Discourse representations guide alternative set activation

Morwenna Hoeks[1], Maziar Toosarvandani[2], and Amanda Rysling[2]

[1]Institute of Cognitive Science, University of Osnabrück

[2]University of California, Santa Cruz

Author Note

The data reported in this paper, as well as the analysis code and materials are made available via the Open Science Framework and can be accessed on https://osf.io/svq6c/?view_only=75184df3a9174f56a640d7c166fe0f9e. The results of these experiments were presented at the 2023 Conference on Human Sentence Processing. We have no conflicts of interests to disclose.

Correspondence concerning this article should be addressed to: Morwenna Hoeks, Institute of Cognitive Science, University of Osnabrück, 49069 Osnabrück. Email: morwenna.hoeks@uni-osnabrueck.de

Abstract

Linguistic focus triggers the activation of contrastive alternatives to the expression in focus (Braun & Tagliapietra, 2010). Priming studies, which did not manipulate the discourse context of sentences containing foci, have suggested a two-stage model of how alternatives are considered in real-time comprehension: first, semantic associates are activated as in normal word recognition, and then alternatives that contrast appropriately with a focus are selected from among them (Husband and Ferreira, 2015). Reading, memory, and visual world studies, however, have shown that comprehenders also utilize discourse information and world knowledge to identify alternatives (Fraundorf et al., 2013; Kim et al., 2015; Sedivy, 2002). This article considers two extensions of Husband and Ferreira's (2015) model that take discourse information into account: one in which discourse information is *only* used at a late stage to select appropriate alternatives after semantic associates are primed, and one in which discourse information may itself be used to activate alternatives during earlier processing stages. Under the first, purely selection-based model, any effect of the discourse (in)appropriateness of alternatives should not be contemporaneous with effects of semantic association, because discourse information should be used only in the selection process, after an initial candidate set has been activated. Three incremental comprehension studies using the Maze task do not support these predictions. The results indicated contemporaneous effects of discourse information and semantic association on comprehension of potential alternatives, pointing instead to a model with effects of discourse at an early stage of activation. Comprehenders access discourse information to differentiate among potential alternatives, whose activation does not depend exclusively on semantic associate priming.

Keywords: focus, contrastive alternatives, time course, discourse, Maze task

Discourse representations guide alternative set activation

Sometimes, to understand a scene or situation, understanding what is not present is just as important as understanding what is. Studies in language processing have demonstrated that comprehenders need to consider *contrastive alternatives*—expressions that contrast with, and substitute for, some expression in a sentence—in order to understand the message it conveys. For example, a comprehender can, at least in some contexts, come to understand the sentence *Sarah remembered to bring only [yogurt]* to mean that Sarah brought nothing else besides yogurt, such as milk or cheese. But the specific inference that the comprehender draws will depend on the context in which it is used. If Sarah is a waiter at an Indian restaurant, a comprehender would likely instead conclude that she forgot to bring another condiment, like chutney, but not that she forgot to bring milk (not an Indian condiment). This article investigates how what has previously been said in a discourse determines the specific alternatives that come to mind when a sentence is comprehended.

The need to consider such alternatives is signaled by varied linguistic devices, including *focus*, which in English is marked both by intonation and by focus particles (e.g., *only*, *even*, *also,...*) or constructions (e.g., a cleft: *it was...that...*). In the example sentence above, the particle *only* and a falling pitch accent on *yogurt* together indicate that a set of alternatives to yogurt needs to be considered. Indeed, a growing body of research has shown that both focus intonation and focus particles lead to the activation of a focused expression's contrastive alternatives in online comprehension, using measures from reading, visual world, priming, and memory tasks (Braun et al., 2018; Braun & Tagliapietra, 2010; Fraundorf et al., 2013; Fraundorf et al., 2010; Gotzner et al., 2016; Hoeks et al., 2023; Kim et al., 2015; Spalek et al., 2014; Yan & Calhoun, 2019).

This evidence has shown that comprehenders use various sources of information to determine the content of a focus' alternative set. Evidence from priming tasks suggests that comprehenders rely on associative relationships between expressions (what we will call *conceptual* information) to first activate a set of semantically associated expressions to the focus, and then select contrastive alternatives from among these activated associates (Husband &

Ferreira, 2015). But to distinguish mere associates from contrastive alternatives, this selection process must also rely on information beyond conceptual relationships. Expressions must be possible replacements of a focus in order to contrast with it: in the example above, *yogurt* contrasts with expressions like *milk* or *chutney* because these are things that Sarah could have brought instead. Since replaceability is governed by the surrounding linguistic material that occurs within the same sentence, we will use the term *sentential* information to refer to this type of information. Finally, in addition to conceptual and sentential information, evidence from memory and visual world tasks suggests that comprehenders also rely on information from the larger discourse context (henceforth: *discourse* information) to differentially encode discourse-appropriate alternatives from implausible ones (Fraundorf et al., 2013; Kim et al., 2015).

However, the evidence across these different tasks has failed to converge on a unified understanding of the time course of alternative set consideration in sentence processing. In Husband and Ferreira's model, conceptual information is understood to shape the alternative set early on, while sentential information determining replaceability is only integrated at a later stage of processing. But, as Husband and Ferreira state themselves, their study did not test for effects of discourse context, and is therefore "unable to speak directly to questions concerning how both discourse and background knowledge can activate candidates for the set of focus alternatives beyond those activated by association with the focused word itself" (Husband and Ferreira, 2015, p. 229). Moreover, studies that investigated the role of discourse context either used offline tasks (Fraundorf et al., 2013), or did not explicitly test the interaction between discourse and conceptual information (Kim et al., 2015). The way that linguistic discourse information affects the activation of alternatives over time has therefore still not been studied independently from effects of conceptual information that is part of our general knowledge.

We present three incremental comprehension studies which aim to shed light on the time-course of the discourse-sensitivity of alternative set processing. These experiments were designed to test the predictions of two potential extensions of Husband and Ferreira's original

model to also incorporate discourse information. On the one hand, this model could be extended to capture the effect of discourse context by suggesting that the process of activating alternatives is temporarily encapsulated from effects of discourse—just like information from the sentential context did not initially affect activation of semantic associates in Husband and Ferreira's (2015) data. Under this view, discourse information is used only to select discourse-appropriate alternatives from among a set of initially activated semantic associates, after those associates are activated (this extension is referred to as Selection Only). On the other hand, another way to extend this model is to suggest that the activation of focus alternatives is guided by the discourse context early on, such that discourse activation may also play a role already during an earlier activation stage (Discourse-guided Activation). Because the first extension involves only a selection mechanism, it predicts that effects of discourse on the activation of potential alternative expressions could only appear after activation on mere associates has been suppressed. The second extension predicts that effects of discourse context could be present contemporaneously with effects of semantic association, and that the differentiation of discourse appropriate alternatives from discourse inappropriate ones does not depend on initial activation of semantic associates.

The studies presented here tested these predictions by measuring response times on a potential alternative expression as comprehenders proceeded incrementally through a sentence, probing the extent to which the processing of a focus facilitated decisions on potential subsequent alternatives while the comprehension of that sentence was still ongoing. These experiments tested discourses like the Indian restaurant example above, in which certain items were appropriate alternatives in the discourse, while others were not. Unlike that example, however, the manipulation of discourse-appropriateness was achieved not by evoking a general type of situation (an Indian restaurant), but by describing a particular event in which some but not all items are expected to be contrasted with the focus. Manipulating the appropriateness of alternatives in this way allowed us to disentangle effects of world-knowledge from effects of discourse information on the activation of alternatives. In discourses like these, world-knowledge alone is insufficient to

determine which alternatives are appropriate. Instead, comprehenders were forced to rely on their representation of the explicit linguistic discourse context to rule in or rule out certain alternatives.

All three experiments used the Maze task (Forster et al., 2009) to test the on-line consideration of alternatives to foci. In this task, comprehenders progress through a sentence by choosing which of two presented words is the most suitable continuation of the sentence up to that point. Such a decision can only be made if preceding material is sufficiently incorporated in the comprehender's representation of the sentence. This task is therefore less prone to spillover effects than other comparable tasks, such as self-paced reading (Witzel et al., 2012). Moreover, decision times in the Maze have been shown to correlate with total reading times (the sum of all fixations on the target region) obtained by monitoring eye movements during naturalistic reading (Forster et al., 2009).

The design of these experiments builds both on Lowder et al. (2021)—in which readers were shown to utilize focus particles as a cue to begin anticipating upcoming sentence continuations in the absence of a discourse context—and on a line of work which manipulated discourse context but which did not use on-line measures (Fraundorf et al., 2013; Fraundorf et al., 2010). Next, we motivate the design of the present studies by providing an overview of the existing literature, which has already demonstrated a role of discourse context in the processing of alternatives to foci.

Evidence for early effects of conceptual information and late effects of sentential information

Studies using cross-modal priming have demonstrated that information from a focus' sentence context is used to winnow down the set of potential alternatives that are activated by that focus, by establishing faster responses to expressions that can replace the focus than for mere semantic associates. Braun and Tagliapietra (2010) showed that a focus accent on a word like *flamingo* leads to facilitation of substitutable expressions of that focus (e.g., words like *pelican*), when compared to expressions that were semantically associated with the focus but could not replace it, and therefore were not possible alternatives (like *pink*). In later work, Husband and Ferreira (2015) probed the time course of this activation of alternatives triggered by a focus, by

manipulating the delay between the auditory presentation of the word that may trigger the activation of alternatives (henceforth: *the trigger*) and the eventual printed targets, as in stimuli like (1). In this example and all other ones we indicate the trigger using square brackets, and the targets on which that activation is measured using boldface.

(1) *Auditory trigger sentence:* The museum thrilled the [sculptor] when they called about his work.

Visual lexical decision targets:

a. **painter** (associated alternative)

b. **statue** (associated non-alternative)

c. **register** (non-associated non-alternative)

Like the Indian restaurant scenario above, the scenario described in (1) suggests the potential involvement of specific objects or individuals. Mention of a museum, in particular, suggests that sculptors or other artists like painters could possibly be relevant. However, unlike in the Indian restaurant example above, the experimental stimuli exemplified in (1) also included words like *statue*, which were not possible alternatives because they are not possible substitutions for the trigger *sculptor* in the sentential context. Nonetheless, in conditions in which the alternative trigger (*sculptor*) received a focus accent, both associated alternatives (*painter*) and associated non-alternatives (*statue*) were found to be facilitated over unrelated controls when presented immediately following the alternative trigger. However, by a 750 ms delay from the offset of the alternative trigger, the effect of semantic association on non-alternative candidates was found to be suppressed.

Husband and Ferreira (2015) took these findings as evidence for a two-stage model of alternative set construction in which all semantic associates first become activated due to general semantic priming mechanisms, before the presence of focus intonation on the alternative trigger later led selection mechanisms to suppress activation of non-alternative associates (see also

Gotzner et al. (2016) for a similar account). In this model, the alternative set is thus derived via a combination of general conceptual representations of lexical expressions and the sentential context, which determines what expressions can be felicitously substituted for a focus in the target sentence. This model of alternative set processing is appealing because it relies on two mechanisms, both of which are known independently to be utilized in the comprehension of language (semantic associate priming and selection).

However, to describe this process as one of activation by general associate priming followed by selection implies that the use of sentential information is contingent on the use of conceptual information, since an initially activated set of semantic associates is winnowed down at a later point using information from the sentence context about potential replacements. It predicts that alternatives cannot be selected from among non-associated expressions, which never become activated in the first place. Notably, Husband and Ferreira were very careful to emphasize that their evidence did not support the interpretation that *only* general associate priming modulated initial activation levels, rather, they had demonstrated that *at least* semantic associate priming played a large role: because they did not test effects of discourse contexts, they could not exclude the possibility that discourse information can also directly constrain the activation stage. In order to confirm this relationship between these types of information, language comprehenders' processing of contrastive alternatives that are *not* associated with the focused prime also have to be tested. Muxica and Harris (2025) and Washburn et al. (2011) investigated the activation of such unassociated alternatives, arguing that semantically unrelated, but potentially substitutable expressions to a focused alternative trigger may also become activated, though only when such expressions are explicitly mentioned in the preceding discourse. This and related work, reviewed next, establishes a role for the discourse context in determining which alternatives are considered in real-time comprehension, suggesting that the mechanisms through which comprehenders activate alternatives must also be sensitive to discourse information.

In what follows, we identify three particular ways in which the plausibility of an alternative can be determined by the discourse context. A potential alternative can become a

discourse appropriate alternative in a particular context by being:

(i) **Mentioned:** an alternative is salient because it was explicitly mentioned in the preceding discourse; or

(ii) **Situationally available:** while an alternative may not have been mentioned, the specific properties of a situation, described in the preceding discourse, make it predictable based on world knowledge (as in the Indian restaurant example on p. 5).

But discourse context may also make a potential alternative a discourse-inappropriate alternative, due to:

(iii) **Discourse exclusion:** even though it is mentioned, information in the preceding discourse context may rule out an otherwise mentioned and/or situationally available alternative as a plausible one within that discourse context.

Below, we discuss behavioral evidence which suggests that the processing of alternatives is affected by discourse in all three ways, though evidence for effects of discourse exclusion is far more limited than evidence for mention and situational availability. We also outline below why studying effects of discourse exclusion is valuable in developing a better understanding of alternative set processing: Since this property can be manipulated while keeping properties like repetition and association constant, it enables us to disentangle effects of the discourse context from more general effects of salience or general knowledge. In the remainder of this article, we therefore specifically tested effects of (iii), and fully crossed this type of discourse-appropriateness of an alternative with the level of semantic association of that alternative to the alternative trigger.

Effects of explicit mention of alternatives

Existing work has established that discourse context determines which expressions are caused to be activated by a focus, by showing that alternatives that may otherwise be unexpected as alternatives to a focus still become activated when explicitly mentioned in the discourse context. Using a priming task, Washburn et al. (2011) found faster responses to targets (*lock*) after

participants read a trigger sentence like (2b), with a potential alternative trigger that was mentioned in a preceding discourse context and that was unassociated to the target, compared to targets preceded by trigger sentences like (2c), with non-associated triggers that had not been mentioned before.

(2) *Discourse Context:* Christina wants to buy a lock, nails, and a bolt. She needs these to fix her front entrance. Two days ago, she went to a store that didn't have a wide selection.

Visual trigger sentence:

a. At the store, she was able to buy (only) a [bolt]. Mentioned associated

b. At the store, she was able to buy (only) [nails]. Mentioned non-associated

c. At the store, she was able to buy (only) a [lamp]. Unmentioned non-associated

Target: **lock**

This effect did not hold when the focus particle was absent in these trigger sentences, indicating that expressions that are not associated with the alternative trigger (as defined using the South Florida Free Association Norms) may still become activated—presumably as contrastive alternatives to that alternative trigger—when these expressions are made salient in the discourse context. Evidence from memory tasks converges on this conclusion, demonstrating increased competition from explicitly mentioned contrastive alternatives (Fraundorf et al., 2010).

Crucially, the targets in Washburn et al.'s study were presented 250ms after the offset of the primes, indicating that merely mentioned but non-associated alternatives may become activated as early as the semantic associates in Husband and Ferreira's (2015) study. In fact, in some studies, discourse mentioned alternatives allowed comprehenders to start reasoning about the alternative set even before the focus was encountered (Kim et al., 2015). Next, we discuss evidence suggesting that, besides explicit mention of contrastive alternatives themselves, other overt material in the discourse context may also guide comprehenders early on during the time course of alternative processing.

Effects of situation-specific information

This section discusses evidence suggesting that discourse context affects the activation of potential alternatives by determining the specific type of situation that is being described, and thus that comprehenders use world knowledge about what objects may be compatible with those situations to activate discourse-appropriate alternatives. In a visual-world study, in which participants listened to trigger sentences like *Neil only wants to buy some [hot dogs]*, Kim et al. (2015) found that the preceding discourse may guide listeners' expectations about the upcoming focus. Fixations in their study converged earlier on a target after biasing discourse contexts that described scenarios compatible with a narrow set of alternatives (e.g., a baseball game), compared to neutral discourse contexts compatible with a wider range of alternatives (e.g., a grocery trip). Kim et al. (2015) therefore proposed that comprehenders generate hypotheses about the discourse-relevant set of focus alternatives, which enables them to more rapidly converge on an appropriate visual target in biasing contexts.

Lowder et al.'s (2021) eye movement data further support the conclusion that focus alternatives may be anticipated based on such situational knowledge. They found that slowdowns on unpredictable nouns were eliminated when they were alternatives to an unambiguously focused alternative trigger compared to cases in which the presence of focus on the alternative trigger was not signaled in advance. That is, only in the absence of a focus particle were expressions with a low predictability based on the trigger (e.g., *the [bride] but also the* ***priest***) read more slowly than more predictable nouns (*..the [bride] but also the* ***groom***). This suggests that the presence of the focus particle may be involved in eliminating such predictability effects via the activation of a set of alternatives to the focus. Together with the results from Kim et al. (2015), this evidence suggests situation-specific information may play a role even before a focus is encountered.

However, the exact mechanisms underlying these early effects of situation-specific information are not entirely clear. One hypothesis is that comprehenders rely on domain-general mechanisms to activate a set of alternatives based on both overt material in the discourse context, and general conceptual knowledge. Perhaps the set of expressions that becomes activated includes

any salient expression in the discourse as well as any expression that becomes primed by other overtly mentioned material. For instance, if the expression *apple* is focused in a scenario that also mentions the word *fruit*, any other term referring to a fruit may simply become activated as a contrastive alternative. But, under another hypothesis, it may be that the set of alternatives that is considered is constrained, not just by salience and world-knowledge, but also by more fine-grained properties of the discourse context that guide comprehenders to consider expressions compatible with the particular scenario described. In that case, comprehenders do not just activate salient and predictable expressions, but instead form a representation of the particular state of affairs in an individual discourse (e.g., as incorporated in a *situation model* in the sense of Zwaan and Radvansky, 1998), and use this representation to determine what expressions are the relevant contrastive alternatives.

Crucially, results from Kim et al. (2015) cannot distinguish between these two hypotheses. In Kim et al.'s (2015) materials, the biasing context sentences included expressions (e.g., *baseball game*) that were arguably more closely related to the target (*hot dogs*) than material in the non-biasing contexts (*supermarket*). It may still be that the benefit for biasing contexts arose, not because fine-grained representations of the discourse context affected the processing of alternatives at an early stage, but because domain-general priming mechanisms allowed comprehenders to (pre-)activate targets in those conditions. There may thus still be a role for semantic priming in explaining these effects, because it is possible that material in the biasing contexts generally primed the targets.

In order to fully understand the role and time-course of the linguistic discourse context in the activation of alternatives, we must therefore look to cases where neither mention nor conceptual information is sufficient to anticipate a set of alternatives. Below, we discuss evidence from a memory study which showed effects of discourse-specific information on the activation of alternatives that cannot be attributed to semantic priming or general predictability based on world knowledge (Fraundorf et al., 2013). The present studies used a similar manipulation as the ones used in these memory tasks, to show that effects of discourse-specific information can also affect

response times while the interpretation of a discourse is still ongoing.

Discourse excluded alternatives

Research by Fraundorf et al. (2013) has shown that the alternatives that become activated in the processing of focus are constrained by discourse information which rules out expressions as plausible alternatives to a focus, despite them being overtly mentioned and situationally available in the discourse context. Their stimuli contained a preceding discourse context like (3a), in which Saturn, Neptune, and Jupiter were all mentioned exactly once in the described event but in which only Saturn was a plausible alternative to Jupiter in the trigger sentence following it, as in (3b).

(3) a. *Discourse context:* Originally, the space probe Cosmo III was designed to fly past Jupiter and Saturn and send photos and measurements back to NASA from both planets. NASA needed this information to guide the videos they were going to take of Neptune on a future mission.

b. *Trigger sentence:* However, due to a glitch in the programming of the Cosmo III, it lost the photos taken of [Jupiter] and put the future mission in trouble.

Although Neptune is mentioned and may generally be predictable in a situation like this, it is a poor alternative to Jupiter in this particular incident, because the discourse context establishes that the mission to Neptune has not yet occurred, and so photos of Neptune could therefore not have been lost instead. In their truth verification task, Fraundorf et al. (2013) indeed found that font emphasis as a cue to focus helped reject false statements about discourse-appropriate target alternatives (*Saturn*), but not about discourse-inappropriate target alternatives that were nevertheless mentioned in the discourse (*Neptune*).

The discourse information manipulated in this experiment—in which discourse contexts always involved the same amount of overt material—crucially differed from that in Kim et al. (2015)—in which the discourse bias was manipulated by including or excluding specific lexical items. Because the identity of the plausible and implausible alternatives was counterbalanced by Fraundorf et al. (2013), participants could not have relied on their conceptual or world knowledge

in ruling out alternatives in this study. These results therefore more clearly suggest that readers encode a narrow set of only those alternatives that are plausible in the particular discourse context that is being described, independently of whether these alternatives were conceptually associated to the focus, mentioned, or situationally available.

However, since this was an offline memory study, it does not provide evidence about the time course over which these discourse excluded expressions are determined to be discourse-inappropriate alternatives. It is therefore more generally unclear how and when comprehenders integrate such discourse-specific information with conceptual information in constructing the relevant set of alternatives. Next, we outline how the present studies aim to test the role of discourse information during alternative set activation.

Testing the role of discourse context in alternative set processing

In their priming studies of the processing of contrastive alternatives, Husband and Ferreira (2015) found early activation of expressions that are semantically associated with focused alternative triggers (even those that are not viable contrastive alternatives), but later maintenance only of the activation of substitutable associates, i.e., expressions that are viable contrastive alternatives to those foci. They proposed a model of alternative set processing in which contrastive alternatives become activated during the processing of a focus, first by spreading activation from the focused trigger to semantically associated expressions, and then by suppressing activation on those initially activated semantic associates that are not substitutable for the focus in the sentence context. This model captures the way that alternatives become activated in out-of-the-blue contexts, but it was not initially developed to specify how discourse information would be integrated in the activation of contrastive alternatives. Together, the evidence from existing work discussed above indicates that the overt mention of alternatives in the discourse, as well as situation-specific information and properties of the linguistic discourse context excluding certain expressions from the alternative set affect the processing of these alternatives in addition to general conceptual knowledge. But existing evidence does not tell us when discourse-specific information affects the activation of the relevant expressions, nor does it

tell us how information from the discourse context interacts with conceptual knowledge.

To further investigate the role of discourse information and its interaction with conceptual information, all experiments crossed semantic association between a potential alternative trigger and a target alternative with a manipulation of discourse-appropriateness inspired by Fraundorf et al. (2013), through which the target alternative was either excluded as an appropriate alternative in the preceding discourse context or not. For instance, in (4), *cheese* is established as a discourse-appropriate alternative to *yogurt*, since it is described as one of the items that was asked for and could therefore be remembered by the waiter to bring back to the table, but *ashtray* is not a discourse-appropriate alternative to *yogurt* because it was not explicitly asked for by the tourist.

(4) a. *Discourse context*: The tourist asked for a variety of items, like some cheese and yogurt. There was already an ashtray on the table.

b. *Trigger sentence*: When the waiter returned, he remembered to bring only some [yogurt] but no **cheese** to the table where the tourist was seated.

Like all the studies discussed above, the present experiments thus used a design in which activation on target alternatives (in bold in the example above) could be measured as a function of the preceding sentence containing a potential alternative trigger (between square brackets), and the discourse context which specified some alternatives as discourse-appropriate (underlined) and some as discourse-inappropriate. Unlike previous studies, however, all present experiments employed the Maze task as a measure of on-line activation of target alternatives—enabling the target's ease of integration, as an alternative to that trigger, to be measured by the time it takes to decide on that target. If semantic associates become activated in the processing of a focused alternative trigger, then targets that are semantically associated to the preceding trigger (as in the example above) should be integrated faster and/or more reliably than targets that are non-associated to the trigger (as in a version of (4) which includes *...only an [ashtray], but no **cheese***). All else being equal, this difference in activation between associated and non-associated alternatives should be realized as a difference in response times on the target alternative that

follows the alternative trigger, such that associated targets are decided on faster than unassociated ones.

However, if comprehenders also initially activate alternatives based on discourse information that is independent of their general conceptual knowledge, and instead rely on the specifics of the event that is described in the discourse context, then non-associated target alternatives may sometimes be integrated more easily than associated target alternatives—for instance, when they are more suitable as alternatives in a particular discourse context. A pattern of results in which non-excluded alternatives are responded to faster than excluded ones would thus suggest that comprehenders rely on information from the discourse to differentially activate discourse-appropriate and inappropriate alternatives at the point in time where the target is encountered.

Similar to Lowder et al. (2021), a focus construction like (4) was used in Experiment 1 in which a potential target alternative always directly followed an alternative trigger. While Experiment 1 thus aimed to probe these effects of discourse exclusion at an early stage of focus processing (with only a few words in between trigger and target and when effects of semantic association are therefore still present when the target is read), Experiment 3 was intended to test these predictions at a later stage of processing (with a target further removed from the trigger). Experiment 2 tested stimuli like (4) in which a focus particle was removed from the trigger sentence, and therefore served to test to what extent effects of discourse exclusion are triggered specifically in the processing of focus.

The results of Experiment 1 will lead us to consider two different types of extensions of the original model proposed by Husband and Ferreira (2015) that take discourse information into account—one in which the activation of contrastive alternatives to a focus is not initially influenced by discourse information (Selection Only), and one in which discourse information can be used early on to activate discourse-appropriate alternatives alongside an initially activated set of semantic associates (Discourse-guided Activation).

Experiment 1

Experiment 1 investigated whether, at an early stage of focus processing, comprehenders rely on information from the linguistic discourse context in addition to semantic association to construct a set of discourse-relevant alternatives. It used on-line measures to probe the time course of comprehenders' sensitivity to these two types of information—in particular, how semantic association and discourse exclusion of an expression jointly affect the integration of that expression as an alternative to a preceding alternative trigger.

Method

Materials. In Experiment 1, every item constituted a short narrative in which two context sentences first introduced three potential alternatives. In all conditions, a target sentence then put one of these explicitly mentioned alternatives in focus using the focus particle *only*, contrasting this focused alternative trigger with a second target alternative, which was previously mentioned in the discourse context. An example item in all four conditions is shown in (5), in which the focused alternative trigger is indicated for presentational purposes (but not in the experiment) using square brackets, the target alternative is bolded, and discourse-appropriate alternatives are underlined. The identity of the alternatives in each context manipulated discourse exclusion, that is, whether the target alternative was explicitly excluded as an alternative to the alternative trigger (excl) or not (non-excl). In the non-excluded conditions, the target alternative inside the target sentence was always mentioned in the first context sentence, which set it up as a discourse appropriate alternative to the focus; in the excluded conditions, the target alternative was always mentioned in the second context sentence which, based on the described incident, ruled it out as a discourse-appropriate alternative to the focus.

(5) a. *Discourse context:* The tourist asked for a variety of items, like some <u>cheese</u> and <u>yogurt</u>. There was already an ashtray on the table.
Trigger sentence: When the waiter returned, he remembered to bring only some [yogurt] but no **cheese** to the table where the tourist was seated. assoc non-excl

b. *Discourse context:* The tourist asked for a variety of items, like an ashtray and yogurt. There was already some cheese on the table.
*Trigger sentence:*When the waiter returned, he remembered to bring only some [yogurt] but no **cheese** to the table where the tourist was seated. assoc excl

c. *Discourse context:* The tourist asked for a variety of items, like some cheese and an ashtray. There was already some yogurt on the table.
Trigger sentence: When the waiter returned, he remembered to bring only an [ashtray] but no **cheese** to the table where the tourist was seated. non-assoc non-excl

d. *Discourse context:* The tourist asked for a variety of items, like an ashtray and yogurt. There was already some cheese on the table.
Trigger sentence: When the waiter returned, he remembered to bring only an [ashtray] but no **cheese** to the table where the tourist was seated. non-assoc excl

Discourse exclusion of the target alternative was achieved by ensuring that the presuppositions of the predicate inside the trigger sentence were satisfied for the first two contextual alternatives but not for the third contextual alternative. In (5), the target sentence's predicate *remember to bring* presupposes, roughly, that whatever its object refers to was being asked for. Since the second sentence in each discourse entails that this item was already on the table, it is unlikely that it was asked for by the tourist, implying that it is not among the things that the waiter should have remembered to bring. Thus, the object mentioned in the second context sentence is always an unlikely, discourse-inappropriate alternative to the focus, because the event described by the trigger sentence contrasted objects that the waiter remembered to bring with objects that the waiter did not remember to bring.

For this reason, the trigger sentence in the excluded conditions, i.e., in (5b) and (5d), was highly likely to be considered unnatural, because it is not coherent with the preceding discourse context. After all, it contrasts an alternative with an alternative trigger which was not relevant as an alternative to that focus. Obtaining longer response times on these discourse-excluded target

alternatives would therefore be generally unsurprising, assuming that comprehenders already integrate the relevant discourse information in comprehending the trigger sentence. However, this is not a confound but an intentional feature of the design: Results from Experiment 1 and subsequent experiments show that comprehenders only slowed down on excluded target alternatives in certain cases. In Experiment 2, for instance, these slowdowns crucially disappeared when the focus particle *only* directly preceding the alternative trigger was removed from the trigger sentence, suggesting that comprehenders are less sensitive to information from the discourse context in the absence of a clear cue to the presence of focus marking.

In addition to discourse exclusion, semantic association between the focus and target alternative was manipulated by varying the nature of the focus that preceded the target alternative. In (5) and all other stimuli, the target sentence was thus identical across conditions except for the alternative trigger, which was either associated with the target alternative or not. Varying the alternative trigger allowed us to keep the target alternative, which was the region of interest, identical in all four conditions, allowing for direct comparisons of the response times on this word. Decision times on the alternative trigger are therefore not analyzed. Note that the location of the potential alternatives in the context sentences also varied across conditions, except for the context sentences in the excluded conditions, which were identical to each other. Since the target alternative should be discourse-inappropriate in these cases, both conditions exemplified in (5b) and (5d) therefore included this alternative (*cheese*) in the second context sentence, with the other two alternatives mentioned in the first.

Association was determined throughout using Latent Semantic Analysis (Landauer et al., 1998), where the average similarity of alternatives and foci was 0.58 (range: 0.4–0.86) in the assoc conditions, and 0.09 (range: 0.18–0.07) in the non-assoc conditions. See Appendix B for a list of all such alternative triplets.

In total, 48 items were constructed, each with the four conditions as illustrated in (5). The total number of items was chosen because it would result in a similar number of items per condition per participant as those used by Husband and Ferreira (2015), who included 36 items in

three conditions. Items used varying types of settings and narration styles to ensure that the obtained effects did not arise due to any particular feature beyond the intended manipulation. All items for Experiment 1 can be found in Appendix A. These experimental items were interspersed with 64 fillers which also consisted of multi-line discourses and included both foci in the trigger sentence and focus alternatives in the preceding discourse context. Using a Latin Square design, all 48 items were counterbalanced over 4 lists, such that each participant saw one condition from every item.

Procedure. All target sentences were presented using the A-Maze task. As in the more commonly used self-paced reading task, this task measures response times using button presses. This task was chosen in particular because it encourages highly incremental processing due to the fact that participants advance through a sentence not by simply pressing a button, but by choosing at each word which of two items is the correct continuation of that sentence. Participants in the Maze task thus see each word in the target sentence presented alongside a distractor word (or *foil*) which would not make a sensical continuation. An example of one Maze sentence is given in (6) below, with the foils presented below each of the corresponding words of the trigger sentence. In the experiment, correct continuations and foils were presented side-by-side, where the relative location of the two was randomized such that participants could not know in advance which of the two words was the correct continuation.

(6) When the waiter returned, he remembered to bring only yogurt but no cheese...
x-x-x arm behave greatest, am democratic on rates ago gone, went or surely...

The first word in each target sentence (*When* in the example above), was always presented alongside a non-word (“x-x-x") to guarantee the correct choice for this initial word. After the correct word has been chosen, the following word was automatically presented alongside its respective foil. In this way, sentences were presented incrementally, and the response time required to make and execute a decision about which word should continue a sentence was measured.

Importantly, this task is also argued to provide a measure of the level to which upcoming

structure is anticipated: A-Maze response times have been shown to be inversely related to noun cloze probabilities, with slower responders showing larger effects of expectation (Husband, 2022). This property is particularly useful here because response times on target alternatives can thus in part be taken to index to what extent these expressions are expected as alternatives to the preceding alternative trigger.

A-Maze foils were automatically generated using the AutoMaze software developed by Boyce, Futrell, and Levy (2020), and manually checked to prevent frequent use of the same foil throughout the materials. This algorithm selects distractor words that are roughly of the same length as the target word, and that are predicted by NLP language models to have a poor fit to the preceding sentence material. For each upcoming word, a conditional probability distribution is determined for potential foils of similar length in the context of the preceding sentence. The words with a predicted probability below a certain threshold (or, *above* a certain *surprisal* threshold) are then selected by the AutoMaze algorithm as the distractor. As can be seen in (6), the length of the foils may therefore deviate slightly from that of the corresponding target words. In those cases, foils of identical length did not reach the intended surprisal threshold. Word frequencies that form the input to these models are obtained from the Google Books Ngrams corpus (Michel et al., 2011).

On every trial, participants first read a context sentence on one screen. On a subsequent screen, participants were presented with the start of the trigger sentence in the format of the Maze task. That is, only the trigger sentence was presented incrementally; the context sentences were presented all at once for normal reading. If participants chose the wrong word in the Maze task, they were directed to the next item and their responses on the rest of the words in the trigger sentence were not recorded. Even though this leads to loss of some data, it also has the advantage that participants who contribute RT data are very likely to have understood the sentence up to that point. Although Maze trials were terminated when the foil was incorrectly chosen over the continuation, observations on the critical region from participants who failed to complete the Maze sentences were still included as long as they did not make an error on or before that critical

region.

Half of the experimental trials were followed by a comprehension question, which probed whether participants had understood the context. This was because there was more cause for concern that participants might not read the contexts than that they might not have understood the trigger sentences. Participants had to comprehend the beginning and all subsequent material of a trigger sentence in order to even make a decision about which word could form a potential continuation as the sentence went on, but participants could successfully go through a whole trigger sentence in the Maze without having read its preceding question. Comprehension questions were therefore included to encourage careful reading of the preceding context. For instance, the example item in (5) was followed by the comprehension question in (7).

(7) What was already on the table where the tourist was seated?

Note that, due to the nature of this task, there are multiple sources of errors participants could make. Participants could provide incorrect answers to comprehension questions, or fail to complete the Maze sentences. In the latter case, they may not be able to answer some of the comprehension questions either. Therefore, a lower accuracy would be expected on this task compared to less taxing tasks such as self-paced or naturalistic reading.

Before being presented with the target stimuli and fillers, participants read a short description of the task, followed by five practice items. Practice items were similar to experimental items in that they involved a short context sentence, followed by a sentence presented in Maze format and a comprehension question. After the short practice phase, the experimental items were presented along with the fillers in a pseudo-random order.

Participants. 81 participants were recruited via the Prolific platform for web-based research and redirected to the PennController for Ibex platform that hosted the experiment (Zehr & Schwarz, 2018). This sample size was chosen so as to exceed the 60 participants recruited by Husband and Ferreira (2015). All participants were paid a $12 hourly rate for their participation. All participants were native speakers of English and gave explicit consent to participate.

Participants who had an accuracy of less than 70% on the comprehension questions or that did not complete more than 75% of the Maze sentences were excluded from analysis. Data from 71 participants were included in the analysis; 10 participants were excluded because they failed to complete more than 75% of the Maze sentences.

Transparency and Openness. All materials, data and analysis code of this and subsequent experiments are made available via the Open Science Framework and can be accessed at https://osf.io/svq6c/?view_only=75184df3a9174f56a640d7c166fe0f9e. This study's design and its analysis were not pre-registered.

Results

The mean comprehension question accuracy for Experiment 1 was 86%, and the mean overall completion rate of experimental and filler Maze trials was 94%. On experimental trials in particular, participants failed to complete 10% of the Maze sentences, resulting in 285 missing observations on the critical region (8% of all trials). The remaining 2% of participants failing to complete critical Maze sentences consisted of cases where an error was made in the Maze task downstream of the critical region. In addition to participant exclusion based on accuracy and Maze completion, individual response times below than 200ms (which is too short to make an informed decision in the Maze task) and above 8 standard deviations above the mean were also excluded. Data trimming based on these cutoffs resulted in the exclusion of four additional observations from the entire data set at the critical region, which were also clear visual outliers.

Mean response times for the target word (*cheese*) and its surrounding regions in all conditions are given in Table 1. They are plotted with 95% confidence intervals in Figure 1.

Data were analyzed using R, version 3.6.3 (R Core Team, 2021). We fit Bayesian (generalized) linear mixed-effect models using Stan, as implemented in the brms package, version 2.18.0 (Bürkner, 2017), with the default (flat) priors. Separate models were fit to log-transformed response times and untransformed response times as dependent measures. For each model, we ran four chains, each with 5000 steps (warmup = 1000 steps). Rhat statistics in all models approached 1.00 and no warnings emerged. Models included fixed effects of semantic association (with the

levels associated vs non-associated coded as -0.5 and 0.5 respectively) and discourse exclusion (with excluded vs non-excluded coded as 0.5 and -0.5 respectively), and random slopes and intercepts for both subjects and items. Bayesian models were chosen here as they allowed for a full random effects structure, which was not possible to implement using the lmer package due to convergence issues (Baayen et al., 2008). Note that these models therefore diverged from those adopted in Husband and Ferreira (2015), but given the likelihood of item- and individual-level random effects a larger random effects structure was chosen here over comparability with models used in previous literature. However, for comparison, the output of the largest converging frequentist models parallel to those reported in Husband and Ferreira (2015) is provided for each experiment in Appendix C.

Tables 2 and 3 present the posterior estimates obtained in the models of Experiment 1 log-transformed response times and untransformed response times on target words, respectively. Posterior model estimates are considered reliable if their 95% credible interval does not overlap with zero. Pairwise comparisons between conditions were carried out using the hypothesis function, with a Bonferroni-style adjustment for the size of the credible intervals.[1]

Two main effects were reliable in both models. Positive estimates of association indicate that semantically associated target alternatives were responded to faster than target alternatives that were not semantically associated with the focus. In addition, positive estimates for discourse exclusion indicated that non-excluded alternatives were responded to faster than discourse-excluded ones. Pairwise comparisons revealed that non-associated excluded target alternatives were responded to more slowly than non-associated non-excluded targets, for log-transformed (β =0.039; 97.5% Cr.I.=[0.015, 0.063]) and untransformed response times (β =130.71; 97.5% Cr.I.=[55.62, 206.27]). Moreover, associated excluded target alternatives were responded to more slowly than associated non-excluded targets, again for both log-transformed (β =0.027; 97.5% Cr.I.=[0.0017, 0.052]) and untransformed response times (β =92.68; 97.5%

[1] Further plots of the by-item random slopes are provided in the OSF storage, available via this link: https://osf.io/svq6c/?view_only=75184df3a9174f56a640d7c166fe0f9e

Cr.I.=[17.07,167.36]).

Post-hoc analysis. Since the distance between the focused alternative trigger and the target alternative varied across items, additional analyses were run to test how response times on the target were affected by the temporal delay between the alternative trigger and that target. On each trial, response times on words that intervened between the trigger and target were summed and this distance measure was added as an additional predictor to the models outlined above. Since models which also included interactions between distance and association and distance and exclusion did not reveal reliable interaction terms, the final models only included a main effect of distance. The model run on untransformed response times revealed a positive estimate for distance (β =0.15; 95% Cr.I.=[0.13, 0.18]]), suggesting that the total time after presentation of the trigger and before presentation of the target indeed affected response times on that target. However, the credible interval for distance in the model run on log-transformed response times overlapped with zero (β =0.00; 95% Cr.I.=[0.0000,0.0001]), suggesting no reliable effect of trigger-target distance on target RTs. In both models, the main effects of discourse exclusion and association remained reliable (see Appendix C for the full model outputs).

Discussion

In Experiment 1, target alternatives that were semantically associated with the focused alternative trigger were read faster than non-associated target alternatives. Moreover, expressions that were discourse-excluded from the alternative set of a focus were read more slowly than expressions that were not excluded as alternatives. This discourse exclusion effect was observed while processing of the target sentence was still ongoing, at material only a couple words following the focus. As such, these findings are in line with previous studies that showed on-line effects of discourse information on the activation of focus alternatives (Kim et al., 2015; Sedivy, 2002; Washburn et al., 2011). But unlike the results from previous on-line studies, alternatives were explicitly mentioned in the preceding discourse context in all the conditions in Experiment 1, and so differences in response times between excluded and non-excluded alternatives cannot be attributed to their relative salience alone.

The obtained results crucially also show that comprehenders were sensitive to discourse information when alternatives were not closely associated to the focused alternative trigger. Distinguishing between discourse-appropriate non-associated alternatives and discourse-inappropriate non-associated ones could not be accomplished by relying on conceptual knowledge, because the difference in their appropriateness was solely determined by information from the discourse context. The effect of discourse exclusion observed in Experiment 1 therefore cannot be explained in terms of a contextual priming mechanism like the one proposed by Kim et al. (2015), where the focus and the discourse context may have jointly primed alternative expressions. Such an account, in which activation would spread from a focused alternative trigger to conceptually related expressions as specified by the context, would have to be augmented by the incorporation of more discourse-specific information independently from any situational information to account for the results obtained here.

Since the task used in Experiment 1 differed from the lexical decision paradigm in Husband and Ferreira (2015), some caveats about the interpretation of the time-course are in order here. Recall that in Husband and Ferreira's (2015) original study, targets were presented at a 0 ms delay from the alternative trigger's offset, whereas target alternatives were presented a couple words following the alternative trigger in the present design. Due to the fact that participants were also required to make a decision between the correct continuation and an incorrect foil at each word, Maze response times are generally longer than those obtained from other incremental reading tasks. Indeed, Maze response times are more closely correlated to total reading times obtained from measuring eye movements during naturalistic reading (Forster et al., 2009), making them less comparable to response times in lexical decision tasks used in Husband and Ferreira (2015).

However, despite the fact that response times measured here are therefore more difficult to directly compare to those from priming tasks, the present task allowed us to measure effects of discourse exclusion while the interpretation of the sentence was still ongoing. The effects of discourse exclusion were not observed as part of sentence-final wrap-up effects, supporting the

idea that such discourse information is used actively during incremental sentence comprehension. The results from Experiment 1 therefore suggest that the model of alternative set processing proposed by Husband and Ferreira (2015) should be extended to incorporate the use of discourse information in the differential activation of potential alternatives to a focus. This could in principle be done in two distinct ways. In the next section, we consider two of these possibilities, where one suggests that discourse information is used only to select discourse-appropriate alternatives from an initially activated set of associates (Selection Only), and the other one holds that discourse information can be used to activate discourse-appropriate alternatives alongside an initially activated set of semantic associates (Discourse-guided Activation).

Interim Discussion

One potential extension of the model proposed by Husband and Ferreira (2015) would incorporate the use of discourse information in the process of alternative activation in a way that is parallel to the proposed role of sentential information in that original model. In Husband and Ferreira's model, the first stage of activating contrastive alternatives does not take sentential information into account, and it is therefore possible that this initial activation stage is not informed by discourse information either. Like sentential information, discourse-specific information would then be used only to select discourse-appropriate alternatives from among a set of initially activated associates later on. While sentential information would allow comprehenders to suppress activation on non-substitutable associates, discourse information then enables suppression of those associates that are discourse-inappropriate alternatives. This Selection Only hypothesis maintains the basic architecture of alternative set construction identified by Husband and Ferreira, but also makes the substantive claim that the process of activating contrastive alternatives is temporarily encapsulated from information provided by the discourse context, and that effects of discourse exclusion are therefore only measurable at a later stage of processing.

Although adopting such a Selection Only extension would be attractive as it only requires a minimal extension of the original model put forward by Husband and Ferreira (2015), its predictions are not consistent with the results found in Experiment 1, because it predicts that

alternatives could only be considered by virtue of their semantic association to the focus. If discourse-appropriate alternatives are only selected from among a set of semantic associates by suppressing activation on discourse-inappropriate alternatives, effects of discourse exclusion of a potential alternative should only be measurable on expressions that become activated in the first place by being semantically associated to the focus. But participants of Experiment 1 were able to distinguish excluded alternatives from non-excluded ones even when these potential alternatives were not associated to the focus, so such an extension would not be compatible with the patterns observed here.

Another way this Selection Only hypothesis is not in line with the results from Experiment 1 is that a suppression mechanism can only operate after initial activation is spread to those associated expressions. In Husband and Ferreira's data, the early lexical decision benefit for non-alternative associates (at 0 ms) was suppressed entirely at the point in time where alternatives were distinguished from non-alternatives (at 750 ms). Extending this type of selection mechanism to the integration of discourse-specific information therefore predicts that effects of discourse exclusion are similarly only measurable after effects of semantic association have been suppressed. Crucially, in Experiment 1, effects of discourse exclusion were found while effects of semantic association were still present, suggesting that semantic association and discourse exclusion could affect the activation of potential alternatives simultaneously.

These findings therefore motivate another class of possible extensions, in which the activation of alternatives is already guided by discourse information early on. In this type of extension, which we call Discourse-guided Activation, the original model from Husband and Ferreira (2015) is supplemented with an additional mechanism that deals with discourse information separately from semantic association. This discourse-specific mechanism would be able to revisit comprehenders' representations of the discourse context to activate discourse-appropriate alternatives directly. Under this view, the initial activation of focus alternatives does not solely rely on the process of spreading activation to semantic associates, which means that the process of activating discourse-appropriate alternatives could be triggered

already before the focus is encountered, e.g., whenever a cue to the presence of focus marking (such as a focus particle or contrastive intonation) has been perceived. This type of extension could be considered a more extensive revision to the foundational model as it involves adding an additional discourse-specific mechanism to the mental architectures employed in alternative set processing.

Despite the independent role of discourse-specific information that is assumed under the Discourse-guided Activation hypothesis, such an extension would not predict effects of semantic association to be entirely absent. Effects of semantic association are still predicted to arise under this extension if these effects are due to a domain-general process of spreading activation that is part of all word recognition. Because of the generality of these semantic associative effects, any extension that incorporates a discourse-specific activation mechanism must therefore also involve some way to suppress activation on inappropriately activated semantic associates. The crucial difference between the Discourse-guided Activation and the Selection Only hypotheses is therefore not whether suppression takes place at all, but whether or not discourse appropriate alternatives may be activated alongside semantic associates: Only if the initial activation of alternatives is guided by discourse context is it possible for effects of discourse exclusion to arise during this initial activation stage.

The predictions of each of these extensions are summarized in Figure 2, where the activation levels for both associated and non-associated, as well as discourse excluded and non-excluded alternatives are plotted over time. As can be seen in this figure, the Discourse-guided Activation extension but not the Selection Only predicts that effects of discourse exclusion can be observed contemporaneously with effects of semantic association. Moreover, since differentiating between discourse excluded and non-excluded alternatives does not solely rely on suppressing activation on semantic associates under the Discourse-guided Activation hypothesis, this extension (but not the Selection Only one) predicts early effects of discourse exclusion to be detectable on alternatives that are not associated to the focus. If decision times on target alternatives are indeed inversely correlated with the activation levels of those

alternatives, as was assumed in Experiment 1, the results of this experiment are more in line with Discourse-guided Activation hypothesis depicted on the right, where, at an early stage, potential alternatives are activated due to their discourse appropriateness, in tandem with activation from semantic association.

Although the results from Experiment 1 are thus only consistent with Discourse-guided Activation, the slowdowns on discourse-excluded alternatives observed in this study may not be surprising on their own, because mention of discourse-inappropriate alternatives likely led to a general incoherence with the preceding discourse context. The fact that such slowdowns indeed arose tells us that, generally, comprehenders form expectations about upcoming expressions based on information that is specific to the particular event described in the discourse context. The discourse exclusion effect observed here could therefore simply be due to properties of the preceding discourse that make the subsequent mention of discourse excluded alternatives generally less natural or less predictable, not to the activation of alternatives in the processing of focus in particular. In Experiment 2, we therefore tested response times on target alternatives in the absence of a clear cue to the presence of focus marking on the alternative trigger, demonstrating that the presence of a focus particle is necessary to obtain the discourse-exclusion effect.

Experiment 3 was furthermore designed to test the predictions that both of these possible extensions make for later stages of processing. As can be seen in Figure 2, both scenarios predict a later difference between associated alternatives that are discourse-appropriate and those that are discourse-inappropriate, but the crucial difference between the two extensions is the relative activation on non-associated alternatives. If discourse-specific information is revisited using a mechanism distinct from semantic priming, as in Discourse-guided Activation, late effects of discourse context should be present among non-associated alternatives. This extension thus predicts that, with a large enough distance between trigger and targets, longer decision times can be observed on target alternatives that are non-associated and excluded, relative to non-associated non-excluded ones. As the effect of semantic association should fade away as time passes, this

extension moreover predicts there to be no reliable difference in decision times on associated non-excluded alternatives compared to non-associated excluded ones. In contrast, if discourse information is *only* used to select alternatives from among semantic associates, as under a SELECTION ONLY hypothesis, there should not be any facilitation for non-associated non-excluded over non-associated excluded ones, because neither were activated at the initial stage. Because initial activation is only maintained on associated non-excluded alternatives, the SELECTION ONLY extension also predicts there to be a difference in response times on associated non-excluded alternatives relative to non-associated non-excluded ones. Experiment 3 was designed to test these predictions.

Experiment 2

To establish that decision time differences on the target alternatives in Experiment 1 arose specifically due to focus marking on the preceding alternative trigger, the alternative trigger was not put in focus with a focus particle in the trigger sentences of Experiment 2. Experiment 2 tested whether the discourse exclusion effect observed in Experiment 1 is still observed when the focus particle is removed. If the RT differences of the previous study were, at least in part, due to activation of alternatives in the processing of focus, then those effects should be diminished or entirely disappear in the absence of a focus particle directly preceding the alternative trigger.

Method

Materials. The materials of Experiment 2 were identical to the materials of Experiment 1, except that the focus particle *only* was removed from all target sentences, as in the example item in (8).

(8) a. *Discourse context:* The tourist asked for a variety of items, like some cheese and yogurt. There was already an ashtray on the table.
Trigger sentence: When the waiter returned, he remembered to bring some [yogurt] but no **cheese** to the table where the tourist was seated. assoc non-excl

b. *Discourse context:* The tourist asked for a variety of items, like an ashtray and

yogurt. There was already some cheese on the table.

Trigger sentence: When the waiter returned, he remembered to bring some [yogurt] but no **cheese** to the table where the tourist was seated. assoc excl

c. *Discourse context:* The tourist asked for a variety of items, like some cheese and an ashtray. There was already some yogurt on the table.

Trigger sentence: When the waiter returned, he remembered to bring an [ashtray] but no **cheese** to the table where the tourist was seated. non-assoc non-excl

d. *Discourse context:* The tourist asked for a variety of items, like an ashtray and yogurt. There was already some cheese on the table.

Trigger sentence: When the waiter returned, he remembered to bring an [ashtray] but no **cheese** to the table where the tourist was seated. non-assoc excl

Participants. 80 participants were recruited via the Prolific platform for web-based research and were paid a $12 hourly rate for their participation. All participants were native speakers of English and gave explicit consent to participate. Participants who had an accuracy of less than 70% on the comprehension questions or who did not complete more than 75% of the Maze sentences were excluded from the analysis. Data from 69 participants were included in the analysis; 11 participants were excluded due to their comprehension question accuracy and/or Maze completion rates.

Procedure. The procedure was the same as that of Experiment 1.

Results

The mean comprehension question accuracy of Experiment 2 was 82%, and the overall mean completion rate of the Maze sentences in experimental trials and fillers was 94%. On experimental trials, 276 observations on the critical region were missing due to participants' failure to complete target sentences (8% of all experimental trials). As in Experiment 1, response times below 200 ms and above 8 standard deviations above the mean were excluded from the analysis (amounting to four observations on the critical region, which were clear visual outliers).

Mean response times for the target word and its surrounding regions in all conditions are given in Table 4. They are plotted with 95% confidence intervals in Figure 3.

The data analysis was analogous to that of Experiment 1, with models including fixed effects of semantic association and discourse exclusion (coded 0.5, -0.5), and with associated and non-excluded conditions treated as reference levels. Tables 5 and 6 present the posterior model estimates results for the log-transformed response times and untransformed response times on target words of Experiment 2, respectively. Both models found only one reliable main effect: positive estimates of association indicate that semantically associated target alternatives were responded to faster than target alternatives that were not semantically associated with the focus. Pairwise comparisons revealed no reliable differences between the non-associated excluded target alternatives and the non-associated non-excluded targets, for both log-transformed (β =0.016; 97.5% Cr.I. =[-0.003, 0.035]) and untransformed response times (β =33.27; 97.5% Cr.I. =[-18.25, 84.36]). Pairwise comparisons between the associated excluded and associated non-excluded targets did not reveal any reliable differences either, again for both log-transformed (β =0.017; 97.5% Cr.I. =[-0.004, 0.037]) and untransformed response times (β =34.86; 97.5% Cr.I. =[-19.85, 88.22]),

Post-hoc analysis. Additional models including the temporal distance (the sum of all response times on each trial between trigger and target alternative) as an independent variable again revealed positive posterior estimates for distance, suggesting that response times on the target were reliably longer with longer distances between trigger and target, for models run on both untransformed (β =29.20; 97.5% Cr.I. =[6.84, 51.31]) and log-transformed response times (β =0.012; 97.5% Cr.I. =[0.003, 0.02]). However, these models again revealed only a reliable main effect of association, but no reliable main effect of exclusion or interaction between association and exclusion.

Discussion

The effect of discourse exclusion found in Experiment 1 was not observed in Experiment 2. Experiment 1 showed that comprehenders integrate discourse-specific information, which rules

out expressions as alternatives to the focus, causing slowdowns on such discourse-excluded alternatives during incremental interpretation. But comprehenders were not sensitive to this type of information in Experiment 2, in which a focus particle was absent from target sentences. This indicates that the effect of discourse exclusion observed on alternatives in Experiment 1 was indeed due to the processing of the focus on the alternative trigger that preceded them.

These findings are consistent with those reported by Lowder et al. (2021), who showed that readers exploit the presence of a focus-sensitive particle as a cue to the location of a focus; this, in turn, allows them to rapidly compute which set of expressions contrast with that focus, affecting reading times on subsequent alternatives. Together with the findings from Experiment 1, they are also consistent with studies which suggest more generally that, in the presence of a focus particle, discourse information is integrated early on in the processing of alternatives (Kim et al., 2015; Sedivy, 2002; Washburn et al., 2011).

To account for the way in which such discourse information is utilized in the processing of focus, above we first considered an extension of the sequential activation-selection model, in which discourse-appropriate alternatives are only selected from among a set of activated associates. This Selection Only hypothesis predicted, first, that the effect of discourse exclusion only appears at a later stage of processing, after the effect of semantic association has already been suppressed; and, second, that discourse exclusion effects could only arise on alternatives that first became activated due to being semantically associated. The results of Experiment 1 falsified both these predictions: Effects of discourse exclusion were observed at a stage of processing where effects of semantic association were still present and slowdowns due to discourse exclusion were observed on non-associated alternatives, suggesting that discourse-appropriate alternatives are not only selected for by suppressing activation on previously activated semantic associates. Because Experiment 2 showed that these discourse exclusion effects were focus-specific, it must be the case that comprehenders generate hypotheses about expressions that can serve as focus alternatives in a specific discourse context, as in the Discourse-guided Activation hypothesis.

This does not require abandoning a role for semantic association altogether. It may be that

alternatives become activated simultaneously by spreading activation from a focus to semantically associated expressions and by accessing representations of the linguistic discourse context. Empirically, semantically associated alternatives were found to be facilitated over non-associated ones in both Experiment 1 and Experiment 2. As Husband and Ferreira (2015) suggest, it seems reasonable to assume that semantic association plays a very general role, giving rise to facilitative effects even in the absence of focus marking. In fact, they argue (p. 229) that comprehenders may take different strategies to activate alternatives in different scenarios. Perhaps in the absence of a discourse context that provides enough information about the nature of the alternative set, comprehenders may rely more on the way their general conceptual knowledge is organized; when foci are interpreted in rich enough discourses, alternatives may also become activated due to a combination of both semantic association and discourse information, as is suggested by the present data.

We should consider the possibility then that these two processes—automatic spreading of activation and reactivation of the discourse information—in fact happen in tandem. Such a model would not only straightforwardly account for the fact that comprehenders' strategies for activating alternatives may depend on their available information, but it would also correctly predict that expressions that are generally associated to the focus become activated early on, alongside discourse-appropriate, non-associated alternatives, as was observed in Experiment 1. Such a model would also be in line with results from Experiment 2, in which the activation of excluded and non-excluded alternatives was not differentiated, but a domain-general activation of semantic associates still gave rise to effects of association.

An extension which combines a general spreading activation mechanism with a distinct discourse reactivation mechanism would also make predictions for later stages of processing. For instance, despite the independent role of discourse information, automatic activation on associates that are non-substitutable or discourse-inappropriate must still be suppressed—as is also the case in the original model. Both the Selection Only and Discourse-guided Activation hypotheses therefore predict a later difference between those associated alternatives that are

discourse-appropriate and those that are discourse-inappropriate. The crucial difference between the two extensions, however, is that if conceptual and discourse information are treated by distinct mechanisms, as under Discourse-guided Activation, late effects of discourse context should be present among non-associated alternatives, too, while there should not be any facilitation for discourse-appropriate non-associated alternatives under the Selection Only hypothesis, because these alternatives were not activated at the initial stage. Thus, Experiment 3 was designed to test how semantic association and discourse exclusion interact when activation is measured on alternatives further removed from the alternative trigger.

Experiment 3

Like Experiment 1, Experiment 3 crossed association between the focus and the target alternative with discourse exclusion of those target alternatives. But unlike Experiment 1, response times were measured on alternatives with a longer distance between those alternatives and the preceding focus, allowing more time between the initial computation of the focus alternatives as triggered by *only* and the later explicit mention of these potential alternatives and their integration into the target sentence. If activation of semantically associated alternatives is only short-lived while the effect of discourse exclusion persists over time, as is suggested above, then Experiment 3 should only show an effect of discourse exclusion because the target alternative occurs in a position in which the effect of semantic association should already have subsided. If, on the other hand, discourse-appropriate alternatives are selected from among initially-activated semantic associates, activation on semantically associated and discourse-appropriate alternatives should be maintained.

Method

Experiment 3 makes use of the same context sentences as those used in Experiment 1 and Experiment 2, but the target sentences in Experiment 3 were constructed such that the distance between the focus and the target alternative inside the target sentences was longer than those in Experiment 1.

Materials. The materials of Experiment 3 were identical to the materials of Experiment 1, except that the target sentence now contained a longer distance between the focus and the target alternative. An example of an target item of Experiment 3 is given in (9).

(9) **Context:**

a. *Discourse context:* The tourist asked for a variety of items, like some cheese and yogurt. There was already an ashtray on the table.
Trigger sentence: When the waiter returned, he remembered to bring only some [yogurt] but he forgot to bring any **cheese** to the table where the tourist was seated.
assoc non-excl

b. *Discourse context:* The tourist asked for a variety of items, like an ashtray and yogurt. There was already some cheese on the table.
*Trigger sentence:*When the waiter returned, he remembered to bring only some [yogurt] but he forgot to bring any **cheese** to the table where the tourist was seated.
assoc excl

c. *Discourse context:* The tourist asked for a variety of items, like some cheese and an ashtray. There was already some yogurt on the table.
Trigger sentence: When the waiter returned, he remembered to bring only an [ashtray] but he forgot to bring any **cheese** to the table where the tourist was seated.
non-assoc non-excl

d. *Discourse context:* The tourist asked for a variety of items, like an ashtray and yogurt. There was already some cheese on the table.
Trigger sentence: When the waiter returned, he remembered to bring only an [ashtray] but he forgot to bring any **cheese** to the table where the tourist was seated.
non-assoc excl

To ensure that the target appeared at a point where (most of) the initial activation of semantic associates would already have decayed, the distance between the focus and the target alternatives

in Experiment 3 was always at least three words longer than that of Experiment 1. Since A-maze response times are generally longer than 250 ms per word, the minimal delay between the focus and the target was at least 750 ms longer than that of Experiment 1. Response times on the target therefore provide a measure of activation at a point in time that is at least as late as the response times measured in Husband and Ferreira's lexical decision task at a 750 ms delay, where initial activation of semantic associates was already found to be suppressed.

Participants. 79 participants were recruited via the Prolific platform for web-based research and were paid a $12 hourly rate for their participation. All participants were native speakers of English and gave explicit consent to participate. Data from 70 participants was included in the analysis. Data from 9 participants, who did not have an accuracy of at least 70% on the comprehension questions and/or completed less than 75% of the Maze sentences, was excluded from the analysis.

Procedure. The procedure was the same as that of Experiment 1 and Experiment 2.

Results

The mean comprehension question accuracy for Experiment 3 was 86%, and the mean completion rate of the maze target sentences of Experiment 3 was 93%. On experimental trials, 373 observations on the critical region were missing due to participants' failure to complete target sentences (11% of all trials). Response times below 200ms and above 8 standard deviations above the mean were excluded from the analysis (amounting to a single observation on the critical region).

Mean response times for the target word and its surrounding regions in all conditions are given in Table 7. They are plotted with 95% confidence intervals in Figure 4.

The data analysis was analogous to that of Experiment 1 and Experiment 2, with models including fixed effects of semantic association and contextual exclusion (deviation-coded), and with associated and non-excluded conditions treated as reference levels. Tables 8 and 9 present the posterior estimates for the models of Experiment 3 log-transformed response times and untransformed response times on target words, respectively. Both models found only one reliable

main effect: positive estimates of exclusion indicate that discourse excluded target alternatives were responded to more slowly than target alternatives that were not discourse excluded. The main effect of association was not reliable. However, the interaction between exclusion and association was reliable, and pairwise comparisons revealed that the difference between the non-associated excluded target alternatives and the associated excluded targets was reliable, for both log-transformed (β =-0.027; 99% Cr.I=[-0.044,-0.010]) and untransformed (β =73.95; 99% Cr.I=[9.61,135.04]) response times, while the difference between non-associated non-excluded and associated non-excluded was not (log-transformed: β =.004; 99% Cr.I=[-.015,.023]; untransformed: β =23.17; 99% Cr.I=[-39.53,65.51]). Moreover, as in Experiment 1, the difference between the non-associated excluded condition and the non-associated non-excluded conditions was again reliable in both log-transformed (β =0.026; 99% Cr.I.=[0.010,0.045]) and untransformed (β =124.84; 99% Cr.I.=[82.37, 167.54]) response times, as was the difference between associated excluded and associated non-excluded conditions, in both log-transformed (β =0.015; 99% Cr.I.=[0.001,0.03]) and untransformed (β =46.27; 99% Cr.I.=[2.21, 90.11]) response times.

Discussion

In Experiment 3, response times on discourse excluded alternatives were longer than those on alternatives that were not discourse-excluded. In addition, the effect of semantic association between the preceding focus and the target alternative was reliable, but crucially only among alternatives that were excluded by the discourse context: response times on discourse-inappropriate alternatives that were not closely associated with the focus were even longer than associated but excluded alternatives.

Experiment 3 thus replicated the effect of discourse exclusion. Although both the Selection Only and Discourse-guided Activation extensions predict some difference between excluded and non-excluded alternatives at this later stage of processing, the overall response time pattern in Experiment 3 is consistent only with a model in which discourse information and semantic association are dealt with separately, for two reasons.

First, as in Experiment 1, the slowdown on non-associated discourse-excluded alternatives, in particular relative to non-associated non-excluded ones, can only be explained if discourse-specific information about the salient expressions that can serve as discourse-appropriate alternatives is encoded and directly re-accessed in the processing of that focus independently from the use of conceptual information. If alternatives were only selected from among initially primed semantic associates, this particular effect would not be expected because such a selection mechanism would only *suppress* activation on discourse-inappropriate associates instead of *increasing* activation on non-associated but discourse-appropriate alternatives.

Second, there was no facilitatory effect of associated discourse-appropriate alternatives over non-associated ones in Experiment 3, not even numerically, despite Experiment 1 showing a reliable effect of semantic association among non-excluded alternatives. For such discourse-appropriate alternatives, any facilitation of associates over non-associates thus entirely disappeared as the distance between the focus and the target alternatives increased. This finding is only in line with a multiple-mechanism model, and not with a Selection Only model: if discourse exclusion could only affect response times via a selection mechanism that suppresses activation of discourse-inappropriate semantic *associates*, then semantically associated and appropriate alternatives should generally be facilitated over semantically non-associated ones. This is because, while activation of discourse-appropriate associates would be maintained over time, non-associated alternatives would not become activated in the first place, even if they were discourse-appropriate.

General discussion

Three on-line comprehension studies tested the nature and time course of the information comprehenders utilize to activate contrastive alternatives in the processing of focus. Target sentences tested how the processing of a focus introduced by a focus particle (*only*) affected how an alternative, explicitly mentioned after the focus, was interpreted and responded to. Preceding discourses always mentioned such alternative expressions, manipulating whether they were either

discourse-appropriate as alternatives to the focus or not.

The results of Experiment 1 and Experiment 3 showed slowdowns on expressions that were specified in the discourse context as inappropriate alternatives to the preceding focus, suggesting that discourse information can guide comprehenders in ruling out alternatives as part of the relevant alternative set, despite their being both salient and closely semantically associated with the focus. Experiment 1, in particular, showed that this type of discourse information can be taken into account during processing of material in close proximity to the focus itself. The results of Experiment 2 verified that these effects of discourse exclusion were due to the processing of the focus, as they disappeared in the absence of a focus particle.

These experiments were designed to test two potential extensions of the model put forth by Husband and Ferreira (2015), in which the process of activating alternatives is either temporally encapsulated from information from both the sentential and discourse context (i.e., Selection Only), or is potentially guided by discourse information already early on (i.e, Discourse-guided Activation). In the Selection Only model, contrastive alternatives to a focus only become activated in two sequential steps. First, when the meaning of the focused lexical material is retrieved, activation is spread to expressions that are semantically associated with the expression in focus. Expressions that are substitutable for the alternative trigger and that are discourse appropriate alternatives are then selected for by suppressing activation of non-substitutable and discourse-inappropriate associates on the basis of information from the sentence context (which determines whether an expression is substitutable for the alternative trigger and therefore a viable alternative) and discourse-specific information (discourse exclusion). This extension predicts, first, that effects of discourse exclusion can only be found after effects of semantic association have already been suppressed; and second, that discourse information only affects those expressions that first become activated via their semantic association with the focused alternative trigger.

The results of Experiments 1-3 were not in line with this Selection Only extension of Husband and Ferreira's original model in the following two ways. First, while results of these

experiments showed early activation of associates, they also indicated that discourse-specific preferences can help facilitate non-associated but discourse-appropriate alternatives alongside semantic associates (Experiment 1). Since Experiment 1 found contemporaneous effects of association and discourse exclusion, this suggests that discourse information can affect decision times while effects of semantic priming are still active. Second, participants' behavior distinguished between non-associated alternatives that were excluded and those that were non-excluded, both at early (Experiment 1) and later stages of focus processing (Experiment 3). Together, this suggests that when foci are interpreted in rich enough discourse contexts, like the ones tested here, comprehenders do not rely solely on conceptual relationships between expressions to initially activate alternatives. They are also able to directly exclude alternatives with the information provided in the discourse context.

The current experimental setup has the advantage that activation could be measured on potential target alternatives that were embedded within the same sentence as the alternative trigger, while comprehension of the discourse was still ongoing. This contrasts with previous work, which has either adopted offline measures (Fraundorf et al., 2013; Fraundorf et al., 2010), or on-line measures such as the visual world paradigm in which looks to the object in focus, but not to potential alternatives could be measured (Kim et al., 2015), or lexical decision tasks which measured response times on words isolated from a sentential context. One shortcoming of the current method, however, was that activation on those potential target alternatives could only be measured several words removed from the expression in focus. The exact time-course of the obtained effects is therefore not completely parallel to those measured by Husband and Ferreira (2015), who visually presented targets at 0 ms after the offset of the alternative trigger. It could therefore be that the current experiments reflect comprehension processes which may not have been triggered in Husband and Ferreira's (2015) studies, and future work is necessary to establish how lexical decision times from cross-modal priming tasks relate to incremental decision times in read sentences.

However, in the present experiment, effects of discourse exclusion were still observed

during the same time window in which effects of semantic association were found, suggesting that it is not the case that activation from semantic priming has fully been suppressed by the point at which effects of discourse can be observed. These results are not straightforwardly compatible with the Selection Only extension, where contrastive alternatives to a focused alternative trigger can become activated by first spreading activation to semantically associated expressions and later suppressing activation on discourse-inappropriate associates. This extension therefore predicts effects of semantic association to have been suppressed by the time effects of discourse exclusion can arise. We take the contemporaneous effects of semantic association and discourse exclusion to support the Discourse-guided Activation extension, in which discourse information affects the activation of focus alternatives because comprehenders re-activate their representations of the discourse context directly, and generate alternatives based on such representations.

The Discourse-guided Activation hypothesis is compatible with several different mechanistic implementations, which still remain to be tested. For instance, the distinct types of discourse information that we identified in the Introduction—mention, situational availability, discourse exclusion—could play more or less active roles or could be used at different points in time to activate candidate sets of alternatives. We consider two possibilities: the present data could be accounted for by a system that re-activates all discourse-specific information directly, or by one that uses discourse exclusion and other types of discourse information in separate processes.

In the first possibility, discourse understanding involves not only the representation of a textbase (i.e., some representation of the linguistic surface structures), but also the activation, encoding, and updating of a *situation model* in episodic memory, i.e., the cognitive representation of the events, actions, individuals and states a discourse is about (Van Dijk and Kintsch 1983). One assumption of this model is that efficient comprehension in discourse is possible because knowledge is used strategically: what information is accessed depends on the goals of the language user, the amount of available information from the discourse context, the level of processing, and the degree of coherence needed for comprehension. To use this knowledge strategically, language users try to establish coherence relationships among the pieces of

linguistic structure that make up a discourse.

Adopting such a model of discourse processing would help with interpreting the pattern of results observed here, too. Since focus marking serves to indicate a relevant contrast between the focused marked expression and its (implicit) set of alternatives, it might be strategic to try to interpret the focus as contrasting with a set of expressions that are provided within the preceding discourse itself. The presence of a focus particle in particular may cue an upcoming focus as well as an upcoming contrast, and may thus trigger comprehenders to access information stored as part of the situation model in order to resolve such potential coherence relationships. In the processing of a focus, comprehenders would then be able to access parts of this situation model in order to re-activate those alternative expressions that are discourse-appropriate as alternatives to the focus, giving rise to both early (Experiment 1) and late (Experiment 3) facilitation of discourse-appropriate alternatives, as well as memory benefits (Fraundorf et al., 2013). In short, as is already suggested by Fraundorf et al. (2013), focus processing may constitute a discourse comprehension process in which both general knowledge and discourse-specific information can be used strategically to construct an alternative set.

Under the second possibility, compatible with previous findings (Kim et al., 2015; Muxica & Harris, 2025; Washburn et al., 2011), mentioned and situationally available information also play distinct roles in the activation of alternatives, such that these types of discourse information allow comprehenders to directly activate alternatives as well. Note that even in such a system, representations of the semantic content of the discourse would need to be revisited in addition to information about which expressions have been mentioned. This is because in the present studies, both discourse-appropriate and discourse-inappropriate alternatives were mentioned in the preceding discourse, so solely reactivating mentioned expressions would not suffice to distinguish between excluded and non-excluded alternatives. Thus, information about what expressions have been mentioned in the preceding discourse could be used to activate alternatives alongside semantic associates, while discourse-specific information could then be used to select discourse-appropriate alternatives from among this candidate set.

Unlike a pure Selection Only model, in which discourse information is only used to suppress activation on semantic associates, a system like this predicts that excluded non-associated alternatives can be distinguished from non-excluded non-associated ones when they are mentioned in the preceding context, as was the case in Experiment 1 and Experiment 3. It also correctly predicts that discourse-mentioned but non-associated alternatives are activated early on. Although most results from Experiments 1-3 can thus straightforwardly be predicted by this implementation, note that the simultaneous effects of semantic association and discourse exclusion obtained in Experiment 1 could only be accounted for in this model if the selection of discourse-appropriate alternatives is assumed to occur already during semantic priming. In other words, these simultaneous effects can only be predicted under the assumption that activation from lexical associates decays *after* discourse-appropriate alternatives have already been selected from among mentioned ones. Thus, even though this implementation adopts a selection mechanism for the use of discourse-specific information, it still crucially differs from the Selection Only model in that activation of alternatives can be guided by discourse information at least as early as activation is spread to semantic associates. Further work in which the provided lexical material is systematically crossed with discourse exclusion is necessary to tease apart these two implementations of the Discourse-guided Activation model. This avenue of future research would also be valuable as it could shed light on the type of discourse representations that are accessed during alternative set processing, and during discourse comprehension more generally.

Both these mechanistic implementations of alternative set processing, in which there are separate mechanisms for semantic association and revisiting discourse representations, are consistent with previous evidence that the presence of a focus particle may cue comprehenders to set in motion the process of activating a set of alternatives even before the focus itself is encountered (Kim et al., 2015; Lowder et al., 2021). Given the lack of discourse exclusion effects in Experiment 2, it in fact seems likely that the focus particle plays an important role in triggering this process. Under the Discourse-guided Activation hypothesis, such pre-emptive use of discourse information may explain the contemporaneous effects of discourse exclusion and

semantic association in Experiment 1, too. The present results should therefore not necessarily be taken to indicate that the integration of discourse information takes place over a shorter temporal interval than the process of decaying semantic association. Instead, these findings are also compatible with a scenario in which the use of discourse-specific information commences before activation is first spread to semantic associates. Although such a scenario could again involve a selection mechanism for distinguishing discourse appropriate alternatives from discourse inappropriate ones, it is still not consistent with the time course assumed by the Selection Only hypothesis, in which it is the suppression of semantic associates in particular that constitutes the selection process. In contrast to the assumptions of the Selection Only hypothesis, these results thus suggest that the process of activating focus alternatives is not temporarily shielded from the use of discourse information, but is instead, like other discourse comprehension processes, guided by such information early on.

As discussed above, the re-activation of discourse information may take place independently from a more general-purpose, automatic spreading of activation to expressions that are conceptually related to the focused expression. Indeed, this general priming mechanism may have driven the effects in the previous studies which only tested semantically associated alternatives (and did not include non-associated alternatives in their designs), as well as some of the data presented here. But since semantic association was fully crossed with discourse appropriateness of an alternative in the present studies, a better picture of the way in which semantic association and discourse information interact was obtained in the current experiments. In short, these experiments indicated that, although both semantic associates and discourse-appropriate alternatives become activated in the comprehension of a focus (and both types of expressions are therefore facilitated in lexical decision, truth verification, memory, reading or Maze tasks), the activation of alternative expressions does not *depend* on the activation of semantic associates. Instead, they suggested that focus alternatives are also activated via a mechanism that accesses discourse-specific information directly.

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.

Boyce, V., Futrell, R., & Levy, R. P. (2020). Maze made easy: Better and easier measurement of incremental processing difficulty. *Journal of Memory and Language*, *111*, 104082.

Braun, B., Asano, Y., & Dehé, N. (2018). When (not) to look for contrastive alternatives: The role of pitch accent type and additive particles. *Language and Speech*, *62*(4), 751–778.

Braun, B., & Tagliapietra, L. (2010). The role of contrastive intonation contours in the retrieval of contextual alternatives. *Language and Cognitive Processes*, *25*(7-9), 1024–1043.

Bürkner, P.-C. (2017). Brms: An r package for bayesian multilevel models using stan. *Journal of Statistical Software*, *80*, 1–28.

Forster, K. I., Guerrera, C., & Elliot, L. (2009). The maze task: Measuring forced incremental sentence processing time. *Behavior Research Methods*, *41*(1), 163–171.

Fraundorf, S. H., Benjamin, A. S., & Watson, D. G. (2013). What happened (and what did not): Discourse constraints on encoding of plausible alternatives. *Journal of Memory and Language*, *69*(3), 196–227.

Fraundorf, S. H., Watson, D. G., & Benjamin, A. S. (2010). Recognition memory reveals just how contrastive contrastive accenting really is. *Journal of Memory and Language*, *63*(3), 367–386.

Gotzner, N., Wartenburger, I., & Spalek, K. (2016). The impact of focus particles on the recognition and rejection of contrastive alternatives. *Language and Cognition*, *8*(1), 59–95.

Hoeks, M., Toosarvandani, M., & Rysling, A. (2023). Processing of linguistic focus depends on contrastive alternatives. *Journal of Memory and Language*, *132*, 104444.

Husband, E. M. (2022). Prediction in the maze: Evidence for probabilistic pre-activation from the english a/an contrast. *Glossa Psycholinguistics*, *1*(1).

Husband, E. M., & Ferreira, F. (2015). The role of selection in generating focus alternatives. *Language, Cognition and Neuroscience*, *31*(2), 217–235.

Kim, C. S., Gunlogson, C., Tanenhaus, M. K., & Runner, J. T. (2015). Context-driven expectations about focus alternatives. *Cognition*, *139*, 28–49.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, *25*(2-3), 259–284.

Lowder, M. W., Ryan, G., Opie, J., & Kaminsky, E. (2021). Effects of contrastive focus on lexical predictability during sentence reading: The case of not only... but also constructions. *Quarterly Journal of Experimental Psychology*, *74*(1), 179–186.

Michel, J.-B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Pickett, J. P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., et al. (2011). Quantitative analysis of culture using millions of digitized books. *Science*, *331*(6014), 176–182.

Muxica, C., & Harris, J. (2025). Constructing focus alternatives from context and the limits of semantic priming. *Experiments in Linguistic Meaning*, *3*, 262–275.

R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Sedivy, J. C. (2002). Invoking discourse-based contrast sets and resolving syntactic ambiguities. *Journal of Memory and Language*, *46*(2), 341–370.

Spalek, K., Gotzner, N., & Wartenburger, I. (2014). Not only the apples: Focus sensitive particles improve memory for information-structural alternatives. *Journal of Memory and Language*, *70*, 68–84.

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. Academic press New York.

Washburn, M. B., Kaiser, E., & Zubizarreta, M. L. (2011). *Focus facilitation and non-associative sets*.

Witzel, N., Witzel, J., & Forster, K. (2012). Comparisons of online reading paradigms: Eye tracking, moving-window, and maze. *Journal of Psycholinguistic Research*, *41*(2), 105–128.

Yan, M., & Calhoun, S. (2019). Priming effects of focus in mandarin chinese. *Frontiers in Psychology*, *10*, 1985.

Zehr, J., & Schwarz, F. (2018). Penncontroller for internet based experiments [https://doi.org/10.17605/OSF.IO/MD832].

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162.

Table 1

Experiment 1: mean RT and standard error of the mean in each condition two words before, at, and two words after the target word.

Condition	Previous -1	Previous	Critical region	Spillover	Spillover +1
assoc non-excl	804.37 (11.61)	810.91 (15.11)	977.66 (17.53)	1086.53 (19.30)	892.66 (17.15)
non-assoc non-excl	813.93 (11.29)	823.43 (17.94)	1028.86 (16.75)	1063.82 (19.18)	947.67 (25.72)
assoc excl	817.55 (14.68)	797.11 (14.95)	1071.86 (21.74)	1077.08 (21.64)	930.20 (17.28)
non-assoc excl	841.73 (14.64)	814.00 (14.83)	1163.81 (22.93)	1118.11 (23.02)	924.25 (20.09)

Table 2

Posterior model estimates of the Bayesian mixed effects model on logRTs of Experiment 1.

Population-level effects

	Estimate	Est.Error	95% CrI	Rhat	Bulk_{ESS}	Tail_{ESS}
(Intercept)	2.98	0.01	[2.96, 3.01]	1.00	2607	5029
Association	0.03	0.01	[0.02, 0.04]	1.00	14789	13609
Exclusion	0.03	0.01	[0.01, 0.06]	1.00	10801	11363
Assoc: Excl	0.01	0.01	[-0.01, 0.03]	1.00	23254	12947

By-item group-level effects

	Est. Std.	Error	95% CrI	Rhat	Bulk_{ESS}	Tail_{ESS}
(Intercept)	0.03	0.00	[0.02, 0.04]	1.00	5954	10542
Association	0.03	0.01	[0.02, 0.05]	1.00	6163	5891
Exclusion	0.05	0.01	[0.03, 0.06]	1.00	8124	10515
Assoc: Excl	0.02	0.01	[0.00, 0.05]	1.00	5415	7828

By-subject group-level effects

	Est. Std.	Error	95% CrI	Rhat	Bulk_{ESS}	Tail_{ESS}
(Intercept)	0.09	0.01	[0.08, 0.11]	1.00	3235	6018
Association	0.02	0.01	[0.00, 0.03]	1.00	3430	4467
Exclusion	0.06	0.01	[0.04, 0.07]	1.00	8765	12061
Assoc: Excl	0.04	0.02	[0.00, 0.07]	1.00	3380	3825

Table 3

Posterior model estimates for the population-level effects of Bayesian mixed effects model on untransformed RTs of Experiment 1.

Population-level effects

	Estimate	Est.Error	95% CrI	Rhat	Bulk_{ESS}	Tail_{ESS}
(Intercept)	1053.74	32.76	[989.07,1118.81]	1.00	2808	5299
Association	70.74	23.56	[23.82, 117.06]	1.00	12464	11962
Exclusion	111.70	33.57	[45.64, 177.30]	1.00	7964	10940
Assoc:Excl	38.03	37.80	[-36.09, 112.55]	1.00	21402	12440

By-item group-level effects

	Est. Std.	Error	95% CrI	Rhat	Bulk_{ESS}	Tail_{ESS}
(Intercept)	93.39	14.12	[68.08, 123.48]	1.00	7471	11206
Association	104.42	27.36	[48.69, 158.90]	1.00	5788	5791
Exclusion	138.29	26.56	[88.71, 194.53]	1.00	7789	9445
Assoc: Excl	59.17	43.19	[2.72, 160.14]	1.00	6360	8770

By-subject group-level effects

	Est. Std.	Error	95% CrI	Rhat	Bulk_{ESS}	Tail_{ESS}
(Intercept)	241.49	22.15	[202.01, 288.98]	1.00	3732	6736
Association	60.86	25.65	[9.63, 111.92]	1.00	4761	4049
Exclusion	178.84	24.57	[133.78, 230.03]	1.00	8647	12031
Assoc: Excl	109.93	62.94	[6.54, 234.54]	1.00	3503	6188

Table 4

Experiment 2: mean RT and standard error of the mean in each condition two words before, at, and two words after the target word.

Condition	Previous -1	Previous	Critical region	Spillover	Spillover +1
assoc non-excl	733.24 (11.31)	796.32 (15.30)	908.02 (14.38)	995.61 (19.85)	878.65 (19.62)
non-assoc non-excl	755.07 (15.62)	753.98 (12.30)	945.13 (15.77)	1003.08 (17.64)	851.83 (16.69)
assoc excl	756.65 (13.54)	788.01 (14.05)	945.85 (14.32)	995.17 (17.93)	873.55 (16.81)
non-assoc excl	733.67 (10.82)	776.01 (14.73)	985.10 (16.14)	987.74 (18.61)	876.46 (20.12)

Table 5

Model estimates for the population-level effects of Bayesian mixed effects model on logRTs of Experiment 2.

Population-level effects

	Estimate	Est.Error	95% CrI	Rhat	Bulk_{ESS}	Tail_{ESS}
(Intercept)	2.9477	0.0107	[2.9267, 2.9685]	1.00	2922	5897
Association	0.0153	0.0063	[0.0031, 0.0278]	1.00	14559	13228
Exclusion	0.0160	0.0085	[-0.0007, 0.0328]	1.00	14416	12511
Assoc: Excl	-0.0004	0.0105	[-0.0210, 0.0202]	1.00	26619	13192

By-item Group-level effects

	Est. Std.	Error	95% CrI	Rhat	Bulk_{ESS}	Tail_{ESS}
(Intercept)	0.0329	0.0044	[0.0252, 0.0424]	1.00	7317	11165
Association	0.0229	0.0069	[0.0094, 0.0369]	1.00	8208	7407
Exclusion	0.0426	0.0074	[0.0291, 0.0582]	1.00	9199	10888
Assoc: Excl	0.0280	0.0156	[0.0019, 0.0596]	1.00	4898	6802

By-subject Group-level effects

	Est. Std.	Error	95% CrI	Rhat	Bulk_{ESS}	Tail_{ESS}
(Intercept)	0.0790	0.0075	[0.0660, 0.0952]	1.00	4075	7891
Association	0.0179	0.0079	[0.0020, 0.0331]	1.00	4371	4015
Exclusion	0.0286	0.0082	[0.0108, 0.0437]	1.00	3252	2566
Assoc: Excl	0.0110	0.0083	[0.0004, 0.0309]	1.00	9580	8504

Table 6

Model estimates for the population-level effects of Bayesian mixed effects model on raw RTs of Experiment 2.

Population-level effects

	Estimate	Est.Error	95% CrI	Rhat	Bulk_{ESS}	Tail_{ESS}
(Intercept)	949.17	27.96	[894.66,1003.71]	1.00	2451	4977
Association	46.24	18.87	[9.78, 83.07]	1.00	9617	11781
Exclusion	34.06	22.69	[-11.14, 78.35]	1.00	12400	12425
Assoc:Excl	-1.59	28.90	[-57.65, 54.65]	1.00	25324	12302

By-item group-level effects

	Est. Std.	Error	95% CrI	Rhat	Bulk_{ESS}	Tail_{ESS}
(Intercept)	81.94	11.47	[61.89,106.36]	1.00	6404	10071
Association	78.22	18.50	[43.47,116.01]	1.00	10599	10087
Exclusion	113.12	20.72	[75.25,156.89]	1.00	8158	11395
Assoc:Excl	64.25	40.87	[3.27,152.42]	1.00	4677	7496

By-subject group-level effects

	Est. Std.	Error	95% CrI	Rhat	Bulk_{ESS}	Tail_{ESS}
(Intercept)	206.39	19.09	[172.88,247.10]	1.00	3588	6742
Association	57.90	17.72	[23.25, 93.62]	1.00	9329	8581
Exclusion	64.60	24.74	[11.40,110.47]	1.00	2990	3566
Assoc:Excl	28.47	21.87	[1.13, 81.13]	1.00	11313	8789

Table 7

Experiment 3: mean RT and standard error of the mean in each condition two words before, at, and two words after the target word.

Condition	Previous -1	Previous	Critical region	Spillover	Spillover +1
assoc non-excl	757.91 (8.74)	804.98 (10.92)	952.71 (15.23)	1043.42 (18.57)	938.64 (23.82)
non-assoc non-excl	752.20 (8.18)	785.82 (11.95)	941.03 (14.67)	1058.23 (18.96)	881.46 (12.55)
assoc excl	763.45 (11.93)	802.52 (11.03)	1001.39 (17.28)	1053.14 (27.11)	994.36 (29.20)
non-assoc excl	754.51 (8.93)	821.71 (12.18)	1068.83 (18.03)	1044.76 (17.86)	923.66 (19.45)

Table 8

Posterior estimates for the population-level effects of Bayesian mixed effects model on logRTs of Experiment 3.

Population-level effects

	Estimate	Est.Error	95% CrI	Rhat	Bulk_{ESS}	Tail_{ESS}
(Intercept)	2.9676	0.0102	[2.9476, 2.9876]	1.00	3263	6210
Association	0.0118	0.0063	[-0.0005, 0.0241]	1.00	16736	12422
Exclusion	0.0322	0.0075	[0.0172, 0.0470]	1.00	13916	12744
Assoc: Excl	0.0321	0.0106	[0.0113, 0.0531]	1.00	26943	13123

By-item group-level effects

	Est. Std.	Error	95% CrI	Rhat	Bulk_{ESS}	Tail_{ESS}
(Intercept)	0.0402	0.0051	[0.0314, 0.0511]	1.00	5409	9739
Association	0.0259	0.0082	[0.0093, 0.0416]	1.00	3968	4558
Exclusion	0.0352	0.0076	[0.0205, 0.0507]	1.00	6483	8447
Assoc: Excl	0.0281	0.0171	[0.0015, 0.0633]	1.00	4176	7497

By-subject group-level effects

	Est. Std.	Error	95% CrI	Rhat	Bulk_{ESS}	Tail_{ESS}
(Intercept)	0.0680	0.0065	[0.0566, 0.0821]	1.00	4097	7005
Association	0.0092	0.0066	[0.0004, 0.0244]	1.00	6198	8971
Exclusion	0.0215	0.0079	[0.0056, 0.0366]	1.00	5063	5269
Assoc: Excl	0.0176	0.0125	[0.0008, 0.0463]	1.00	6466	7943

Table 9

Posterior estimates for the population-level effects of Bayesian mixed effects model on raw RTs of Experiment 3.

Population-level effects

	Estimate	Est.Error	95% CrI	Rhat	Bulk_{ESS}	Tail_{ESS}
(Intercept)	993.94	25.90	[943.87, 045.23]	1.00	4435	6886
Association	27.45	17.28	[-6.49, 61.39]	1.00	17351	12751
Exclusion	88.81	22.35	[44.29, 133.02]	1.00	12335	12534
Assoc:Excl	80.86	30.81	[20.51, 140.30]	1.00	22978	12691

By-item group-level effects

	Est. Std.	Error	95% CrI	Rhat	Bulk_{ESS}	Tail_{ESS}
(Intercept)	102.99	13.66	[79.15, 132.82]	1.00	6114	9935
Association	64.22	24.67	[12.52, 111.52]	1.00	4357	4044
Exclusion	100.98	22.93	[56.38, 147.03]	1.00	6351	7405
Assoc: Excl	57.16	40.91	[2.20, 149.97]	1.00	5639	7713

By-subject group-level effects

	Est. Std.	Error	95% CrI	Rhat	Bulk_{ESS}	Tail_{ESS}
(Intercept)	165.45	16.23	[136.81, 200.91]	1.00	4946	8708
Association	24.71	18.56	[0.97, 68.30]	1.00	6826	8594
Exclusion	69.67	21.37	[27.95, 112.11]	1.00	7119	6924
Assoc: Excl	46.83	34.29	[1.90, 127.75]	1.00	7585	8560

Figure 1

mean RT in each region per condition of Experiment 1. Error bars represent 95% confidence intervals.

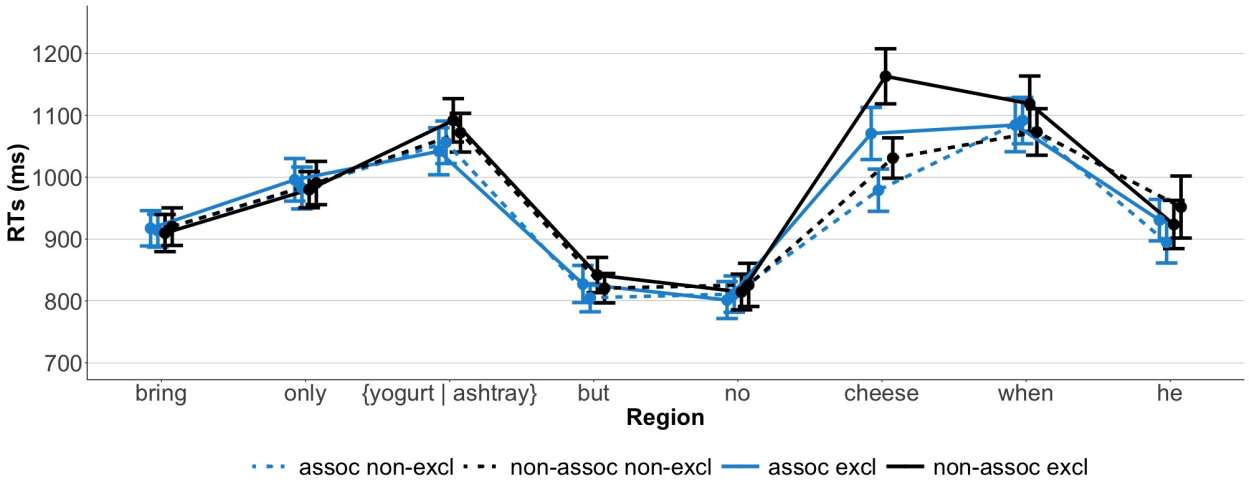


Figure 2

Predicted activation levels for Selection Only *(left) and* Discourse-guided Activation *(right)*

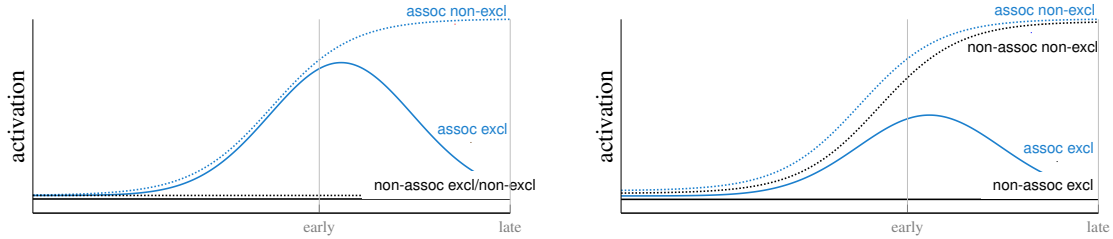


Figure 3

Experiment 2: mean RT in each region in each condition. Error bars represent the 95% confidence interval.

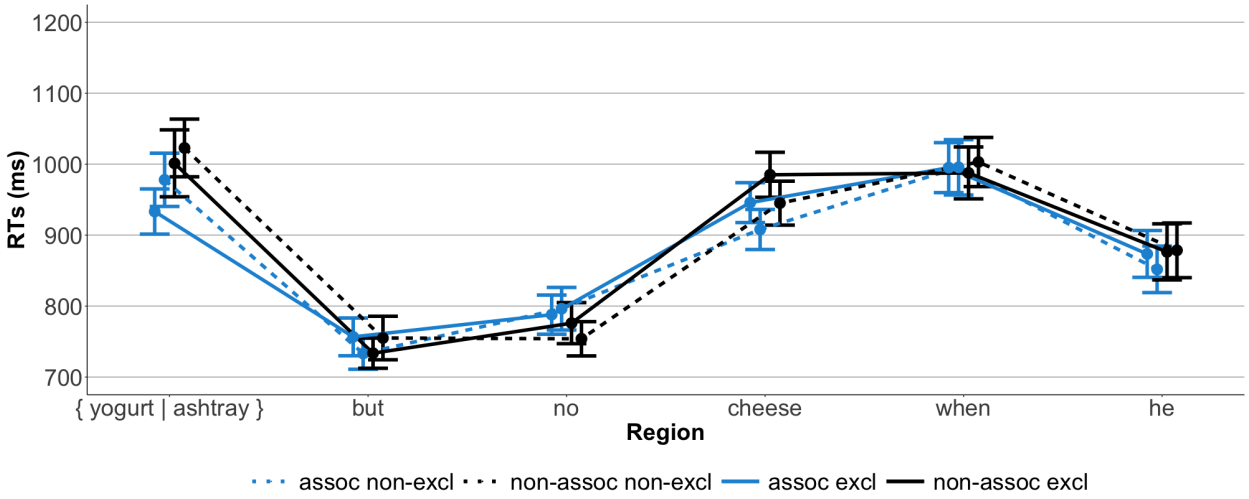
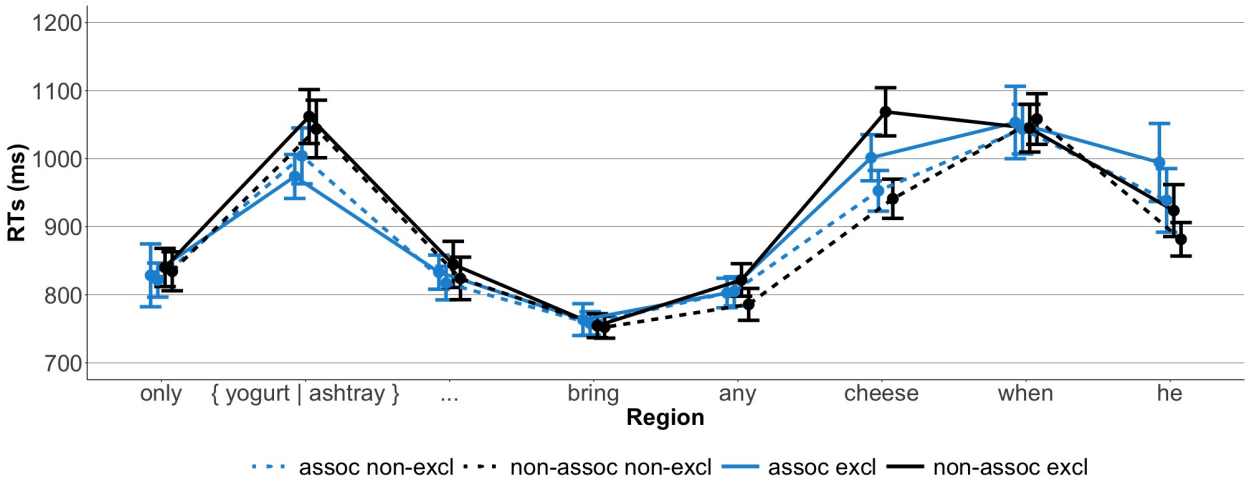


Figure 4

Experiment 3: mean RT in each region in each condition. Error bars represent the 95% confidence interval.



Appendix A

		Materials
1	**Context:**	The tourist had asked for a variety of items, such as cheese and some yogurt. There was already an ashtray on the table.
	Target:	When the waiter returned, he remembered to bring only some yogurt but no cheese to the table.
2	**Context:**	At the bar yesterday, John had ordered some wine and some beer. He didn't order any nuts the whole evening.
	Target:	All of a sudden, he had to leave quickly and he finished only his beer but not his wine when he stood up.
3	**Context:**	This weekend I made a few phone calls, including to my uncle and to my aunt. I couldn't call my bank until next week.
	Target:	I got a call back only from my aunt but not from my uncle, before the end of the weekend.
4	**Context:**	Erin's neighbor has all kinds of interesting objects, such as an antique violin and a piano. She was hoping to get some clocks at some point in the future.
	Target:	She showed Erin only a piano but not a violin, while Erin was at her house.
5	**Context:**	Ben loves to help out on his uncle's farm, for example by taking care of the ponies and the horses. His uncle doesn't keep any ducks at the farm anymore.
	Target:	This summer, Ben's uncle will keep raising only some horses but no ponies, even though Ben liked them very much.
6	**Context:**	The corner store sells a bunch of things, such as magazines and newspapers. They never sold any cigarettes.
	Target:	Last summer, they only stopped selling newspapers but not magazines due to supply chain issues.
7	**Context:**	The city council had big plans to improve the neighborhood, including the construction of a bus station and a metro station. A few years ago, they had already built a swimming pool.
	Target:	It will be difficult to get construction permits, but only for a metro station, not for a bus station, until they meet with the mayor.
8	**Context:**	In the cabinet underneath the sink, Stephanie stored various things, like a hammer and a screwdriver. She didn't keep her soap there.
	Target:	While cleaning out the cabinet, she threw away only her screwdriver but not her hammer because she wanted to use it later.

9 **Context:** The stylist added some sofas and chairs to the hotel lobby. She decided not to put any flowerpots there.

Target: When the owner saw the results, she noticed only some chairs but no sofas in the corner of the room.

10 **Context:** After breaking up with his girlfriend, Peter put her tv and her radio on Craigslist. He decided to keep her bike for himself.

Target: In the end, he managed to sell only her radio but not her tv, because potential buyers were looking for a lower price.

11 **Context:** In her first year of college, Monique was very good at biology and chemistry. She had dropped her sports class earlier that year.

Target: After winter break, she kept performing well only in chemistry but not in biology for a while.

12 **Context:** At the zoo, they used to have tigers and lions. They did not have the right permits to add some pelicans to their new exhibit.

Target: After animal rights activists discovered how some of the animals were treated, they kept only some lions but no tigers until they improved their living conditions.

13 **Context:** Jess went into town to get some new shoes and socks. She was also planning to order some new pencils on Amazon later that week.

Target: The stores were almost empty, and she found only some socks but no shoes before she had to go home.

14 **Context:** The artist who has a booth at the local fair sells bracelets made with different materials, like wool and cotton. She has never used metal before.

Target: People bought her bracelets, but only those with cotton, not with wool this time.

15 **Context:** Magda still needed some things to finish her new tiny house, such as some windows and a door. She couldn't find any tape anywhere in the store.

Target: At the hardware store, she thought the prices were reasonable, but only of the doors and not of the windows, even though she brought a lot of cash.

16 **Context:** The concierge was busy fixing the damages from the storm, including the broken fence and the gate. Luckily, the camera on the other side of the property didn't need any repairs.

Target: He managed to fix only the gate but not the fence before his workday was over.

17 **Context:** Isabel had only eyes for her new project, and she ignored important emails and some letters. She did, however, respond to the software updates that came in.

Target: When she finished the project, she finally took care of some things, but only of the letters and not of any emails in her backlog.

18 **Context:** Aron was getting some final things for his son's pirate-themed birthday party, like a pie and some cake. Not knowing what his son's friends could handle, he had decided that he wouldn't buy any swords for the party.

Target: It was already late, so he managed to buy only the cake but no pie, before the stores closed

19 **Context:** Brenda was busy packing, and in her suitcase there were some boots and sandals. She completely forgot to pack her toothbrush because she was in such a hurry.

Target: When she was going through security at the airport the next day, they would let her bring only her sandals but not her boots in her hand luggage

20 **Context:** Because that region is extremely remote, there are only a few ways to get there, such as by airplane or by helicopter. There is no way you can reach the area by bus at all.

Target: Jonathan knew how to get there, but he had gone there only by helicopter and not by airplane, even though it takes less time.

21 **Context:** Owen and Chris are organizing a big picnic at the park, and they had asked people to bring some tomatoes and some cucumbers. They had already brought a cooler with some beers in it.

Target: Their friends managed to get only some cucumbers but no tomatoes before it started raining.

22 **Context:** In his bag, David had packed a few things for the weekend, like jeans and a pair of shorts. He forgot to bring a book on his trip.

Target: By the end of the weekend, he had unpacked only his shorts but not his jeans, because it was unexpectedly warm.

23 **Context:** The organizers of the workshop had invited some painters and some sculptors. They deliberately didn't ask a lawyer this time.

Target: A few weeks before the event, they heard back only from a sculptor but not from any painters, even though they had sent a number of emails.

24 **Context:** Daniel was planning a deep-clean of his house this weekend, including the stove and the oven. His wife had already cleaned the garage last week.

Target: He managed to clean only the oven but not the stove, before going to bed.

25 **Context:** Sophia made a quick run to the grocery store to get some flour and some milk. Dan texted her that she didn't need to bring any sponges from the store.

Target: At the store, she remembered only to buy milk but no flour for some reason.

26 **Context:** Lily loves to go to thrift stores, and this Saturday she was hoping to find a sweater or a jacket. Last week she found a great puzzle so she didn't need one of those anymore.

Target: At the store she managed to find only a jacket but not a sweater, to her chagrin.

27 **Context:** The photographer made a list of things to bring to the shoot the next day, including an extra charger and some batteries. Her colleague was bringing some snacks so she didn't have to.

Target: The next morning she remembered to bring only some batteries but no charger because it slipped her mind.

28 **Context:** This 3-bedroom apartment has a few nice perks. For example, it comes with a large attic and a basement. It doesn't come with a year-long lease, unfortunately.

Target: The advertisement mentioned only the basement but not the attic on the website.

29 **Context:** Jim had gone to Ikea to get spoons and some knives. He wanted to go to an antique store later to find a nice bed.

Target: Later, his partner returned only the knives but no spoons, even though they were quite expensive.

30 **Context:** Maria wanted to be a hairdresser, and she used to practice on her mother and her sister. Unfortunately, her dog 's hair was too brittle for her to work with.

Target: Nowadays, Maria still likes to cut only her sister's hair but not her mother's, for some reason.

31 **Context:** At the lost-and-found of the museum, there were a range of things that people had left behind, such as a purse and a bag. There weren't any cameras that had been found recently.

Target: After the museum had made some announcements, people picked up only the bag but no purse, even though the purses looked more expensive.

32 **Context:** At the donation center, several volunteers were helping to gather supplies for the homeless, like pillows and blankets. Despite many requests, no one had dropped off any food at the donation center.

Target: At the end of the day, the volunteers were happy only with the blankets but not with any pillows, which was unfortunate.

33 **Context:** Most of the people who were at the protest were unionized, like the doctors and the nurses. Unfortunately, none of the carpenters had shown up to the protest.

Target: At the end of the day, only some nurses but no doctor had left the protest.

34 **Context:** Lauren told James that there were several things they still needed for the recipe they wanted to make, like some pears and some apples. For allergy reasons, they had decided to leave out the salmon.

Target: At the supermarket, James could find only some apples but no pears, because they were in the wrong aisle.

35 **Context:** Nicole went to the pharmacy where she bought a number of things, like soap and shampoo. She forgot to buy the vitamins that her mother had asked her to get.

Target: When she looked at the receipt, it listed only shampoo but not soap to her surprise.

36 **Context:** At the organic supermarket, a number of things were on sale this week, including parsley and thyme. The store had been out of candy for a few weeks now.

Target: The store managers noticed that customers bought only some thyme but no parsley even though it was on sale.

37 **Context:** Allie's housemates had asked her to bring back a few things for the house, like cherries and strawberries. She couldn't find any toilet paper at the store.

Target: Allie's bag could fit only the strawberries but no cherries on her way home.

38 **Context:** Before the contractors came, a lot of decisions still had to be made, like the choice of the sink and the faucet. The house already came with an alarm system when Judith bought it.

Target: Judith managed to choose only a faucet but no sink, before they started building the kitchen.

39 **Context:** At the recreation center, you can rent stuff like canoes and kayaks. Camper vans are not available here.

Target: Lately, people continue to rent only the kayaks but no canoes, because of numerous safety concerns.

40 **Context:** In her handbag Alex always carries a lot of stuff, such as her lipstick and her eyeshadow. She knew her keys were in her pocket.

Target: After searching for a while, she found only her eyeshadow but not her lipstick in her bag.

41 **Context:** At the zoo, there used to be a lot of different kinds of animals, like goats and sheep. There had never been any fish there.

Target: Jason's nephew enjoyed seeing only the sheep but not the goats during his visit to the zoo.

42 **Context:** Briana ordered a few things for the table, like pizza and some pasta. She didn't want to order any water.

Target: At the end of the evening, Briana had touched only some of her pasta but no pizza, even though she was very hungry.

43	**Context:**	The local farmer's market sells various crafts such as handmade gloves and scarves. As always, the teapots were already sold out before 10am.
	Target:	When Hannah went there at noon, she liked only some scarves but no gloves at the market.
44	**Context:**	Along the driveway to the hotel, the gardener had planted some pines and some palms. He had always wanted to plant some tulips but he had never gotten around to it.
	Target:	After the storm, he trimmed only some palms but no pines on the right side of the driveway
45	**Context:**	Sabrina was on her way to the hardware store where she wanted to buy a few things like nails and screws. She didn't need any lightbulbs anymore.
	Target:	In the end, she managed to buy only some screws but no nails at the store.
46	**Context:**	After Linda got back to the hostel, she noticed that some of her belongings were gone, like her necklace and her bracelet. Luckily, she saw that her computer was still there.
	Target:	After searching for hours, she found only her bracelet but not her necklace, abandoned in the dumpster.
47	**Context:**	It's really hard to concentrate in this office because there are always lots of sounds, like that of an alarm or a phone. Fortunately, though, you can never hear the wind in this place.
	Target:	Noise-cancelling headphones cancel out some of the noise, but only of the phones, not of any alarms going off all the time.
48	**Context:**	Last summer, this neighborhood was under heavy construction because they're building some new apartments and houses. The existing bridge had to be demolished.
	Target:	By fall, they had finished building only a few houses but no apartments on time.

Appendix B

	alt1 (target)	alt2 (assoc)	alt3 (non-assoc)	alt1-alt2	alt1-alt3
1	yogurt	cheese	ashtray	0.86	-0.0
2	beer	wine	nuts	0.85	0.07
3	aunt	uncle	bank	0.82	0.01
4	piano	violin	clocks	0.80	0.10
5	horses	ponies	ducks	0.79	0.06
6	newspapers	magazines	cigarettes	0.75	0.06
7	metro station	bus station	swimming pool	0.75	0.11
8	screwdriver	hammer	soap	0.59	0.08
9	chair	sofa	flowerpot	0.73	0.05
10	radio	tv	bike	0.70	0.01
11	chemistry	biology	sports	0.68	0.04
12	lions	tigers	pelicans	0.68	0.08
13	socks	shoes	pencils	0.67	0.09
14	cotton	wool	metal	0.67	0.0
15	doors	windows	tape	0.66	0.0
16	gate	fence	camera	0.65	0.0
17	letters	emails	software updates	0.63	0.0
18	cake	pie	swords	0.62	0.0
19	sandals	boots	toothbrush	0.45	0.0
20	helicopter	airplane	bus	0.62	0.1
21	cucumbers	tomatoes	beers	0.58	0.0
22	shorts	jeans	book	0.58	0.1
23	sculptor	painter	lawyer	0.58	0.0
24	oven	stove	garage	0.57	0.1

	alt1 (target)	alt2 (assoc)	alt3 (non-assoc)	alt1-alt2	alt1-alt3
25	milk	flour	sponges	0.56	-0.01
26	jacket	sweater	puzzle	0.56	0.07
27	batteries	charger	snacks	0.54	0.00
28	basement	attic	lease	0.55	0.08
29	knives	spoons	bed	0.55	0.11
30	sister	mother	dog	0.54	0.09
31	bag	purse	camera	0.54	0.06
32	blankets	pillows	food	0.54	0.03
33	nurses	doctors	carpenters	0.52	0.02
34	apples	pears	salmon	0.51	0.05
35	shampoo	soap	vitamins	0.50	0.02
36	thyme	parsley	candy	0.49	0.12
37	strawberries	cherries	toilet paper	0.48	0.04
38	faucet	sink	alarm system	0.48	0.04
39	kayaks	canoes	camper vans	0.47	0.04
40	eyeshadow	lipstick	keys	0.46	0.03
41	sheep	goats	fish	0.46	0.04
42	pasta	pizza	water	0.44	0.02
43	scarves	gloves	teapots	0.44	0.1
44	palms	pines	tulips	0.43	0.04
45	screws	nails	lightbulb	0.43	-0.03
46	bracelet	necklace	computer	0.42	0.02
47	phone	alarm	wind	0.41	0.11
48	houses	apartments	bridge	0.40	0.06

Appendix C

Table C1

*Model output for the largest LME model comparable to that of H&F without convergence issues, on logRTs of Experiment 1 (Model code: logRT ~ Assoc * Excl + (1 + Excl | ItemNo) + (1 | Subj)).*

Fixed effects

	Estimate	Std. Error	t-value
(Intercept)	2.984751	0.011574	257.883
Association	0.027539	0.005135	5.362
Exclusion	0.035091	0.008254	4.251
Assoc: Excl	0.011478	0.010273	1.117

Random effects

		Variance	Std. Dev	Corr
Subject	Intercept	0.007719	0.08786	
Item	Intercept	0.000893	0.02988	
	Exclusion	0.002004	0.04477	-0.01

Table C2

*Model output for the largest LME model comparable to that of H&F without convergence issues, on untransformed RTs of Experiment 1 (Model code: RT ~ Assoc * Excl + (1 + Excl | ItemNo) + (1 | Subj)).*

Fixed effects

	Estimate	Std. Error	t-value	
	(Intercept)	1057.41	32.14	32.900
	Association	70.40	16.96	4.150
	Exclusion	116.21	26.07	4.458
	Assoc: Excl	38.31	33.93	1.129

Random effects

		Variance	Std. Dev	Corr
Subject	Intercept	56665	238.04	
Item	Intercept	7801	88.32	
	Exclusion	18806	137.14	0.27

Table C3

*Model output for the largest LME model without convergence issues, on logRTs of Experiment 2 (Model code: logRT ~ Assoc * Excl + (1 + Excl + Assoc | ItemNo) + (1 + Excl | Subj)).*

Fixed effects

	Estimate	Std. Error	t-value	
	(Intercept)	2.947474	0.010505	280.586
	Association	0.015639	0.005599	2.793
	Exclusion	0.016533	0.008354	1.979
	Assoc: Excl	-0.001302	0.009251	-0.141

Random effects

		Variance	Std. Dev	Corr	
Subject	Intercept	0.0057603	0.07590		
	Exclusion	0.0008917	0.02986	0.07	
Item	Intercept	0.0010393	0.03224		
	Exclusion	0.0004778	0.02186	0.92	
	Association	0.0017363	0.04167	-0.47	-0.36

Table C4

*Model output for the largest LME model without convergence issues, on untransformed RTs of Experiment 2 (Model code: RT ~ Assoc * Excl + (1 + Excl | ItemNo) + (1 | Subj)).*

Fixed effects

	Estimate	Std. Error	t-value	
	(Intercept)	950.112	27.623	34.396
	Association	44.979	13.037	3.450
	Exclusion	36.555	20.007	1.827
	Assoc: Excl	-4.902	26.081	-0.188

Random effects

		Variance	Std. Dev	Corr
Subject	Intercept	40141	200.35	
Item	Intercept	6718	81.96	
	Exclusion	11369	106.63	-0.46

Table C5

*Model output for the largest LME model without convergence issues, on logRTs of Experiment 3 (Model code: logRT ~ Assoc * Excl + (1 + Assoc + Excl | ItemNo) + (1+ Assoc | Subj)).*

Fixed effects

	Estimate	Std. Error	t-value	
	(Intercept)	2.967825	0.009941	298.553
	Association	0.011803	0.006157	1.917
	Exclusion	0.032410	0.006856	4.727
	Assoc: Excl	0.032230	0.009357	3.444

Random effects

		Variance	Std. Dev	Corr	
Subject	Intercept	4.395e-03	0.066291		
	Association	4.605e-05	0.006786	-0.38	
Item	Intercept	1.464e-03	0.038261		
	Exclusion	7.364e-04	0.027137	0.37	
	Association	1.203e-03	0.034687	0.04	0.13

Table C6

*Model output for the largest LME model without convergence issues, on untransformed RTs of Experiment 3 (Model code: RT ~ Assoc * Excl + (1 + Assoc + Excl | ItemNo) + (1 | Subj)).*

Fixed effects

	Estimate	Std. Error	t-value	
	(Intercept)	994.72	25.16	39.530
	Association	27.86	17.43	1.598
	Exclusion	89.49	20.26	4.417
	Assoc: Excl	81.04	28.04	2.890

Random effects

		Variance	Std. Dev	Corr	
Subject	Intercept	26540	162.91		
Item	Intercept	9809	99.04		
	Exclusion	5145	71.73	0.44	
	Association	10243	101.21	0.15	0.23