in target-match predicted also by standard ACT-R, the extended model additionally predicts that the fan effect only arises for sufficiently activated distractors (cf. the rising inhibition in target-match configurations in the figure). This could explain the right-shift of the inhibitory interference effect in target-match configurations of non-agreement subject-verb dependencies due to a retroactive distractor (cf. Table 2) in addition to the stimuli-dependent explanation given in the previous section.

The above equations define the probability of a memory item i being needed given cue j, $P(i|j)$, with respect to the item's general saliency and its association with cue j, M_{ji}. The equations ensure that cues can be of variable selectivity (can be associated with one or more features), i.e., cue confusion, and that more salient (or useful) items are more strongly associated with the cues and, hence, receive more spreading activation, i.e., prominence. Since $P(i|j)$ is a probability that takes into account all memory items, both the selectivity of cues and the saliency of an item affect the fan effect, i.e., the strength of inhibitory interference. The equations for the total spreading activation for item i (Eq. 2) and the retrieval latency (Eq. 6) remain the same as in the original implementation.

For the extended model we present below, we assume that factors such as subject position and topicalization increase the prominence of a distractor, making it more salient in memory. Saliency in ACT-R is expressed by an item's base-level activation level, which also reflects its recency. Therefore, the effect of interference type is predicted to add to the effect of prominence: A distractor that interferes retroactively and is the discourse topic and in subject position has the highest activation. Our account of distractor prominence predicts that distractor activation due to prominence and recency systematically increases the interference effect in target-match and target-mismatch configurations and leads to facilitation in target-match configurations at very high activation levels.

Simulations

We simulated all studies from the meta-analysis of Jäger et al. (2016) with both the original LV05 model and the new model that was extended with prominence and associative cues as described above. The aim of the simulations was to demonstrate in which way distractor prominence and cue confusion change the model predictions when accounting for distractor position and different co-occurrence patterns between dependency types.

Method

The original LV05 model and the extended model were both implemented and run in R (R Core Team, 2016). The model code is publicly available on GitHub.[10] In addition, we provide an application for running simulations with the extended model online.[11] The model simulated retrieval latencies with two or more memory items present (some studies used more than one distractor) and feature settings according to the target-match and target-mismatch conditions in Example 5. Two cues were specified at retrieval. The first cue was matched by one memory item in all conditions, which distinguished this item as the target. The second cue was matched by the target in conditions a and b (target-match) and by the distractor in conditions b and c (distractor-match).

Data. Figure 5 shows the number of target-match and target-mismatch comparisons that were studied in the meta-analysis and included in our simulations, arranged by dependency type and level of distractor prominence. The data included studies on target-match and target-mismatch configurations in subject-verb agreement, reflexive-antecedent, and reciprocal-antecedent dependencies. Studies on non-agreement subject-verb dependencies

[10] The model code is available at `https://github.com/felixengelmann/inter-act`.

[11] The application was built using shiny (Chang, Cheng, Allaire, Xie, & McPherson, 2016) and is located at `https://engelmann.shinyapps.io/inter-act/`.

only included target-match configurations. All three prominence levels (*other*, *subject OR topic*, and *subject AND topic*) were available only in reflexive-antecedent studies. For the other dependency types, only the levels *other* and *subject OR topic* were available. In order to ensure common parameter settings for target-match and target-mismatch comparisons when they were part of the same experiment, the 73 data points used in the meta-analysis were modeled in 49 *experimental sets*, such that physically identical experimental conditions presupposed identical model conditions in the simulations.

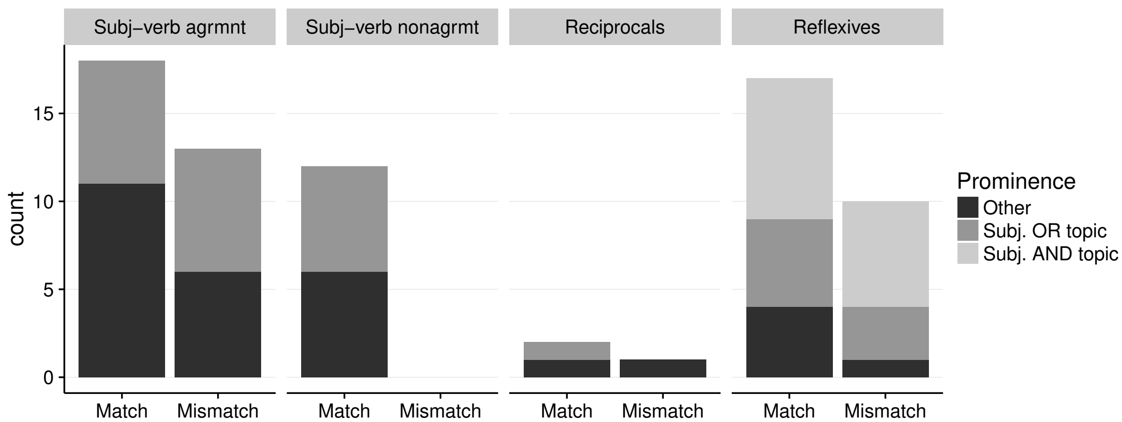


Figure 5. Number of studies included in the meta-analysis and in the simulations by dependency type and prominence.

Parameter estimation. In the extended model, we introduce two new parameters that are not present in the original LV05 model: distractor prominence p_{dstr} and cue confusion level C. In order to avoid arbitrary degrees of freedom and overfitting, we took the following measures: (i) In the extended model, the introduction of varying p_{dstr} was compensated by removing the estimation of the latency factor F; (ii) the range of p_{dstr} was restricted with respect to experimental design as explained below; (iii) cue confusion level C was estimated only for the two cases motivated above: reciprocals and the Chinese reflexive *ziji*. It was set to 0 otherwise.

As is common practice in ACT-R, we estimated the latency factor F (see Equation 6) for each experiment in the LV05 model to scale the predictions into a range that is comparable with the data. For the simulations with the extended model, we estimated the distractor prominence parameter p_{dstr} by experiment *instead* of F. The latency scaling factor F was estimated by dependency type only, resulting in 0.15 for reflexive-antecedent dependencies and 0.2 for the other dependency types. The range of possible values for p_{dstr} was shifted according to the prominence level of the distractor (*OTHER*, *OR*, *AND*) as follows. The parameter p_{dstr} was restricted to three ordered ranges of values such that medium prominence (*subject OR topic*) was constrained to be close to the target prominence ($p_{trgt} = 0$) with $-1 \leq p_{dstr} \leq 2$, whereas low prominence (*OTHER*) was constrained to be smaller than the target prominence ($-3 \leq p_{dstr} \leq 0$), and high prominence (*subject AND topic*) was bound to higher values ($1 \leq p_{dstr} \leq 4$). Thus, the full range of predictions shown in Figure 4B can be generated but these are restricted to specific properties of the distractor. The ordering of the levels is motivated by our assumption that distractors that are *subject*

AND topic are most prominent and distractors that are neither subject nor topic (labeled *OTHER*) are least prominent. We allowed the value ranges to overlap, because we had no specific assumptions about the prominence parameter values except for the ordering of the levels. Since the relevant manipulation was the difference between distractor prominence and target prominence, only distractor prominence was estimated and target prominence stayed fixed at the default prominence value of 0. Note that, by substituting distractor prominence for the latency factor in the extended model, we replaced a scaling parameter with a meaningful parameter which is interpretable within our cognitive model.

Interference type (retro vs. proactive interference) was reflected in the model by manipulating the order of target and distractor. For retroactive designs, the target was more distant from the retrieval site than the distractor and vice versa. Hence, interference type affects the model through the memory decay component, which reduces the activation of an item as a function of time.

Results

We ran simulations with the LV05 model and the extended model. Because Lewis and Vasishth (2005) speculated that model fit might improve without the decay component of ACT-R,[12] we also ran variants of both models without the decay component. Interference effects were computed within target-match and target-mismatch conditions as the difference between distractor-match (high interference) and distractor-mismatch (low interference) conditions, averaged over 5000 iterations per simulation.

Table 4
Root-mean-square deviation between model predictions and observed data, averaged within dependency type and model.

Dependency	LV05	LV05+decay	Extended	Extended+decay
Subj-verb agreement	15.92	17.25	**13.52**	14.12
Subj-verb non-agrmnt	8.25	7.31	8.39	**5.75**
Reflexives/Reciproc.	11.76	11.97	**6.17**	6.82

Table 4 summarizes the fit of all four model configurations, averaged within dependency type. Overall, the extended model fit the available data better than the original model of LV05. Except for non-agreement subject-verb dependencies, the use of decay did not improve the fit with the data. With respect to the extended model, decay only improved the fit for non-agreement subject-verb dependencies but, for the other dependency types, produced a worse fit compared to the model without decay. This suggests that the information about the linear order of target and distractor (pro- vs. retroactive interference) is not useful as a predictor for all dependency types when it is accounted for by time-based decay. We come back to this in the Discussion section.

For a report of all simulated data points, see Figure 6, which shows the predictions of LV05 and the extended model with decay and the data ordered by increasing distractor

[12] Lewis and Vasishth (2005) write on p. 408: “Any structural or quantitative change to the model that moves in the direction of decreased emphasis decay and increased emphasis on interference would likely yield better fits.”

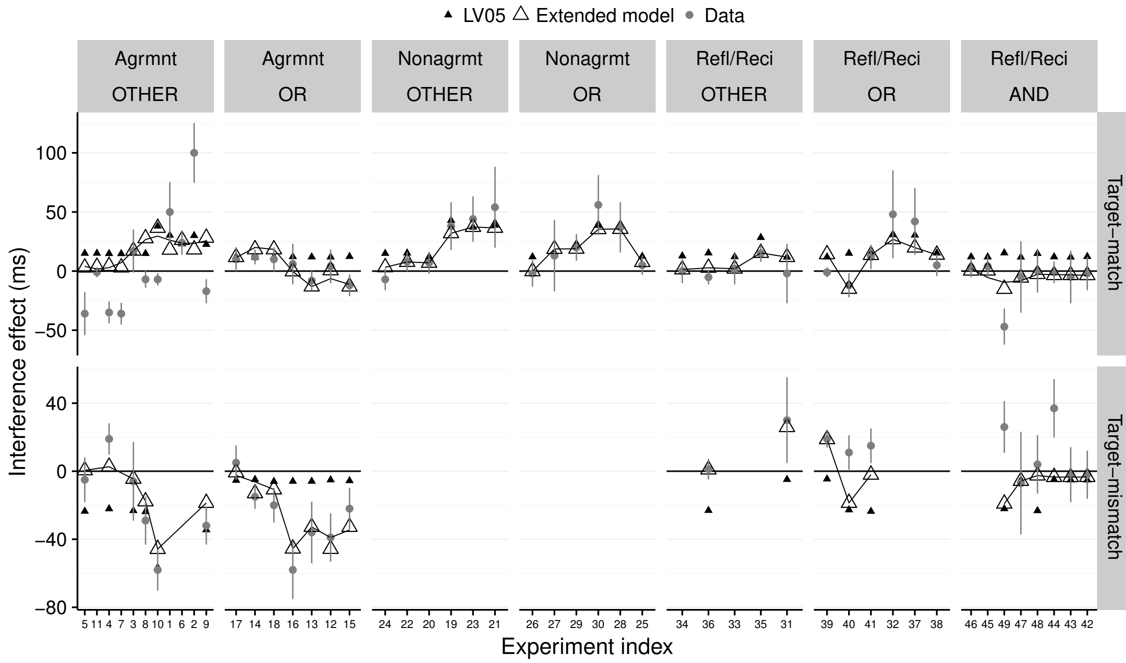


Figure 6. Data (circles) and predictions of the models LV05+decay (closed triangles) and Extended+decay (open triangles), arranged by increasing distractor prominence parameter p_{dstr} within each prominence level (*OTHER*, *subject OR topic*, *subject AND topic*), dependency type (subject-verb agreement, non-agreement, reciprocals/reflexives), and target-type (match or mismatch). A thin smoothed line highlights the relation between prominence and the interference effect in the model. The numbers on the x-axis refer to the experiment index in Table 5 in the Appendix.

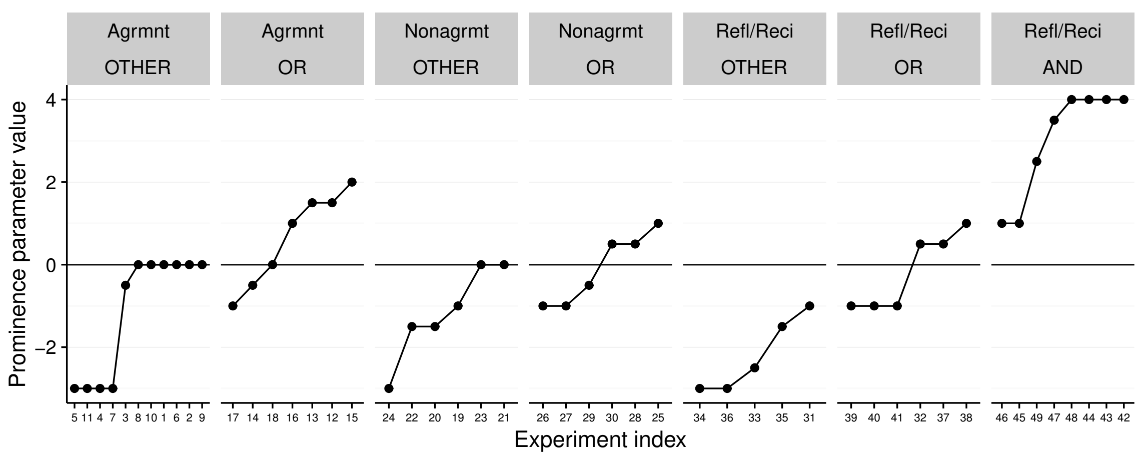


Figure 7. Estimated distractor prominence parameter values p_{dstr} for all simulations with the extended model (with decay), arranged by increasing parameter value within prominence level and dependency type. The numbers on the x-axis refer to the experiment index in Table 5 in the Appendix.

prominence parameter values. The values of the distractor prominence parameter p_{dstr} are plotted in Figure 7. In target-match configurations, most improvements of the extended model over LV05 in the fit with the data are due to *qualitative* rather than *quantitative* differences in the predictions due to the estimation of the prominence parameter p_{dstr} instead of the latency factor F. For example, the extended model, but not LV05, predicts near zero effects or facilitatory effects in target-match configurations. In target-mismatch configurations, the replacement of F by p_{dstr} leads to some quantitative improvements in the extended model. In addition, cue confusion predicts qualitative changes for two studies in target-mismatch configurations (31 and 39).

Reflexive-/reciprocal-antecedent dependencies. The biggest improvements in comparison with the original model were achieved for reflexive and reciprocal dependencies. As can be seen in Figure 6, the predictions according to the three prominence levels *OTHER*, *OR*, and *AND* (especially in target-match configurations) fit the data quite well, including some *qualitative* exceptions such as facilitatory instead of inhibitory effects in target-match configurations and inhibitory instead of facilitatory effects in target-mismatch configurations. The pattern of effects in target-match configurations of reflexives/reciprocals across prominence levels in Figure 6 (three rightmost panels) resembles the curve in Figure 4B that describes target-match effects as a function of increasing distractor prominence: It shows an at first increasing, then decreasing inhibitory effect that becomes facilitatory in the high-prominence range and finally returns towards zero. As is shown in Figure 7, the estimated distractor prominence values for reflexive/reciprocal dependencies are quite evenly distributed along the three prominence levels.

Non-agreement subject-verb dependencies. Also for non-agreement subject-verb dependencies (third and fourth panel in Figure 7), the estimated distractor prominence values are evenly distributed. These data could also be fit well within the two prominence value ranges *OTHER* and *OR* (third and fourth panel from the left in Figure 6). However, because the data for non-agreement subject-verb dependencies only contain target-match configurations and the results — mainly inhibitory interference — are compatible with LV05, there was no improvement in the fit with the data for the extended model over the LV05 model (except for a slight improvement in the extended+decay model).

Subject-verb agreement dependencies. The target-match and target-mismatch data of agreement dependencies at a medium prominence level (*subject OR topic*) were fit remarkably well. The improvements over the LV05 model are particularly obvious in target-mismatch configurations. With a varying latency factor in LV05, the prediction of a small target-match effect also restricts the target-mismatch effect to a small magnitude, whereas the relation between target-match and target-mismatch effects are more complex in the extended model, where distractor prominence rather than the latency factor varied by experiment. The fit for subject-verb agreement data is, however, downgraded by the target-match results for low prominence (*OTHER*), where six out of eleven data points show facilitatory interference. Facilitatory target-match effects cannot be predicted by the model for low prominence. This is the reason why most estimated values of p_{dstr} for low-prominence agreement data are at the outer boundaries -3 and 0 (leftmost panel of Figure 7).

Figure 8 reports simulation results for five exemplary cases where the extended model fit the data *qualitatively* better than LV05. Reciprocals in Kush and Phillips (2014) and the

Chinese reflexive in Jäger et al. (2015) are two cases where the extended model accounts for the *inhibitory* effect in target-mismatch configurations due to cue confusion, whereas facilitation is predicted by LV05. The model parameter for the cue confusion level C was estimated for reciprocals (Kush & Phillips, 2014) as 1 and for *ziji* (Jäger et al., 2015) as 0.7.

Lago et al. (2015), Cunnings and Felser (2013), and Sturt (2003) are examples of *facilitatory* effects in target-match configurations, which only the extended model accounts for due to high distractor prominence values. The inhibitory target-mismatch effect in Cunnings and Felser (2013) is, however, not predicted by either model, because no increased cue confusion is assumed in English reflexives. Notably, even if the cue confusion level was assumed to be elevated in this case, it would be impossible to predict an inhibitory target-mismatch effect and a facilitatory target-match effect at the same time. We return to this point in the General Discussion.

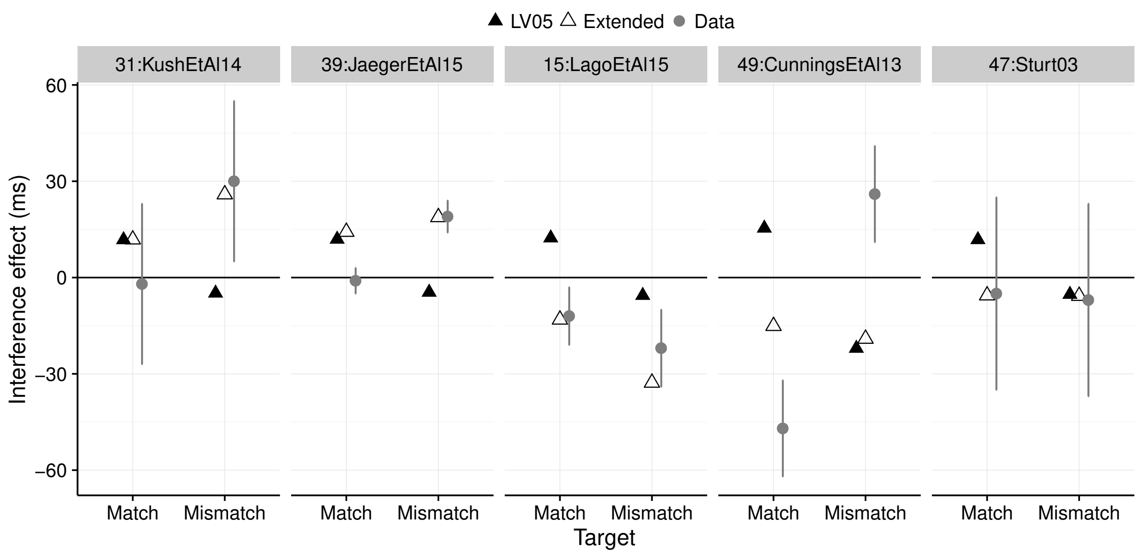


Figure 8. Data (circles) and predictions of LV05+decay (closed triangles) and the Extended+decay (open triangles) for interference effects of Kush and Phillips (2014), Jäger et al. (2015, Exp. 1), Lago et al. (2015, Exp. 3b), Cunnings and Felser (2013, Exp. 2, low working memory capacity), and Sturt (2003, Exp. 1). The numbers in the panel titles refer to the experiment index in Table 5 in the Appendix and Figures 6 and 7.

Discussion

The results show that the predictions of the LV05 cue-based retrieval model can be improved by extending it with prominence and associative cues while restricting parameters within bounds motivated by experimental design. In addition, by estimating distractor prominence instead of the latency factor in the extended model, the estimated parameter values are interpretable with respect to the strength of the distractor's representation in memory, whereas the latency factor in the original model does not have a direct cognitive interpretation.

Interestingly, the model's performance differed systematically between dependency types. The data on non-agreement subject-verb dependencies concurs overall with the general LV05 predictions — inhibition in target-match configurations — and was thus fit well by both models. The picture is, however, incomplete since there is no data on target-mismatch configurations for this dependency type. Agreement studies show a lot of variation in the effects, which is in most cases not predicted by distractor prominence or cue confusion. In order to test the predictions of prominence, more experiments that manipulate distractor position would be necessary. In particular, very prominent distractors that are subject and topic were not tested in this dependency. The available data on reflexive/reciprocal-antecedent dependencies involves all three prominence levels and special cue-feature co-occurrence contexts, i.e., reciprocals and Chinese reflexives. In these dependencies, the predictions of prominence and associative cues agree well with the data and thus improved the model fit.

Taken together, it is not clear whether the reasons for the differences between dependency types are of methodological or cognitive nature. On the one hand, between dependency types, the studies differed in the amount of available data, the testing of different prominence levels, and potentially involved confounds. For instance, a confound that may affect a number of subject-verb agreement studies was pointed out by Wagers et al. (2009): As a singular distractor is both shorter and morphologically less complex than a plural-marked distractor, it might cause shorter reading times that spill-over to successive regions. This is particularly problematic in studies that used singular verbs and had the verb directly follow the distractor. In target-match configurations with a singular verb, the distractor is singular in the high-interference condition where it matches the number feature of the target, whereas it is plural in the low-interference condition. Thus, in the high-interference condition, the inhibitory effect of similarity-based interference and the facilitatory spill-over effect of a singular distractor could potentially cancel each other out.

On the other hand, the processing of different dependency types might involve different cognitive mechanisms that are not accounted for by the model. These cognitive processing differences can be rooted in the dependency context itself or in the cues that were manipulated in the respective studies or in both. For example, subject-verb agreement studies exclusively focused on number interference, whereas reflexive studies mostly investigated gender, and non-agreement subject-verb dependency studies again manipulated other semantic or syntactic cues. In addition, Dillon (2011) has proposed that, in contrast to subject-verb dependencies, a syntactically guided search mechanism is necessary for accessing a reflexive's antecedent because it relies on c-command, which is a structural relation rather than a feature. An additional mechanism could also be present in number agreement studies that involve complex noun phrases where the distractor is located inside a prepositional phrase attached to the target. Agreement attraction in complex noun-phrases has been explained by several authors through *feature percolation*, which causes the target's number feature to be overwritten by the distractor's (Bock & Eberhard, 1993; Nicol, Forster, & Veres, 1997; Pearlmutter et al., 1999; Staub, 2009). In target-match configurations with a singular verb, the distractor would be plural in the low-interference condition and singular in the high-interference condition. The low-interference condition would, thus, be slower because the target's number feature is overwritten and perceived as being plural, which does not fit the verb. Hence, feature percolation predicts the high-interference condition to be faster, i.e., facilitatory interference, whereas cue-based retrieval predicts inhibitory interference, because target and distractor share the same number feature in the high-interference condition, which causes similarity-based interference (see also Nicenboim, Engelmann, Suckow, & Vasishth, 2016). If both mechanisms, cue-based retrieval and feature percolation, exist as independent processes, they would in some cases cancel each other out.

Thus, target-match configurations in subject-verb agreement studies exhibit two potential issues — feature percolation and a complexity-based spill-over effect — that predict opposite effects. This could explain why the meta-analysis did not find conclusive results in target-match configurations of subject-verb agreement and why distractor prominence does not much improve the fit of the extended model compared to LV05 for subject-verb agreement studies.

A confound could also be the explanation for the generally insignificant or detrimental effect of decay in the model fit seen in Table 4. It is possible that interference type and prominence are confounded in the available data: Studies with prominent distractors more often used a proactive rather than a retroactive interference design (17 vs. 10), while studies with non-prominent distractors more often used a retroactive rather than a proactive design (15 vs. 7). Hence, the two factors prominence and interference type, which both influence the distractor activation in memory, might tend to cancel each other out due to experimental design. This could explain the finding of the meta-analysis that retroactively interfering distractors cause a lower inhibitory effect in target-match conditions: This effect might be caused by prominence and not interference type.

However, decay could also play a smaller role than usually assumed. Indeed, independent work in psychology argues that interference rather than decay is the more important construct (Berman, Jonides, & Lewis, 2009; Lewis & Badecker, 2010; Oberauer & Lewandowsky, 2013, 2014). Because interference type and distractor prominence are confounded in the literature, we cannot conclude whether decay has actually no impact or is only disguised by a counteracting effect of prominence. This question can only be addressed by running an experimental design that crosses pro- and retroactive distractor position with the prominence of the distractor.

General Discussion

The aim of this work was to show the quantitative constraints of the Lewis and Vasishth (2005) model and investigate the consequences of memory accessibility and context-dependent cue-feature associations in the light of the available evidence on interference effects in dependency resolution. We have presented an implemented account of prominence and associative cues as an extension to the cue-based retrieval model of LV05. Our simulations show that distractor prominence and cue confusion predict a range of data points that were previously outside the scope of the model. This suggests that the assumptions of the original LV05 model were not entirely correct: it is important to account for different aspects of memory accessibility, for individual study design, and context-based feature-selectivity in order to generate accurate predictions of a model of cue-based memory retrieval such as LV05. We therefore believe that these independently motivated extensions help to more precisely interpret individual empirical results as being evidence in favor of or against the model. The simulations presented here thus provide new insights into the cognitive mechanisms behind interference effects. Moreover, the model assumes the same cognitive mechanisms and the same parameter values for simulating target-match and target-mismatch configurations within a given experiment. This restricts the predictions of the model considerably; for example, the model cannot predict, for a given experiment, an *inhibitory* effect in target-mismatch as well as a *facilitatory* effect in target-match configurations, which was found in gaze durations of readers with low working memory capacity by Exp. 2 of Cunnings and Felser (2013) as shown in Figure 8. This is because a facilitatory target-match effect is caused by a high distractor activation that overrides the fan effect. Consequently, the fan effect must be eliminated in both the target-match *and* target-mismatch configurations in the presence of a highly prominent distractor even if we assumed a high cue confusion level. Hence, the effect pattern of Cunnings and Felser (2013) is impossible to predict under any parameter setting. If the model simulations had involved separate parameter fits for target-match and -mismatch within the same experiment, the model would have been able to predict this and other patterns that are implausible under the model's cognitive assumptions. Our simulations therefore restrict the model's prediction space. If any outcome were possible, the model would not be very useful.

It is important to point out that the extended model is not an ad-hoc model aiming at an improved fit with the available data but that the extensions are based on independently motivated assumptions. The comparison of the model predictions with the data thus not only tests the model but at the same time helps to interpret the available data. As the meta-analysis (Jäger et al., 2016) points out, low power and publication bias could be important factors that weaken the empirical base. In addition to the evaluation of the evidence through a Bayesian meta-analysis, an evaluation through simulation helps to carve out the solid regions of all the evidence available. The model does so because of three main features: (i) The presented model is built on independently motivated — and, in terms of ACT-R, domain-independently validated — assumptions about memory retrieval, prominence, and associative cues, which are sensitive to experimental design choices; (ii) the model predictions are restricted by interactions between variables such as prominence, recency, and cue confusion; and (iii) the parameters are fixed within a given experiment, thus ruling out certain patterns of target-match and target-mismatch effects. An important prediction of the model in this respect is that the previously unexplained observations of facilitation in target-match or inhibition in target-mismatch can be explained under certain conditions, but seeing both *in the same experiment* is impossible according to the model assumptions as explained above. Gaps in the prediction space like this one are an important part of a theory which makes falsification possible and which should be taken into account when discussing the implications of experimental results.

As we have discussed above, the conclusions to be drawn about prominence and cue confusion are preliminary because (i) the available data is sparse with respect to the levels of distractor prominence studied within dependency types and different levels of feature selectivity, (ii) there may be confounds between prominence and other factors, and (iii) there may be different cognitive processes involved in certain dependency types that the model does not account for. In the following, we further discuss the implications of our approach for distractor prominence and cue-feature associations and potential alternatives.

Distractor activation

In the model we have presented, the prominence of a distractor is a function of its syntactic position and discourse saliency. An alternative account of how distractor position could affect the magnitude of interference has been discussed in Van Dyke and McElree (2011). By way of a weighting mechanism, a mismatching syntactic feature would lower the consideration of a distractor as a retrieval candidate—or, with gating rather than weighting, even rule it out completely, irrespective of any matching semantic or pragmatic features. This account predicts that interference effects are very small or absent if a distractor does not match the syntactic requirement, e.g., of being a grammatical subject. The predictions of syntactic weighting are consistent with our prominence account and are also compatible with ACT-R and LV05. Because of its reduced activation, a distractor that mismatches the subject would have a very low probability of being retrieved instead of the target, and, thus, no facilitatory interference is expected in target-mismatch configurations. The fan effect in target-match configurations would not be directly affected, because the fan effect in ACT-R is a consequence only of the feature that is manipulated between two conditions: The difference in the target activation between the distractor-match and the distractor-mismatch conditions is the same no matter how many additional cues the distractor matches across conditions. However, an effect of syntactic match in target-match configurations would nevertheless be predicted on the basis of a generally lower activated target: Because the relation between activation and latency in ACT-R is a negative exponential function (cf. Equation 6), differences in activation have less impact on the retrieval speed for items with a higher activation than for items with a lower activation. In case distractor and target both match the subject cue, the fan effect reduces both in activation across conditions compared to the case when only the target matches the subject cue. In consequence, when the distractor matches the subject cue, the retrieval latency of the target is more affected by the fan effect of a feature manipulation, i.e., a greater inhibitory interference effect is predicted in target-match configurations.

Hence, the predictions of the syntactic weighting account regarding syntactic position are similar to the predictions of our prominence account: A distractor in subject position compared to object position increases the inhibitory interference effect in target-match configurations and the facilitatory effect in target-mismatch configurations. However, the predictions of syntactic weighting are only valid when it can be assumed that grammatical position is part of the retrieval cues. In contrast, the predictions of our prominence account are independent of cue combinatorics and the match quality of the distractor at retrieval. Instead, it rests on the assumption that items in subject position have a higher relevance for interpreting a sentence and are, thus, maintained more actively in memory (Brennan, 1995; Chafe, 1976; Grosz et al., 1995; Keenan & Comrie, 1977). In the same way, this account of prominence due to relevance can be extended to discourse saliency such as topic or other contributing factors that we have not considered here: For example, thematic role (Arnold, 2001), contrastive focus (Cowles, Walenski, & Kluender, 2007), first mention (Gernsbacher & Hargreaves, 1988), and animacy (Fukumura & van Gompel, 2011) are known to affect discourse saliency and might thus influence distractor prominence. The results of our Bayesian meta-analysis in Jäger et al. (2016) and the simulations in the current article both suggest that the interference effect is more affected by discourse saliency than grammatical position.

This is outside the scope of an account based on cue weighting. Furthermore, a facilitatory effect in target-match configurations as a consequence of distractor prominence cannot be explained in terms of cue combinatorics.

Cue associations

The principle of associative cues states that cues can be associated with multiple features to different degrees depending on experience with the linguistic context. Crossed cue-feature associations between two cues lead to cue confusion and predict inhibitory interference in target-mismatch condition for dependency environments with high feature-co-occurrence in comparison to environments with low feature-co-occurrence. This is based on the assumption that cue-feature associations are the result of associative learning through exposure to different dependency types and their grammatical antecedents. The learning process would be best described along the lines of the *naive discriminative learning* model developed by Baayen, Milin, Đurđević, Hendrix, and Marelli (2011). Their model is an implementation of the Rescorla & Wagner equations for classical conditioning based on the presence and absence of cues and outcomes and has been applied to a range of morphological effects in the context of language acquisition.

A possible way to test the cue confusion hypothesis for English in a controlled experiment would be to directly compare reflexives and reciprocals, manipulating the number cue in both. An example design we have also suggested in Jäger et al. (2015) is shown in Example 6.

(6) a. *Reflexive; distractor-match*
The *nurse* who cared for the *children* had pricked *themselves* ...

b. *Reflexive; distractor-mismatch*
The *nurse* who cared for the *child* had pricked *themselves* ...

c. *Reciprocal; distractor-match*
The *nurse* who cared for the *children* had pricked *each other* ...

d. *Reciprocal; distractor-mismatch*
The *nurse* who cared for the *child* had pricked *each other* ...

Under the cue confusion hypothesis, a reduced facilitatory effect or an inhibitory effect is predicted for the reciprocal *each other* compared to the reflexive *themselves*. In order to derive a finer-grained metric that predicts differences in cue confusion levels between different dependency environments, co-occurrence frequencies could be computed from a corpus in which sufficient dependency information is available.

Our theory of cue confusion predicts a higher confusion level for both reciprocals and the Chinese reflexive *ziji* compared to English reflexives. This could explain the result of Kush and Phillips (2014), who found inhibitory interference in target-mismatch conditions in Hindi reciprocals, as well as our finding of an inhibitory target-mismatch effect for *ziji* in Experiment 1 of Jäger et al. (2015). However, the specific account of cue confusion pursued here does not account for the finding that the overall interference effect in target-mismatch configurations studies of reflexive- and reciprocal-antecedent dependencies is inhibitory according to the meta-analysis in Jäger et al. (2016), shown here in Table 1. Importantly, this overall inhibitory effect was found even when excluding the Chinese reflexives study of Jäger et al. (2015), which had a large sample size. The extended model did not predict an overall inhibitory effect in target-mismatch configurations in the presented simulations, because we used a rather conservative account of cue confusion by setting the confusion level greater than zero only for specifically motivated cases. A less conservative simulation with a freely varying cue confusion parameter would, however, result in an overall increased confusion level for reflexives compared to subject-verb agreement dependencies (subject-verb agreement showed an overall facilitatory effect in target-mismatch configurations). In support for a theory of a higher cue confusion level in reflexive-antecedent than in subject-verb dependencies in general, one could argue that reflexive-antecedent dependencies have a rather restrictive set of cues that define the target, whereas subject-verb dependencies occur in a wide range of contexts in which various semantic cues in addition to grammatical ones might be used (cf. Van Dyke & McElree, 2006).

Under a theory of cue confusion, an interesting question is whether categorically distinguishing two cues requires cognitive effort. If so, one would expect an additional variation of the confusion level that depends on task demands and individual differences. There is evidence that the depth of linguistic processing is influenced by task-specification (Logačev & Vasishth, 2016; Swets, Desmet, Clifton, & Ferreira, 2008) and individual differences (von der Malsburg & Vasishth, 2013; Nicenboim, Logačev, Gattei, & Vasishth, 2016; Traxler, 2007), resulting in underspecification of sentence representations or "good-enough processing" (Ferreira, Ferraro, & Bailey, 2002). In the same way, cue confusion could be part of a dynamically adapted resource-preserving strategy. This assumption predicts elevated confusion levels for readers with less cognitive resources in order to compensate for slower processing. It also predicts increased cue confusion for experiments with little task demand, like easy comprehension questions, because the effort of a precise cue specification would not be necessary. There is one experiment on reflexives that controlled for participants' working memory capacity: Cunnings and Felser (2013) found in their Experiment 2 on English reflexives an inhibitory effect on the critical region in target-mismatch conditions only for low-capacity readers. The effect is only marginally significant (mean 22 ms, SE 26 ms) but would be in line with the assumption of an individual-level variation of cue confusion due to adaptive processes. Note, however, that, even if it was the case that low-capacity readers experience higher cue confusion, for reasons explained above, the current model could not predict an *inhibitory* target-mismatch effect at the same time as a *facilitatory* target-match effect as is the case in Cunnings & Felser, 2013. Since there is only one experiment testing low-capacity readers on target-mismatch configurations, a hypothesis of cue confusion being adaptive to individual capacity limits is currently speculative, and planned experiments are needed in order to test it.

Other factors besides cue confusion that affect the strength of cue associations have not been considered here. Most prominently, it has been claimed that syntactic cues are weighted more strongly than semantic cues (e.g., Nicol, 1988; Sturt, 2003; Van Dyke, 2007; Van Dyke & McElree, 2011). A stronger weighting for syntactic cues might actually be subsumed by co-occurrence, assuming that syntactic cues are more reliable (i.e., have a higher co-occurrence) in a certain construction than semantic cues.

Other associations may, however, go beyond pure co-occurrence. For example, an experiment conducted by Van Dyke and McElree (2006) showed interference effects based on similarities between nouns that tap into world knowledge, such as the property of being *fixable*. Some cues may be stronger than others based on their semantics and pragmatics: Carminati (2005) have proposed a hierarchy between features, such that *person* > *number* > *gender*. Additionally, in English, *number* is morphologically overt while *animacy* and *gender* are not. The effects of semantically, pragmatically, or morphologically motivated differences between retrieval cues remain to be investigated.

Conclusion

The extended model of cue-based retrieval provides, for the first time, quantitative predictions with respect to systematic variability in experimental design across studies. The presented model is therefore an important step forward in helping us interpret results in the context of previous findings and for formulating computationally informed predictions for future experiments.

The two principles of prominence and associative cues that constitute our extended model are compatible with the general ACT-R theory of cue-based retrieval as the essential mechanism underlying dependency resolution in sentence processing. Both principles are independently motivated and should be considered as domain-general mechanisms and as extensions to the current ACT-R architecture. Looking beyond ACT-R, future work should also investigate whether inhibitory interference in target-mismatch configurations can be explained in terms of other computational/mathematical models of memory, such as the well-known drift-diffusion model account of Ratcliff (1978). A further, very productive line of inquiry would be a systematic study of the quantitative predictions of other theories of dependency completion in language comprehension (e.g., Rizzi, 2004) relative to published data.

We strongly encourage further research in order to support or falsify our claims about distractor prominence and cue confusion in language processing and in general cognition. Researchers are invited to use the extended model presented here to conduct further simulations. In order to facilitate this, we provide an online application of the model at `https://engelmann.shinyapps.io/inter-act`.

Acknowledgements

For valuable comments and discussions, we are grateful to Colin Phillips, Hedderik van Rijn, Niels Taatgen, Patrick Sturt, Ian Cunnings, Brian Dillon, Dan Parker, Dave Kush, Sol Lago, Bruno Nicenboim, and Dario Paape. We also want to thank the audiences of AMLaP 2014, CUNY 2015, and ICCM 2015 for commenting on our work.

References

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*(4), 1036–60.

Anderson, J. R., & Lebiere, C. (1998). *Atomic components of thought*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Ariel, M. (1990). *Accessing noun-phrase antecedents*. London, UK: Routledge.

Arnold, J. E. (2001). The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse Processes*, *31*, 137–162.

Baayen, R. H., Milin, P., Đurđević, D. F., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review*, *118*(3), 438.

Badecker, W., & Straub, K. (2002). The processing role of structural constraints on the interpretation of pronouns and anaphors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(4), 748–769.

Berman, M. G., Jonides, J., & Lewis, R. L. (2009). In search of decay in verbal short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(2), 317.

Bock, K., & Eberhard, K. M. (1993). Meaning, sound and syntax in English number agreement. *Language and Cognitive Processes*, *8*(1), 57–99.

Brennan, S. E. (1995). Centering attention in discourse. *Language and Cognitive Processes*, *10*(2), 137–167.

Carminati, M. N. (2005). Processing reflexes of the feature hierarchy (person> number> gender) and implications for linguistic theory. *Lingua*, *115*(3), 259–285.

Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (Ed.), *Subject and topic* (pp. 25–56). New York: Academic Press.

Chang, W., Cheng, J., Allaire, J., Xie, Y., & McPherson, J. (2016). shiny: Web Application Framework for R [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=shiny (R package version 0.14)

Chen, Z., Jäger, L. A., & Vasishth, S. (2012). How structure-sensitive is the parser? Evidence from Mandarin Chinese. In B. Stolterfoht & S. Featherston (Eds.), *Empirical approaches to linguistic theory: Studies of meaning and structure* (pp. 43–62). Berlin, Germany: Mouton de Gruyter.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, The Netherlands: Foris.

Cowles, H. W., Walenski, M., & Kluender, R. (2007). Linguistic and cognitive prominence in anaphor resolution: topic, contrastive focus and pronouns. *Topoi*, *26*, 3–18.

Cunnings, I., & Felser, C. (2013). The role of working memory in the processing of reflexives. *Language and Cognitive Processes*, *28*(1-2), 188–219.

Cunnings, I., & Sturt, P. (2014). Coargumenthood and the processing of reflexives. *Journal of Memory and Language*, *75*, 117–139.

Dillon, B. W. (2011). *Structured access in sentence comprehension* (PhD thesis). University of Maryland, College Park, MD.

Dillon, B. W., Mishler, A., Sloggett, S., & Phillips, C. (2013). Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*, *69*, 85–103.

Du Bois, J. W. (2003). Argument structure: Grammar in use. In J. W. Du Bois, L. E. Kumpf, & W. J. Ashby (Eds.), *Preferred argument structure: Grammar as architecture for function* (Vol. 14, pp. 11–60). Amsterdam, The Netherlands: John Benjamins.

Engelmann, F. (2016). *Toward an integrated model of sentence processing in reading* (Unpublished doctoral dissertation). University of Potsdam, Potsdam, Germany.

Felser, C., Sato, M., & Bertenshaw, N. (2009). The on-line application of Binding Principle A in English as a second language. *Bilingualism: Language and Cognition*, *12*, 485–502.

Ferreira, F., Ferraro, V., & Bailey, K. G. D. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*, 11–15.

Franck, J., Colonna, S., & Rizzi, L. (2015). Task-dependency and structure-dependency in number interference effects in sentence comprehension. *Frontiers in Psychology*, *6*(349).

Fukumura, K., & van Gompel, R. P. G. (2011). The effect of animacy on the choice of referring expression. *Language and Cognitive Processes*, *26*(10), 1472–1504.

Gernsbacher, M. A., & Hargreaves, D. J. (1988). Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language*, *27*, 699–717.

Grosz, B. J., Weinstein, S., & Joshi, A. K. (1995). Centering: A framework for modeling the local

coherence of discourse. *Computational Linguistics*, *21*(2), 203–225.
Jäger, L. A., Engelmann, F., & Vasishth, S. (2015). Retrieval interference in reflexive processing: Experimental evidence from Mandarin, and computational modeling. *Frontiers in Psychology*, *6*(617).
Jäger, L. A., Engelmann, F., & Vasishth, S. (2016). Similarity-based interference in sentence comprehension: Literature review and Bayesian meta-analysis. *Journal of Memory and Language*. (Submitted following revision)
Keenan, E. L., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, *8*(1), 63–99.
King, J., Andrews, C., & Wagers, M. (2012). Do reflexives always find a grammatical antecedent for themselves? In *25th Annual CUNY Conference on Human Sentence Processing* (p. 67). New York, NY: The CUNY Graduate Center.
Kush, D. (2013). *Respecting relations: Memory access and antecedent retrieval in incremental sentence processing* (Unpublished doctoral dissertation). University of Maryland, College Park, MD.
Kush, D., & Phillips, C. (2014). Local anaphor licensing in an SOV language: Implications for retrieval strategies. *Frontiers in Psychology*, *5*(1252).
Lago, S., Shalom, D. E., Sigman, M., Lau, E. F., & Phillips, C. (2015). Agreement processes in Spanish comprehension. *Journal of Memory and Language*, *82*, 133–149.
Lewis, R. L., & Badecker, W. (2010). Short-term memory in sentence production and its adaptive control. In *The CUNY conference of human sentence processing.*
Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, *29*(3), 375–419.
Lewis, R. L., Vasishth, S., & Van Dyke, J. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, *10*(10), 447–454.
Logačev, P., & Vasishth, S. (2016). A multiple-channel model of task-dependent ambiguity resolution in sentence comprehension. *Cognitive Science*, *40*(2), 266–298.
von der Malsburg, T., & Vasishth, S. (2013). Scanpaths reveal syntactic underspecification and reanalysis strategies. *Language and Cognitive Processes*, *28*(10), 1545–1578.
McElree, B. (1993). The locus of lexical preference effects in sentence comprehension: A time-course analysis. *Journal of Memory and Language*, *32*, 536–571.
McElree, B. (2000). Sentence comprehension is mediated by content-addressable memory structures. *Journal of Psycholinguistic Research*, *29*(2), 111–123.
McElree, B. (2006). Accessing recent events. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 46, pp. 155–200). San Diego, CA: Elsevier.
McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language*, *48*, 67–91.
Nicenboim, B., Engelmann, F., Suckow, K., & Vasishth, S. (2016). *Number interferenence in German.* (Under revision)
Nicenboim, B., Logačev, P., Gattei, C., & Vasishth, S. (2016). When high-capacity readers slow down and low-capacity readers speed up: Working memory differences in unbounded dependencies. *Frontiers in Psychology*, *7*(280).
Nicol, J. (1988). *Coreference processing during sentence comprehension* (PhD thesis). Massachusetts Institute of Technology, Cambridge, MA.
Nicol, J., Forster, K. I., & Veres, C. (1997). Subject–verb agreement processes in comprehension. *Journal of Memory and Language*, *36*, 569–587.
Oberauer, K., & Lewandowsky, S. (2013). Evidence against decay in verbal working memory. *Journal of Experimental Psychology: General*, *142*(2), 380.
Oberauer, K., & Lewandowsky, S. (2014). Further evidence against decay in working memory. *Journal of Memory and Language*, *73*, 15 - 30.
Patil, U., Vasishth, S., & Lewis, R. L. (2016). Retrieval interference in syntactic processing: The case

of reflexive binding in English. *Frontiers in Psychology*, *7*(329). (Special Issue on Encoding and Navigating Linguistic Representations in Memory)

Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language*, *41*, 427–456.

R Core Team. (2016). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from https://www.R-project.org/

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning ii: Current research and theory* (Vol. 2, pp. 64–99). New York, NY: Appleton-Century-Crofts.

Rizzi, L. (2004). Locality and left periphery. *Structures and beyond: The cartography of syntactic structures*, *3*, 223–251.

Schneider, D. W., & Anderson, J. R. (2012). Modeling fan effects on the time course of associative recognition. *Cognitive Psychology*, *64*(3), 127–160.

Staub, A. (2009). On the interpretation of the number attraction effect: Response time evidence. *Journal of Memory and Language*, *60*(2), 308–327.

Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, *48*, 542–562.

Swets, B., Desmet, T., Clifton, C., & Ferreira, F. (2008). Underspecification of syntactic ambiguities: Evidence from self-paced reading. *Memory and Cognition*, *36*(1), 201–216.

Traxler, M. J. (2007). Working memory contributions to relative clause attachment processing: A hierarchical linear modeling analysis. *Memory and Cognition*, *35*(5), 1107–1121.

Tucker, M. A., Idrissi, A., & Almeida, D. (2015). Representing number in the real-time processing of agreement: Self-paced reading evidence from Arabic. *Frontiers in Psychology*, *6*(347).

Van Dyke, J. (2002). *Parsing as working memory retrieval: Interference, decay, and priming effects in long distance attachment* (Unpublished doctoral dissertation). University of Pittsburgh, PA.

Van Dyke, J. (2007). Interference effects from grammatically unavailable constituents during sentence processing. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *33*(2), 407–430.

Van Dyke, J., & Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language*, *49*, 285–316.

Van Dyke, J., & McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of Memory and Language*, *55*(2), 157–166.

Van Dyke, J., & McElree, B. (2011). Cue-dependent interference in comprehension. *Journal of Memory and Language*, *65*(3), 247–263.

Van Gompel, R. P., Pickering, M. J., & Traxler, M. J. (2001). Reanalysis in sentence processing: Evidence against current constraint-based and two-stage models. *Journal of Memory and Language*, *45*(2), 225–258.

Wagers, M., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, *61*, 206–237.

Watkins, O. C., & Watkins, M. J. (1975). Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning and Memory*, *104*(4), 442–452.

Table 5

List of experiments included in the simulations.

Dependency	Prominence	ID	Publication	Int. type	Lang.	Distr. pos.
S-V agreement	OTHER	1	Franck et al. (2015, E1, Compl)	pro	FR	obj
		2	Franck et al. (2015, E1, RC)	pro	FR	obj
		3	Dillon et al. (2013, E1)	retro	EN	obj
		4	Pearlmutter et al. (1999, E1)	retro	EN	PP
		5	Pearlmutter et al. (1999, E2)	retro	EN	PP
		6	Pearlmutter et al. (1999, E3, plur)	retro	EN	PP
		7	Pearlmutter et al. (1999, E3, sing)	retro	EN	PP
		8	Tucker et al. (2015)	retro	AR	obj
		9	Wagers et al. (2009, E4, PP)	retro	EN	PP
		10	Wagers et al. (2009, E5)	retro	EN	PP
		11	Wagers et al. (2009, E6)	retro	EN	PP
	OR	12	Lago et al. (2015, E1)	pro	SP	subj
		13	Lago et al. (2015, E2)	pro	EN	subj
		14	Lago et al. (2015, E3a)	pro	SP	subj
		15	Lago et al. (2015, E3b)	pro	SP	subj
		16	Wagers et al. (2009, E2)	pro	EN	subj
		17	Wagers et al. (2009, E3, RN, plur)	pro	EN	subj
		18	Wagers et al. (2009, E3, RN, sing)	pro	EN	subj
S-V non-agrmnt	OTHER	19	VanDyke et al. (2006)	pro	EN	3x memory
		20	VanDyke et al. (2011,E2b)	pro	EN	obj
		21	VanDyke (2007, E1, LoSyn)	retro	EN	PP
		22	VanDyke (2007, E3, LoSyn)	retro	EN	PP
		23	VanDyke (2007, E2, LoSyn)	retro	EN	PP
		24	VanDyke et al. (2011, E2b)	retro	EN	obj
	OR	25	VanDyke et al. (11E1bpro)	pro	EN	subj
		26	VanDyke et al. (11E1bretro)	pro	EN	subj
		27	VanDyke (2007, E1, LoSem)	retro	EN	PP, subj
		28	VanDyke (2007, E2, LoSem)	retro	EN	PP, subj
		29	VanDyke (2007, E3, LoSem)	retro	EN	PP, subj
		30	VanDyke et al. (2003, E4)	retro	EN	PP, subj
Reciprocals	OTHER	31	Kush et al. (2014)	retro	HI	prepobj
	OR	32	Badecker et al. (2002, E4)	pro	EN	subj
Reflexives	OTHER	33	Badecker et al. (2002, E5)	pro	EN	gen
		34	Badecker et al. (2002, E6)	pro	EN	prepobj
		35	Jäger et al. (2015, E2)	pro	CN	3x memory
		36	Dillon et al. (2013, E1)	retro	EN	obj
	OR	37	Badecker et al. (2002, E3)	pro	EN	subj
		38	Chen et al. (2012, local)	retro	CN	subj
		39	Jäger et al. (2015, E1)	retro	CN	subj
		40	Patil et al. (2016)	retro	EN	subj
		41	Sturt (2003, E2)	retro	EN	obj, topic
	AND	42	Cunnings et al. (2013, E1, HiWMC)	pro	EN	subj, topic
		43	Cunnings et al. (2013, E1, LoWMC)	pro	EN	subj, topic
		44	Cunnings et al. (2014, E1)	pro	EN	subj, topic
		45	Felser et al. (2009, inaccMism)	pro	EN	subj, topic
		46	Felser et al. (2009, noCcom)	pro	EN	subj, topic
		47	Sturt (2003, E1)	pro	EN	subj, topic
		48	Cunnings et al. (2013, E2, HiWMC)	retro	EN	subj, topic
		49	Cunnings et al. (2013, E2, LoWMC)	retro	EN	subj, topic

Note. The experiments are ordered by dependency type (subject-verb agreement, subject-verb non-agreement, reciprocals, and reflexives), prominence level (*OTHER*, *subject OR topic*, and *subject AND topic*), and interference type (proactive vs. retroactive). The experiments are further classified by language (AR = Arabic, CN = Mandarin Chinese, EN = English, FR = French, HI = Hindi, SP = Spanish) and by syntactic position of the distractor (subject, object, genitive attribute, prepositional phrase, sentence external memory load, discourse topic).