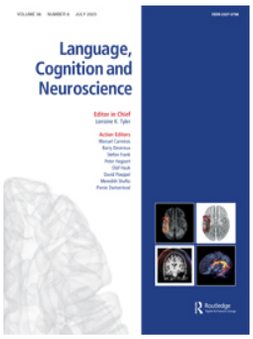
Illusions of plausibility in adjuncts and co-ordination



Routledge
Taylor & Francis Group

Language, Cognition and Neuroscience

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/plcp21

Illusions of plausibility in adjuncts and co-ordination

Ian Cunnings & Patrick Sturt

To cite this article: Ian Cunnings & Patrick Sturt (2023): Illusions of plausibility in adjuncts and co-ordination, Language, Cognition and Neuroscience, DOI: 10.1080/23273798.2023.2235033

To link to this article: https://doi.org/10.1080/23273798.2023.2235033

Published online: 22 Jul 2023.

LANGUAGE, COGNITION AND NEUROSCIENCE
https://doi.org/10.1080/23273798.2023.2235033

Routledge
Taylor & Francis Group

REGULAR ARTICLE

OPEN ACCESS Check for updates

Illusions of plausibility in adjuncts and co-ordination

Ian Cunnings [a] and Patrick Sturt[b]

[a]School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK; [b]Psychology, University of Edinburgh, Edinburgh, UK

ABSTRACT

Illusions of grammaticality, where ungrammatical sentences are misperceived as grammatical (e.g. The key to the cabinets were rusty), have been widely studied during language comprehension. Such grammatical illusions have been influential in debate surrounding so-called representational and retrieval-based accounts of linguistic dependency resolution. Whether analogous illusions of plausibility occur at the level of semantic interpretation has only recently begun to be examined, and thus far, these illusions have been restricted to a narrow range of linguistic phenomena. In two eye-tracking during reading experiments (n = 48 in each) and two self-paced reading experiments (n = 192 in each) we examined the possibility of semantic illusions during the processing of adjuncts and co-ordination. Across experiments, our results suggest illusions of plausibility during dependency resolution, though interference effects were clearer in adjuncts than co-ordination. We argue that our findings are more compatible with retrieval-based rather than representational accounts of linguistic dependency resolution.

ARTICLE HISTORY
Received 30 August 2022
Accepted 5 July 2023

KEYWORDS
Sentence processing; memory retrieval; interference; linguistic dependencies; Reading

Introduction

A key aspect of language comprehension is understanding sentences that contain linguistic dependencies between non-adjacent constituents. A growing body of research has investigated the memory operations involved in resolving such dependencies (for review, see Jäger et al., 2017; Parker et al., 2017; Vasishth et al., 2019). Two examples of linguistic dependencies involve adjuncts as in (1a) and co-ordination as in (1b). In (1a), the temporal adjunct is missing an overt subject, but the sentence subject "the boy" is typically interpreted as being the agent of the verb "walking". A similar issue arises in (1b), where correct interpretation requires that "the boy" is interpreted as the agent of "walked". In both cases, a linearly closer constituent ("the girl") could potentially interfere in how these dependencies are resolved, if this constituent is wrongly interpreted to be the agent of the verb in either case.

(1a) The boy spoke to the girl after walking to school.
(1b) The boy spoke to the girl and then walked to school.

Cue-based models of memory retrieval during language processing (e.g. Jäger et al., 2017) would predict that dependency resolution in adjuncts and co-ordination like (1a/b) should be susceptible to interference, but to date very little research has examined the processing of such structures. In research on linguistic dependencies, a number of studies have reported so-called illusions of grammaticality, where ungrammatical sentences are perceived as being grammatical (e.g. The key to the cabinets were rusty), as evidence of interference (e.g. Lago et al., 2015; Wagers et al., 2009). Research in this area has focused on whether grammatical illusions are best explained by so-called representational models (e.g. Hammerly et al., 2019), which argue that illusions result from how information is encoded and represented in memory, and retrieval-based models (e.g. Jäger et al., 2017), which argue that grammatical illusions arise during memory retrieval. Illusions of plausibility, where implausible sentences are misperceived as being plausible (e.g. Julia saw the beer that the lady with the food quite happily ate) have also recently been reported (Cunnings & Sturt, 2018), but much less research has examined these semantic illusions as compared to the considerable body of research on grammatical illusions.

In this study, we utilised adjuncts and co-ordination to further investigate semantic illusions during sentence processing. We aimed to conceptually replicate

CONTACT Ian Cunnings i.cunnings@reading.ac.uk School of Psychology and Clinical Language Sciences, University of Reading, Reading RG6 7BE, UK

previously observed semantic illusions in new linguistic phenomena, to assess the extent to which illusions of plausibility are a general property of language comprehension. We report two eye-tracking and two self-paced reading experiments investigating memory retrieval during the processing of adjuncts and co-ordination. In so doing, we also aim to contribute to debate concerning representational and retrieval-based models of linguistic illusions. We begin below by discussing different theoretical accounts of linguistic illusions, before discussing existing research on processing adjuncts and co-ordination.

Linguistic illusions in language comprehension

Perhaps the most well-studied linguistic illusion in the psycholinguistics literature is that of agreement attraction, as in (2), taken from Wagers et al. (2009).

(2a) The key to the cell unsurprisingly was rusty from many years of disuse.
(2b) The key to the cells unsurprisingly was rusty from many years of disuse.
(2c) The key to the cells unsurprisingly were rusty from many years of disuse.
(2d) The key to the cell unsurprisingly were rusty from many years of disuse.

(2a/b) are grammatical as the auxiliary verb ("was") matches in number with the head of the sentence subject ("the key"), while (2c/d) are ungrammatical as the head ("the key") and verb ("were") mismatch. The number properties of a distractor noun ("the cell/s") are also manipulated and matches the verb in (2a/c) but mismatches in (2b/d). A number of studies have reported so-called *grammatical illusions* during the processing of such sentences, with shorter reading times in (2c) than (2d) (e.g. Dillon et al., 2013; Lago et al., 2015; Wagers et al., 2009). Here, ungrammatical sentences like (2c) are sometimes misperceived as being grammatical, because the distractor noun "cells" matches the number properties of the plural verb.

There are two main accounts of this finding in comprehension. According to cue-based parsing (e.g. Jäger et al., 2017), subject-verb dependencies as in (2) involve retrieving a representation of the subject noun from memory at the verb. Retrieval is achieved via a cue-based mechanism that matches a set of retrieval cues with all items in memory in parallel. Items that match the retrieval cues become activated to varying degrees based on their degree of match to the set of retrieval cues, and the item with the highest level of activation is retrieved to complete the dependency. In sentences like (2), retrieval cues will involve syntactic information about the required constituent ([+ SUBJECT HEAD]) along with the number properties of the verb (e.g. [+PLURAL]). In ungrammatical (2c/d) no item in the sentence fully matches the retrieval cues, as the retrieval target ("the key") is [+SUBJECT HEAD] but does not match the verb in number. The distractor ("the cabinets") is not the syntactic head of the subject but does match the number properties of the verb in (2c). Cue-based parsing thus predicts that on some proportion of occasions, the partially matching distractor is activated and retrieved in (2c), leading to a grammatical illusion. This predicts shorter reading times in (2c) than (2d). We refer to this as *facilitatory interference* (Jäger et al., 2017).

Representational accounts explain differences between (2c) and (2d) based on how the number properties of the complex subject ("the key/s to the cabinet/s") are encoded in memory. Thus, these accounts focus on how information is represented in memory, rather than retrieval. Representational accounts were originally formulated to account for findings in sentence production research (Bock & Cutting, 1992; Bock & Miller, 1991), but have been extended to comprehension (e.g. Hammerly et al., 2019; Pearlmutter et al., 1999). Although there are different representational accounts, such as feature percolation (Bock & Eberhard, 1993; Franck et al., 2002) and marking and morphing (Bock & Middleton, 2011; Eberhard et al., 2005), they share the idea that in a complex noun phrase like "the key to the cabinets", as in (2c), the number properties of the phrasal head become misspecified. According to feature percolation, in "the key to the cabinets" the plural number properties of "cabinets" can percolate up to the head noun ("the key"). According to marking and morphing, a noun phrase number is not categorically singular or plural, but has a gradient value between the two. Here, a complex noun phrase with a singular head but plural distractor, will have a gradient number specification compared to a complex noun phrase with two singular nouns. Although these representational accounts thus differ, in both cases it is featural misspecification that explains reduced reading times for (2c) compared to (2d).

Both retrieval-based and representational models also predict differences in grammatical sentences, like (2a/b). Here however, the models make opposing predictions. Cue-based models of memory retrieval would predict longer reading times in (2a) than (2b), as evidence of so-called *inhibitory interference* (Jäger et al., 2017).[1] That is, in (2a), the fact that both nouns match the number properties of the verb should lead to competition compared to (2b), where only the head

matches the verb in number. Alternatively, as representational accounts predict that attraction effects result from how complex noun phrase number is represented, such that a singular noun phrase with a plural distractor is always represented as being more plural than a noun phrase with two singular nouns (whether through feature percolation or some other mechanism), such accounts would predict the opposite, with longer reading times predicted in (2b) than (2a).

Despite these predictions, a number of studies have failed to observe effects in either direction in grammatical sentences (e.g. Dillon et al., 2013; Lago et al., 2015; Wagers et al., 2009), a finding Wagers et al. dubbed the *grammatical asymmetry*. This asymmetry has typically been taken as evidence in favour of retrieval-based accounts rather than representational accounts, though how to account for the apparent lack of effects in grammatical conditions is debated. Wagers et al. considered two possible accounts of this finding. Firstly, the lack of effects in grammatical sentences may indicate that cue-based retrieval only occurs in ungrammatical sentences. They argued that readers predict the number properties of an upcoming verb based on head noun number, and only require retrieval when this prediction clashes with the number properties of the actual verb. Alternatively, Wagers et al. considered that retrieval always occurs, but that interference is highly unlikely when the target fully matches the cues to retrieval, as in grammatical sentences. Indeed, Nicenboim et al. (2018) claimed that inhibitory interference effects are small and may have been difficult to detect in previous studies due to lack of power. In an analysis combining data from 184 participants, they reported inhibitory effects in grammatical sentences consistent with cue-based parsing.

Additionally, Hammerly et al. (2019) claimed that the asymmetry observed in judgement studies may result from response bias, such that participants in judgement studies typically respond more accurately to grammatical rather than ungrammatical conditions. They argued that when response bias is taken into consideration, a pattern of results across grammatical and ungrammatical sentences emerges that is more consistent with representational accounts.

Although grammatical illusions in ungrammatical sentences have been well examined, much less research has examined whether analogous semantic illusions exist in sentences that are implausible. Cunnings and Sturt (2018) tested for semantic illusions in sentences containing filler gap dependencies such as (3). In (3), upon encountering the critical verb "shattered", the reader must retrieve a constituent from memory as its non-adjacent object. Cunnings and Sturt manipulated the plausibility of both a target constituent ("the plate"/"the letter") and also a distractor ("the cup"/"the tie") as a potential direct object of the verb. Note in each case the retrieval target is the only constituent that can grammaticality be interpreted as the verb's direct object.

(3a) Sue remembered the plate that the butler with the cup accidently shattered today in the dining room.
(3b) Sue remembered the plate that the butler with the tie accidently shattered today in the dining room.
(3c) Sue remembered the letter that the butler with the cup accidently shattered today in the dining room.
(3d) Sue remembered the letter that the butler with the tie accidently shattered today in the dining room.

At and after the critical verb ("shattered") reading times were significantly longer for implausible (3c/d) than plausible (3a/b). There was also a significant target by distractor interaction, such that reading times were significantly shorter in (3c), when the distractor was a plausible direct object, compared to (3d), when it was not. No significant differences were observed in (3a/b). Cunnings and Sturt interpreted these findings as indicating facilitatory interference in implausible sentences. The asymmetry in interference effects, with illusions of plausibility in implausible sentences but not illusions of implausibility in plausible sentences, Cunnings and Sturt interpreted as being similar to the previously observed asymmetry in illusions of grammaticality in subject-verb agreement. They interpreted their results as being compatible with cue-based models. In particular, they argued that semantic cues derived from the lexical properties of individual lexical items must have been driving the semantic illusions that they observed (see Smith & Vasishth, 2020 for further discussion).

These findings have since been replicated in both native and non-native English readers (Fujita & Cunnings, 2022). Laurinavichyute and von der Malsburg (2022) also recently reported semantic illusions in a series experiments examining sentences such as (4), where participants had to judge whether the capitalised word was a plausible continuation of the sentence. (4a/b) are both implausible, as "the bakery" is an implausible agent of the verb "sprays'. Laurinavichyute and von der Malsburg found however that participants were more likely to wrongly say that the continuation was plausible when the distractor ("the office building"/"the fire hydrant") was itself a plausible agent of the verb. This semantic illusion suggests that illusions of plausibility are a general property of language comprehension, as would be predicted by cue-based models.

(4a) The bakery near the office building rarely … SPRAYS
(4b) The bakery near the fire hydrant rarely … SPRAYS

It is not clear that representational models can straightforwardly account for these semantic illusions. Such models were formulated to account for attraction from grammatical features, such as number, and as such they predict effects in a narrower range of contexts than retrieval-based models. While representational models can explain semantic effects related to different ways in which number is expressed, either morphosyntactically or conceptually (e.g. Schlueter et al., 2018), they do not, in their current formulation, predict the types of semantic illusions as discussed above. Consider the semantic properties of number for example. The majority of English nouns have two number values, either singular or plural, and thus it is conceivable that such values may become mis-specified. Nouns however typically only have one value for the types of semantic features manipulated in the studies on semantic illusions above however. For example, nouns are either [+ ANIMATE] or [+ INANIMATE] (i.e. English nouns do not have separate animate and inanimate forms) and, as such, it is less clear how such features could become mis-specified under a representational account. Cue-based parsing on the other hand predicts that interference should be a general property of language comprehension, and as such can explain both grammatical and semantic illusions. It might be possible to extend representational accounts to semantic features, though this would be a considerable extension of their original remit. The asymmetry in semantic illusions in implausible but not plausible sentences, as reported by Cunnings and Sturt (2018), would still be difficult to reconcile from a representational point of view however.

Some representational accounts, most notably feature percolation, predict that attraction should only occur when a distractor is inside a complex noun phrase, and predict that the amount of attraction is dependent on how embedded the distractor is within the noun phrase (Bock & Cutting, 1992; Franck et al., 2002). Laurinavichyute and von der Malsburg's tested materials with a distractor embedded directly in a complex noun phrase ("the bakery near the office"). Cunnings and Sturt (2018) and Fujita and Cunnings (2022) however both tested a distractor, "the tie"/"the cup" in (3), that is not directly embedded within the target noun phrase, ("the plate"/"the letter") and is instead embedded within another noun phrase ("the butler") in a different clause. However in both these studies, although the distractor is not directly embedded within the target noun phrase, it is embedded within a relative clause that modifies the retrieval target. By examining semantic interference in adjuncts and co-ordination as in (1), we can test whether semantic illusions are also observed in cases where the distractor is not in any way embedded within the target noun phrase. Testing sentences such as (1) thus constitutes an important test of the generalisability of semantic illusions in comprehension.

While the original implementation of the activation-based cue-based model (Lewis et al., 2006; Lewis & Vasishth, 2005) predicts that interference can occur irrespective of a distractor's the syntactic position, some proponents of cue-based models have claimed that interference is influenced by the syntactic position of a distractor. For example, in their study on subject retrieval in sentences such as "The attorney who the judge realised had declared that the witness was inappropriate compromised", Van Dyke and McElree (2011) claimed that either syntactic cues ([+ SUBJECT]) may "gate" interference to distractors that also match such cues, or alternatively that syntactic cues may be more highly weighted than other non-syntactic cues. In a study on facilitatory interference in subject-verb agreement, Parker and An (2018) argued that facilitatory interference occurs only when a distractor is in an oblique argument position ("The waitress who sat near the girls unsurprisingly were unhappy") but not when it is a core argument of a verb ("The waitress who sat the girls unsurprisingly were unhappy"). Parker and An claimed that core argument distractors are encoded in memory with more distinct representations, that lead to decreased interference. In sentences like (1a/b), the distractor ("the girl") is in a core argument position and does not match the syntactic [+ SUBJECT] cue required for the missing subject at the critical verb ("walked/walked"). Testing interference in sentences like (1) thus allows us to assess claims regarding the syntactic position of a distractor.

Interference in adjuncts and co-ordination

Very little research has investigated the memory encoding and retrieval processes involved in processing adjuncts and co-ordination as in (1). There has been considerable research in the child language acquisition literature on the development of adjuncts, and a number of studies have indicated non-adultlike interpretation for sentences similar to (1a) in ages up to and above 6 years of age, with children interpreting constituents such as "the girl" in (1a) as the agent of "walking" instead of the sentence subject (Cairns et al., 1994; Goodluck, 2001; Hsu, 1985; Janke, 2018).

Although there are different accounts of this finding (for discussion, see Goodluck, 2001; Janke, 2018), interference has been shown to play at least some role in explaining children's interpretation in such cases (Gerard et al., 2017).

We are aware of only one study to have examined retrieval of the sentence subject during the processing of adjuncts in adults. In a judgement study, Parker et al. (2015) found that sentences containing adjuncts were more acceptable with animate rather than inanimate subjects (e.g. The doctor was certified after debunking the hypothesis vs. The discovery was certified after debunking the hypothesis). They also conducted a self-paced reading experiment that manipulated both animacy of the sentence subject ("the doctor" / "the discovery") and additionally a distractor ("the researcher" / "the report"), as in (5). Reading times at both the verb "debunking" and after the reflexive "himself" in sentences like (5b) were significantly longer than in (5a) only when the distractor was inanimate. The shorter reading times for sentences containing inanimate subjects when the distractor was itself animate indicates facilitatory interference. No significant effects of distractor animacy were observed in sentences like (5a), when the sentence subject was animate. Parker et al. interpreted their results as indicating that a similar asymmetry in interference effects is observed in the processing of adjuncts as has been observed for subject-verb agreement.

(5a) The doctor that the researcher/report described meticulously was certified after debunking the urban myth himself in the new scientific journal.
(5b) The discovery that the researcher/report described meticulously was certified after debunking the urban myth himself in the new scientific journal.

Given that in this study the distractor is in a subject position, these results are difficult to reconcile with the claim that distractors that are core arguments resist interference (Parker & An, 2018). However, they are compatible with the claim that distractors that match syntactic cues at retrieval, in this case [+SUBJECT], can cause interference (Van Dyke & McElree, 2011). Additionally, while Parker et al.'s results indicate interference during the processing of adjuncts when the retrieval target is inanimate, to our knowledge interference has not previously been observed in such constructions when the sentence subject is itself also animate. Indeed, in a further self-paced reading experiment, Parker et al. tested similar materials but instead of manipulating animacy, manipulated the stereotypical gender of the sentence subject and a distractor (e.g. "The harpist/drummer that the diva/guitarist liked very much was congratulated after playing the beautiful song herself at the brand new recording studio"). In this experiment, they observed longer reading times when the sentence subject mismatched in gender with the reflexive, but the distractor did not cause any significant interference. Parker et al. considered two possibilities for these different results across experiments. Firstly, it might be because animacy is a particularly reliable and heavily weighted cue to memory retrieval. In this case, interference might only be observed in adjuncts when the retrieval target is inanimate. Alternatively, Parker et al. also noted that the degree of target feature mismatch was higher in their experiment that manipulated animacy than the one which manipulated gender. Specifically, in their experiment which manipulated animacy, inanimate subjects did not match the reflexive in either animacy or gender, while in their gender manipulation experiment the sentence subject mismatched in gender only. Parker et al. hypothesised that facilitatory interference may only arise in cases when the target is a particularly poor match (see also Parker & Phillips, 2017). In the study presented in this paper, we utilise a plausibility manipulation to further test whether interference can be observed in cases where a retrieval target is animate. If we do observe interference in such cases, this would limit any claims about the prioritised role of animacy information in guiding memory retrieval.

To our knowledge, only one study has examined co-ordination as in (1b) in the child language acquisition literature. Similar to what has been observed for adjuncts, children sometimes assign non-adultlike interpretations to such sentences (Friedmann & Costa, 2010). We are unaware of any study to have investigated memory retrieval during the processing of co-ordination in adults.

The present study

Against this background, we report four experiments investigating interference in adjuncts and co-ordination. We used a plausibility manipulation similar to Cunnings and Sturt (2018) in an attempt to conceptually replicate previously observed illusions of plausibility. In so doing, our study contributes to debate regarding retrieval-based and representational theories of interference during dependency resolution. In Experiments 1 and 2 we monitored participant's eye-movements as they read sentences containing adjuncts (Experiment 1) and co-ordination (Experiment 2), while Experiments 3 and 4 used web-based self-paced reading with considerably larger sample sizes. All experiments reported here received

ethical approval from the School of Psychology and Clinical Language Sciences Research Ethics Committee at the University of Reading (approval codes 2015-118-IC and 2018-103-IC) before participants were recruited.

Experiment 1

In Experiment 1, we monitored participants' eye-movements as they read texts as in (6). The critical second sentence in (6a/b) is plausible because the retrieval target ("the detective") is a plausible agent of the very "arresting". (6c/d) are implausible as the retrieval target ("the criminal") is an implausible agent of this verb. We also manipulated the animacy of a distractor constituent. In (6a/c) the distractor "the cop" is animate and is also a plausible agent of "arresting", while the inanimate noun "car" is an implausible agent of this verb in (6b/d).

A bad crime had been committed.

(6a) Plausible Sentence, Animate Distractor
The detective stood by the cop very calmly after arresting the robber outside the city bank.

(6b) Plausible Sentence, Inanimate Distractor
The detective stood by the car very calmly after arresting the robber outside the city bank.

(6c) Implausible Sentence, Animate Distractor
The criminal stood by the cop very calmly after arresting the robber outside the city bank.

(6d) Implausible Sentence, Inanimate Distractor
The criminal stood by the car very calmly after arresting the robber outside the city bank.

A lot of money had been stolen.

We expected longer reading times at and after the critical verb ("arresting") in (6c/d) compared to (6a/b) as evidence of plausibility effects. The original implementation of the activation-based cue-based model (Lewis et al., 2006; Lewis & Vasishth, 2005) would also predict interference. Facilitatory interference should be observed in implausible sentences, such that (6c) should have shorter reading times than (6d). This would conceptually replicate the illusions of plausibility reported by Cunnings and Sturt (2018). Cue-based models also predict inhibitory interference in plausible sentences, that is, reading times should be longer in (6a), when both retrieval target ("the detective") and distractor ("the cop") are plausible agents of "arresting", than in (6b), when only the retrieval target is plausible.

A different set of predictions can be made if the syntactic position of a distractor influences interference. Given that the critical verb "arresting" is missing a local subject, we assume that retrieval cues include a syntactic cue such as [+SUBJECT]. If interference is either restricted to constituents that match the syntactic [+SUBJECT] cue (Van Dyke & McElree, 2011) or if distractors that are core arguments resist interference (Parker & An, 2018), we should not observe interference effects.

Traditional representational accounts would not predict distractor effects, given that they predict featural misspecification can occur for morphosyntactic features such as number but not semantic features as manipulated here, and also because the distractor is not embedded within the target noun phrase in any way. Finding main effects of distractor plausibility, with longer reading times in (6b/d) than (6a/c), could potentially be compatible with an extension of representational accounts to semantic features, but this finding would be difficult to explain in terms of cue-based retrieval. Thus, interactions between sentence plausibility and distractor animacy would be most consistent with retrieval-based models.

Participants

48 native English speakers from the University of Reading took part for course credit or were paid a nominal sum. All participants had normal or corrected-to-normal vision.

Materials

An initial set of 40 items was constructed as in (6), each containing an introduction sentence, critical sentence and wrap-up sentence. We included a wrap-up sentence to minimise end-of-trial artefacts from influencing reading times of the critical second sentences. The initial set of 40 items was pre-tested to ensure the retrieval target and distractor nouns displayed the intended range of plausibilities. The pre-test was completed by 24 native English speakers who did not take part in any of the main experiments.

In the pre-test, the 40 items appeared in four conditions as in (7) that tested the plausibility of each target and distractor noun as an agent of the critical verb. The 40 items (including introduction and wrap-up sentences) were pseudo-randomised with 40 fillers, which also displayed a range of plausibilities, in a Latin-square design. Forward and reverse orders of each list were completed by the same number of participants. Participants were instructed to rate each text on a scale from 1 (highly implausible) to 7 (highly plausible).

(7a) The detective arrested the robber outside the city bank.
(7b) The cop arrested the robber outside the city bank.
(7c) The criminal arrested the robber outside the city bank.
(7d) The car arrested the robber outside the city bank

From these results, we selected 32 items that displayed the intended range of plausibilities (for a full list of experimental items, see https://osf.io/fqrh6/). For these 32 items, plausible target nouns ("the detective") and animate distractors ("the cop") received similarly high mean plausibility ratings, with mean ratings of 6.42 (SD = 0.47) and 6.51 (SD = 0.50) respectively. Implausible target nouns ("the criminal") and inanimate distractors ("the car") received lower ratings, with mean ratings of 2.33 (SD = 0.71) and 1.68 (SD = 0.88) respectively.

In the eye-tracking experiment, the 32 experimental items were randomised with 96 fillers. Experimental items appeared across three lines onscreen, with line breaks after the distractor noun and the end of the second sentence. Comprehension questions that required a yes-no push-button response, which did not probe interpretation of the critical dependency, were asked after all critical trials and two thirds of the fillers.

Procedure

Items were pseudo-randomised across four lists in a Latin-square design, with a different order being presented to each participant. Each list was completed by the same number of participants. While viewing was binocular, eye-movements were recorded from the right eye using an EYELINK 2000 sampling at 1000 Hz. Each session began with calibration on a nine-point grid. Recalibration between trials was conducted if required. Before each trial, participants fixated on a marker above the first word of the upcoming trial. Upon fixation on this marker, the text appeared. Participants read each text silently, pressing a button on a control pad once completed. The experiment lasted 30-45 min and began with six practice trials.

Table 1. Summary of eye-movement measures in milliseconds in Experiment 1 (standard errors in parentheses).

	First Pass Time	Regression Path Time	Total Viewing Time
Critical Region			
Plausible Sentence, Animate Distractor	269 (8)	413 (26)	418 (16)
Plausible Sentence, Inanimate Distractor	256 (6)	410 (27)	414 (14)
Implausible Sentence, Animate Distractor	269 (9)	368 (17)	452 (17)
Implausible Sentence, Inanimate Distractor	273 (7)	371 (19)	489 (18)
Spillover Region			
Plausible Sentence, Animate Distractor	748 (22)	1225 (46)	1110 (30)
Plausible Sentence, Inanimate Distractor	807 (25)	1275 (46)	1129 (31)
Implausible Sentence, Animate Distractor	789 (26)	1480 (63)	1228 (37)
Implausible Sentence, Inanimate Distractor	794 (25)	1644 (67)	1302 (38)

Data analysis

We report three reading time measures. *First pass times* sum fixations within a region when it is first entered from the left, while *regression path times* sum fixations, starting when a region is first entered from the left, up until it is exited to the right. Both measures index "early" processing, and while there is some debate about whether regression path times index "later" processing, it must index difficulty that was recognised when a region was first entered, even if it does include additional processing during regressions to earlier regions of text (Clifton et al., 2003). We also report *total viewing times*, the total time spent in a region of interest, as a global index of processing load. We report these three measures at two regions of text. The *critical region* consisted of the verb that triggered subject retrieval ("arresting"), while the *spillover region* consisted of the rest of the second sentence ("the robber outside the city bank").

Short fixations of 80 ms or below within one degree of visual arc of another fixation were merged. All other fixations of 80 ms or below, as well as those above 800 ms, were removed before analysis. Trials in which a region was skipped were treated as missing data for that region. Trials with excessive track loss were also removed before analysis. This accounted for less than 0.1% of the data.

Analysis was conducted using linear mixed-effects models with crossed random effects for subjects and items (Baayen et al., 2008). Reading times were log-transformed to minimise skew (see Vasishth & Nicenboim, 2016). Models included sum coded (−1, 1) fixed main effects of "region" (critical region vs. spillover region), "plausibility" (plausible sentence vs. implausible sentence) and "distractor" (animate distractor vs. inanimate distractor), along with their interactions. Each model was fit using the "maximal" random effects structure (Barr et al., 2013) that converged.[2] The maximal model included by-subject and by-item random intercepts, and random slopes for all fixed effects. As including region in the analysis involves two non-independent datapoints from the same trial, we also included a random intercept for trial. Given that "region" is the only repeated measure at trial level, "region" was the only random slope included under trial (see Cunnings & Sturt, 2018).

In the case of plausibility by distractor interactions, nested contrasts were conducted examining distractor effects at the two levels of plausibility. We do not discuss main effects of region below as these merely reflect differences in lexical material across the critical and spillover regions. However, any interactions with region provide insight into the time-course of processing and were also followed-up using appropriate nested contrasts. Datasets and analysis code for all experiments reported here can be found at https://osf.io/fqrh6/.

Results

Mean comprehension accuracy rates were 89% (all above 71%), indicating that participants paid attention during the task. Summaries of the eye-movement data and statistical analysis can be found in Tables 1 and 2 respectively.

No effects were significant in first pass times. In regression path times, which are illustrated in Figure 1a, there was a significant main effect of plausibility, with longer reading times in implausible sentences. There was also a significant region by plausibility interaction. Nested contrasts examining plausibility effects at each region indicated a significant plausibility effect at the spillover region only (critical region estimate = −0.015, SE = 0.02, $t = -0.97$, $p = .330$; spillover region estimate = 0.089, SE = 0.01, $t = 6.36$, $p < .001$). There was also a significant main effect of plausibility, with longer reading times in implausible than plausible sentences, in total viewing times in the absence of any further significant effects.

Table 2. Summary of the statistical analysis in Experiment 1.

	Model Estimate (SE)	*t*	*p*
First Pass Time			
Region	0.497 (0.03)	26.40	<.001
Plausibility	0.007 (0.02)	0.62	.538
Distractor	0.009 (0.01)	0.85	.398
Region * Plausibility	−0.004 (0. 01)	−0.44	.662
Region * Distractor	0.007 (0. 01)	−0.64	.520
Plausibility * Distractor	−0.001 (0. 01)	−0.05	.960
Region * Plausibility * Distractor	−0.016 (0. 01)	−1.40	.161
Regression Path Time			
Region	0.662 (0.04)	27.43	<.001
Plausibility	0.037 (0.02)	3.53	<.001
Distractor	0.017 (0.01)	1.36	.174
Region * Plausibility	0.052 (0.01)	5.18	<.001
Region * Distractor	0.019 (0.01)	1.67	.094
Plausibility * Distractor	0.010 (0.01)	0.88	.380
Region * Plausibility * Distractor	0.006 (0.01)	0.46	.647
Total Viewing Time			
Region	0.532 (0.04)	26.55	<.001
Plausibility	0.056 (0.02)	4.73	<.001
Distractor	0.019 (0.01)	1.85	.064
Region * Plausibility	−0.001 (0.01)	−0.15	.878
Region * Distractor	−0.001 (0.01)	−0.08	.938
Plausibility * Distractor	0.015 (0.01)	1.44	.151
Region * Plausibility * Distractor	−0.004 (0.01)	−0.49	.627

Discussion

The results of Experiment 1 yielded significant plausibility effects in both regression path times and total viewing times. We did not observe any significant effects of the distractor, nor any significant interactions between plausibility and the distractor in any measure. We note however that descriptively, Figure 1a illustrates a numerical pattern at the spillover region that is suggestive of facilitatory interference in implausible sentences, with numerically shorter reading times when the distractor is animate at the spillover region. However, we fully acknowledge that the relevant interactions are not significant. We return to this after reporting Experiment 2.

Experiment 2

The aim of Experiment 2 was to investigate memory retrieval during the processing of co-ordination rather than adjuncts. The materials in Experiment 2 were identical to those in Experiment 1, except the adjunct was replaced with a co-ordinated construction as in (8). As in Experiment 1, we factorially manipulated the plausibility of the retrieval target ("the detective"/"the criminal") and distractor noun ("the cop"/"the car") as an agent of the critical verb ("arrested").

(8) A bad crime had been committed. The detective/criminal stood by the cop/car very calmly and then arrested the robber outside the city bank. A lot of money had been stolen.

Our predictions were the same as in Experiment 1. We expected longer reading times for implausible then plausible sentences. The activation-based cue-based model (Lewis et al., 2006; Lewis & Vasishth, 2005) would predict facilitatory in implausible sentences, with shorter reading times when the distractor is animate ("the cop") rather than inanimate ("the car"), and inhibitory interference in plausible experiments, with longer reading times following animate rather than inanimate distractors.

Participants

48 native English speakers from the University of Reading, none of whom took part in Experiment 1, took part in Experiment 2 for course credit or a nominal sum.

Materials

The 32 experimental items from Experiment 1 were adapted as in (8). All items were identical to Experiment

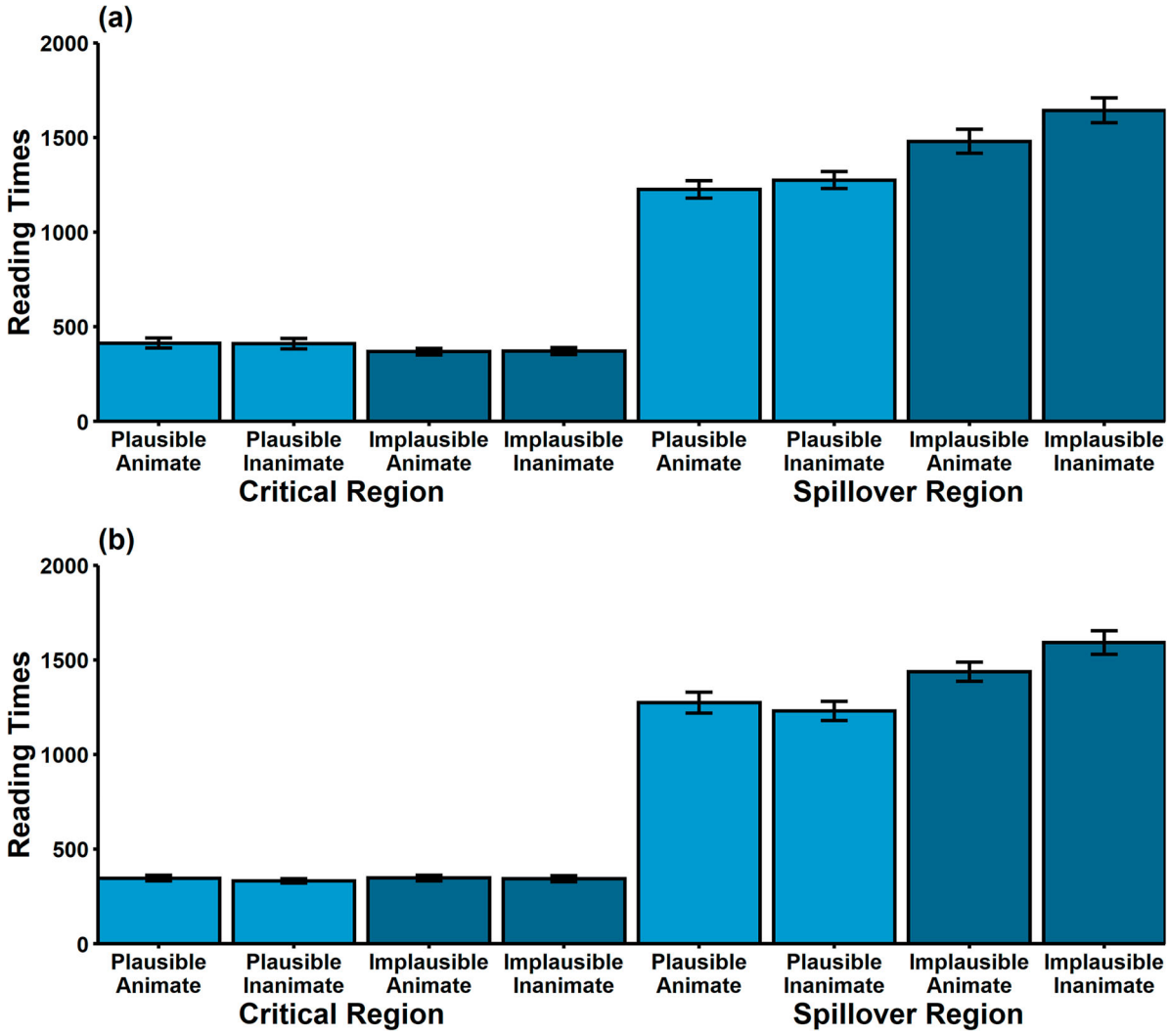


Figure 1. Regression path duration in milliseconds in (a) Experiment 1 and (b) Experiment 2 (Im/plausible = Im/plausible sentence; In/animate = In/animate distractor).

1 except that the critical verb that triggered retrieval was now co-ordinated with the rest of the critical sentence rather than being in an adjunct. Experimental items were again interspersed with 96 fillers.

Procedure and data analysis

The procedure and data analysis was the same as in Experiment 1. Less than 0.1% of trials were removed due to excessive track loss.

Results

Average comprehension question accuracy was 91% (all above 79%). Summaries of the eye-movement data and statistical analysis are provided in Tables 3 and 4.

No effects were significant in first-pass times. In regression path times, which are illustrated in Figure 1b, there was a significant main effect of plausibility and a significant interaction between region and plausibility. Nested contrasts examining plausibility effects at each region indicated a significant plausibility effect, with longer reading times in implausible than plausible sentences, at the spillover region only (critical region estimate < −0.001, SE = 0.01, $t = -0.03$, $p = .979$; spillover region estimate = 0.091, SE = 0.01, $t = 6.64$, $p < .001$). Total viewing times indicated a significant plausibility effect in the absence of any further significant effects.

Discussion

The results of Experiment 2 in many ways mirror the results of Experiment 1. We observed significant plausibility effects, with longer reading times in implausible sentences, in both regression path times and total viewing times. Descriptively, as shown in Figure 1b,

Table 3. Summary of eye-movement measures in milliseconds in Experiment 2 (standard errors in parentheses).

	First Pass Time	Regression Path Time	Total Viewing Time
Critical Region			
Plausible Sentence, Animate Distractor	263 (8)	346 (15)	384 (14)
Plausible Sentence, Inanimate Distractor	277 (9)	332 (11)	390 (13)
Implausible Sentence, Animate Distractor	265 (8)	347 (15)	439 (24)
Implausible Sentence, Inanimate Distractor	279 (8)	344 (16)	446 (18)
Spillover Region			
Plausible Sentence, Animate Distractor	839 (28)	1274 (56)	1170 (35)
Plausible Sentence, Inanimate Distractor	819 (23)	1230 (51)	1138 (34)
Implausible Sentence, Animate Distractor	862 (25)	1437 (51)	1298 (36)
Implausible Sentence, Inanimate Distractor	852 (27)	1591 (62)	1319 (39)

regression path times at the spillover region in Experiment 2 are suggestive of facilitatory interference in implausible sentences. However, as in Experiment 1, we acknowledge that no effects of distractor or interactions with distractor were significant.

Combined analysis of experiments 1 and 2

Despite the lack of significant distractor effects in our individual analyses of Experiments 1 and 2, numerically both experiments are suggestive of facilitatory interference in implausible sentences in regression path times at the spillover regions. Based on these descriptive observations in Experiments 1 and 2, we ran an additional, exploratory analysis of regression path times at the spillover region that combined the data from both experiments to maximise statistical power. To estimate plausibility and distractor effects Experiments 1 and 2, we used Bayesian mixed-effects models. We chose a Bayesian rather than frequentist analysis here as the Bayesian approach allows us to estimate the possible range of effect sizes for plausibility and distractor effects across experiments, given the data and priors (see e.g. Nicenboim & Vasishth, 2016; Vasishth et al., 2018).

Table 4. Summary of the statistical analysis in Experiment 2.

	Model Estimate (SE)	*t*	*p*
First Pass Time			
Region	0.524 (0.03)	19.07	<.001
Plausibility	0.013 (0.01)	1.41	.159
Distractor	0.009 (0.01)	0.82	.411
Region * Plausibility	0.005 (0.01)	0.56	.577
Region * Distractor	−0.180 (0.01)	−1.88	.060
Plausibility * Distractor	−0.010 (0.01)	−0.97	.332
Region * Plausibility * Distractor	−0.010 (0.01)	−0.90	.369
Regression Path Time			
Region	0.678 (0.02)	29.49	<.001
Plausibility	0.045 (0.01)	4.26	<.001
Distractor	0.008 (0.01)	0.60	.549
Region * Plausibility	0.046 (0.01)	4.74	<.001
Region * Distractor	0.004 (0.01)	0.33	.741
Plausibility * Distractor	0.008 (0.01)	0.85	.394
Region * Plausibility * Distractor	0.013 (0.01)	1.06	.284
Total Viewing Time			
Region	0.573 (0.03)	21.85	<.001
Plausibility	0.055 (0.01)	5.02	<.001
Distractor	0.010 (0.01)	0.81	.416
Region * Plausibility	0.006 (0.01)	0.87	.386
Region * Distractor	−0.013 (0.01)	−1.81	.070
Plausibility * Distractor	0.003 (0.01)	0.26	.789
Region * Plausibility * Distractor	0.003 (0.01)	0.49	.625

The analysis was conducted using the brms package in R (Bürkner, 2018). The dependent variable was the (log-transformed) spillover region regression path times. The model included sum-coded (−1/1) fixed effects of experiment (Experiment 1 vs. Experiment 2), plausibility (plausible sentence vs. implausible sentence), distractor (animate distractor vs. inanimate distractor) and their interactions. By-subject and by-item random effects were included using the maximal random effects structure in which Experiment was treated as a between-subject but within-item manipulation. The model was fit assuming a Gaussian distribution with weakly informative priors, including a normal(0,10) prior for the intercept and normal(0,1) priors for fixed effects and standard deviations, along with a so-called LKJ prior with a value of 2 for the correlation of random effects (Sorensen et al., 2016). The model was run for four sampling chains each with 2000 iterations. Convergence was checked by visual inspection of model chains and ensuring R-hats were close to 1.00. We report model estimates, 95% credible intervals (CrI) and the probability that each effect is in the observed direction. We interpret effects in which the 95% credible interval excludes zero as providing strong support for that effect. For effects in which the 95% credible interval includes zero, the probability allows us to quantify in a graded manner how much support there is for the effect being in the observed direction, given the model and data.

Figure 2a plots the posterior distributions of the effects of most theoretical interest across experiments, namely the effect of sentence plausibility, distractor animacy, their interaction and the 3-way interaction with experiment. In the analysis, the estimate for the main effect of experiment provided no clear support for overall differences in reading times across experiments (estimate = −0.013, SE = 0.033, 95% CrI [−0.80, 0.051], $P(\beta < 0) = .60$). As shown in Figure 2a, there was a clear main effect of plausibility (estimate = 0.089, SE = 0.011, 95% CrI [0.068, 0.111], $P(\beta > 0) = .98$), with longer reading times in implausible sentences. The effect of plausibility did not vary across experiments

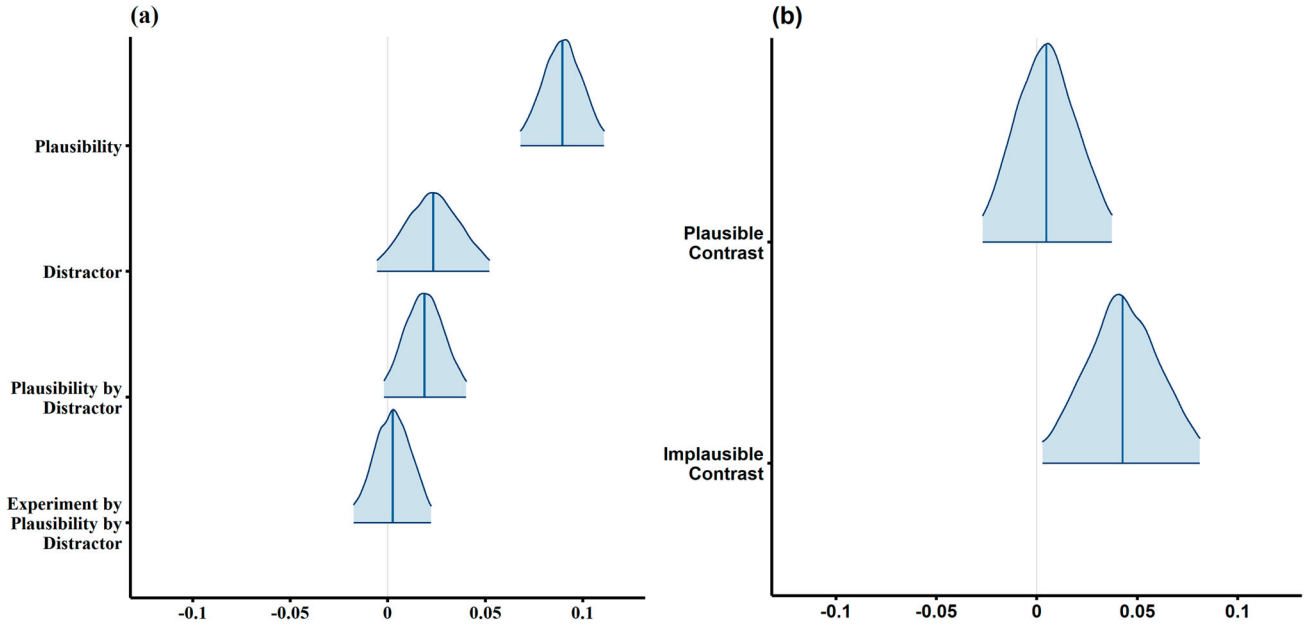


Figure 2. Posterior distributions for combined analysis of Experiments 1 and 2 showing (a) main effects and interactions of main theoretical interest; (b) nested contrasts for plausible and implausible sentences.

(estimate < 0.001, SE = 0.010, 95% CrI [−0.020, 0.021], $P(\beta > 0) = .55$). As also shown in Figure 2a, though the credible interval included zero, there was also some support for a main effect of distractor (estimate = 0.023, SE = 0.015, 95% CrI [−0.006, 0.052], $P(\beta > 0) = .94$), with longer reading times when the distractor was inanimate. Though the credible interval again crossed zero, there was also some support for an experiment by distractor interaction (estimate = −0.013, SE = 0.011, 95% CrI [−0.034, 0.010], $P(\beta < 0) = .87$), with the main effect of distractor tending to be larger in Experiment 1 than Experiment 2 (compare Figure 1a and b). Importantly, as illustrated in Figure 2a, although the credible interval just crossed zero, there was also support for a plausibility by distractor interaction (estimate = 0.019, SE = 0.011, 95% CrI [−0.002, 0.040], $P(\beta > 0) = .96$). The lack of a 3-way interaction (estimate = 0.002, SE = 0.010, 95% CrI [−0.018, 0.022], $P(\beta > 0) = .60$) suggests there was no evidence to indicate that this two-way interaction varied across experiments.

Based on this interaction between plausibility by distractor, we ran an additional model containing nested contrasts examining distractor effects in plausible and implausible sentences. These nested contrasts are illustrated in Figure 2b. For implausible sentences, reading times were shorter when the distractor was animate compared to when it was inanimate, with the credible interval of this difference excluding zero (estimate = 0.043, SE = 0.020, 95% CrI [0.003, 0.081], $P(\beta > 0) = .98$). For plausible sentences, there was no clear difference between the animate and inanimate distractor conditions (estimate = 0.005, SE = 0.016, 95% CrI [−0.027, 0.037], $P(\beta > 0) = .61$). These results suggest facilitatory interference in implausible sentences only, as evidence of an illusion of plausibility.

In sum, we contend that this combined analysis provides evidence of illusions of plausibility in adjuncts and co-ordination, that conceptually replicate the findings reported by Cunnings and Sturt (2018) for filler-gap dependencies. We acknowledge however that this is only observed in an additional combined analysis of regression path times at the spillover region. To further assess the replicability of these findings, we ran two additional web-based self-paced reading studies using the experimental materials from Experiments 1 and 2 but with considerably larger sample sizes.

Experiment 3

Experiment 3 examined interference during the processing of adjuncts using the same experimental materials as used in Experiment 1. The materials were adapted to phrase-by-phrase self-paced reading as in (9). Here, [] denote the phrases as presented to participants. We again manipulated sentence plausibility and distractor animacy. Our predictions were the same as Experiments 1 and 2.

(9) [A bad crime had been committed.][The detective/ criminal stood by the cop/car][very calmly][after

arresting][the robber outside the city bank.][A lot of money had been stolen.]

Participants

Participants included 192 native English speakers that were recruited online via Prolific and paid a small fee. We ran the experiment until we had 192 participants who scored over 70% on comprehension questions, as a threshold to indicate that they paid attention during the task. An additional four participants were tested but not included in the analysis due to comprehension accuracy rates below 70%.

Materials

The experimental materials were identical to Experiment 1. Experimental texts were displayed across three lines on-screen, with line breaks as in Experiments 1 (after the distractor noun and the second sentence). As noted above, the [] marks in (9) denote how the texts were divided into phrases for presentation. 80 filler texts of varying lengths were also constructed. Comprehension questions requiring a yes/no push button response were asked after each experimental item and two-thirds of the fillers. The comprehension questions to experimental items were the same as those used in Experiment 1.

Procedure

The experiment was administered online using IbexFarm. Participants read each sentence in a non-cumulative phrase-by-phrase self-paced reading fashion. At the start of each trial, the text was masked by a series of underscores. By pressing the space-bar, the first phrase appeared, and subsequent phrases appeared by pressing the space bar again. Comprehension questions were then shown, which participants answered by pressing either the "1" or "2" keys, which corresponded to "yes' and "no" responses. The experiment began with four practice trials.

Experimental and filler items were pseudo randomised in a Latin-square design across four experimental lists. Each list was completed by the same number of participants. The entire experiment lasted approximately 30 min.

Data analysis

We analysed reading times at the critical region ("after arresting") and spillover region ("the robber outside the city bank"). Before analysis, extremely short reaction times (< = 100 ms) and extremely long reaction times (> = 10,000 ms) were removed, as these likely index lapses in attention. This affected less than 0.01% of the data. As in Experiments 1 and 2, reading times were log-transformed to remove skew, and the analysis used mixed-effects models with sum coded fixed effects of region, plausibility, distractor and all relevant interactions, fit using the maximal random effects structure that converged.

Results

Average comprehension accuracy rates for the 192 participants included in the analysis was 92% (all above 70%). Reading times at the critical and spillover regions are illustrated in Figure 3a, while a summary of the statistical analysis is provided in Table 5.

Analysis revealed a significant main effect of plausibility and also a significant region by plausibility interaction. Nested contrasts indicated significantly longer reading times for implausible than plausible sentences at the spillover region only (critical region estimate = 0.001, SE < 0.01, $t = 0.22$, $p = .828$; spillover region estimate = 0054, SE = 0.01, $t = 6.83$, $p < .001$). The main effect of distractor was not significant, but the region by distractor interaction was, however nested contrasts testing the main effect of distractor at each region were not significant (critical region estimate = −0.005, SE = < 0.01, $t = -1.18$, $p = .237$; spillover region estimate = 0.010, SE = 0.01, $t = 1.55$, $p = .120$). Importantly, there was a significant plausibility by distractor interaction. Nested contrasts showed that reading times were significantly shorter in implausible sentences when the distractor was animate rather than inanimate (estimate = 0.011, SE = 0.01, t = 2.02, $p = .044$), while the two plausible conditions did not differ significantly (estimate = −0.006, SE = 0.01, t = −1.02, $p = .309$). These findings indicate facilitatory interference in implausible sentences. Although descriptively these effects are largely carried by differences at the spillover region (see Figure 3a), the 3-way interaction was not significant.

Discussion

The results of Experiment 3 provide clear evidence for illusions of plausibility in implausible sentences, as evidenced by the significant interaction between sentence plausibility and distractor animacy. The pattern of results is very similar to the combined analysis of Experiments 1 and 2, and conceptually replicate the illusions of plausibility reported by Cunnings and Sturt (2018). While we observed significant facilitatory interference in implausible sentences, we did not find evidence of any

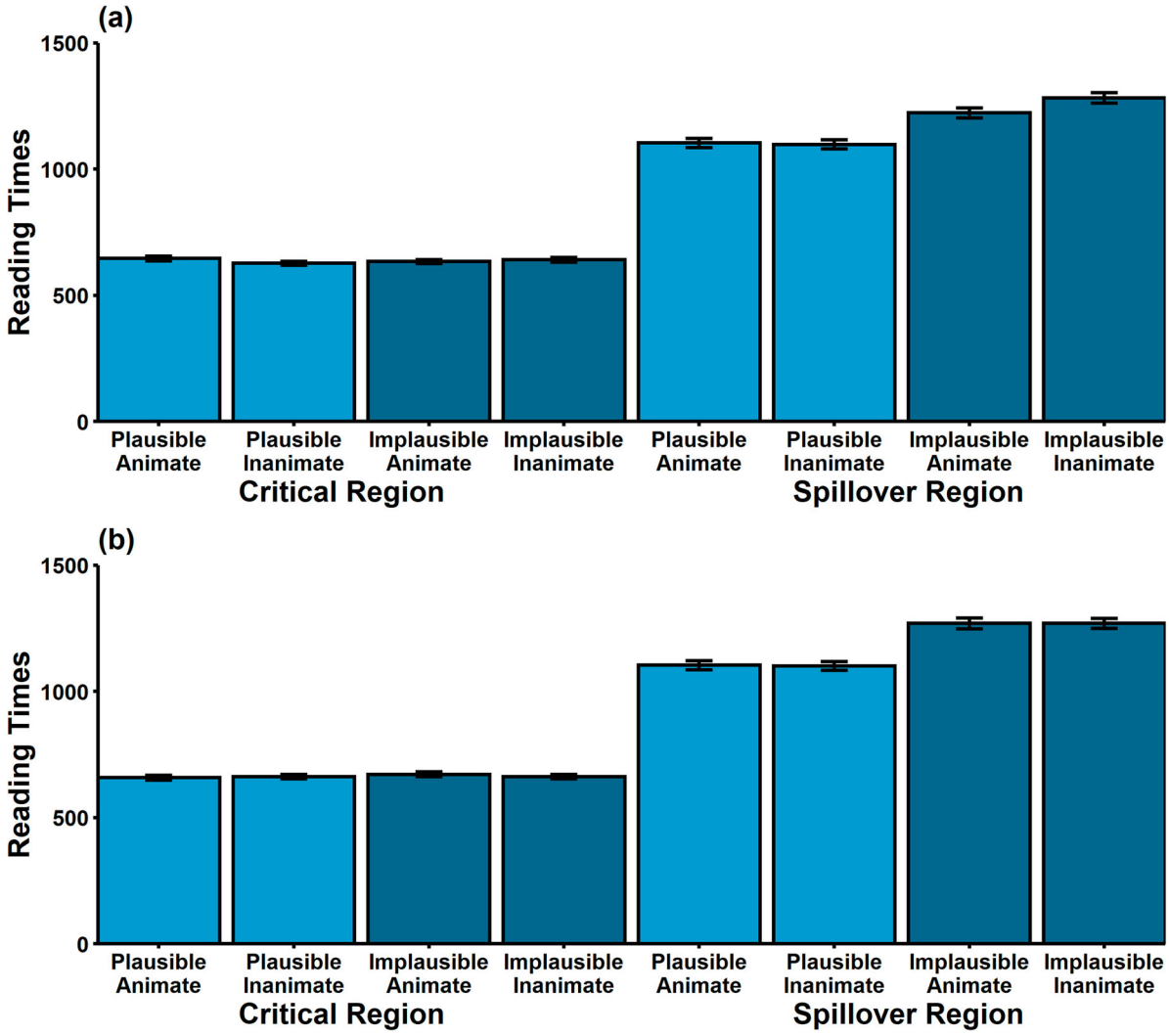


Figure 3. Self-paced reading times in milliseconds in (a) Experiment 3 and (b) Experiment 4 (Im/plausible = Im/plausible sentence; In/animate = In/animate distractor).

significant differences in plausible sentences. We return to this in the General Discussion, but first report Experiment 4.

Experiment 4

Experiment 4 tested interference during the processing of co-ordination using the same experimental items as in Experiment 2. As in Experiment 3, phrase-by-phrase self-paced reading was adopted. Phrases were divided in the same way as Experiment 3, as illustrated in (10).

(10) [A bad crime had been committed.][The detective/criminal stood by the cop/car][very calmly][and then arrested][the robber outside the city bank.][A lot of money had been stolen.]

Table 5. Summary of the statistical analysis in Experiment 3.

	Model Estimate	*(SE)*	*t*	*p*
Region	0.277	(0.01)	19.30	<.001
Plausibility	0.027	(< 0.01)	5.72	<.001
Distractor	0.002	(< 0.01)	0.54	.592
Region * Plausibility	0.026	(< 0.01)	7.18	<.001
Region * Distractor	0.008	(< 0.01)	2.37	.018
Plausibility * Distractor	0.009	(< 0.01)	2.35	.019
Region * Plausibility * Distractor	0.003	(< 0.01)	0.29	.290

Our design and analysis plan for Experiment 4, which was conducted after Experiments 1-3, were pre-registered before data collection began (see https://osf.io/ent8j). As discussed below, our pre-registration aimed to copy the procedure and analysis used in Experiment 3.

Participants

As was pre-registered, we ran the experiment until we had recruited 192 native English speakers who met the threshold of answering the comprehension questions with an average accuracy of at least 70%. One participant was recruited but not analysed as they failed to meet this threshold. Participants were recruited online via Prolific and paid a small fee. None of the participants had taken part in Experiment 3.

Materials

The materials included 32 experimental items as in (10) and 80 fillers. As in Experiment 3, all experimental items and two-thirds of the fillers were followed by yes/no comprehension questions.

Procedure and data analysis

All aspects of the procedure and data analysis were pre-registered to be the same as in Experiment 3. Less than 0.01% of the data were removed due to trimming of extremely short (< = 100 ms) and extremely long (> = 10,000 ms) reaction times.

Results

Average comprehension accuracy rates for the 192 participants included in the analysis was 92% (all above 70%). Reading times at the critical and spillover regions are illustrated in Figure 3b, and a summary of the statistical analysis is provided in Table 6.

Analysis revealed a significant main effect of plausibility and a significant region by plausibility interaction, and nested contrasts indicated a significant plausibility effect at the spillover region only (critical region estimate = 0.002, SE = < 0.01, $t = 0.53$, $p = .596$; spillover region estimate = 0.062, SE = 0.01, $t = 8.56$, $p < .001$). Neither the main effect of distractor, nor any further interactions were significant. Indeed, as illustrated in Figure 3b, there is little, if any, discernible effect of distractor at either region in Experiment 4.

Table 6. Summary of the statistical analysis in Experiment 4.

	Model Estimate (SE)		*t*	*p*
Region	0.260	(0.01)	18.88	<.001
Plausibility	0.032	(< 0.01)	9.42	<.001
Distractor	0.001	(< 0.01)	0.15	.882
Region * Plausibility	0.030	(< 0.01)	7.62	<.001
Region * Distractor	< 0.001	(< 0.01)	−0.17	.862
Plausibility * Distractor	< 0.001	(< 0.01)	< 0.01	.999
Region * Plausibility * Distractor	0.003	(< 0.01)	1.06	.290

Combined analysis of experiments 3 and 4

Our pre-registration of Experiment 4 also included a combined Bayesian analysis of the already collected data from Experiment 3, and the to-be-collected data in Experiment 4. This pre-registered analysis aimed to mimic the analysis we had already conducted for regression path times at the spillover region in Experiments 1 and 2. Specifically, we pre-registered a Bayesian analysis of self-paced reading times at the spillover region in Experiments 3 and 4. All other aspects of the analysis were the same as the combined analysis of spillover region regression path times in Experiments 1 and 2.[3]

The posterior distributions of the effects of most theoretical interest are shown in Figure 4a. The estimate for the main effect of experiment suggested similar overall reading times in Experiments 3 and 4 (estimate < 0.001, SE = 0.019, 95% CrI [−0.037, 0.037], $P(\beta > 0) = .39$). There was clear support for a main effect of plausibility (estimate = 0.058, SE = 0.007, 95% CrI [0.044, 0.071], $P(\beta > 0) = .99$), which did not vary across experiments (estimate = 0.004, SE < 0.004, 95% CrI [−0.004, 0.012], $P(\beta > 0) = .48$), indicating longer reading times in implausible than plausible sentences across both experiments. There was also some support for a main effect of distractor, though the credible interval crossed zero (estimate = 0.005, SE = 0.005, 95% CrI [−0.005, 0.015], $P(\beta > 0) = .85$), and also an experiment by distractor interaction, though again the credible interval crossed zero (estimate = −0.005, SE = 0.004, 95% CrI [−0.013, 0.003], $P(\beta < 0) = .89$). Most importantly, there was clear support of a plausibility by distractor interaction, with the credible interval excluding zero (estimate = 0.007, SE = 0.004, 95% CrI [< 0.001, 0.014], $P(\beta > 0) = .98$), and, although the credible interval crossed zero, some support for the experiment by plausibility by distractor interaction (estimate = −0.004, SE = 0.004, 95% CrI [−0.012, 0.003], $P(\beta < 0) = .87$).

Based on these findings, we conducted an analysis with nested contrasts testing for distractor effects in plausible and implausible sentences, for each experiment separately. These are shown in Figure 4b. In Experiment 3, reading times were shorter in implausible sentences when the distractor was animate compared to inanimate (estimate = 0.022, SE = 0.008, 95% CrI [0.005, 0.038], $P(\beta > 0) > .99$), while reading times for plausible sentences were similar (estimate = −0.001, SE = 0.008, CrI [−0.018, 0.0015], $P(\beta < 0) = .57$). In Experiment 4, there was no support for any effects of distractor in either plausible or implausible sentences (for plausible sentences estimate = −0.003, SE = 0.008, CrI [−0.018, 0.012], $P(\beta < 0) = .64$; for implausible sentences estimate = 0.003, SE = 0.008, CrI [−0.014, 0.020], $P(\beta > 0) = .65$). These findings indicate an illusion of plausibility

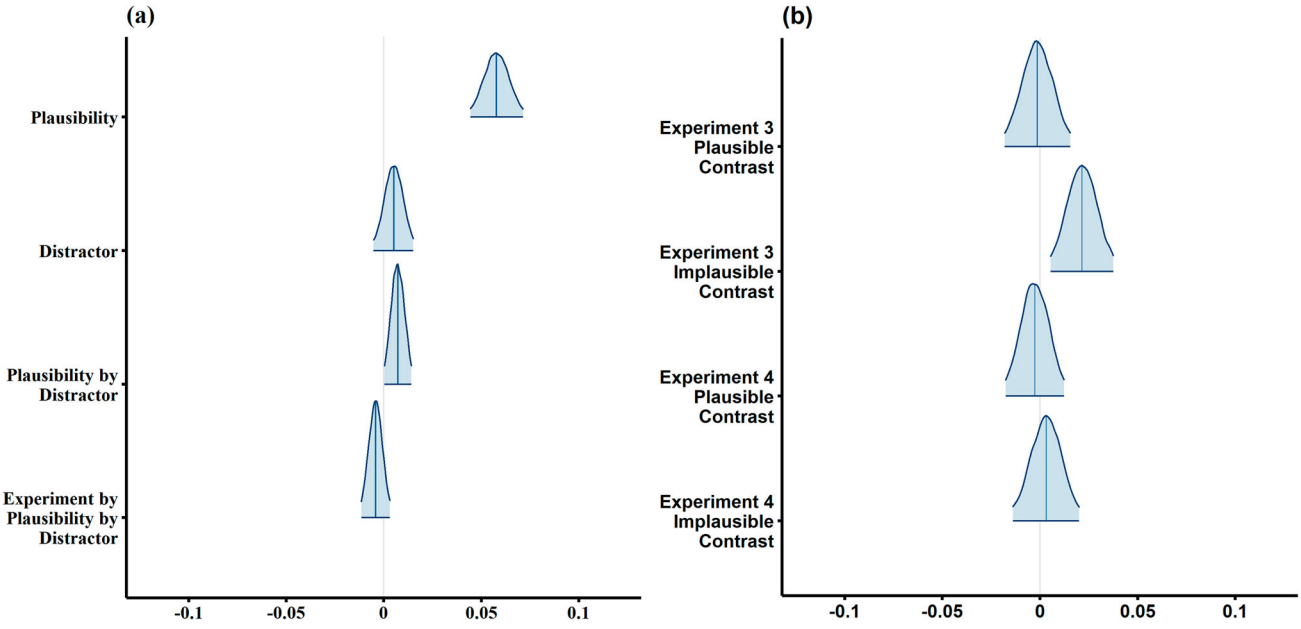


Figure 4. Posterior distributions for combined analysis of Experiments 3 and 4 showing (a) main effects and interactions of main theoretical interest; (b) nested contrasts for plausible and implausible sentences in each experiment.

in implausible sentences in Experiment 3 but not Experiment 4. We discuss these findings, along with Experiments 1 and 2, in more detail below.

General discussion

In four experiments, we examined the processing of adjuncts and co-ordination to test the extent to which linguistic dependency resolution during language comprehension is susceptible to interference. Although we acknowledge that effects indicative of interference were not significant in each experiment individually, we maintain that our combined analyses suggest illusions of plausibility that conceptually replicate and extend previous findings of such effects in other dependencies (Cunnings & Sturt, 2018; Fujita & Cunnings, 2022; Laurinavichyute & von der Malsburg, 2022). These illusions however appear more robust in adjuncts than in co-ordination. Below, we discuss these results in relation to models of interference effects during sentence processing, and memory retrieval during the processing of adjuncts and co-ordination.

Illusions of plausibility and retrieval interference

In Experiments 1 and 2, participants' eye-movements indicated longer reading times when a sentence subject was an implausible, rather than plausible, agent of a verb in a temporal adjunct or co-ordinated construction. Although the plausibility by distractor interaction was not significant in either experiment, regression path times at the spillover region were numerically shorter in implausible sentences when the distractor was plausible in both experiments. Our combined Bayesian analysis of Experiments 1 and 2 supported this observation and indicated an illusion of plausibility when assessed across the two constructions. This illusion was robust in Experiment 3, which used self-paced reading and a larger sample, but we did not observe it in Experiment 4. Thus, though we acknowledge the effect was not observed in each experiment individually, we maintain that the overall pattern of results across experiments suggests illusions of plausibility. This facilitatory interference effect in implausible sentences is compatible with the predictions of the activation-based cue-based model of memory retrieval during sentence processing (Jäger et al., 2017; Lewis et al., 2006; Vasishth et al., 2019).

More specifically, we interpret this facilitatory interference effect as indicating that structural and semantic cues combine to guide memory retrieval during the resolution of adjuncts and co-ordination. We assume the critical verbs cue retrieval for an item in memory that matches the relevant required syntactic property (or properties) of the verb. In our experiments, the critical verbs required a subject, and as such we assume a syntactic cue for [+SUBJECT]. We also assume that semantic retrieval cues, derived from the lexical properties of individual verbs, also guide retrieval. For a verb such as "arresting" or "arrested" as in (6/8), this will include a semantic cue [+ARRESTER], which guides towards a semantically plausible agent. Whether the

relevant semantic information should be considered a single cue or is best described as a bundle of semantic cues, is a question we leave open (for discussion, see Cunnings & Sturt, 2018; Smith & Vasishth, 2020). For present purposes we suffice to say that a semantic cue or cues, derived from the lexical properties of individual verbs, interact with the syntactic cue [+ SUBJECT].

In implausible sentences, the sentence subject matches the syntactic cue [+ SUBJECT] but mismatches the semantic cue derived from the verb. The distractor never matches the syntactic cue, but in the animate condition satisfies the semantic cues. It is this partial match which leads to facilitatory interference in implausible sentences when the distractor is animate (and a plausible agent of the verb) compared to when it is inanimate.

The activation-based cue-based model also predicts inhibitory interference, such that reading times of plausible sentences in our experiments are predicted to be longer when the distractor is animate rather than inanimate. That is, when the sentence subject is a plausible agent of the critical verb (e.g. "detective"), inhibitory competition should occur between it and the distractor when the distractor itself is also plausible ("cop"). We however did not observe this effect in any of our experiments.

Nicenboim et al. (2018) argued that many studies do not report inhibitory interference due to small samples and lack of power. However, we did not observe inhibitory interference in Experiments 1 and 2, either in the individual analyses or combined analysis of both experiments. We also did not find inhibitory interference in our analysis of Experiments 3 and 4. Note that our combined analysis of Experiments 3 and 4 includes data from 384 participants, which is over double the 184 participants tested by Niceonboim et al., and yet we still did not observe reliable inhibitory interference. As such, we contend that our null effects here are unlikely to result from lack of power.

As mentioned in the Introduction, in their study on subject-verb agreement attraction, Wagers et al. (2009) argued that the lack of effects in their grammatical conditions may be retrieval is only initiated as a reanalysis process once ungrammaticality is detected. Our results are potentially compatible with this proposal, if one assumes that detection of implausibility triggers retrieval as a reanalysis process in a similar way to ungrammaticality detection (see also Fujita & Cunnings, 2022).

Some cue-based accounts have argued that syntactic cues may gate or highly bias interference to cases when a distractor matches such cues (Van Dyke & McElree, 2011) or that interference is restricted to cases when a distractor is an oblique argument (Parker & An, 2018). In our experiments, the distractor did not match the syntactic [+ SUBJECT] cue and was a core rather than oblique argument. The facilitatory interference effects we observed, especially in Experiment 3, we argue are not compatible with strong versions of these claims, that interference is restricted to constituents that match syntactic retrieval cues and/or are oblique arguments. We acknowledge however that although interference was robustly observed in Experiment 3, across experiments the interference effects we observed were quite weak. This might support a weaker version of such claims, that syntactic cues are highly weighed during retrieval, or that core arguments resist, but are not impervious to, interference effects.

We also believe that our results are difficult to explain from the perspective of representational models (Bock & Eberhard, 1993; Bock & Middleton, 2011; Hammerly et al., 2019). Representational models typically predict attraction based on morphological features such as number, and thus these models do not readily predict semantic interference at all. One might consider whether it is possible to extend representational accounts to include semantic features, but this would be a considerable rework of the theory. If it were possible to extend representational accounts to semantic features, they would also most obviously predict longer reading times for conditions with inanimate distractors in both plausible and implausible sentences, however the semantic illusions we observed were restricted to implausible sentences. Finally, some representational accounts, such as feature percolation, typically require distractors to be embedded within a noun phrase to cause attraction (e.g. "the key to the cabinets …"), but in our materials the distractor was not embedded within the target noun phrase and did not modify it in any way. For these reasons, we believe the illusions of plausibility we found are difficult to explain from the perspective of representational models, and instead maintain that they are more compatible with retrieval-based accounts.

Recently, Yadav et al. (2023) claimed that the null effects observed in grammatical sentences in studies on subject-verb agreement may indicate that models of linguistic dependency resolution need to incorporate both cue-based interference as predicted by retrieval-based accounts, and feature misspecification as predicted by representational models. That is, in grammatical sentences retrieval-based models predict inhibitory interference ("The key to the cabinet was ….") should have longer reading times than "The key to the cabinets was …") while representational models predict attraction effects in the opposite direction. Yadav et al. claimed these two effects cancel each other out, yielding

null effects in grammatical sentences. Applying this logic to our study, the null effects in plausible sentences might be a result of conflicting inhibitory interference and (semantic) attraction effects. We are cautious in drawing this conclusion however. As discussed above, representational accounts were developed to account for misspecification of morphosyntactic features, and it is not clear that such models can be extended to the semantic features manipulated in this study. Additionally, some representational accounts, such as feature percolation, require distractors to be embedded within the target noun phrase to cause attraction, which was not the case in our study. For Yadav et al.'s account to be extended to our results, we would need to posit semantic attraction effects compatible with representational accounts that are as strong as the inhibitory interference effects predicted by cue-based models. Given it is unclear if representational models should be generalised to semantic features, and given some representational accounts predict little or no attraction in cases when a distractor is not embedded within a target noun phrase, it is unclear whether Yadav et al.'s proposal can account for the null effects we observed in plausible sentences.

Memory retrieval in adjuncts and co-ordination

To date, very few studies have investigated linguistic dependency resolution during the processing of adjuncts and co-ordination, and our study contributes to our understanding of memory retrieval operations during the processing of these constructions. Parker et al. (2015) previously reported facilitatory interference in the processing of temporal adjuncts using an animacy manipulation. The facilitatory interference we observed using a plausibility diagnostic, especially in Experiment 3, provides further support that the resolution of temporal adjuncts is susceptible to interference, as predicted by cue-based models.

While Parker et al. observed facilitatory interference in an experiment that manipulated the animacy of the sentence subject and a distractor ("The doctor/discovery that the researcher/the report described meticulously was certified after debunking the urban myth himself in the new scientific journal"), they did not find evidence of interference in a further experiment that involved only animate subjects/distractors and instead manipulated gender (mis)match between constituents and a subsequent reflexive ("The harpist/drummer that the diva/guitarist liked very much was congratulated after playing the beautiful song herself at the brand new recording studio"). Parker et al. considered two possible accounts of their cross-experiment findings. Firstly, they considered that the lack of interference in their second experiment might be because animacy is a strongly weighted cue to memory retrieval, such that facilitatory interference is only observed when the sentence subject does not match the [+ANIMATE] requirements of the verb, as in their first experiment where facilitatory interference was observed when the sentence subject was inanimate. Our results go against this interpretation of their results. Given that we observed facilitatory interference in experimental sentences when the sentence subject was always animate, it cannot be the case that [+ANIMATE] is such a strongly weighted retrieval cue that facilitatory interference is only observed when it is violated. Parker et al. also considered that facilitatory interference only occurs when the retrieval target (the sentence subject in this case), is a particularly poor match to the set of retrieval cues. Our results are more compatible with this claim, such that facilitatory interference in the resolution of temporal adjuncts is possible even if the sentence subject is animate, if it provides a particularly poor match with the semantic properties of the verb.

Although we maintain our results indicate facilitatory interference in the resolution of temporal adjuncts, our results for co-ordination in particular warrant further discussion, given the lack of facilitatory interference, even descriptively, in Experiment 4. We are unaware of any existing study examining memory retrieval during the processing of co-ordination, and as such comparisons with existing literature are not possible. Our combined Bayesian analysis of Experiments 1 and 2 suggests semantic interference across adjuncts and co-ordination. Our combined Bayesian analysis of Experiments 3 and 4 however provides support for facilitatory interference in adjuncts in Experiment 3, but we did not observe this effect in co-ordination in Experiment 4. Here, we consider two possible interpretations of these conflicting findings.

Firstly, it might be that co-ordinated constructions resist interference in a way that temporal adjuncts do not. Indeed, the extent to which different linguistic dependencies are similarly susceptible to interference has been debated. This is most evident in research comparing subject-verb agreement and reflexives, where some have claimed that syntactic constraints are more highly weighted in reflexive resolution such that reflexive dependencies resist interference in a way that agreement does not (Dillon et al., 2013). Although more recent research has contested this claim about reflexives (Jäger et al., 2020), it is at least possible that syntactic and semantic retrieval cues are weighted differently across dependencies. In this case, it might be that the syntactic [+SUBJECT] cue is more highly

weighted in co-ordination than adjuncts, such that co-ordination resists interference from distractors that do not match this cue. Why this might be is an open question. In linguistic theory, although different accounts have been proposed, the interpretation of non-finite verbs in temporal adjuncts ("The boy spoke to the girl after walking to school") is often predicted to be resolved via a null anaphoric element PRO (Chomsky, 1981). PRO is however typically not assumed to be possible as the subject of a finite verb, as in the co-ordinated construction "The boy spoke to the girl and then walked to school", but how to account for interpretation in such cases is debated (Burton & Grimshaw, 1992; Van Valin, 1986). Burton and Grimshaw (1992) argued that interpretation here is resolved via movement. While we are unaware of principled theoretical reasons for why anaphoric PRO but not movement should be susceptible to interference, it is at least possible that the weightings of syntactic retrieval cues may differ between these two dependencies.

Another second possibility is that both adjuncts and co-ordination are indeed susceptible to interference, as would be predicted by cue-based parsing. This would be compatible with our combined Bayesian analysis of Experiments 1 and 2, which found facilitatory interference in an analysis across constructions, and would be compatible with previously reported illusions of plausibility during reading (Cunnings & Sturt, 2018; Fujita & Cunnings, 2022). In this case, to be compatible with cue-based parsing, even if interference effects are weaker in co-ordination than adjuncts, we would need to contend that the lack of facilitatory interference in Experiment 4 may be a Type II error. Alternatively, if facilitatory interference is indeed small in co-ordination, self-paced reading may not be sensitive enough to observe it. We acknowledge however that further research is required to tease these two possible accounts apart, and to assess the extent to which subject retrieval in co-ordinated constructions in particular is susceptible to interference.

Conclusion

We investigated linguistic dependency resolution during the processing of adjuncts and co-ordination. We observed plausibility effects indicating attempted retrieval of the sentence subject in both dependencies. These plausibility effects were however attenuated when a distractor constituent was itself a plausible subject of the verb, indicating an illusion of plausibility. This illusion of plausibility was however more robust in adjuncts than in co-ordination, which might indicate that structural cues are weighted differently in the two dependencies. Though we acknowledge this effect was not robust across each experiment, we maintain that our results across experiments support cue-based models of memory retrieval during sentence processing. Specifically, we argue that our results indicate that memory retrieval during language comprehension involves the interaction of both structural cues derived from the local syntactic context, and semantic cues derived from the lexical properties of individual verbs.

Notes

1. Note that different instantiations of cue-based retrieval make different predictions regarding the source of effects in grammatical sentences (for discussion see Parker et al., 2017; Vasishth et al., 2019). In this paper, we focus on the predictions of the activation-based model (Jäger et al., 2017; Lewis et al., 2006; Lewis & Vasishth, 2005).
2. In case of non-convergence, we first removed the random correlation parameters and refit the model. If this model still didn't converge, we removed all random effects with an estimated variance of 0. If the model still did not converge, the random effect that accounted for the least variance was iteratively removed until convergence was achieved.
3. We increased the number of iterations for each chain to 8000 in the analysis of Experiments 3/4, due to low bulk/tail effective sample size (ESS) and an Rhat of 1.02 for some model estimates in an initial model fit using the default number of iterations (2000).

Acknowledgements

We thank the editor and two reviewers for comments on previous drafts.

Disclosure statement

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

Funding

This work was supported by British Academy/Leverhulme Trust Small Research: [Grant Number SG150015].

ORCID

Ian Cunnings http://orcid.org/0000-0002-5318-0186

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(*4*), 390–412. http://dx.doi.org/10.1016/j.jml.2007.12.005

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(*3*), 255–278. http://dx.doi.org/10.1016/j.jml.2012.11.001

Bock, K., & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, *31*(1), 99–127. https://doi.org/10.1016/0749-596X(92)90007-K

Bock, K., & Eberhard, K. M. (1993). Meaning, sound and syntax in English number agreement. *Language and Cognitive Processes*, *8*(1), 57–99. https://doi.org/10.1080/01690969308406949

Bock, K., & Middleton, E. L. (2011). Reaching agreement. *Natural Language & Linguistic Theory*, *29*(4), 1033–1069. https://doi.org/10.1007/s11049-011-9148-y

Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology*, *23*(1), 45–93. https://doi.org/10.1016/0010-0285(91)90003-7

Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, *10*(1), 395. https://doi.org/10.32614/RJ-2018-017

Burton, S., & Grimshaw, J. (1992). Coordination and VP-internal subjects. *Linguistic Inquiry*, *23*(2), 305–313. JSTOR. https://www.jstor.org/stable/4178769

Cairns, H. S., McDaniel, D., Hsu, J. R., & Rapp, M. (1994). A longitudinal study of principles of control and pronominal reference in child English. *Language*, *70*(2), 260. https://doi.org/10.2307/415829

Chomsky, N. (1981). *Lectures on government and binding: The Pisa lectures* (7th ed). Mouton de Gruyter.

Clifton, C., Traxler, M., Mohamed, M., Williams, R., Morris, R., & Rayner, K. (2003). The use of thematic role information in parsing: Syntactic processing autonomy revisited. *Journal of Memory and Language*, *49*, 317–334. https://doi.org/10.1016/S0749-596X(03)00070-6

Cunnings, I., & Sturt, P. (2018). Retrieval interference and semantic interpretation. *Journal of Memory and Language*, *102*, 16–27. https://doi.org/10.1016/j.jml.2018.05.001

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

Franck, J., Vigliocco, G., & Nicol, J. (2002). Subject-verb agreement errors in French and English: The role of syntactic hierarchy. *Language and Cognitive Processes*, *17*(4), 371–404. https://doi.org/10.1080/01690960143000254

Friedmann, N., & Costa, J. (2010). The child heard a coordinated sentence and wondered: On children's difficulty in understanding coordination and relative clauses with crossing dependencies. *Lingua. International Review of General Linguistics. Revue internationale De Linguistique Generale*, *120*(6), 1502–1515. https://doi.org/10.1016/j.lingua.2009.10.006

Fujita, H., & Cunnings, I. (2022). Interference and filler-gap dependency formation in native and non-native language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48*(5), 702–716. https://doi.org/10.1037/xlm0001134

Gerard, J., Lidz, J., Zuckerman, S., & Pinto, M. (2017). Similarity-based interference and the acquisition of adjunct control. *Frontiers in Psychology*, *8*, 1822. https://doi.org/10.3389/fpsyg.2017.01822

Goodluck, H. (2001). The nominal analysis of children's interpretations of adjunct PRO clauses. *Language*, *77*(3), 494–509. https://doi.org/10.1353/lan.2001.0159

Hammerly, C., Staub, A., & Dillon, B. (2019). The grammaticality asymmetry in agreement attraction reflects response bias: Experimental and modeling evidence. *Cognitive Psychology*, *110*, 70–104. https://doi.org/10.1016/j.cogpsych.2019.01.001

Hsu, J. (1985). The development of grammars underlying children's interpretation of complex sentences. *Cognition*, *20*(1), 25–48. https://doi.org/10.1016/0010-0277(85)90003-4

Jäger, L. A., 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. A., 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*, 104063. https://doi.org/10.1016/j.jml.2019.104063

Janke, V. (2018). Discourse effects on older children's interpretations of complement control and temporal adjunct control. *Language Acquisition*, *25*(4), 366–391. https://doi.org/10.1080/10489223.2017.1359271

Lago, S., Shalom, D. E., Sigman, M., Lau, E. F., & 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., & von der Malsburg, T. (2022). Semantic attraction in sentence comprehension. *Cognitive Science*, *46*(2), e13086. https://doi.org/10.1111/cogs.13086

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, *29*(3), 375–419. https://doi.org/10.1207/s15516709cog0000_25

Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, *10*(10), 447–454. https://doi.org/10.1016/j.tics.2006.08.007

Nicenboim, B., & Vasishth, S. (2016). Statistical methods for linguistic research: Foundational Ideas-Part II. *Language and Linguistics Compass*, *10*(11), 591–613. https://doi.org/10.1111/lnc3.12207

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.12589

Parker, D., & An, A. (2018). Not all phrases are equally attractive: Experimental evidence for selective agreement attraction effects. *Frontiers in Psychology*, *9*, 1566. https://doi.org/10.3389/fpsyg.2018.01566

Parker, D., Lago, S., & Phillips, C. (2015). Interference in the processing of adjunct control. *Frontiers in Psychology*, *6*, Article no. 1346. https://doi.org/10.3389/fpsyg.2015.01346

Parker, D., & Phillips, C. (2017). Reflexive attraction in comprehension is selective. *Journal of Memory and Language*, *94*, 272–290. https://doi.org/10.1016/j.jml.2017.01.002

Parker, D., Shvartsman, M., & Van Dyke, J. A. (2017). The cue-based based retrieval theory of sentence comprehension: New findings and new challenges. In L. Escobar, V. Torrens, & T. Parodi (Eds.), *Language processing and disorders* (pp. 121–144). Cambridge Scholars Publishing.

Pearlmutter, N. J., 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

Schlueter, Z., Williams, A., & Lau, E. (2018). Exploring the abstractness of number retrieval cues in the computation of subject-verb agreement in comprehension. *Journal of Memory and Language*, *99*, 74–89. https://doi.org/10.1016/j.jml.2017.10.002

Smith, G., & Vasishth, S. (2020). A principled approach to feature selection in models of sentence processing. *Cognitive Science*, *44*, https://doi.org/10.1111/cogs.12918

Sorensen, T., Hohenstein, S., & Vasishth, S. (2016). Bayesian linear mixed models using Stan: A tutorial for psychologists, linguists, and cognitive scientists. *The Quantitative Methods for Psychology*, *12*(3), 175–200. https://doi.org/10.20982/tqmp.12.3.p175

Van Dyke, J. A., & McElree, B. (2011). Cue-dependent interference in comprehension. *Journal of Memory and Language*, *65*(3), 247–263. https://doi.org/10.1016/j.jml.2011.05.002

Van Valin, R. D. (1986). An empty category as the subject of a tensed S in English. *Linguistic Inquiry*, *17*(3), 581–586. JSTOR. https://www.jstor.org/stable/4178506

Vasishth, S., & Nicenboim, B. (2016). Statistical Methods for Linguistic Research: Foundational Ideas - Part I. *Language and Linguistics Compass*, *10*(*8*), 349–369. http://dx.doi.org/10.1111/lnc3.12201

Vasishth, S., Nicenboim, B., Beckman, M. E., Li, F., & Kong, E. J. (2018). Bayesian data analysis in the phonetic sciences: A tutorial introduction. *Journal of Phonetics*, *71*, 147–161. https://doi.org/10.1016/j.wocn.2018.07.008

Vasishth, S., Nicenboim, B., Engelmann, F., & Burchert, F. (2019). Computational models of retrieval processes in sentence processing. *Trends in Cognitive Sciences*, *23*(11), 968–982. https://doi.org/10.1016/j.tics.2019.09.003

Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, *61*(2), 206–237. 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*, 104400. https://doi.org/10.1016/j.jml.2022.104400