



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Agreement attraction in comprehension

Citation for published version:

Sturt, P & Kwon, N 2023, 'Agreement attraction in comprehension: Do active dependencies and distractor position play a role?', *Language, Cognition and Neuroscience*.
<https://doi.org/10.1080/23273798.2023.2269282>

Digital Object Identifier (DOI):

[10.1080/23273798.2023.2269282](https://doi.org/10.1080/23273798.2023.2269282)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

Language, Cognition and Neuroscience

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Agreement attraction in comprehension: do active dependencies and distractor position play a role?

Patrick Sturt^a and Nayoung Kwon^b

^aPPLS, University of Edinburgh, Edinburgh, UK; ^bEALL, University of Oregon, Eugene, OR, USA

ABSTRACT

Across four eye-tracking studies and one self-paced reading study, we test whether attraction in subject-verb agreement is affected by (a) the relative linear positions of target and distractor, and (b) the active dependency status of the distractor. We find an effect of relative position, with greater attraction in retro-active interference configurations, where the distractor is linearly closer to the critical verb (Subject...Distractor...V) than in pro-active interference where it is more distant (Distractor...Subject...V). However, within pro-active interference configurations, attraction was not affected by the active dependency status of the distractor: attraction effects were similarly small whether or not the distractor was waiting to complete an upcoming dependency at the critical verb, with Bayes Factor analyses showing evidence in favour of a null effect of active dependency status. We discuss these findings in terms of the decay of activation, and whether such decay is affected by maintenance of features in memory.

ARTICLE HISTORY

Received 23 September 2022
Accepted 3 October 2023

KEYWORDS

Comprehension; number agreement; attraction; reading

Introduction

Agreement attraction in comprehension occurs when the computation of a dependency is affected by items in memory that partially match features that are required by the agreement relation. For example, (1) is easier to process than (2), even though both of these sentences are ungrammatical (Wagers et al., 2009)¹:

- (1) *The musicians who the reviewer praise so highly will probably win a Grammy.
- (2) *The musician who the reviewer praise so highly will probably win a Grammy.

According to the *cue-based retrieval* model (Lewis & Vasishth, 2005), the reason for the relatively facilitated processing of (1) is the partial match of features between the plural distractor *the musicians* and the plural requirement of the verb *praise*, as the match facilitates the activation level of the distractor, resulting in it being mis-retrieved on a proportion of trials. Since the overall retrieval process ends when one item has been retrieved, this situation leads to relatively short retrieval times associated with the processing of the verb, when averaged across trials. In contrast, in (2), neither *the musician* nor *the reviewer* matches the plural feature required

by the verb, leading to a prediction of greater processing difficulty due to the low activation level of the mismatching agreement target *the reviewer*, and relatively low probability of misretrieval of the distractor, thus longer retrieval times. This phenomenon has been supported by a large number of studies that have found similar processing patterns for ungrammatical sentences that are analogous to (1) and (2) (e.g. Dillon et al., 2013; Jäger et al., 2020; Lago et al., 2015; Pearlmutter et al., 1999).

As should be clear from the above discussion, the cue-based retrieval model assumes an important role for features, such as grammatical number, and features are often assumed to correspond directly to retrieval cues. Features have been the focus of several previous studies on attraction in comprehension – such studies have asked, for example, how features are combined (Parker, 2019), or whether different types of features are used for different types of dependencies (Dillon et al., 2013). Another research question, which has arguably been less widely studied, is the role of the sentence structure in modulating attraction. This question has been more frequently considered in studies of language production. For example, Franck et al. (2006) argued that attraction effects are modulated by structural relations like c-command.

CONTACT Patrick Sturt  patrick.sturt@ed.ac.uk  University of Edinburgh, Edinburgh EH8 9JZ, UK
 Supplemental data for this article can be accessed online at <http://dx.doi.org/10.1080/23273798.2023.2269282>.

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

One of the relatively small number of comprehension studies that examines the role of sentence structure in attraction is reported by Franck et al. (2015). In their Experiment 2, Franck et al. (2015) examined sentences like (3,4) below, in a speeded grammaticality judgement experiment:

(3) **Relative clause:**

- a* Jérôme parle à la prisonnière que le gardien sortent.
Jérôme talks to the prisoner who the guard take out.
 b* Jérôme parle aux prisonnières que le gardien sortent.
Jérôme talks to the prisoners who the guard take out.

(4) **Complement clause:**

- a* Jérôme dit à la prisonnière que le gardien sortent.
Jérôme tells the prisoner that the guard go out.
 b* Jérôme dit aux prisonnières que le gardien sortent.
Jérôme tells the prisoners that the guard go out.

All experimental sentences in the experiment were ungrammatical, as the licit subject, *le gardien* (“the guard”) does not agree with its verb *sortent* (“take out”/“go out”) in the number feature. The conditions differed in whether the distractor (the indirect object of the main clause) “la/aux prisonnière(s)” (the prisoner (s)) also serves as the direct object within the subordinate clause (3) or not (4).

The authors found a greater degree of attraction, as measured by reaction times in the grammaticality judgement task, for the relative clauses (3) compared with the complement clauses (4). Specifically, for the relative clause examples (3) reaction times for the grammaticality judgements were longer in (b), where the distractor (*prisonnières*) was plural (and was thus compatible with the plural requirement of its local verb (*sortent*)), relative to (a), where the number of the distractor was not compatible with the verb’s requirements. However, for the complement clauses (4), this difference was not found, resulting in an interaction. This shows that attraction effects can be influenced by the structure of the containing sentence. The authors interpret this difference as evidence of cyclic movement. According to this interpretation, a trace, coindexed with the distractor, intervenes structurally (through c-command) between the verb and its subject in (3), while this is not the case in (4). However, another possible interpretation is in terms of activation – In the relative clause example, (3), the distractor is effectively an argument of the verb *sortent*, as it is coindexed with the object gap. In a cue-based retrieval model, this would presumably mean that the distractor would be re-activated at the verb at the same time as the subject, as two different retrieval processes would need to be launched. This reactivation

may increase the level of attraction. In the complement clause example (4), by contrast, there is no direct dependency between the critical verb (*sortent*) and the distractor (*la prisonnière*), and thus no reactivation would be triggered by the verb.

The study that we report in this paper will also examine how agreement attraction is affected by sentence structure, with the specific aim of investigating the role of *maintenance* of relevant information in working memory on attraction. In the cue-based retrieval model as presented by Lewis and Vasishth (2005), the degree of attraction is affected by relative activation of target and distractor phrases, and also activation levels are assumed to decay over time, relative to the point where a given phrase is first encountered, or reactivated, in the input string.² One straightforward prediction of this is that attraction effects should be greater when the attractor linearly intervenes between the target and the verb (*retro-active interference*) relative to when the target linearly intervenes between the attractor and the verb (*pro-active interference*). For example, consider (5) and (6), which show retro- and pro-active interference respectively:

(5) **Retro-active interference:**

*The nurse who *the widows* relied on definitely were reluctant to work long shifts.

(6) **Pro-active interference:**

**The widows* said that the nurse most definitely were reluctant to work long shifts.

In each of the examples above, *the nurse* is the target of the agreement dependency, while *the widows* is the distractor. In (5), the distractor is linearly closer to the verb *were* than the target, while this is reversed in (6). Other things being equal, the activation level of the distractor, and thus the degree of attraction, should therefore be higher in (5) than in (6). Note, incidentally, that (6) does not involve a direct dependency between the critical verb *were* and the distractor *the widows*, so no reactivation of the distractor is involved before the critical retrieval event, and thus, levels of attraction would be expected to be relatively weak, as in the complement clause examples examined by Franck et al. (2015). In addition, in a recent development of the cue-based retrieval account Engelmann et al. (2019) have claimed that baseline activation levels can be affected by information structure, with a boost to activation for items that are typically discourse topics, such as main clause subjects. If this is the case, then this would affect the distractor (*the widows*) in (6), as this phrase is a main clause subject. Therefore, the Engelmann et al. (2019) version of the cue-based retrieval model predicts greater levels of

attraction for these pro-active interference conditions relative to the original Lewis and Vasishth (2005) model. These predictions are consistent with previous results. For example, while Kwon and Sturt (2019) failed to find clear differences in the levels of attraction between sentences with retro- and pro-active interference, this lack of clear difference could have been because their target sentences differed in the structural role of a distractor. On the other hand, the authors found a stronger attraction effect when a distractor was a main clause subject than when it was a dative object, and this is consistent with the claims of Engelmann et al. (2019).

In the experiments reported below, given that attraction effects for pro-active interference (as in (6)) are expected to be relatively low (modulo possible effects of topicality, as we keep the distractor as the main clause subject), we ask specifically whether the active dependency status of a distractor has an effect on its activation level, as measured by the attraction effect.

Compare (6) (repeated below) with (7):

(6) **Non-active dependency:**

**The widows* said that the nurse most definitely were reluctant to work long shifts.

(7) **Active dependency:**

**The widows* who said that the nurse most definitely were reluctant to work long shifts had become quite annoyed.

Sentences (6) and (7) are both ungrammatical (due to the number mismatch between *were* and *the nurse*) and include a plural distractor *the widows* occurring in the main clause subject position. However, these two sentences differ, as in (6), the distractor *the widows* forms a dependency with the immediately following word *said*. Thereafter, the distractor does not participate in dependency formation at the point where *were* is processed, nor does it participate in any other dependencies throughout the sentence. Thus, we would expect the baseline activation of this phrase to be relatively low, due to decay, at the point where *were* is processed. In contrast, in (7), the distractor is modified by a relative clause, and later forms a dependency with the main auxiliary verb *had* (henceforth: *active dependency*). We assume that the use of the relative clause in (7) therefore leads to the requirement to maintain relevant features of the distractor in working memory, while the expectation of the main verb is active. If this maintenance requires the features of the distractor to remain activated, the rate of decay may be lower. This would lead to a prediction of greater attraction in (7) than in (6). Note, as discussed above, unlike in the relative clause examples tested by Franck et al. (2015) (see 3 above), the sentence

structure of our active dependency condition does not allow for the reactivation of the distractor at the critical verb *were*, because there is no direct dependency between these two elements. The crucial question is whether or not a dependency involving the distractor is waiting to be resolved. We should note that cue-based retrieval models, and other related accounts of content-addressable memory typically do not incorporate a storage component, and thus, the effect of maintenance on decay, if supported by the evidence, would need to be a new addition to such models. However, as we discuss below and in the general discussion, there is empirical evidence suggesting that maintaining unresolved dependencies has a measurable cost, resulting in slower processing times and a distinct ERP signature. In our view, it is reasonable to test whether such maintenance also affects attraction.

As we have seen in the discussion above, in the *cue-based retrieval* model, memory retrieval plays a crucial role in explaining attraction effects. However, *cue-based retrieval* is not the only theoretical account of attraction, and other theories instead highlight the importance of *representation distortion*, whereby attraction effects arise as a result of previously encoded items in memory becoming corrupted (see Nairne, 1990; Oberauer & Kliegl, 2006, for domain-general memory-based implementations of this idea). This can occur either through similarity-based encoding interference affecting target and distractor noun phrases that share features (e.g. Gordon et al., 2001; Jäger et al., 2015; Laurinavichyute et al., 2017; Smith et al., 2021; Villata & Franck, 2020; Villata et al., 2018), or through feature percolation, where the number feature of a distractor can distort or overwrite the analogous feature of a target (Eberhard et al., 2005; Yadav et al., 2023).

Above, we discussed potential effects of active dependencies solely in terms of the *cue-based retrieval* model. However, we believe that similar predictions could be applied to other models, including those based on representation distortion, given appropriate assumptions. The key point is that active dependencies should increase the size of attraction effects in any model as long as the following conditions hold: (a) an open active dependency increases the activation of a distractor with which it is associated; and (b) greater relative activation of the distractor leads to larger attraction effects, other things being equal. For example, Yadav et al. (2023) implement a number of models of attraction in a Bayesian framework, including a feature percolation model, which is derived from earlier proposals from the production literature (e.g. Bock & Eberhard, 1993). In such a model, the number feature of a plural distractor can overwrite the representation of the

target subject noun phrase, so that a target that is singular in the input becomes represented as plural. Attraction in ungrammatical sentences like (1) (repeated below) occurs because the plural feature on *the musicians* leads to the target *the reviewer* being represented as plural, therefore matching the features of the verb *praise*, resulting in the sentence being perceived as grammatical³:

- (1) *The musicians who the reviewer praise so highly will probably win a Grammy.

In contrast, no such facilitation is predicted for ((2), repeated below), because the distractor is singular, so that, even if feature percolation occurs, the target will still be represented as singular, therefore mismatching the required features of the verb:

- (2) *The musician who the reviewer praise so highly will probably win a Grammy.

It is possible to envisage a model where the probability of feature percolation is correlated with the activation of the distractor. If this is the case, and, crucially, if active dependencies also increase activation, as we speculated in the discussion above, then active dependencies would also be expected to be associated with larger attraction effects in such a percolation model, even if such a model does not rely on cue-based retrieval.⁴

Across the four eyetracking experiments and one self-paced reading experiment reported below, we will test for attraction effects in a number of different configurations, crucially including pro-active interference configurations where the distractor either participates in an active dependency (7) or does not (6). For the non-active dependency sentences (6), attraction effects are expected to be relatively small, due to the low levels of activation of the distractor in this pro-active interference configuration, for reasons explained above. The question then is whether the active dependency in (7) then leads to a larger attraction effect over this baseline.

To our knowledge, only a small number of previous studies have examined the effect of maintenance on number agreement, and we briefly discuss two of those studies below, before turning to describe our own studies.

Since the active dependency that we will test in the studies below is a subject-verb dependency, it is important to consider prior evidence for storage in such dependencies. A study by Ristic et al. (2022) showed such evidence for subject-verb dependencies. In their

study Experiment 2, the authors tested sentences like (8), in which an adverbial clause (e.g. *a professor finishes the class*) linearly intervenes (a–d) or does not intervene (e–h) in a subject-verb dependency. Thus, in the intervening cases, the subject-verb dependency is active during the processing of the adverbial clause, while in the non-intervening cases it is not.

(8) **Intervening conditions:**

- a. That student, as soon as a professor finishes the class, leaves the classroom.
- b. Those students, as soon as a professor finishes the class, leave the classroom.
- c. Those students, as soon as professors finish the class, leave the classroom.
- d. That student, as soon as professors finish the class, leaves the classroom.

Non-intervening conditions:

- e. I watched that student, and as soon as a professor finishes the class, she leaves the classroom.
- f. I watched those students, and as soon as a professor finishes the class, they leave the classroom.
- g. I watched those students, and as soon as professors finish the class, they leave the classroom.
- h. I watched that student, and as soon as professors finish the class, she leaves the classroom.

The authors found that first pass times and go-past times at *leaves* were longer in the intervening conditions than in the non-intervening conditions, suggesting that there is a processing cost for maintaining the active subject-verb dependency before the main clause verb is received in the input. As mentioned above, this is consistent with our assumption that the subject-verb dependency is actively maintained during the processing of intervening material, assuming that the elevated reading times are due to this cost. Also relevant to our studies is the fact that, as well as intervention, Ristic et al. (2022) also manipulated whether the main clause and embedded clause subjects matched in number (8a,c,e,g) or mismatched (8b,d,f,h). They did not find clear evidence for an interaction between number matching and intervention. This aspect of Ristic et al.'s results may be considered in the light of the research question of the present paper, which asks whether number attraction is greater when the relevant retrieval is launched during the maintenance of an active dependency (as in Ristic et al.'s intervening conditions) than when there is no such active dependency (as in the non-intervening conditions). The lack of such an interaction in Ristic et al.'s study may be seen as evidence against such an effect of active dependency

maintenance. However, note that in all of the conditions in (8), the dependency between *professor(s)* and *leave(s)* is grammatical, and, as we will discuss below, interference effects involving grammatical subject-verb agreement dependencies are quite variable across studies, in contrast to the more consistent and larger effects in ungrammatical sentences (Jäger et al., 2017, 2020; Yadav et al., 2023). Thus, any interference effect in this part of the sentence may have been difficult to detect. Therefore, the lack of an interaction in Ristic et al.'s (2022) study is not necessarily evidence against the idea that attraction is modulated by active dependency maintenance.

In a series of acceptability judgment and self-paced reading studies, Kim et al. (2020) investigated the role of both re-activation and maintenance in English number agreement attraction. In contrast to Ristic et al. (2022), these authors did include ungrammatical conditions.

In their Experiment 3b, Kim et al. (2020) presented participants with sentences like (9) in a self-paced reading task:

(9) a. **Active filler: grammatical (singular/plural distractor):**

Which mistake in the program/programs that will be disastrous for the company certainly is harmful for everyone involved?

b. **Active filler: ungrammatical (singular/plural distractor):**

*Which mistake in the program/programs that will be disastrous for the company certainly are harmful for everyone involved?

c. **Re-activated filler: grammatical (singular/plural distractor):**

Which mistake in the program/programs will be disastrous for the company and certainly is harmful for everyone involved?

d. **Re-activated filler: ungrammatical (singular/plural distractor):**

*Which mistake in the program/programs will be disastrous for the company and certainly are harmful for everyone involved?

In each of the conditions (9a–d), a dependency is formed between *which mistake* and the predicate *is/are harmful*. Thus, according to cue-based retrieval, a retrieval is launched at *is/are*, involving a retrieval cue for number. In the *active filler* conditions, the authors assume that the filler *which mistake in the program(s)* has been maintained in memory over the intervening words without participating in further dependencies, because the phrase requires a main verb, which does

not appear in the input until *is/are harmful*. In contrast, in the *re-activated filler* conditions, the filler is released from maintenance at *will*, allowing the subject-verb dependency to be formed, and is subsequently re-activated at *and*. The reason for the reactivation is that the *across-the-board constraint* requires *which mistake* to be coindexed with a gap in each of the two conjuncts. The word *and*, then provides bottom-up evidence for the second conjunct, thus leading to reactivation. Kim et al. (2020) also manipulated whether a distractor (*program/programs*) matched or mismatched the number retrieval cue, and whether the overall sentence was grammatical or ungrammatical. In the self-paced reading data, Kim et al. (2020) found the expected number attraction effect, i.e. a smaller ungrammaticality cost in conditions where the distractor matched the retrieval cue than where it mismatched, resulting in a two-way interaction of local noun with grammaticality. At the word position immediately following the critical verb (i.e. at the word *harmful* in 9a–d), this two-way interaction was significant only for the *active filler* conditions, but not for the *re-activated filler* conditions, resulting in a 3-way interaction. However, at the next word *for* in 9a–d), the two-way interaction was not modulated by a significant three-way interaction. Thus, the attraction effect appeared earlier in the active filler conditions than in the reactivated filler conditions. This difference in timing of the attraction effect might be due to differences in the memory representation of the filler phrase *which mistake in the program(s)*. Since the sentence structure of the active filler requires this phrase to be maintained over a long span of the sentence, this could increase the availability of the contents of the phrase, crucially including the distractor *the program(s)*, leading to larger (or earlier) attraction effects. In contrast, for the reactivated filler conditions, the filler phrase has been released from maintenance in the first conjunct, allowing decay of its contents in memory before being reactivated at *and*, thus leading to less (or later) attraction effects.

The goals of the experiments that we report below are similar in some ways to those of Kim et al. (2020). However, there are some major differences. While Kim et al. (2020) concentrated on retro-active interference, we examine both retro- and pro-active interference. A second difference is that our active-dependency manipulation is focussed on the distractor, while keeping the active status of the retrieval target constant. Specifically, in (7) (but not in 6) above, we assume that the distractor (*the widows*) is maintained in memory, and that maintenance is active at the point where *were* is processed. At the same time, the retrieval target (*the nurse*) does not differ in active status

between (6) and (7). In contrast, in the study of Kim et al. (2020), the retrieval target (*which mistake*) and the distractor *the programs* form part of the same phrase, and this phrase either participates in an active dependency (9a,b) or is reactivated (9c,d). This difference means that our manipulation can be expected to affect the activation of the distractor specifically, while that of Kim et al. (2020) may have affected the activation of both the distractor and the retrieval target, both in terms of maintenance and reactivation. This difference simplifies the interpretation of any effects of active dependency status in our experiments, relative to those of Kim et al. (2020). Arguably, it also provides a clearer test of the modulating effect of maintenance on attraction, particularly as attraction in the cue-based retrieval model is linked to the activation of the distractor, as a factor in its misretrieval.

In the four eye-tracking experiments and one self-paced reading experiment reported below, we test the prediction that attraction in subject-verb number agreement is modulated by active dependencies. We also compare attraction effects between pro-active and retro-active interference. In Experiment 1, we obtain a baseline estimate for attraction effects in retro-active interference using sentences similar to (5). In Experiments 2 and 3, we test pro-active interference using non-active dependencies (as in 6) and active dependencies (as in 7) respectively. In Experiment 4, we replicate the attraction effect for retro-active interference, with a small change from the conditions of Experiment 1, namely, having the distractor in object position instead of subject position. The stimuli for all four experiments were very similar, with identical critical regions, to allow for cross-experiment comparisons. To test the effect of active dependency status, we compare the attraction effect in Experiment 2 with that of Experiment 3. To test the effect of retro-active vs. pro-active interference, we compare Experiments 1 and 4 on the one hand, with Experiments 2 and 3 on the other. Finally, in Experiment 5, we report a self-paced reading experiment that combines the conditions of Experiments 2 and 3 in a single within-participant design, in order to provide a further test of the effect of active dependencies on pro-active interference.

Experiment 1

Participants

Thirty-nine native speakers of English from the University of Edinburgh community were paid to participate in the experiment.

Stimuli

Forty-eight stimuli were constructed on the model of (10a–c):

- (10) a. **Ungrammatical; Matching distractor:**
*The nurse who the widows relied on definitely were reluctant to work long shifts.
- b. **Ungrammatical; Mismatching distractor:**
*The nurse who the widow relied on definitely were reluctant to work long shifts.
- c. **Grammatical:**
The nurses who the widow relied on definitely were reluctant to work long shifts.

The stimuli always used the plural auxiliary verb *were*, and we manipulated grammaticality by changing the number of the head noun (*nurse(s)*), and manipulated attraction by changing the number of the distractor (*widow(s)*), leading to a design with one grammatical baseline condition (10c) and two ungrammatical conditions, with either a distractor that matched (10a) or mismatched (10b) the number of the verb.

We adapted the stimuli from those of Dillon et al. (2013), by making the following changes: (a) we adapted the content of some of the sentences to make them suitable for readers of British English (the original study used North American participants); (b) we used object relative clauses instead of subject relative clauses, and (c) we used only three conditions instead of Dillon et al.'s eight. Of the three changes, (b) and (c) were designed to maximise the chances of finding an attraction effect. Using object relative clauses means that the distractor (*widow(s)*) is a subject, and subjecthood presumably corresponds to a retrieval cue for verb-subject agreement. Using only three conditions with 48 items yields a greater number of items per condition (i.e. 16) than Dillon et al. (2013).

Note that our design is set up to examine attraction effects specifically in ungrammatical sentences, and the lack of a second grammatical condition precludes the possibility of looking for attraction effects among grammatical sentences. However, attraction effects among grammatical conditions are variable, with some studies showing a reading time cost after the verb when a distractor matched the number retrieval cue (Franck et al., 2015; Nicenboim et al., 2018), and other studies showing the opposite pattern, with a reading time cost in the context of a *mismatching* distractor (na Fariña et al., 2014), while still other studies have shown null differences among grammatical conditions (Dillon et al., 2013; Lago et al., 2015; Wagers et al., 2009). In contrast, attraction effects among

ungrammatical conditions appear to be larger and with a more consistent direction (Dillon et al., 2013; Jäger et al., 2017; Lago et al., 2015; Wagers et al., 2009) – see also Yadav et al. (2023) for further discussion. Moreover, we did not have any theoretical predictions that applied specifically to attraction in grammatical sentences. We therefore took the decision to use three conditions instead of four, as way to maximise the chance of finding a reliable attraction effect. We will return to discuss our decision to use ungrammatical sentences in the General Discussion.

Given that we are measuring attraction only in ungrammatical sentences, this effect is measured by comparing the reading times of the two ungrammatical conditions (see below for an explanation of this contrast, and its predicted outcome). This raises the question of whether the grammatical condition is in fact needed. In our view, number attraction in ungrammatical sentences equates to the reduction of the cost of ungrammatical agreement. Therefore, it seems desirable to show both that ungrammatical sentences have a cost, and also that this cost is reduced in the context of a distractor matching the retrieval cue. Given this, the grammatical condition is needed as a baseline to demonstrate the cost of ungrammatical sentences.

A further question is whether it would have been more desirable to have a fully factorial design, with two grammatical conditions, in which the distractor's match with the verb's retrieval cue is manipulated in a way that parallels the two ungrammatical conditions. In such a design, the evidence for attraction would be found in the interaction of grammaticality by distractor match – facilitation would be expected in the ungrammatical sentences where the distractor matched the verb's retrieval cue, relative to sentences where it did not, while this effect might be expected to be absent among the grammatical conditions (Wagers et al., 2009). Such a fully factorial design would undoubtedly provide a higher level of experimental control than is possible with the design that we employ in the current experiments, as the interaction would rule out alternative explanations of the attraction effect based solely on the processing of the distractor. For example, if reading times at the verb differ as a function of distractor matching in the ungrammatical conditions, then this could potentially be explained as a delayed effect of processing differences between singular and plural distractors, rather than as a genuine attraction effect. Note that, in a fully factorial design, if such an effect is found for the ungrammatical conditions but is absent, reduced or reversed in the grammatical conditions, then the resulting interaction must be explained by factors that are over and above simple processing differences between

singular and plural distractors. However, we believe that such a confound is unlikely given the current design. In all our experiments, the distractor is separated from the verb by at least one agreement-neutral word, making an influence of the distractor on reading times at or following the verb unlikely. Moreover, although it is possible that a plural distractor could cause local processing difficulty due to its conceptual complexity relative to a singular distractor (see discussion in Wagers et al., 2009), in our design, the verb is always plural (i.e. *were*), so the expected attraction effect would equate to a facilitation in the context of a plural distractor, not the inhibitory effect that would be predicted on the basis of the alternative explanation. Finally, as mentioned above, although the fully factorial design would provide more experimental control, this would likely require more observations to reach the same level of power, given the larger number of conditions.

Procedure

Eye movements were recorded using an SR Research EyeLink 1000 eye tracker, with a sampling rate of 1 KHz. The stimuli were displayed on a 19 inch Viewsonic monitor, at a viewing distance of 81 cm, using Times New Roman 18 pt, with black text on a white background. Stimuli were rotated in a Latin-square design, so that each participant was exposed to an equal number of trials for each condition, and no participant saw the same item in more than one condition. The 48 stimuli in each list were combined with 100 filler sentences, of which 36 were part of an unrelated experiment on parallelism effects (example item: *The footballer who Trevor admired, and Angela imitated, scored in the championship.*) Fillers and stimuli were combined so that no two items from the current experiment appeared adjacent to each other. Approximately 1/3 of all stimuli were followed by a Yes/No comprehension question, which the participants answered by pressing the appropriate button on a gamepad. The correct answer was “yes” on 50% of the trials. Each experimental and filler item fitted on one line. Mean comprehension accuracy was 92%.

At the start of each trial, a black square was displayed on the left of the screen. When a stable fixation was detected on this square, it was replaced by the text.

Data analysis

Fixation data were screened and corrected for vertical drift. Following this, each fixation that was shorter than 80 msec was combined with the previous (or next) fixation, provided that the two fixations were within 1

character of each other. We then removed any remaining fixations that were shorter than 80 msec or longer than 1200 msec.

The analysis regions are illustrated in (11) below:

- (11) The nurses who the widow relied on definitely/ were/ reluctant/ to work/ long shifts.

The Verb region consisted of the word *were*. The (Verb+1) region consisted of the next word after *were*. Finally, the Verb+2 region consisted of the next two words after the Verb+1 region. In all cases, the region includes the immediately preceding space. As the Verb region was a short function word, and was expected to be skipped on a high proportion of trials, an iterative procedure extended its left region boundary one character to the left in cases where no initial fixation was detected within the originally defined Verb region. If there was still no fixation on the region with the adjusted boundary, the region was extended by one further character, and this was done iteratively, until either a fixation was detected, or the region was extended a maximum of four characters to the left of the original region boundary, in which case the trial was not included in the analysis (see also Sturt, 2003; Sturt & Lombardo, 2005). This procedure only applied to the Verb region. The procedure was used in the calculation of means for the Verb region in Experiments 1-4, and in the analysis of the Verb region for Experiment 1.

We concentrated on two first-pass measures within these regions, namely *first-pass reading time* (the sum of fixation durations from the reader's first entry into the region from the left, until the first exit from the region, either to left or right), and *go-past times* (the sum of fixation durations from the reader's first entry into the region from the left, until the first rightward exit from the region). Note that first-pass time includes only fixations in the region, and does not include any fixations following either a leftward or rightward exit of the region. However, go-past time may include fixations to the left of the region, in cases where a first-pass regression is made. Neither of these measures includes trials where the region was skipped. Note that Dillon et al. (2013) reported attraction effects in *Total Time*, a measure consisting of the summed duration of all fixations on the region. We do not report Total Time, for several reasons. The first reason is that we wanted to be able to compare reading times on the target regions across all four experiments reported in this paper. The four experiments have identical content in the analysis regions, but differ in content towards the end of the sentence. In particular, Experiments 2 and 3, reported below, use sentences that are

longer by several words, relative to Experiments 1 and 4. In contrast to the first-pass measures that we use in this paper, Total Time is affected by regressions that are launched from beyond the analysis regions. Thus, if we had included Total Time as a measure, comparisons across experiments could have been confounded by regressions launched from the later part of the sentences. A second reason for excluding Total Time is that we wanted to avoid reporting a large number of measures and regions, as this can lead to problems associated with familywise error (von der Malsberg & Angele, 2017).

For data analysis, Linear Mixed effect models were computed using Bayesian inference. The models were implemented using the brms R package (Bürkner, 2017, 2018, 2021), which provides an interface to the Stan programming language (Carpenter et al., 2017). The models included the three-level factor of condition, using dummy coding, and treating the ungrammatical-mismatching distractor condition as the reference level. Thus, there were two fixed effect contrasts, which we will call *Grammaticality* (contrasting the ungrammatical-mismatching distractor condition with the grammatical condition) and *Attraction* (contrasting the ungrammatical-mismatching distractor condition with the ungrammatical-matching distractor condition). Both of these contrasts are expected to show negative coefficients: For the grammaticality contrast, the negative coefficient reflects the facilitation of a grammatical sentence relative to an ungrammatical sentence, and for the attraction contrast, the negative coefficient reflects the facilitation of a number-matching distractor relative to a number-mismatching distractor, among the two ungrammatical conditions. The models also included crossed random effects for participants and items, with random slopes for both participant and item for both fixed effect contrasts, and the full correlation matrix. The reading times were log-transformed in the analysis. The model used uninformative priors. The prior for the intercept was Normal(0, 10). Given the contrasts for our model, this implies that we are 68% certain that the mean reading time for the ungrammatical condition (i.e. the intercept) will be between a number very close to zero (i.e. $\exp(-10)$) and 22,026 milliseconds (i.e. $\exp(10)$). The prior for the fixed effect contrasts was set to Normal(0, 1). Here, if we assume a mean of 400 milliseconds for the baseline (ungrammatical mismatch) condition, this would equate to an intercept of around 6 on the log-millisecond scale. Given this intercept and the specified fixed effect prior, this implies that we are 68% certain that the mean for the comparison (either grammatical or ungrammatical match) condition will range between 148msec (i.e. $\exp(6-1)$) and 1097 (i.e. $\exp(6+1)$). The prior for the

random effect correlation matrix was set to $I_{kj}(2)$. This provides a bias against overestimating random effect correlations. Remaining priors used brms defaults.

It should be noted that Bayesian inference does not involve the calculation of p -values. In the analyses below, we report model coefficients (β), along with a term $P(\beta < 0)$, representing the probability that the coefficient in question is less than zero. The probability estimates were obtained by sampling coefficient estimates from the posterior distribution of the LME model, and calculating the proportion of these estimates that fell below zero. Thus, for a coefficient whose predicted sign is negative, evidence for the hypothesis is deemed to be strong to the extent that the value of $P(\beta < 0)$ approaches 1. Conversely, for a coefficient whose predicted sign is positive, evidence for the hypothesis will be strong to the extent that the value of $P(\beta < 0)$ approaches 0.

We also report 95% credible intervals for each coefficient. These indicate the range of the coefficient's possible values that cover 95% of the probability density, based on the posterior sample. Coefficient estimates were calculated using the mean of the posterior samples, which is the brms default.

Results

Stimuli, data and scripts for Experiments 1–4 are available at <https://osf.io/ahfrw/>.

For Experiment 1, we report analyses of all three regions for first-pass reading time and go-past time.

The means for each region are given for both eye-movement measures in Table 1, and results of the Bayesian linear mixed effect model are given in Table 2. Means for go-past time are displayed graphically in Figure 1.

The results for first-pass reading time showed the predicted negative coefficient for the grammaticality effect in the Verb and Verb+1 regions, but no strong evidence for a grammaticality effect in the Verb+2 region. There was no strong evidence for an attraction effect in First-pass reading time.

Table 1. Means and standard errors (aggregated by participants) for Experiment 1.

	First pass reading time		
	Ungram (mismatch)	Ungram (match)	Grammatical
Verb	267 (11)	267 (9)	250 (8)
Verb+1	310 (14)	303 (13)	296 (11)
Verb+2	328 (11)	326 (11)	330 (10)
	Go-Past Time		
	Ungram (mismatch)	Ungram (match)	Grammatical
Verb	394 (24)	384 (25)	314 (18)
Verb+1	545 (31)	454 (24)	418 (18)
Verb+2	505 (34)	413 (21)	421 (20)

In contrast, go-past time showed clear evidence for a grammaticality effect in all three regions, and for an attraction effect in the Verb+1 and Verb+2 regions, with $P(\beta < 0) > .99$ in both of these regions.

Since the present paper is specifically concerned with the attraction effect, it was decided to restrict analyses in the remaining experiments to the Verb+1 and Verb+2 regions, for go-past time only. This also serves to alleviate statistical concerns with using multiple regions and measures (von der Malsberg & Angele, 2017).

In summary, Experiment 1 examined sentences with retro-active interference. The distractor used in Experiment 1 was also the subject of its clause. Both of these features of the design were designed to maximise the chance of finding an attraction effect. The results showed an attraction effect in go-past time, on the Verb+1 and Verb+2 regions: ungrammatical sentences had shorter go-past times when the distractor matched the number cue of the verb than when it did not.

Experiments 2 and 3

In Experiments 2 and 3, we examine attraction effect in sentences with pro-active interference, such that the distractor appears before the retrieval target, unlike in Experiment 1. However, as in Experiment 1, the distractor was the subject of its clause in both experiments. Experiment 2 and 3 differ in terms of active dependency status. In Experiment 3, the distractor is involved in an active dependency at the point where retrieval takes place, but it is not in Experiment 2. By comparing the attraction effects in Experiment 2 and 3, we aim to test the effect of active dependency status.

Participants

Experiments 2 and 3 each used 39 native speakers of English, who were paid to participate in the study.

Stimuli

The 48 experimental stimuli for Experiment 1 were adapted to create 48 stimuli for Experiment 2 and 48 for Experiment 3, with designs summarised below:

(12) EXPERIMENT 2.

a. Ungrammatical; Matching distractor:

*The widows said that the nurse most definitely/ were/ reluctant/ to work/ long shifts.

b. Ungrammatical; Mismatching distractor:

*The widow said that the nurse most definitely/ were/ reluctant/ to work/ long shifts.

Table 2. Statistical analyses of first-pass and go-past reading time for Experiment 1.

		First pass reading time				
		Estimate	Error	Lower 95%	Upper 95%	$P(\beta < 0)$
Verb	Intercept	5.50	0.03	5.43	5.57	
	Attraction	0.00	0.03	-0.05	0.05	.57
	Grammaticality	-0.05	0.02	-0.10	-0.01	.99
Verb+1	Intercept	5.61	0.04	5.53	5.69	
	Attraction	-0.02	0.03	-0.08	0.03	.79
	Grammaticality	-0.03	0.03	-0.09	0.02	.90
Verb+2	Intercept	5.67	0.04	5.58	5.75	
	Attraction	0.00	0.03	-0.06	0.05	.54
	Grammaticality	0.00	0.03	-0.05	0.05	.53

		GO-PAST READING TIME				
		Estimate	Error	Lower 95%	Upper 95%	$P(\beta < 0)$
Verb	Intercept	5.72	0.05	5.63	5.81	
	Attraction	-0.04	0.04	-0.11	0.04	.84
	Grammaticality	-0.16	0.04	-0.23	-0.09	>.99
Verb+1	Intercept	5.99	0.05	5.90	6.08	
	Attraction	-0.12	0.04	-0.20	-0.03	>.99
	Grammaticality	-0.19	0.04	-0.27	-0.11	>.99
Verb+2	Intercept	5.93	0.05	5.84	6.02	
	Attraction	-0.11	0.04	-0.19	-0.03	>.99
	Grammaticality	-0.11	0.04	-0.18	-0.03	>.99

Notes: Model estimates, error, lower and upper 95% credible interval limits (in log-millisecond units) and probability that the coefficient is smaller than zero, based on posterior samples.

c. Grammatical:

The widow said that the nurses most definitely/ were/ reluctant/ to work/ long shifts.

definitely/ were/ reluctant/ to work/ long shifts had become quite annoyed.

c. Grammatical:

The widow who said that the nurses most definitely/ were/ reluctant/ to work/ long shifts had become quite annoyed.

(13) **EXPERIMENT 3.**

a. Ungrammatical; Matching distractor:

*The widows who said that the nurse most definitely/ were/ reluctant/ to work/ long shifts had become quite annoyed.

b. Ungrammatical; Mismatching distractor:

*The widow who said that the nurse most

All other things being equal, the Lewis and Vasishth (2005) model would predict a smaller interference effect for Experiments 2 and 3, where interference is pro-active, relative to Experiment 1, where interference

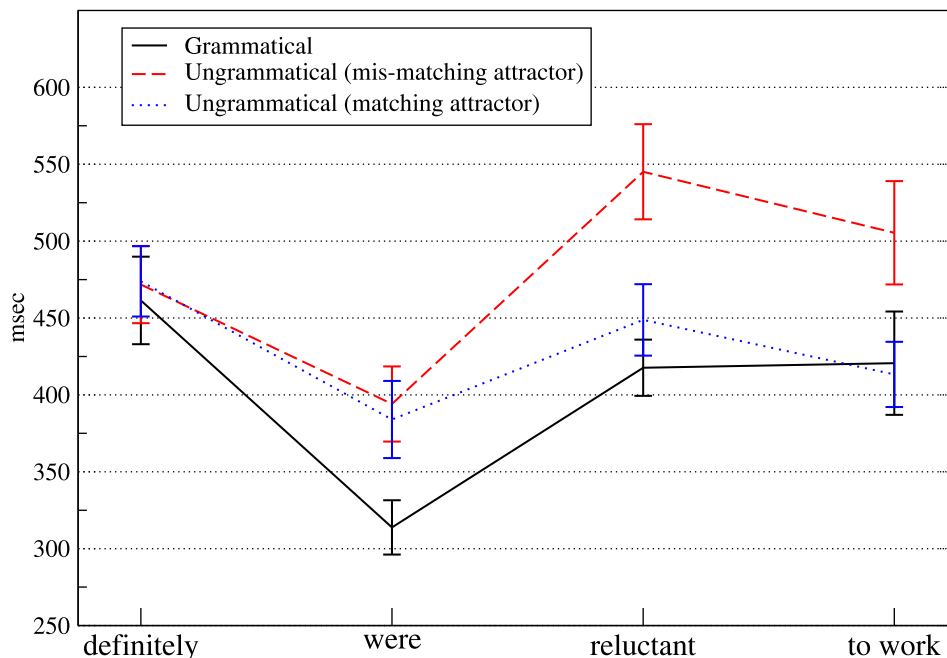


Figure 1. Mean go-past time for Experiment 1.

is retro-active. This is because, in the pro-active case, the activation of the distractor *the widow* will have decayed to a greater extent at the point where the retrieval occurs at *were*. Experiment 2 therefore serves as a baseline for Experiment 3, where, although the interference is pro-active, the distractor *the widow(s)* is involved in an active (i.e. unresolved) dependency at the point where retrieval takes place. If maintaining the distractor in memory for the purpose of later resolution involves a slower rate of decay, then we would expect increased interference in Experiment 3 relative to Experiment 2.

Procedure

The procedure was identical to that of Experiment 1, except that the screen was placed at a distance of 70 cm, instead of 81 cm, due to a change in the laboratory set-up. At the same time, the font size was reduced from 18 pt to 16 pt, maintaining a similar size of visual angle per character.⁵

Mean comprehension accuracy was 94% for Experiment 2 and 92% for Experiment 3.

Results

As mentioned above, for Experiment 2 and remaining experiments, we restricted our analyses on two regions of go-past time (namely Verb+1 and Verb+2). Based on the results of Experiment 1, this gives the best possibility of observing an attraction effect, without the statistical problems introduced by reporting multiple measures and regions (von der Malsberg & Angele, 2017).

Table 3 shows the means for go-past time for Experiment 2, and Table 4 shows the results of the statistical analysis for Experiment 2. Means are displayed graphically in Figure 2.

Table 3. Means and standard errors (aggregated by participants) for Experiment 2.

	Go-past time		
	Ungram (mismatch)	Ungram (match)	Grammatical
Verb	455 (39)	413 (28)	322 (16)
Verb+1	543 (38)	504 (36)	420 (29)
Verb+2	472 (25)	489 (33)	382 (27)

Table 4. Model estimates for go-past time, for Experiment 2.

		Estimate	Error	Lower 95%	Upper 95%	$P(\beta < 0)$
Verb+1	Intercept	5.96	0.06	5.84	6.07	
	Attraction	-0.06	0.05	-0.15	0.04	.87
	Grammaticality	-0.17	0.05	-0.26	-0.08	>.99
Verb+2	Intercept	5.86	0.06	5.76	5.97	
	Attraction	-0.02	0.05	-0.12	0.08	.68
	Grammaticality	-0.15	0.04	-0.23	-0.06	>.99

Experiment 2 again showed strong evidence for a grammaticality effect ($P(\beta < 0) > .99$ for both regions). However, there was no strong evidence for an attraction effect in this experiment, although the means are in the predicted direction ($P(\beta < 0) = .87$ for Verb+1, and $P(\beta < 0) = .68$ for Verb+2).

Experiment 3: results

Means for Experiment 3 are presented in Table 5, and results of statistical analysis in Table 6. Means are displayed graphically in Figure 3

As in Experiments 1 and 2, there was again a large grammaticality effect. However, as in Experiment 2, there was again no strong evidence of attraction ($P(\beta < 0) = .87$ for Verb+1, and $P(\beta < 0) = .53$ for Verb+2), although the means were again in the predicted direction for an attraction effect.

Combined analysis of Experiments 2 and 3

We conducted a combined analysis of Experiments 2 and 3, in order to quantify the degree to which the attraction effect depended on the active status of the dependency in which the distractor participated. This is measured by the attraction-by-experiment interaction, which should show a larger attraction effect in Experiment 3, where there was an active dependency, than in Experiment 2, where there was not. In this analysis, we combined data for the Verb+1 and Verb+2 regions into one analysis, including region as a factor (Cunnings & Sturt, 2014). We argue that this combined analysis simplifies the interpretation of results, due to having only one single coefficient for each of the effects of Attraction, Experiment, and the crucial Attraction-by-Experiment interaction, as opposed to the six separate coefficients that would be expected with the two separate analysis regions. The region factor was coded as a centred predictor, with a mean of zero and a range of 1. The experiment factor was also coded in the same way. Therefore, the effect of attraction is evaluated at the average across the two regions and across the two experiments, while the interaction of attraction by experiment or by region measures the degree to

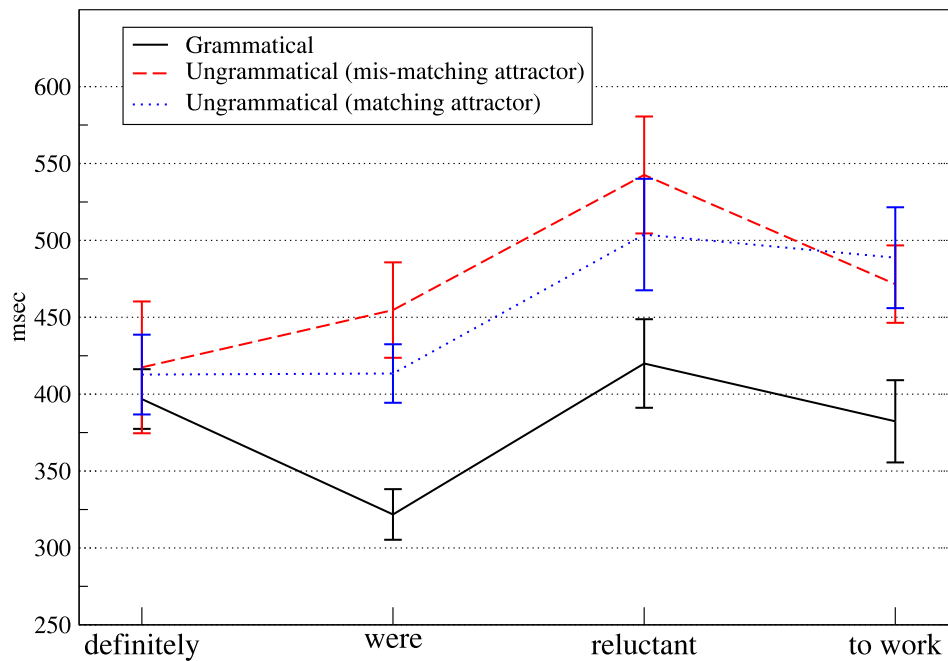


Figure 2. Mean go-past time for Experiment 2.

which attraction differs by experiment or by region. To account for the statistical dependency between the data points across the pair of regions in any given trial, this model also included a random intercept for trial. This random factor had a single level for each unique trial in the combined two experiments.⁶

The results of this analysis are given in Table 7:

As expected, the combined analysis shows clear evidence for the grammaticality effect ($P(\beta < 0) > .99$). However, there was no strong evidence for an Attraction effect, even though the coefficient was in the expected direction ($P(\beta < 0) = .94$). There was also no evidence for the predicted interaction of the attraction effect by experiment ($P(\beta < 0) = .4$), and indeed, the

coefficient for this effect is positive, whereas it would have been expected to be negative if Experiment 3 had shown more attraction than Experiment 2. Because of the theoretical importance of the Attraction-by-experiment interaction, we computed Bayes Factors to evaluate the evidence that this effect was null (i.e. BF_{01}). In addition, because of the unexpectedly small effect of Attraction, we decided to compute Bayes Factors for this effect as well, given that previous studies have not investigated this type of pro-active interference sentence structure, and it is theoretically informative to test which types of sentence structure lead to attraction and which do not.

The Bayes Factor analyses used the Savage-Dickey method, and in order to obtain stable Bayes Factor estimates, the models were run with a large number of iterations (12,000, of which 2000 were warmup). Also, because of the sensitivity of Bayes Factors to the choice of prior, the models were run with a range of priors for the beta coefficients: In addition to the $\text{normal}(0, 1)$ prior that we used to obtain the coefficient estimates, the Bayes Factor models

Table 5. Means and standard errors of go-past time (aggregated by participants) for Experiment 3.

	Go-past time		
	Ungram (mismatch)	Ungram (match)	Grammatical
Verb	365 (25)	350 (24)	322 (24)
Verb+1	596 (51)	560 (42)	442 (27)
Verb+2	463 (38)	483 (36)	394 (25)

Table 6. Model estimates for go-past time, for Experiment 3.

		Estimate	Error	Lower 95%	Upper 95%	$P(\beta < 0)$
Verb+1	Intercept	6.01	0.06	5.89	6.13	
	Attraction	-0.05	0.05	-0.14	0.03	.87
	Grammaticality	-0.19	0.05	-0.29	-0.10	>.99
Verb+2	Intercept	5.82	0.06	5.71	5.93	
	Attraction	0.00	0.04	-0.09	0.09	.53
	Grammaticality	-0.13	0.04	-0.21	-0.05	>.99

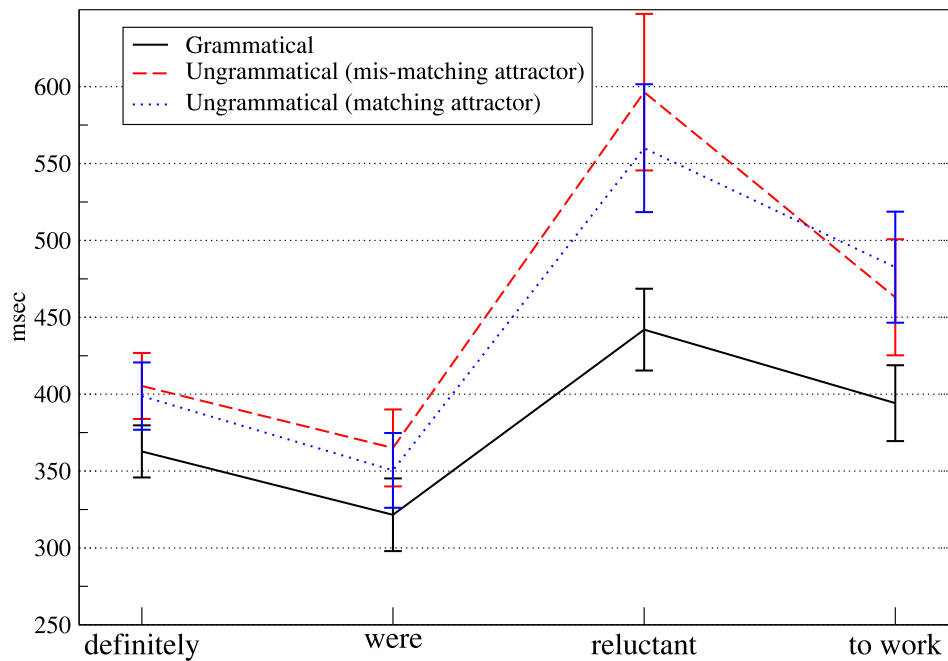


Figure 3. Mean go-past time for Experiment 3.

were also run with `normal(0,0.25)`; `normal(0,0.5)`; and `normal(0,2)`. Given an intercept of 6 on the log-millisecond scale, corresponding to an overall mean of around 400msec for the ungrammatical-mismatch conditions, these priors would imply that we are 68% confident that mean of the grammatical condition, or that of the ungrammatical match condition, would range between 312 and 514 msec for the `normal(0,0.25)` prior, between 243 and 659 msec for the `normal(0,0.5)` prior, and between 54 and 2956 msec for the `normal(0,2)` prior. The outcomes of these Bayes Factor analyses are reported in Table 8.

The Bayes Factors analyses suggest moderate to strong support for a null effect of both Attraction and the Attraction-by-experiment interaction, across the range of priors.

Table 7. Model estimates for combined analysis of Experiments 2 and 3.

	Estimate	Error	Lower 95%	Upper 95%	$P(\beta < 0)$
Intercept	5.91	0.04	5.84	5.98	
Attraction	-0.03	0.02	-0.08	0.01	.94
Grammaticality	-0.16	0.02	-0.20	-0.12	>.99
Region	-0.14	0.05	-0.24	-0.05	>.99
Experiment	0.01	0.06	-0.11	0.13	.42
Attr × Reg	0.04	0.05	-0.05	0.13	.18
Gram × Reg	0.04	0.04	-0.04	0.12	.17
Attr × Exp	0.01	0.05	-0.08	0.10	.40
Gram × Exp	-0.01	0.04	-0.09	0.08	.56
Reg × Exp	-0.10	0.08	-0.25	0.06	.88
Attr × Reg × Exp	0.02	0.09	-0.15	0.20	.44
Gram × Reg × Exp	0.04	0.08	-0.13	0.20	.29

Discussion of Experiments 2 and 3

Experiment 2 was a baseline experiment, designed to quantify the degree of attraction in a situation where interference is pro-active, and where the distractor does not participate in an active dependency at the point of retrieval. Experiment 3 was closely matched to Experiment 2 in that it involved pro-active interference, but it differed from Experiment 2 in that the distractor participated in an active dependency at the point where the crucial retrieval event occurred. We predicted that attraction effects would be larger in Experiment 3 than in Experiment 2, due to the active status of the dependency. However, the posterior distributions of the Bayesian analyses did not support this prediction, and Bayes Factor analyses instead supported the absence of such an effect, as well as the absence of an overall attraction effect.

Experiment 4

Experiment 1, which used retro-active interference, showed a robust attraction effect, while Experiments 2 and 3, which used pro-active interference, did not. We designed Experiment 4 as a partial replication of

Table 8. Bayes factors (BF_{01}) evaluating the evidence for a null effect of attraction, and of the attraction-by-experiment interaction, for Experiments 2 and 3, computed with four priors.

	<code>normal(0,0.25)</code>	<code>normal(0,0.5)</code>	<code>normal(0,1)</code>	<code>normal(0,2)</code>
Attraction	3.38	6.65	13.78	27.91
Attr × Exp	5.38	11.28	21.87	41.99

Experiment 1, with the aim of confirming that our method was sensitive to attraction, and that this effect could be reliably detected using retro-active interference. Experiment 4 was similar to Experiment 1 except that the distractor was an object in Experiment 4, while it had been a subject in Experiment 1. Although there have been demonstrations that subject-based attraction is stronger than object-based attraction for certain types of subject-verb dependencies (Van Dyke, 2007), our purpose here is not to compare these two types of attraction directly. Rather, our purpose is to confirm that the retro-active attraction effect that we found in Experiment 1 generalises to a different type of sentence.

Participants

Thirty-nine native speakers of English from the University of Edinburgh community participated in the experiment. The data for an additional 11 participants were collected, but were not included in the analysis, due to an error in the script, resulting in incorrect counterbalancing.⁷

Stimuli

There were forty-eight stimuli similar to 14:

(14) EXPERIMENT 4:

- a. **Ungrammatical; Matching distractor:**
*The nurse who cared for the widows definitely/ **were**/ reluctant/ to work/ long shifts.
- b. **Ungrammatical; Mismatching distractor:**
*The nurse who cared for the widow definitely/ **were**/ reluctant/ to work/ long shifts.
- c. **Grammatical:**
The nurses who cared for the widow definitely/ **were**/ reluctant/ to work/ long shifts.

Procedure

Procedure was identical to that of Experiments 2 and 3. Mean comprehension accuracy was 92%.

Results

Mean go-past times are presented in Table 9, and results of statistical analysis are presented in Table 10. Means are displayed graphically in Figure 4.

As in all the other experiments, Experiment 4 showed a robust grammaticality effect in both Verb+1 and Verb+2 regions ($P(\beta < 0) > .99$). As in Experiment 1, Experiment 4 showed strong evidence for an attraction effect in the Verb+1 region ($P(\beta < 0) > .99$). However, unlike Experiment 1, there was only weak

Table 9. Means and standard errors (aggregated by participants) for Experiment 4.

	Go-past time		
	Ungram (mismatch)	Ungram (match)	Grammatical
Verb	359 (25)	341 (21)	301 (13)
Verb+1	543 (33)	470 (27)	398 (24)
Verb+2	512 (43)	468 (26)	404 (23)

evidence for the attraction effect in the Verb+2 region ($P(\beta < 0) > .87$).

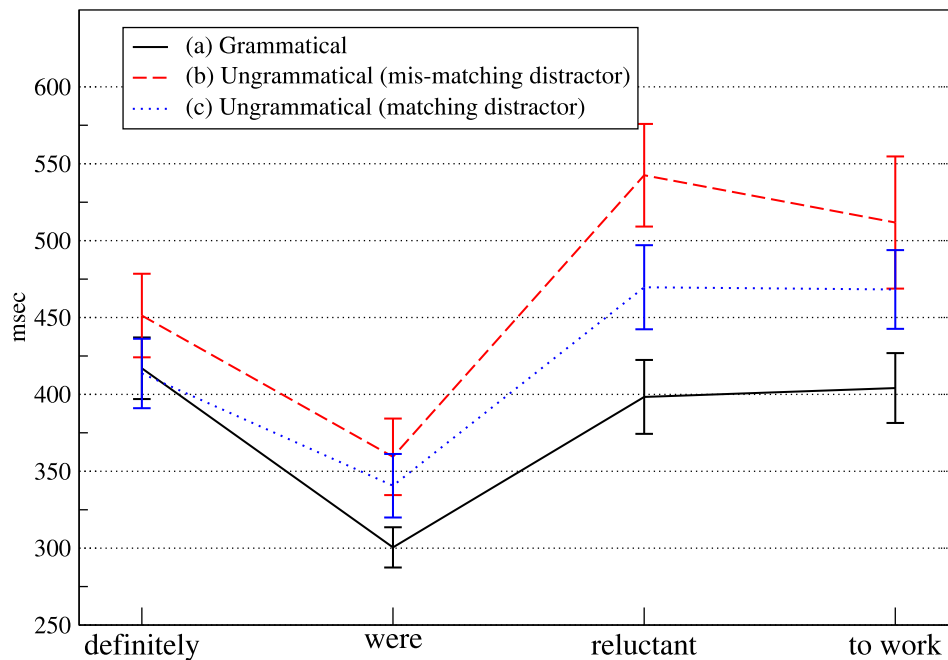
Experiment 4 confirmed that our method was able to detect an attraction effect for retro-active interference. This effectively replicates (Dillon et al., 2013), who also found evidence for interference effects where the distractor is an object. Although it was not our aim to compare Experiments 1 and 4 directly, we note incidentally that Experiment 1, which used a subject distractor, showed a large attraction effect in both the Verb+1 and Verb+2 regions, while Experiment 4, which used an object distractor, showed clear evidence for the effect only in the Verb+1 region. If this reflects a difference in the true effect, it may support subjecthood as a retrieval cue in this type of dependency, and this would fit with findings using other subject-verb dependencies (Van Dyke, 2007).

The relatively large attraction effect for our retro-active interference experiments (1 and 4) appears to contrast with the smaller interference effect for our pro-active interference experiments (2 and 3). We tested this difference by combining all four experiments into one large analysis, grouping the relevant pairs of experiments together to form a factor of interference type (Pro-active [Expts 2,3] vs. Retro-active [Expts 1,4]). This factor was treated as a centred predictor variable, with a mean of zero and a range of 1. As with the combined analysis of Experiments 2 and 3, we used both analysis regions, Verb+1 and Verb+2, including region as a centred factor. As well as allowing the comparison of pro-active and retro-active interference effects, the combined analysis also provides the opportunity to observe estimates of the grammaticality and attraction effects based on a large sample of 156 participants. The results of this analysis are summarised in Table 11:

The results of the analysis are consistent with a larger attraction effect for the retro-active interference experiments relative to the pro-active interference experiments (Attraction \times Pro/retro-interference: $P(\beta < 0) = .01$). Interestingly, there was no evidence for a modulation of the grammaticality effect as a function of interference type ($P(\beta < 0) = .39$). The analysis also confirmed overall main effects of both grammaticality and attraction (both $P(\beta < 0) > .99$)

Table 10. Model estimates for go-past reading time, for Experiment 4.

		Estimate	Error	Lower 95%	Upper 95%	$P(\beta < 0)$
Verb+1	Intercept	6.02	0.05	5.91	6.12	
	Attraction	-0.12	0.04	-0.20	-0.05	>.99
	Grammaticality	-0.24	0.04	-0.32	-0.15	>.99
Verb+2	Intercept	5.92	0.06	5.81	6.04	
	Attraction	-0.05	0.04	-0.14	0.04	.87
	Grammaticality	-0.14	0.05	-0.23	-0.05	>.99

**Figure 4.** Mean go-past time for Experiment 4.

Experiment 5

The results of Experiments 2 and 3 suggest that (a) attraction is weak or non-existent in the types of proactive interference sentences that we examined, and (b) that the degree of attraction in these sentences is not increased by an active dependency involving the distractor. However, one possible objection to this argument is that the power of the experiments may have

been too low, given that each experiment used only 39 participants. The power of the combined analysis of Experiments 2 and 3 may also have been compromised by the fact that one of the crucial factors, active-dependency status, was a between-participant variable, meaning that the error term for the interaction between active dependency and attraction would be inflated, due to the effect of individual differences in reading time, relative to a fully within-participant design where this variance would be partialled out.

Table 11. Results of combined analysis of Experiments 1–4.

	Estimate	Error	Lower 95%	Upper 95%	$P(\beta < 0)$
Intercept	5.94	0.03	5.88	5.99	
Attraction	-0.07	0.02	-0.10	-0.04	>.99
Grammaticality	-0.16	0.02	-0.19	-0.13	>.99
Region	-0.11	0.04	-0.18	-0.04	>.99
Pro/Retro	-0.05	0.04	-0.14	0.03	.89
Att × Reg	0.04	0.03	-0.02	0.10	.08
Gram × Reg	0.07	0.03	0.01	0.12	.01
Att × Pro/Retro	0.07	0.03	0.01	0.12	.01
Gram × Pro/Retro	0.01	0.03	-0.05	0.07	.39
Reg × Pro/Retro	-0.06	0.05	-0.17	0.04	.88
Attr × Reg × Pro/Retro	0.00	0.06	-0.11	0.11	.48
Gram × Reg × Pro/Retro	-0.05	0.06	-0.16	0.06	.8

Experiment 5, was a web-based self-paced reading experiment that was designed as a partial replication of Experiments 2 and 3, with an increased sample size (198 participants), and a fully within-participant design.

Sample size, data exclusion criteria, and analysis methods were pre-registered.⁸

Participants

The participants were 198 native speakers of English, recruited and paid via Prolific.co. One further participant was excluded due to low accuracy on comprehension

(below 70%), in accordance with the pre-registered analysis plan. Participants were paid GBP 6 for participation.

Before the experiment took place, a separate group of eighteen participants, again recruited and paid via Prolific, ran through the experiment. The purpose of this pilot study was to test the procedure, and to establish reasonable data exclusion criteria for the pre-registration. The analyses reported below do not include these eighteen pilot participants.

Stimuli

The 48 stimuli were adapted from Experiments 2 and 3 (see 15 for an example).

(15) NON-ACTIVE DEPENDENCY:

- a. **Ungrammatical; Matching distractor:**
*The widows/ said that/ the nurse/ most/ definitely/ were/ reluctant/ to work/ long shifts/ and seemed quite annoyed.
- b. **Ungrammatical; Mismatching distractor:**
*The widow/ said that/ the nurse/ most/ definitely/ were/ reluctant/ to work/ long shifts/ and seemed quite annoyed.
- c. **Grammatical:**
The widow/ said that/ the nurses/ most/ definitely/ were/ reluctant/ to work/ long shifts/ and seemed quite annoyed.

ACTIVE DEPENDENCY:

- d. **Ungrammatical; Matching distractor:**
*The widows/ who said that/ the nurse/ most/ definitely/ were/ reluctant/ to work/ long shifts/ had become quite annoyed.
- e. **Ungrammatical; Mismatching distractor:**
*The widow/ who said that/ the nurse/ most/ definitely/ were/ reluctant/ to work/ long shifts/ had become quite annoyed.
- f. **Grammatical:**
The widow/ who said that/ the nurses/ most/ definitely/ were/ reluctant/ to work/ long shifts/ had become quite annoyed.

The *non-active dependency conditions* were adapted from the stimuli of Experiment 2. An extra region was added to the end of each non-intervening item (e.g. “and seemed quite annoyed” in (15)), in order allow a rough match in length to the *active dependency conditions*. The *active dependency conditions* were adapted from the stimuli of Experiment 3. In some cases, changes were made to the final region, in order to allow a closer match of the content with the *non-active dependency conditions*.

Given that there were forty-eight items, and six within-participant conditions, each participant would be exposed to eight examples of each condition. This is a smaller number than the 16 examples per condition for Experiments 2 and 3. However, due to the increased number of participants, overall there were around 2.5 times more observations per condition in Experiment 5 than in there were in Experiments 2 and 3 (1584 vs. 624, before data exclusion).

Procedure

The experiment was implemented on PCIBex (Zehr & Schwarz, 2018), and participants took part on the web. Each trial began with a row of underscores, indicating the positions of the words of the sentence, preserving spaces between words. The participant pressed the spacebar on his/her computer to reveal each new segment of the sentence. The segments are indicated in (15), and include segments that match the analysis regions of the eye-tracking experiments. The self-paced reading function required the participant to release the spacebar between each button-press, to prevent participants from simply holding down the spacebar to reveal each segment in quick succession. Time between button presses was recorded in milliseconds.

The six conditions of the experimental sentences were distributed among six latin square groups, such that a participant in any given group was exposed to only one condition of each item, and an equal number of the six conditions. An equal number of participants was assigned to each group.

The experimental sentences were interleaved with 62 filler sentences, adapted from the fillers of Experiments 2 and 3. Around one third of stimuli were followed by comprehension questions, which were answered by selecting one of two possible answers using the 1 and 2 keys on the keyboard. Mean comprehension accuracy was 93%.

Results

Stimuli, data, scripts and pre-registration for Experiment 5 are available at <https://osf.io/g6tdn/>

Before analysis, we removed any data point for an individual segment that was greater than 3000 msec or less than 150 msec, affecting less than 0.5% of the data. This criterion was pre-registered.

Given the sentence segmentation (see (15) above), segment 6 (*were*) corresponds to the Verb analysis region of the eye-tracking experiments, segment 7 (*reluctant*) corresponds to the Verb+1 region, and

Table 12. Means and standard errors (aggregated by participants) for Experiment 5

		Self-paced reading time (msec)		
		Ungram (mismatch)	Ungram (match)	Grammatical
Verb	Non-Active	458 (9)	454 (9)	440 (7)
	Active	450 (8)	446 (8)	445 (8)
Verb+1	Non-Active	477 (10)	475 (11)	465 (12)
	Active	484 (12)	484 (12)	459 (9)
Verb+2	Non-Active	494 (11)	479 (9)	475 (10)
	Active	484 (10)	486 (10)	482 (10)

segment 8 (*to work*) corresponds to the Verb+2 region. Henceforth, we will use these region names, to facilitate comparison with the eye-tracking results. Table 12 shows the mean self-paced reading times for the Verb, Verb+1 and Verb+2 regions, and these are also displayed graphically in Figure 5.

As Experiment 5 was pre-registered, we will report the outcomes of pre-registered and exploratory analyses separately.

Pre-registered analyses

The main pre-registered analysis was analogous to the combined analysis of Experiments 2 and 3. A Bayesian linear mixed effects model was computed to analyse the reading times of the Verb+1 and Verb+2 regions, combined into one analysis, treating region as a factor. The fixed effects were Region (Verb+1 vs. Verb+2, centred), Active dependency status (Active vs. Non-Active, also centred), and condition (Ungrammatical-match vs. Ungrammatical-mismatch vs. Grammatical, treatment coded with the Ungrammatical-mismatch condition as the reference). The model also included all interactions among these factors. Full random effects were included for participant and item, with random intercepts as well as random slopes for all fixed effect interactions and main effects. As in the combined analysis of Experiments 2 and 3, a random intercept was also included for trial-id, whose value was unique to each trial. This was to account for the lack of independence between two regions of the same trial. Random correlation parameters were not included in the model.

The Bayesian model used a prior of $\text{normal}(0, 1)$ on the fixed effects, and $\text{normal}(0, 10)$ on the intercept, with brms default priors for all other parameters. The model was run with 6000 iterations, of which 1000 were warmup. The outcome of the analysis is given in Table 13.

The analysis did not show strong evidence of an attraction effect, although the direction of the effect was in the expected direction ($P(\beta < 0) = .86$). There was also no clear evidence for the crucial interaction

between attraction and active dependency, and in any case, the direction of the effect was opposite to what would have been predicted if active dependencies increased attraction relative to non-active dependencies ($P(\beta < 0) = .86$).

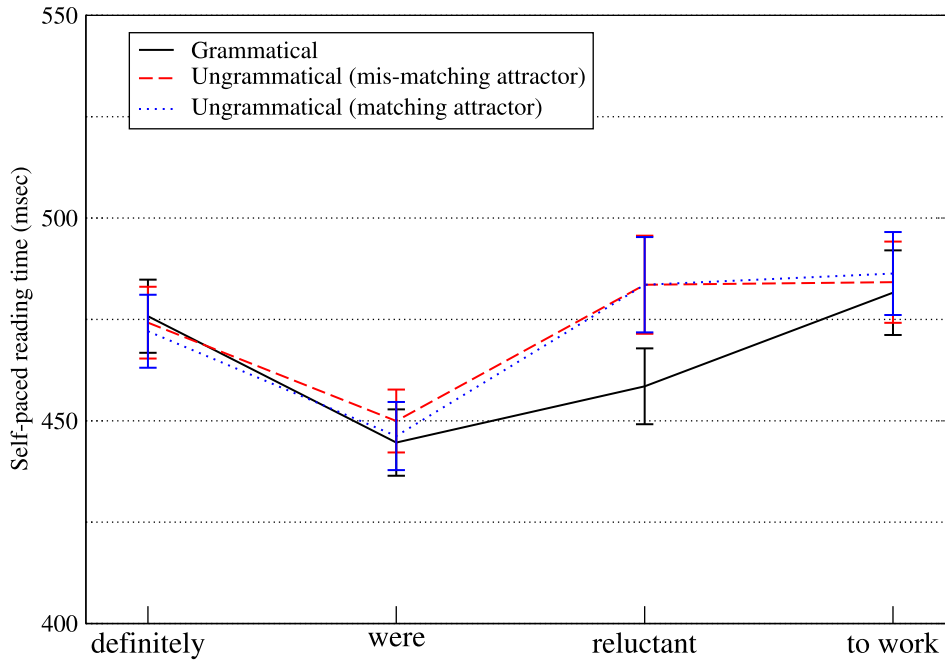
The pre-registration included computation of Bayes Factors for the attraction effect and for the attraction-by-active-dependency interaction, to test the null hypothesis that each of these effects were equal to zero, in comparison to the alternative hypothesis that the effect was not zero. These bayes factors yielded strong evidence in favour of the null hypothesis (Attraction: $\text{BF}_{01} = 96.22$; Attraction \times Active-dependency: $\text{BF}_{01} = 56.3$). Thus Bayes Factor results replicate those of the combined analysis of Experiments 2 and 3 in supporting a null effect of attraction, and of the effect of active dependencies on attraction.

Exploratory analyses

In addition to the pre-registered analysis reported above, we ran two exploratory analyses that were not part of the pre-registration plan. The first pre-registered analysis was run in order to counter a possible objection to the outcomes of our Bayes Factor analysis reported above. In the pre-registered analysis, we used a prior of $\text{normal}(0, 1)$, which, in the context of log-based self-paced reading times is quite a flat, unconstraining distribution. It could be the case that the Bayes Factor support for the null hypothesis that we found in the pre-registered analysis was amplified by using an unconstraining prior. The Bayes Factors that we previously reported in the combined analysis of Experiments 2 and 3 showed that BF_{01} values were reduced for more constraining priors centred on zero (i.e. priors with smaller standard deviation parameters), compared to the less constraining priors.

In the first exploratory analysis, we therefore repeated the Bayes Factor analyses as described above, but using a fixed effect prior of $\text{normal}(0, 2.5)$ instead of the pre-registered prior of $\text{normal}(0, 1)$. The prior of $\text{normal}(0, 2.5)$ corresponds to the most constrained prior that we ran in the combined analysis of Experiments 2 and 3, and in those analyses, it yielded the smallest value for BF_{01} among the four priors tested. If we continue to find strong evidence for the null hypothesis with this more constrained prior, it will therefore strengthen the support for the null effects of attraction and attraction-by-active-dependency. The analysis with the $\text{normal}(0, 2.5)$ prior continued to find strong evidence in favour of the null hypothesis, both for

ACTIVE DEPENDENCIES



NON-ACTIVE DEPENDENCIES

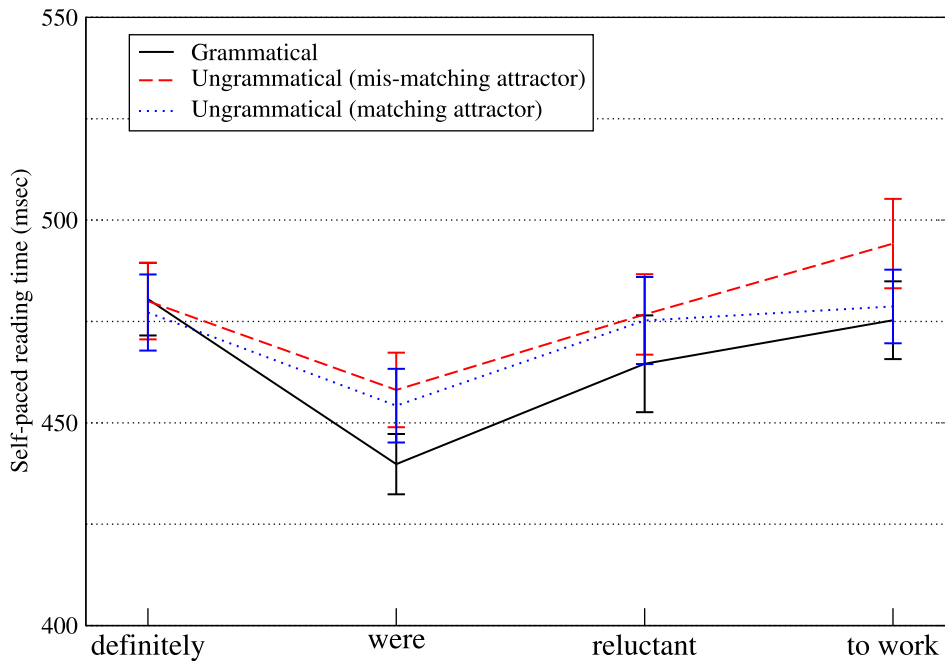


Figure 5. Mean self-paced reading time for Experiment 5 (Upper panel: active dependencies; Lower panel: non-active dependencies).

attraction ($BF_{01} = 23.77$), and for the interaction of attraction with active dependency status ($BF_{01} = 13.45$). Thus, the support for the null hypothesis that we found in the pre-registered analysis generalises to a more specific prior.

The second exploratory analysis was an analysis of the Verb region. We had originally chosen the combined analysis of Verb+1 and Verb+2 as the target of the pre-registration, on the basis of the results of the Eye-tracking experiments. However, it remains possible

Table 13. Results of the pre-registered analysis of Experiment 5.

	Estimate	Error	Lower 95%	Upper 95%	$P(\beta < 0)$
Intercept	6.11	0.02	6.07	6.15	
Attraction	-0.01	0.01	-0.02	0.01	.86
Grammaticality	-0.03	0.01	-0.04	-0.02	>.99
Region	0.02	0.01	0.01	0.04	.01
Active	0.00	0.01	-0.01	0.02	.79
Dependency					
Attr × Region	-0.00	0.01	-0.02	0.02	.59
Gram × Region	0.02	0.01	-0.00	0.04	.03
Attr × Active	-0.01	0.01	-0.03	0.01	.86
Gram × Active	-0.01	0.01	-0.03	0.01	.78
Region × Active	0.01	0.01	-0.01	0.03	.14
Attr × Reg × Act	-0.01	0.02	-0.05	0.02	.81
Gram × Reg × Act	-0.02	0.02	-0.06	0.01	.91

Notes: Note that, given the condition labelling in this experiment, the interaction between attraction and active dependencies is expected to be positive, if active dependencies increase the size of the attraction effect.

that evidence for a role of active dependencies would be found on the verb region, which we may have missed by concentrating on the pre-registered analysis. We report the outcome of the analysis of the Verb region below. The model that we used to analyse the Verb region included the crossed factors of active-dependency and condition, with the same contrast coding as in the pre-registered analysis. Since only one region was being analysed, the region factor was not included in the fixed effects, and the trial-id factor was not included in the random effects. The model included random effects for both participants and items, with random intercepts and random slopes corresponding to the full crossing of the fixed effects, and, like the pre-registered analysis, excluded random correlation parameters. The analysis used the same priors as the pre-registered analysis.

The coefficients and credible intervals for the exploratory analysis of the Verb region are given in Table 14:

The exploratory analysis of the Verb region clearly confirmed a grammaticality effect, with faster reading times for grammatical sentences than for ungrammatical sentences with mismatching distractor ($P(\beta < 0) = .99$). However, the evidence for an attraction effect was not strong, even though the coefficient was in the expected direction ($P(\beta < 0) = .9$). Similarly, the interaction of attraction with active dependencies did not correspond to a clear effect in the analysis, although it was in the

Table 14. Results of the exploratory analysis of the verb region in Experiment 5.

	Estimate	Error	Lower 95%	Upper 95%	$P(\beta < 0)$
Intercept	6.05	0.02	6.02	6.09	
Attraction	-0.01	0.01	-0.02	0.00	.9
Grammaticality	-0.02	0.01	-0.03	-0.00	.99
Active	0.01	0.01	-0.01	0.03	.2
Attr × Act	0.01	0.01	-0.02	0.03	.3
Gram × Act	-0.02	0.01	-0.05	0.00	.95

expected direction attraction and active-dependency ($P(\beta < 0) = .3$). Bayes Factor analyses confirmed strong support for the null hypothesis, both for attraction ($BF_{01} = 62.08$), and for its interaction with active-dependency ($BF_{01} = 65.89$). A further analysis using the more constraining prior of $\text{normal}(0, 2.5)$ again confirmed strong support for the null hypothesis for both of these effects (Attraction: $BF_{01} = 15.67$; Attraction × Active-dependency: $BF_{01} = 15.54$).

General discussion

Across five experiments, we examined attraction effects in subject-verb agreement. We found stronger attraction effects for retro-active interference relative to pro-active interference. This may reflect effects of decay of activation of the distractor, which was at a greater distance from the critical verb in the pro-active experiments than in the retro-active experiments. However, within pro-active interference, we did not find clear evidence that the degree of attraction was modulated by whether the distractor was participating in an active dependency at the point where the agreement computation took place.

One important question to consider is that of why the number attraction was relatively weak in both of the pro-active interference configurations that we tested (Experiments 2, 3 and 5).

As noted in the introduction to this paper, other things being equal, standard accounts of cue-based retrieval (Lewis & Vasisht, 2005) would have predicted relatively weak effects of attraction for pro-active interference, due to the decay of activation of the distractor, and this is consistent with the fact that we only found clear evidence for attraction using retro-active interference (Experiments 1 and 4), while for pro-active interference (Experiments 2, 3 and 5), we found evidence against the presence of an attraction effect. On the other hand, more recent accounts of cue-based retrieval assume that activation is also affected by the information theoretic status of the distractor, with a boost of baseline activation being assumed for main clause subjects, due to their high topicality (Engelmann et al., 2019). If topicality does have an influence, however, it failed to yield reliable attraction effects in Experiments 2 and 3 reported above, where the distractor was a main clause subject. A second factor that we should consider is whether the distractor is reactivated in pro-active interference before the retrieval event occurs. In the introduction to this paper, we discussed two studies that showed evidence of pro-active interference in language comprehension, namely those reported by

Wagers et al. (2009) (examples (1) and (2), repeated below as (16) and (17), and by Franck et al. (2015) (examples (3a,b), repeated below as (18a,b)):

- (16) *The musicians who the reviewer praise so highly will probably win a Grammy.
 (17) *The musician who the reviewer praise so highly will probably win a Grammy.
 (18) a*Jerôme parle à la prisonnière que le gardien sortent.
Jerôme talks to the prisoner who the guard take out.
 b*Jerôme parle aux prisonnières que le gardien sortent.
Jerôme talks to the prisoners who the guard take out.

Both studies (Franck et al., 2015; Wagers et al., 2009) showed evidence of attraction, though with different measures, despite the fact that the distractor (*the musicians, the prisoners*) appears before the retrieval target (*the reviewer, the guard*). However, in both cases, the distractor forms a dependency with the verb that signals the critical retrieval event (*praise, take out*). Thus, arguably, the distractor is reactivated at the point where the number features of the target phrase have to be retrieved – in fact, a second retrieval event (that of the verb retrieving its object) is taking place at the same time. This situation is in contrast with the sentence types that we used in our study, where we wished to test for effects of maintenance without the complicating factor of reactivation. Nevertheless, it may be the case that reactivation is a necessary condition for finding measurable attraction effects with pro-active interference.⁹ Moreover, reactivation may be a stronger influence on attraction effects than topicality. Note that the distractor in Wagers et al.'s (2009) study is a main clause subject (*the musicians* in 16 and 17), but that of Franck et al.'s (2015) study is not (*the prisoner(s)* in 18). The relative influence of reactivation and topicality should be reassessed in future studies that manipulate both of these factors.

Given the discussion above, one potential criticism of our studies is that we were unable to observe an effect of maintenance simply because the sentence structures that we were using in Experiments 2 and 3 were not appropriate to yield reliable attraction effects, whether or not an active dependency was being maintained, or that we would have needed more power to detect interference in our sentence structures. Although we cannot discount this possibility on the basis of the data reported here, our results may instead indicate that the types of active subject-verb dependencies that we examined, or perhaps active dependencies in general, do not affect the activation of items in memory, or that they do not

require the maintenance of relevant features in memory. However, this view would need to be reconciled with a substantial literature showing ERP effects of maintenance of active dependencies, which we review below.

A number of ERP studies have shown a sustained ERP response for the sequence of words within a sentence where an active long-distance dependency remains unresolved, relative to control sentences (Cruz Heredia et al., 2022; Fiebach et al., 2002; Hagiwara et al., 2007; King & Kutas, 1995; Phillips et al., 2005). For example, in their ERP experiment (Phillips et al., 2005) examined sentences like (19), among other conditions:

- (19) a. The lieutenant knew which accomplice the detective hoped that the shrewd witness would recognize in the lineup.
 b. The lieutenant knew that the detective hoped that the shrewd witness would recognize the accomplice in the lineup.

In (19a), there is a long distance dependency linking *which accomplice* with the most deeply embedded verb *recognize*. In the control condition (19b), there is no such wh-dependency. In their ERP experiment, Phillips et al. (2005) found a sustained negativity in (19a), where the dependency was active, relative to the control condition (19b), for word positions between the two end-points of the wh-dependency. This, along with other similar demonstrations, shows that the brain is sensitive to active dependencies. However, there are different possible ways in which active dependencies might be represented. According to the view that we have assumed in this paper, the relevant features of the left-hand element of the dependency are actively maintained in working memory until the right-hand element of the dependency is encountered. As an example, for the distractor phrase in our Experiment 3, the maintained features would include its number features. In this *maintenance* view, it would be natural to expect that attraction would be affected by the active status of the dependency, due to the maintenance of the features in working memory. However, according to another possible account discussed by Cruz Heredia et al. (2022), rather than the maintenance of the left-hand element, active dependencies might instead be characterised by the prediction of the right-hand element itself, or features of it. If this is the case, then activation of the distractor may not be affected by active dependency status, and therefore, a difference may not be expected in the size of any attraction effect for active versus non-active dependencies.

One possibility that should be considered is that dependency types may differ in whether they show effects associated with active dependencies. In an ERP study, Cruz Heredia et al. (2022) examined sentences like (20a–d):

- (20) a. What did the commentary from the spokesman interrupt?
 b. Did the commentary from the spokesman interrupt the game?
 c. After Charlotte quit her boring job, she traveled to Europe like she always wanted to.
 d. Today Charlotte quit her boring job (and traveled to Europe like she always wanted to)

In (20a), there is a dependency between *what* and the verb *interrupt*, while in (20c), the initial subordinate clause (*after...*) indicates the upcoming presence of the main clause, making prediction possible. Thus, there is a sense in which we can think of (20c) as also involving an active dependency. In both (20a) and (20c), the relevant dependencies are arguably active while the comprehender is processing the intervening words. Sentences (20b) and (20d) are control conditions that lack the active dependencies of (20a) and (20c) respectively. If sustained negativity is a general ERP signature of unresolved dependencies then such an effect would be expected for both (20a) and (20c) relative to their respective control conditions. However, Cruz Heredia et al. (2022) found the expected effect only in (20a). This might be taken to suggest that the expectancy-based processing mechanisms underlying the sustained negativity effect are limited to *wh*-dependencies.¹⁰ If this reflects a general difference between *wh*-dependencies and other active dependencies, then this might also explain why we failed to find modulation of number attraction as a function of active subject-verb dependencies, as our studies did not use *wh*-dependencies. In addition, this may also explain why Kim et al. (2020) were able to observe active dependency effects, given that their active dependencies used *wh*-dependencies (see example 9 above) – in other words, it may be the case that number agreement attraction is affected by the active status of *wh*-dependencies, but not other forms of dependency.¹¹ However, recall above that Ristic et al. (2022) detected a processing cost for active (non-*wh*) subject-verb dependencies in their self-paced reading experiment, so it seems that active subject-verb dependencies do have some measurable impact on processing, even if this does not modulate attraction, or lead to a sustained negativity ERP signature.

Finally, it is worth commenting on our decision to test attraction effects only in ungrammatical conditions, without also manipulating distractor number in grammatical sentences. As mentioned above, previous comprehension studies of verb-subject number agreement have shown a grammaticality asymmetry, such that more consistent effects have been found in ungrammatical sentences than in grammatical sentences,¹² and therefore, it seemed reasonable to test for attraction only in ungrammatical sentences. However, it should be acknowledged that the processing of ungrammatical sentences may be subject to individual strategies, and be atypical of normal sentence processing. Also, it is always possible that by concentrating on ungrammatical sentences, we missed finding effects of active dependencies, that we would have found if we had also examined grammatical sentences, although we are not aware of any theoretical framework that would predict larger effects of active dependencies in grammatical than in ungrammatical sentences.

Conclusion

Across five experiments, we investigated the factors that affect number attraction in the comprehension of subject-verb dependencies. We found evidence for greater attraction in retro-active interference, where the distractor linearly intervenes between the verb and retrieval target, relative to pro-active interference, where the distractor precedes the target. This is compatible with the notion of decay of activation, as proposed in the original version of the cue-based retrieval model (Lewis & Vasishth, 2005). We also asked whether the maintenance of a dependency affected attraction within pro-active interference conditions, as would be expected if active dependencies required maintenance of the relevant features in memory. However, this result was not obtained. This might indicate that active dependencies do not involve the maintenance of features in memory, or otherwise might indicate that our pro-active interference conditions were not suitable for examining attraction effects.

Notes

1. In this paper, we will indicate ungrammatical sentences with an asterisk.
2. Although the cue-based retrieval model assumes the decay of activation, it should be noted that the notion of decay is controversial in the domain-general memory literature (see Lewandowsky et al., 2009).
3. Note that, in the production literature, it is usually assumed that percolation can only act upwards in a syntactic tree, which would rule out attraction occurring in

cases like (1), and in our pro-active interference experiments, because in these cases the distractor is above the target in the tree. However, in Bayesian simulations of a model that combined feature percolation with cue-based retrieval, Yadav et al. (2023) find good fits with data from a number of experiments, regardless of whether the distractor was lower or higher than the target in the tree.

4. However, note that among the models tested by Yadav et al. (2023), the best fit to experimental data across both grammatical and ungrammatical sentences was provided by a model that combined both feature percolation and cue-based retrieval.
5. We cannot give a precise measurement for the visual angle per character, because the experiments used Times Roman font, which is a variable width font.
6. An initial model also included a random slope for trial by region, but inspection of the trace plot indicated that the estimate for this slope did not converge.
7. Further analyses including these extra participants yielded very similar results to those reported below. These extra analyses, along with the data, are available on the OSF site for this experiment.
8. See <https://osf.io/g6tdn/>.
9. Another possibility is that the referent of the distractor must participate in the same event as the referent of the target, as is the case in the relative clause examples of both Franck et al. (2015) and Wagers et al. (2009), but not in our Experiments 2 or 3.
10. In fact Cruz Heredia et al.'s full results suggest that the sustained negativity effect is limited to *matrix* wh-questions, as they did not find an effect for embedded wh-questions. However, Phillips et al. (2005) did find sustained negativity for embedded questions.
11. Note, however, that Kim et al. (2020) tested the effect of active dependency status only in retro-active interference, so their results are not strictly comparable to ours.
12. Explaining the grammaticality asymmetry is beyond the scope of this paper, though one interesting recent suggestion is that, in a model that uses both cue-based retrieval and feature percolation, these two processes both push in the same direction for ungrammatical sentences, resulting in a facilitatory attraction effect, while the two processes push in opposite directions in grammatical sentences, resulting in a small net effect, of variable direction (Yadav et al., 2023).

Acknowledgments

We thank Lauren Hadley and Lihua Xia for assistance in data collection. We thank two anonymous reviewers for insightful comments on a previous version of this manuscript. Remaining errors are our own. Ethical Approval for the studies was obtained from the PPLS Research Ethics committee, University of Edinburgh.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The research reported here was supported by a University of Edinburgh PPLS Small Project grant, and a Global Research Network grant awarded by the National Research Foundation of Korea [grant number NRF-2014S1A2A2028232].

References

- Acuña Fariña, J. C. A., Meseguer, E., & Careirras, M. (2014). Gender and number agreement in comprehension in Spanish. *Lingua*, *143*, 108–128. <https://doi.org/10.1016/j.lingua.2014.01.013>
- Bock, J. K., & Eberhard, K. (1993). Meaning, sound and syntax in English number agreement. *Language and Cognitive Processes*, *8*(1), 57–99. <https://doi.org/10.1080/01690969308406949>
- Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using stan. *Journal of Statistical Software*, *80*(1), 1–28. <https://doi.org/10.18637/jss.v080.i01>
- Bürkner, P. C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, *10*(1), 395–411. <https://doi.org/10.32614/RJ-2018-017>
- Bürkner, P. C. (2021). Bayesian item response modeling in R with brms and stan. *Journal of Statistical Software*, *100*(5), 1–54. <https://doi.org/10.18637/jss.v100.i05>
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Brubaker, M. A., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1). <https://doi.org/10.18637/jss.v076.i01>
- Cruz Heredia, A. A. L., Dickerson, B., & Lau, E. (2022). Towards understanding sustained neural activity across syntactic dependencies. *Neurobiology of Language*, *3*(1), 87–108. https://doi.org/10.1162/nol_a_00050
- Cummings, I., & Sturt, P. (2014). Coargumenthood and the processing of reflexives. *Journal of Memory and Language*, *75*, 117–139. <https://doi.org/10.1016/j.jml.2014.05.006>
- Dillon, B., Mishler, A., Sloggett, S., & Phillips, C. (2013). Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*, *69*(2), 85–103. <https://doi.org/10.1016/j.jml.2013.04.003>
- Eberhard, K. M., Cutting, J. C., & Bock, K. (2005). Making syntax of sense: Number agreement in sentence production. *Psychological Review*, *112*(3), 531–559. <https://doi.org/10.1037/0033-295X.112.3.531>
- Engelmann, F., Jäger, L. A., & Vasishth, S. (2019). The effect of prominence and cue association on retrieval processes: A computational account. *Cognitive Science*, *43*, e12800. <https://doi.org/10.1111/cogs.v43.12>
- Fiebach, C. J., Schlesewsky, M., & Friederici, A. D. (2002). Separating syntactic memory costs and syntactic integration costs during parsing: The processing of German WH-questions. *Journal of Memory and Language*, *47*(2), 250–272. [https://doi.org/10.1016/S0749-596X\(02\)00004-9](https://doi.org/10.1016/S0749-596X(02)00004-9)
- Franck, J., Colonna, S., & Rizzi, L. (2015). Task-dependency and structure-dependency in number interference effects in sentence comprehension. *Frontiers in Psychology*, *6*, Article 349.
- Franck, J., Lassi, G., Frauenfelder, U. H., & Rizzi, L. (2006). Agreement and movement: A syntactic analysis of attraction. *Cognition*, *101*(1), 173–216. <https://doi.org/10.1016/j.cognition.2005.10.003>

- Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(6), 1411–1423.
- Hagiwara, H., Soshi, T., Ishihara, M., & Imanaka, K. (2007). A topographical study on the event-related potential correlates of scrambled word order in Japanese complex sentences. *Journal of Cognitive Neuroscience*, 19(2), 175–193. <https://doi.org/10.1162/jocn.2007.19.2.175>
- Jäger, L., Benz, L., Roeser, J., Dillon, B., & Vasishth, S. (2015). Teasing apart retrieval and encoding interference in the processing of anaphors. *Frontiers in Psychology*, 6, Article 506.
- Jäger, L., Engelmann, F., & Vasishth, S. (2017). Similarity-based interference in sentence comprehension: Literature review and Bayesian meta-analysis. *Journal of Memory and Language*, 94, 316–339. <https://doi.org/10.1016/j.jml.2017.01.004>
- Jäger, L., Mertzen, D., Van Dyke, J. A., & Vasishth, S. (2020). Interference patterns in subject-verb agreement and reflexives revisited: A large-sample study. *Journal of Memory and Language*, 111, Article 104063. <https://doi.org/10.1016/j.jml.2019.104063>
- Kim, N., Brehm, L., Sturt, P., & Yoshida, M. (2020). How long can you hold the filler? Maintenance and retrieval. *Language, Cognition and Neuroscience*, 35, 17–42. <https://doi.org/10.1080/23273798.2019.1626456>
- King, J. W., & Kutas, M. (1995). Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, 7(3), 376–395. <https://doi.org/10.1162/jocn.1995.7.3.376>
- Kwon, N., & Sturt, P. (2019). Proximity and same case marking do not increase attraction effect in comprehension: Evidence from eye-tracking experiments in Korean. *Frontiers in Psychology (Language Sciences)*, 10, Article 1320. <https://doi.org/10.3389/fpsyg.2019.01320>
- Lago, S., Shalom, D., Sigman, M., Lau, E., & Phillips, C. (2015). Agreement attraction in Spanish comprehension. *Journal of Memory and Language*, 82, 133–149. <https://doi.org/10.1016/j.jml.2015.02.002>
- Laurinavichyute, A., Jäger, L., Akinina, Y., & Roß, J. (2017). Retrieval and encoding interference: Cross-linguistic evidence from anaphor processing. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.00965>
- Lewandowsky, S., Oberauer, K., & Brown, G. (2009). No temporal decay in verbal short-term memory. *Trends in Cognitive Sciences*, 13(3), 120–126. <https://doi.org/10.1016/j.tics.2008.12.003>
- Lewis, R., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29, 1–45. https://doi.org/10.1207/s15516709cog0000_25
- Nairne, J. (1990). A feature model of immediate memory. *Memory and Cognition*, 18(3), 251–269. <https://doi.org/10.3758/BF03213879>
- Nicenboim, B., Vasishth, S., Engelmann, F., & Suckow, K. (2018). Exploratory and confirmatory analyses in sentence processing: A case study of number interference in German. *Cognitive Science*, 42, 1075–1100. <https://doi.org/10.1111/cogs.2018.42.issue-S4>
- Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory and Language*, 55, 601–626. <https://doi.org/10.1016/j.jml.2006.08.009>
- Parker, D. (2019). Cue combinatorics in memory retrieval for anaphora. *Cognitive Science*, 43, e12715. <https://doi.org/10.1111/cogs.2019.43.issue-3>
- Pearlmutter, N., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language*, 41(3), 427–456. <https://doi.org/10.1006/jmla.1999.2653>
- Phillips, C., Kazanina, N., & Abada, S. H. (2005). ERP effects of processing syntactic long-distance dependencies. *Cognitive Brain Research*, 22, 407–428. <https://doi.org/10.1016/j.cogbrainres.2004.09.012>
- Ristic, B., Mancini, S., Molinaro, N., & Staub, A. (2022). Maintenance cost in the processing of subject–verb dependencies. *Journal of Experimental Psychology: Learning Memory and Cognition*, 48(6), 829–838.
- Smith, G., Franck, J., & Tabor, W. (2021). Encoding interference effects support self-organized sentence processing. *Cognitive Psychology*, 124, Article 101356. <https://doi.org/10.1016/j.cogpsych.2020.101356>
- Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, 48(3), 542–562. [https://doi.org/10.1016/S0749-596X\(02\)00536-3](https://doi.org/10.1016/S0749-596X(02)00536-3)
- Sturt, P., & Lombardo, V. (2005). Processing coordinate structures: Incrementality and connectedness. *Cognitive Science*, 29, 291–305. https://doi.org/10.1207/s15516709cog0000_8
- Van Dyke, J. A. (2007). Interference effects from grammatically unavailable constituents during sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(2), 407–430.
- Villata, S., & Franck, J. (2020). Similarity-based interference in agreement comprehension and production: Evidence from object agreement. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46(1), 170–188.
- Villata, S., Tabor, W., & Franck, J. (2018). Encoding and retrieval interference in sentence comprehension: Evidence from agreement. *Frontiers in Psychology*, 9, Article 2.
- von der Malsberg, T., & Angele, B. (2017). False positives and other statistical errors in standard analyses of eye movements in reading. *Journal of Memory and Language*, 94, 119–133. <https://doi.org/10.1016/j.jml.2016.10.003>
- Wagers, M., Lau, E., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, 61, 206–237. <https://doi.org/10.1016/j.jml.2009.04.002>
- Yadav, H., Smith, G., Reich, S., & Vasishth, S. (2023). Number feature distortion modulates cue-based retrieval in reading. *Journal of Memory and Language*, 129. <https://doi.org/10.1016/j.jml.2022.104400>
- Zehr, J., & Schwarz, F. (2018). *PennController for internet based experiments (IBEX)*. <https://doi.org/10.17605/OSF.IO/MD832>