Journal of Memory and Language 132 (2023) 104444

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Processing of linguistic focus depends on contrastive alternatives

Morwenna Hoeks *, Maziar Toosarvandani, Amanda Rysling

University of California Santa Cruz, 1156 High Street, Santa Cruz, 95064, CA, United States

ARTICLE INFO

Dataset link: https://osf.io/k6tbw/?view_only=71d86431090046929d56f1ba94dcc38b

Keywords:
Linguistic focus
Contrastive alternatives
Newness/givenness
Maze task
Reading

ABSTRACT

Readers progressed through a sentence in the Maze task (Forster et al., 2009), deciding at each word between a sensical and a non-sensical continuation. Contexts presented before these sentences manipulated whether words were linguistically focused and whether they were given or new (Experiment 1); focused targets were read more slowly even when they were given, and new targets were read slowly in general. This both replicated earlier results in which slowdowns were found in the reading of focus (Benatar and Clifton, 2014; Birch and Rayner, 1997; Lowder and Gordon, 2015), and demonstrated that focus slowdowns are not reducible to newness. To clarify earlier results in which speed-ups were found on focused words (Birch and Rayner, 2010; Morris and Folk, 1998), contexts manipulated whether contrastive alternatives to focused words were presented with a focus particle (Experiment 2) or in a cleft construction (Experiment 3). Focused targets were read less slowly when a contrastive alternative was present in the context. This effect of contrastive alternatives cannot be reduced to simple semantic associate priming: Contexts also manipulated whether a semantically associated expression was present independently of the presence of a contrastive alternative (Experiment 4). Readers slowed down less when an alternative was present in the context, even when this alternative was not semantically associated to the target. These results indicate that the processing of focus depends on contrastive alternatives, in their interaction with newness, semantic association, and focus construction.

Introduction

Perceivers must continually use evidence from context to determine where, when, and how greatly to allocate their processing resources. Human languages have systematized devices to highlight or background information in discourse, and so language processing provides a case of the real-time deployment of expert knowledge in resource allocation, one that is common to nearly all humans. Studies that have investigated the processing of linguistic *focus* – sometimes characterized as the most "prominent" or "important" information in an utterance – by monitoring eye movements during reading have found mixed results. Some have reported a decrease in reading times on focused material (Birch & Rayner, 2010; Morris & Folk, 1998), while others observed an increase in reading times (Benatar & Clifton, 2014; Birch & Rayner, 1997; Lowder & Gordon, 2015). Explanations for the differences in these effects have acknowledged that focus is a complex construct and that its effects are likely to be modulated by many factors. For example, focus often covaries with informational newness; foci are frequently prosodically prominent; and a word or phrase can be focused using a variety of different syntactic constructions, which have their own specific properties and may require different resource allocations.

This paper reports four experiments that clarify how these different factors modulate the processing of focus. Experiment 1 showed slowdowns on foci even when those foci were not new information, demonstrating that focus slowdowns cannot be reduced to the cost of incorporating new material. Instead, there must be a cost for processing focus itself, in line with the results of Benatar and Clifton (2014), Birch and Rayner (1997), and Lowder and Gordon (2015). In order to clarify why some previous studies also found speed-ups on target foci (Birch & Rayner, 2010; Morris & Folk, 1998), Experiments 2 and 3 manipulated the presence or absence in the context of *contrastive alternatives* —- linguistic expressions which could have taken the focus's place. Using either a focus particle (*only*) or a cleft construction (*it was the…*), these experiments showed a reduction of the focus slowdown in the presence of contextually mentioned alternatives. Experiment 4 demonstrated that this facilitatory effect of contrastive alternatives cannot be reduced to an effect of simple semantic associate priming. Alternatives that were not close semantic associates of foci attenuated the focus slowdown more than close semantic associates that were not alternatives. Together, these findings indicate that a comprehensive theory of the processing of focus must incorporate the concept of contrastive alternatives, as well as an explanation for their systematic interactions with newness (Experiment 1), different focus constructions (Experiments 2 and 3), and semantic association (Experiment 4).

* Corresponding author.
E-mail address: mhoeks@ucsc.edu (M. Hoeks).

https://doi.org/10.1016/j.jml.2023.104444
Received 17 August 2022; Received in revised form 5 June 2023; Accepted 8 June 2023
Available online 24 June 2023

All four experiments were run using the Maze task (Forster, Guerrera, & Elliot, 2009), due to the global Covid-19 pandemic. In this task, readers 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 reader's representation of the sentence. This task is less prone to spillover effects than alternative tasks, such as self-paced reading (Witzel, Witzel, & Forster, 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), our originally-intended method. While a close comparison with previous studies was lost, this change in method does allow conceptual replications of findings from the earlier literature to demonstrate the robustness of those effects across tasks, and so may increase our confidence in those effects' generality beyond task-specific conditions.

The rest of this section describes the earlier findings from eye movement data, and motivates Experiment 1 by explaining the need for a study that manipulates newness independent of focus, using question/answer pairs and without changing the sentential position of focused words across conditions. In subsequent sections, a systematic review of the findings from eye movements during the naturalistic reading of linguistic focus reveals that the only studies that reported speed-ups in the processing of foci were ones in which a contrastive alternative to a focus – a concept from the formal semantics literature which will be explained – was presented before that focus was encountered. This motivates Experiments 2–4.

A proposed role for newness

The most detailed proposal about the comprehension of focus in sentence processing was put forward by Benatar and Clifton (2014). They suggested that, while many factors seem to cooccur with focus, the distinction that best captures the processing of focus is the dimension of newness versus givenness of information. Information is *given* if it is semantically entailed by what has come before in the discourse; all information that is not given is *new* (Schwarzschild, 1999). New information would be expected to require more effort to process than given information, because that new information has not already been integrated into a comprehender's understanding of the world. This might explain the longer reading times on foci than non-foci that Benatar and Clifton and others found. If new information is always in focus, then any slowdown on foci might be attributed to their status as discourse new.

Under Benatar and Clifton's newness/givenness-based proposal about focus processing effects, the studies which instead found *shorter* reading times on focused words (Birch & Rayner, 2010; Morris & Folk, 1998) may have done so for a variety of reasons. Perhaps these shorter reading times were found because these studies compared foci that occurred later in sentences with baseline words that occurred earlier, and so natural occulomotor differences in reading the positions of words on a line caused foci to be read more quickly. Perhaps these shorter reading times were found because these studies provided greater contextual support for the focus-containing target sentence, in the form of a preceding context sentence. Or perhaps these factors interacted with the syntactic devices that were used to focus words in these studies: all of them used clefts (*It was… that…*), which introduce extra inferences into sentence comprehension. For these reasons, the best study under this account would: (i) hold target sentences constant across conditions, (ii) provide stable contextual support for all target sentences, and (iii) use questions to manipulate the location of focus as the part of an answer target sentence that provided the asked-for information, so that the focus manipulation did not introduce extra inferencing during the processing of the focused word.

Benatar and Clifton's studies accordingly compared new focused with given non-focused target words. Sentence position and inferencing requirements were equalized across conditions by employing question/answer pairs and holding target sentences constant within each item. In their first experiment, target sentences were answers to preceding questions. The baseline condition consisted of completely given information, as in (1a), in which the preceding context question introduced both the existence of the characters and their possible relationship. In test conditions, longer reading times were then observed on target focused words, such as *Natalie*, when either the information about their relationships were new (here, the information that Kyle cared about Natalie) but the target word itself was not, as in (1b), or when the target word was entirely new, as in (1c). This held for first fixations (i.e., duration of the first fixation on a word during the first pass through the text), gaze durations (i.e., the sum of all first-pass fixations on a word), and total reading times.

(1) a. A: I'm confused, does Kyle care about Natalie?
B: Kyle cares about **Natalie**, but he doesn't show it. given
b. A: Natalie is confused, does Kyle care about someone?
B: Kyle cares about **Natalie**, but he doesn't show it. new, rep
c. A: Isabella is confused, does Kyle care about someone?
B: Kyle cares about **Natalie**, but he doesn't show it. new, no rep

This pattern of longer reading times in (1b) and (1c), which both contained new information pertaining to the focused word, than (1a), which contained only given information pertaining to the focused word, was interpreted as support for Benatar and Clifton's account.

Separating sentence position from focus construction

Support for the suggestion that sentence positional differences contributed to the findings of shorter reading times on foci was found by Lowder and Gordon (2015), who demonstrated that longer reading times on focused material generalize to syntactically-focused pseudo-clefts (e.g., *what the secretary typed was…*). This construction permitted control of the sentence position of target words across conditions, something which had been lacking in the earlier studies that used simple clefts (*it was the…*) and found shorter reading times on foci (Birch & Rayner, 2010; Morris & Folk, 1998). A sample item from Lowder and Gordon's study is shown in (2).

(2) a. What the secretary typed was the official **memo** about… focus
b. Yesterday the secretary typed the official **memo** about… neutral
c. It was the secretary that typed the official **memo** about… defocus

Lowder and Gordon found longer reading times on focused target words as compared to defocused target words (which were defocused, because another word in their sentences were overtly focused) in both gaze durations and regression-path durations (i.e., the sum of all fixation durations beginning with an initial fixation on a region and ending when the reader exits this region to the right). But they were further able to demonstrate that as target words became more focused – that is, comparing reading times in order across (2c), (2b), and (2a) – reading times increased. This increase with degree of focus supports the suggestion that variation in the effect sizes observed in previous work could also have been due in part to the differences in baseline conditions that were employed, i.e., neutral vs. non-focused.

A possible role for contextual support

Lowder and Gordon's results clarified that the effect of syntactically-cued focus is one of longer reading times once sentence position is controlled. But differences in sentence position between foci and non-foci could not, by themselves, account for the difference between Birch and Rayner's (1997) finding of longer reading times on foci than non-foci versus Birch and Rayner's (2010) finding of *shorter* reading

times on foci than non-foci. Birch and Rayner's (1997) first experiment employed both new foci and new baselines, as shown in the sample item in (3), where *suburb* was the target word that was compared across conditions and in neither condition was it mentioned before or entailed by anything preceding it. Birch and Rayner (1997) found longer second-pass reading times on foci and greater probability of regression from foci than non-foci, even though *suburb* was newly mentioned in both conditions.

(3) a. It was the **suburb** that received the most damage from the ice storm. focus
b. Workers in the **suburb** hurried to restore power after the ice storm. non-focus

This pattern of longer reading times contrasted with the findings of Birch and Rayner (2010), who used the same cleft focusing structure as Birch and Rayner (1997), but with a context sentence presented before each one, as in their item shown in (4).

(4) Context: The tenants at the complex were sick an tired of all the noise coming from #204.
a. It was the **landlady** who confronted the woman who lived there. focus
b. The **landlady** confronted the woman who lived there. non-focus

It thus seems that contextual support played a role in Birch and Rayner's (2010) findings of shorter reading times on foci versus non-foci, but this property may not have held of Morris and Folk's (1998) stimuli, which were not preceded by separate context sentences, but nonetheless yielded shorter reading times on foci. An example item from Morris and Folk's study is shown in (5), in which the target word *accountant* was compared across conditions. Notably, Morris and Folk's conditions compared *accountant* in focus and *de-focus*, that is, with a different word from accountant in the syntactically focused position in (5b), unlike the simple declarative sentence with no special focus on the target word that served as a baseline for Birch and Rayner.

(5) a. While the waiter watched, it was the **accountant** who balanced the ledger a second time. focus
b. It was the waiter who watched while the **accountant** balanced the ledger a second time. defocus

It would be surprising if Morris and Folk's items provided enough contextual support to facilitate the processing of their focused words, but Lowder and Gordon's items did not. The pseudocleft structure employed by Lowder and Gordon likely provides more contextual support before its focus than a simple cleft; in the case of the sentence beginning *What the secretary typed was the official **memo** about…*, for example, a secretary typing was already introduced before the focused word *memo*, making that focused word much more expected in context than if it were early in the sentence and out-of-the-blue. Longer reading times on focused *memo* in Lowder and Gordon's study would be mysterious under an account which appealed to contextual support to explain the shorter reading times on foci that previous studies had found. If newness is the primary driver of focus processing costs, but it can be overcome by contextual support for focused words, then Lowder and Gordon's effects might be expected to be more like Birch and Rayner's (2010) and Morris and Folk's (1998) pattern of shorter reading times on clefted foci with contexts, instead of Birch and Rayner's (1997) pattern of longer reading times on clefted foci without contexts. We take up the question of why Morris and Folk (1998) and Birch and Rayner (1997) found speed-ups again in the General Discussion.

While Lowder and Gordon agreed with Benatar and Clifton that focus is a complex conjunction of many different properties, they suggested that focused material generally is more deeply encoded than non-focused material, with more effort expended to integrate it during language processing due to its greater importance in its sentences. This greater effort expended on focus would be expected to require more time, and would account for why focused material is advantaged in other tasks, for example, it is reliably better remembered than non-focused material (Birch & Garnsey, 1995; Gernsbacher & Jescheniak, 1995; McKoon, Ratcliff, Ward, & Sproat, 1993; Singer, 1976). In principle, Lowder and Gordon's suggestion is not logically incompatible with the proposal that the processing profile of focus is generally due to a greater cost of processing new versus given material. But their pattern of longer reading times on foci with greater contextual support when newness was a property of both foci and baselines is not straightforwardly explained under Benatar and Clifton's proposal.

Manipulating focus independent of newness

There is a test case which would separate Lowder and Gordon's account of the focus cost as due to general deeper processing from Benatar and Clifton's proposal that the focus cost is largely due to the cost of new information. If focus effects are generally reducible to the newness/givenness distinction, then longer reading times for focused material than non-focused material when both the focus and its comparator non-focus were given would be unexpected (in the absence of confounding factors like different sentence positions between focus and comparator or the introduction of extra inferences by a specific focusing structure). In contrast, Lowder and Gordon's proposal that focused material is simply more deeply encoded and effortfully processed than non-focused material would account for longer reading times on focused material even when both are given and all else is held equal. This motivated the present Experiment 1.

Experiment 1

In Experiment 1, two factors – focus (focus vs. not) and newness (given vs. new) – were fully crossed, with focus controlled by preceding questions, which do not introduce the kind of extra inferences that clefts as focusing devices do. This provided a test of whether focus has an effect on reading in the absence of newness, an outcome that would not be expected if the cost of focus processing is generally due to the cost of integrating new material. This was the first study to compare focus and baseline conditions that were both given information.

Method

Participants. 51 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 80% on the comprehension questions or that did not complete more than 70% of the Maze sentences were excluded from analysis. Data from 48 participants were included in the analysis; 3 participants were excluded because they failed to complete more than 70% of the Maze sentences.

Materials. In all the experiments presented here, every item took the form of a short dialogue between two speakers, Speaker A and Speaker B. Speaker A first introduced a short premise, followed by a question. Speaker B's utterance formed a response to the question from Speaker A. Speaker A's utterance was considered the context sentence and was presented all at once on a single screen; it then disappeared when participants advanced to the next part of the trial. Speaker B's utterance was considered the target sentence and was presented using the Maze task. Within one item, the same sentence was the target for every condition, in order to ensure that differences across conditions would only be due to preceding context sentences. Within each target sentence, measurements on a single target word were expected to particularly reflect the effects of preceding contexts.

Preceding context questions determined whether a target word was new or given by either mentioning that target word in the question or not. Orthogonal to this manipulation of newness, preceding questions

determined whether target words received narrow focus (NF) or broad focus (BF) by asking for differently specific information. Narrow focus questions were ones to which a following target word on its own would provide a complete answer; broad focus questions were ones to which a target word alone would not seem a complete answer. We employ the distinction between narrow and broad focus, rather than the distinction between focused and neutral words, because focus in this study is manipulated by which parts of a target sentence provide the answer to a question. Our narrow versus broad focus conditions are analogous to Lowder and Gordon's (2015) focused versus neutral conditions. Here, the narrow versus broad distinction captures Lowder and Gordon's point that differences in degree of focus matter; the degrees of focus in the answers to questions can be understood as the proportion of the focus of a sentence that single word encompasses. In narrow focus conditions, the single focused word would be a complete answer to a preceding question and is the entirety of the focus. In broad focus conditions, the words that must be included in the focus are more numerous; in the case of these stimuli, they are the entirety of a phrase. This is illustrated in the example experimental item shown in (6) below. In (6), the target word is *lawyer*.

(6) *Speaker A:* This company often makes bad decisions, but...

a.	Did they hire a **lawyer** last fall, or an accountant?	NF, given
b.	Did they hire a **lawyer** last fall?	BF, given
c.	Did they hire an accountant last fall?	NF, new
d.	What did they announce last time?	BF, new

Speaker B: I think they announced they hired a **lawyer** last fall, but I'm not sure.

In response to narrow focus questions, as in (6a) and (6c), *lawyer* would be a complete answer. Across all items, for creating narrow focus and givenness on the target word, alternative questions were used (i.e., questions in which two alternatives are given in the form of a disjunction). Since the answer to such a question is expected to be one of the mentioned alternatives, the answer was either *accountant* or *lawyer* in the case of (6a). Therefore, the questions in the NF given conditions put only the target word *lawyer* in focus in the target sentence.

The NF new items always employed polar questions (i.e., questions whose expected answers are either confirmative or negative) that mentioned a different alternative from the one mentioned in the target sentence. The target sentence would therefore be unambiguously interpreted with corrective narrow focus on the target word.

After broad focus questions, *lawyer* in the target sentence would be part of a larger focused phrase, because a whole phrase from the target sentence would be required in order to provide a complete answer to the preceding questions. The BF given condition always used polar questions as well, but in these questions the alternative was the same as in the target sentence. This puts the target sentence as a whole in broad focus, as is the case in (6b). This had the result that both the wh-question in (6d) and the polar question in (6b) put at least the whole phrase *they hired a lawyer last fall* in focus, because this is the phrase that forms a congruent answer to each of these questions. Although it would be less natural, it is still technically possible that (6b) could be interpreted with narrow focus on the target word; there is nothing that *prevents* a reader from interpreting this as a narrowly focused phrase. However, evidence from interpretation and completion studies supports the assumption that comprehenders default to the broadest possible focus that is supported by the context (Harris & Carlson, 2014, 2017); this accords with theoretical semantic assumptions as well. But, even assuming that narrow focus is more costly than broad focus, and that a narrow focus parse was maintained in at least some of our items' broad focus given conditions, the estimated effect of focus from our study would be, if anything, slightly diminished and so decrease the likelihood that we would find an effect of focus in given conditions, because we expect narrow focus to be more costly to process than broad focus.

It is worth addressing a concern raised by an anonymous reviewer of an earlier version of this work: in order to manipulate newness/givenness in these stimuli, more material immediately before the target word was repeated in the given conditions (6a) and (6b) than in the new conditions (6c) and (6d). Any potential effect of the newness/givenness difference in these stimuli was thus perfectly confounded with repetition differences, with the result that expected shorter response times on given conditions relative to new conditions could be due to the simple repetition effects; this was the problem that led Benatar and Clifton to adopt hypernyms in their second experiment. This inherent covariation of the newness/givenness distinction with repetition in our items was not a confound for the particular effect that this study seeks to demonstrate: the presence of a focus effect even when *both* a focus and its non-focus comparator are given. The present study employed simple repetition to manipulate givenness because repeating material is the most straightforward and the most unambiguous way to make material discourse given. In our study, the goal was to show that *given* foci indeed slowdown relative to *given* non-foci, while we were less interested in the givenness effect itself. It was therefore crucial that given material was indeed undeniably given, which was most easily obtainable if both conditions involved repetition. Since it is pragmatically marked to use a different term as the one that is already established in the context to refer to an individual, the use of a synonym could always be interpreted as introducing new, contrastive information.

In each item, the target word was always followed by an adverbial phrase (*last fall*) which served as a spillover region. This spillover region was also always followed by a second clause (*but I'm not sure*), to ensure naturalness of the target sentence in the BF given condition.

In total, 48 items were constructed, each with the four conditions illustrated in (6). All items for Experiment 1 can be found in "Appendix B"; these were first tested in an acceptability judgment study, the results of which can be found in "Appendix A". Another 96 filler items which also consisted of multi-line discourses were interspersed with test stimuli. Using a Latin Square design, all 48 items were counterbalanced over 4 lists, such that each participant saw one condition from every item.

Procedure. The Maze task is similar to the more commonly used self-paced reading task in that response times are measured using button presses. But instead of simply pressing a button to advance to a following word each time a participant has read the current word, participants in the Maze task see each word in the target sentence presented alongside a distractor word (or *foil*). Participants must at every new word choose the correct continuation between the intended item and its foil, which would not make a sensical continuation.

Foils were automatically generated using the AutoMaze software developed by Boyce, Futrell, and Levy (2020). This algorithm selects distractor words that are 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 the same length in the context of the preceding sentence. The words with a predicted probability below a certain threshold (or, *above* a certain *suprisal* threshold) are then selected by the AutoMaze algorithm as the distractor. Word frequencies that form the input to these models are obtained from the Google Books Ngrams corpus (Michel, Shen, Aiden, Veres, Gray, Pickett, Hoiberg, Clancy, Norvig, Orwant, et al., 2011).

An example of the AutoMaze output for one target sentence is given in (7) below. On the second line, the distractor word is presented below its corresponding word of the target sentence.

(7) I think Sarah said she wanted cake for dessert, but I am not sure.
x-x-x goods Runes blue sum bottom knee sum classed, tax Sin far sat send.

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
NF (alt), given	686.71 (9.33)	643.62 (7.87)	774.52 (11.62)	741.49 (10.10)	711.16 (11.63)
BF (no alt), given	720.57 (11.53)	664.96 (8.81)	724.32 (10.66)	736.01 (11.00)	701.30 (10.12)
NF (alt), new	732.08 (12.91)	663.40 (8.42)	952.54 (15.64)	790.33 (12.16)	718.12 (11.38)
BF (no alt), new	893.66 (17.01)	745.91 (11.08)	937.43 (15.83)	867.97 (15.97)	770.00 (13.22)

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.

On every trial, participants first read a context sentence on one screen. On a subsequent screen, participants were presented with the start of the target sentence in the format of the Maze task. That is, only the utterance of Speaker B was presented incrementally; the utterance of Speaker A was presented all at once for normal reading. The context sentence disappeared from the screen when participants moved on to the target sentence. All experimental trials were followed by a comprehension question, which probed whether participants had read the context preceding the target sentence. This was because there was more cause for concern that participants might not read the contexts than that they might not read the target sentences. Participants had to read the beginning and all subsequent material of a target sentence in order to even make a decision about which word could form a potential continuation as the sentence went on. If they 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 target sentence were not recorded. But participants could successfully go through a whole target sentence in the Maze without having read its preceding question, and so comprehension questions were included after each trial that encouraged careful reading of the preceding context. For instance, the example item in (6) was followed by the comprehension question in (8).

(8) Is the company known for its strategic actions?

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.

Analysis. Data were analyzed using R, version 3.6.3 (R Core Team, 2021). Linear mixed effects models were fit using the *lme4* package (version 1.1–28; Bates, Mächler, Bolker, and Walker 2014). Models included fixed effects of focus and newness (coded as simple effects), with broad focus and given conditions treated as reference levels, and random slopes and intercepts for both subjects and items (Baayen, Davidson, & Bates, 2008). Separate models were fit with log-transformed response times and untransformed response times as dependent measures. Due to the fact that multiple comparisons were performed on the same data, an absolute value of t of 2.29 was considered to be the threshold for significance. This value was chosen instead of the traditional value of $t = 2$, because it corresponded to the significance threshold if instead of setting $\alpha = 0.05$, $\alpha = 0.025$ with the same degrees of freedom.

Results

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

The mean comprehension question accuracy was 88%, and the mean completion rate of the maze target sentences of Experiment 1 was 87%.

Table 2
Parameter values for fixed effects in mixed linear regression model of LogRTs in Experiment 1.

	Estimate	Std. Error	t value
(Intercept)	6.54506	0.02175	300.890
Focus	0.06801	0.01822	3.732
Newness	0.23841	0.03005	7.933
Focus:Newness	−0.04306	0.02986	−1.442

Table 3
Parameter values for fixed effects in mixed linear regression model of raw RTs in Experiment 1.

	Estimate	Std. Error	t value
(Intercept)	725.07	18.13	39.994
Focus	50.36	17.05	2.954
Newness	210.88	30.63	6.886
Focus:Newness	−32.21	30.49	−1.057

Tables 2 and 3 present the fixed effects results for the models of Experiment 1 log-transformed response times and untransformed response times on target words, respectively (See "Appendix D" for results on spillover regions). Both models found two significant effects. First, significantly positive estimates of focus indicate that given targets in narrow focus were responded to more slowly than given targets that were part of a broad focused phrase. Second, significantly positive estimates of newness indicate that responses were slower in the BF new conditions compared to the BF given conditions. The interaction estimates did not reach significance in either model. On the word following the target word, models revealed a main effect of newness ($t = 6.24$ for log RTs) and a significant interaction between newness and focus ($t = -3.58$). Pairwise comparisons on log RTs revealed that the focus effect was significant in the new conditions ($t = 3.65$), but not in the given conditions ($t = 0.53$).

Discussion

In addition to being affected by newness versus givenness, response times on the target word in Experiment 1 also depended on whether the target word was in narrow or broad focus. Most importantly, comparing only target words that were given, response times were longer when targets were narrowly focused than when they were only a part of a broadly focused phrase. This is unexpected if focus effects were primarily driven by the newness/givenness difference in the absence of either inferences demanded by syntactic constructions or the contextual support provided by preceding material. The stimulus sentences of the present study held the target sentence identical across all conditions, and so no condition's target sentence introduced meaning inferences that the others' did not. All of the conditions were preceded by a question which contained much of the same linguistic material in the target sentence, thereby providing contextual support for the target word in narrow focus. Indeed, in the given narrow focus conditions, the preceding questions provided more contextual support than in the given broad focus conditions, because the narrow focusing questions also provided an alternative to the target word. Nonetheless, participants took longer to respond to the given narrow focused target word than the given broad focused target word, exactly the opposite of the pattern predicted by an account appealing to a combination of newness and contextual support to explain previous findings. Instead, the difference

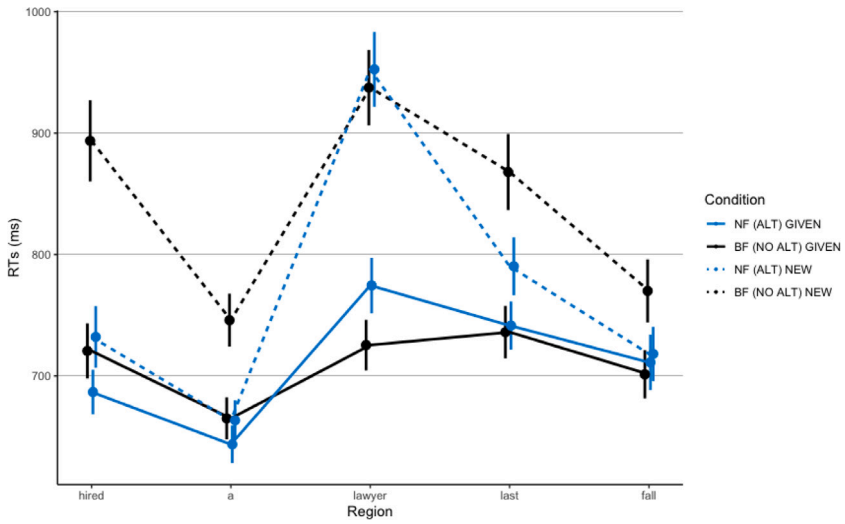

Fig. 1. Experiment 1: mean RT in each region in each condition. Error bars represent the 95% confidence interval.

between the broad versus narrow given conditions found here conceptually replicates Lowder and Gordon's observation that words take longer to read as they become more focused once sentence position is held constant, but at present with a different syntactic structure and task.

Experiment 1 was thus the first study to manipulate newness versus givenness and narrow versus broad focus independently of each other and to find an effect of focus within entirely given material. In the process, the narrow versus broad focus distinction was achieved by the inclusion or exclusion of contrastive alternatives in the context, e.g., *accountant*, to the ultimate target word, e.g., *lawyer*. The role of contrastive alternatives in focus processing has been little explored in eye movement studies of the reading of focused words themselves (studies that have examined the role of focus in the subsequent processing of alternatives are discussed after the presentation of novel results). Experiment 1 was also the first study to present contrastive alternatives to target focused words in preceding context sentences and find longer reading times on foci than non-foci. We argue that this is likely due to the use of question/answer pairs, not clefts. The next section presents a review of the previous reading studies on focus processing, in which the only studies that found faster reading times on foci than non-foci were ones that presented alternatives to words that were focused by cleft constructions. This suggests that the role of alternatives in the processing of focus is one of those as-yet unaccounted for factors referred to by Lowder and Gordon and Benatar and Clifton that will be crucial for the construction of a complete understanding of focus processing.

Contrastive alternatives in the processing of focus

The *contrastive alternatives* to a focus are all those linguistic expressions which could have coherently taken the place of that focus in a sentence (Rooth, 1985). For example, in a sentence like (9), the alternatives to *cake* include, among other things, *steak* and *cookies*, since both expressions could be substituted for it to form a grammatical sentence. But, because steak is rarely eaten for dessert, after having read the entire sentence comprehenders would generally not find it to be a *relevant* alternative. In other words, it is not a relevant alternative to *cake* for this sentence in most contexts.

(9) It was **cake** that Sarah wanted for dessert.

At the same time, some expressions may seem more natural or expected, and thus, more likely to be relevant alternatives than others. For example, if one were asked out of the blue to name "things like cake", one would more likely name *cookies* and other desserts or baked goods before one named *steak*. These more expected alternatives (*cookies, brownies, pie*) contrast with a target expression (*cake*) on its more salient dimension(s), and are likely to be substitutable for that target expression in a larger set of contexts than a less-closely-related word that can still be an alternative in many contexts (*steak*).

Similarly, although an expression like *pastry chef* may be closely associated to the word *cake* and is possibly also relevant to the broader scenario described by (9), this expression cannot be an alternative to the focus *cake*, because we cannot substitute *pastry chef* for *cake* in this sentence. Further, in an out-of-the-blue context, *pastry chef* is not as expected an alternative for *cake* as many other expressions, because we would often not expect them to be contrasted with each other as options that would fulfill the same role in one situation. But an alternative expression such as *cookie* would likely be both highly relevant and replaceable with *cake*, and therefore would count as a relevant contrastive alternative, unless the comprehender of (9) also had access to some situational knowledge that ruled cookies out for other reasons.

Contrastive alternatives are well-known to play a key role in the inferences that clefts and focus particles like *only* give rise to. In (9), the comprehender is likely to understand that Sarah did *not* want cookies for dessert. Whether this particular inference is derived depends on whether *cookies* is considered a relevant contrastive alternative in this particular context, which in turn depends on world knowledge, the content of the sentence itself, and the information provided by the preceding discourse context.

Notably, question/answer pairs also give rise to a similar inference. In (10), the answer implies that Sarah did not want anything else besides cake for dessert, including possibly cookies.

(10) a. What did Sarah want for dessert?
b. Sarah wanted cake for dessert.

Perhaps importantly for our understanding of the processing of focus, this inference in simple question/answer pairs is defeasible, with comprehenders less likely to draw it in every context (Groenendijk & Stokhof, 1984; Hintikka, 1976; van Rooy, 2003). This is unlike the alternative-dependent inferences involved in particles and clefts that are non-defeasible. Thus, in order to compute the final interpretation of the sentences containing such *bound foci* – foci which are signaled by a syntactic device in the same sentence – comprehenders are required to reason about contrastive alternatives.

Non-reading psycholinguistic studies have employed priming, lexical decision, and memory tasks to show that, when a focused expression

is encountered, linguistic expressions that contrast with that focus (such as *cookies* for the sentence in (9)) become more strongly activated compared to expressions that are semantically associated with the focus but do not contrast with it (such as *pastry chef* when *cake* is focused; e.g., Braun and Tagliapietra, 2010; Fraundorf, Benjamin, and Watson 2013, Fraundorf, Watson, and Benjamin 2010, Gotzner, Wartenburger, and Spalek 2016; i.a.). These studies strongly suggest that integrating the meaning of foci requires comprehenders to not only represent what that sentence described, but also to calculate alternatives to what was asserted. These non-reading experimental results accord well with the theoretical semantic literature in linguistics. In the standard theory, foci are the word(s) in a sentence that must be contrasted with alternative expressions in order to understand what the sentence means (Rooth, 1985, 1992).

A survey of earlier reading studies in terms of the alternatives to foci that were present or absent in stimuli reveals that this dimension of focus perfectly demarcates studies by their reading time patterns. All of the studies which report decreased reading times on focused material are ones in which potential alternatives to focused words were presented before those foci. This was true of Birch and Rayner (2010), whose stimuli are repeated in (11).

(11) Context: The tenants at the complex were sick an tired of all the noise coming from #204.

a. It was the **landlady** who confronted the woman who lived there. focus

b. The **landlady** confronted the woman who lived there. non-focus

What (11a) conveys is that it was the landlady and not one of the tenants who confronted the woman who lived in the noise-making apartment. The word *tenants* can thus serve as an alternative expression to the focused *landlady* and was presented in the preceding context sentence; this was systematic throughout Birch and Rayner's (2010) items. This property also held of Morris and Folk's (1998) stimuli.

In contrast, the studies which report reading time slowdowns on focus in eye movements did not present alternatives to the focused material in preceding context sentences. This was true of Birch and Rayner's (1997) first experiment, which did not employ preceding context sentences altogether. This property was also true of Lowder and Gordon's (2015) stimuli, even though these provided general contextual support for the target word. The absence of alternatives to target foci in the context also held for previous studies employing question/answer pairs as focusing devices (Birch and Rayner, 1997, experiments 2 and 3).

Even in Benatar and Clifton's (2014) studies, where stimuli were presented with preceding discourse context, the target words were difficult or impossible to understand as having contrastive alternatives. An example item from their first study is repeated in (12).

(12) a. A: I'm confused, does Kyle care about Natalie?
B: Kyle cares about **Natalie**, but he doesn't show it. given

b. A: Natalie is confused, does Kyle care about someone?
B: Kyle cares about **Natalie**, but he doesn't show it. new, rep

c. A: Isabella is confused, does Kyle care about someone?
B: Kyle cares about **Natalie**, but he doesn't show it. new, no rep

For a comprehender of (12) to establish that *Isabella* in (12c) was a contrastive alternative to a person named *Natalie*, either more contextual support or additional world knowledge would have been necessary, because proper names do not convey descriptive context. This is an easy intuition to grasp when it is compared with the results from semantic priming for contrastive alternatives to foci (e.g., Braun and Tagliapietra 2010, Fraundorf et al. 2013, 2010; i.a.): the name *Natalie* would not be expected to generally prime the name *Isabella* (it would only do so for people who know a person named Natalie and a person named Isabella from the same context). In their second study, Benatar and Clifton used hypernyms for target words, which contrasted with hyponyms in the preceding context. These expressions would not qualify alternatives, because alternatives must be exclusive of each other (consider the infelicity of *#I own a **poodle**, but not a **dog***, unless the speaker were somehow claiming that a poodle is not a dog). Thus, Benatar and Clifton's studies, too, fit the pattern across all investigations of eye movements in the reading of focus: faster reading times on foci which were presented after contrastive alternatives, but slower reading times on foci in the absence of (unambiguous) contrastive alternatives.

The results of the entire literature are summarized in Table 4. It is only the difference between the presence versus absence of alternatives to foci that demarcates the faster from the slower reading of foci across this earlier literature. We take this as suggestive evidence that the alternatives-based understanding of focus that is employed in the theoretical linguistic literature may be useful for building theories of language processing as well.

Upon inspection of Table 4, several other patterns are apparent. All reading studies of focus before the present Experiment 1 had tested only new foci. All reading studies that had employed question/answer pairs as focusing mechanisms did not present alternatives to foci. None of these studies investigated the reading of foci that were marked by focus-sensitive particles, such as *only*, which obligatorily focus an element in their scope. And, while the closest comparisons in the literature so far were between studies that employed clefts or (pseudo)clefts with or without the presence of possible alternatives to the foci, there still is not a minimal comparison of clefts that are always preceded by contexts, which themselves differ only in whether alternatives are present versus absent. This is because Birch and Rayner's two cleft studies differed in the presence versus absence of entire contexts and Lowder and Gordon's pseudoclefts differed from Morris and Folk's clefts in sentential positions as well as not having contexts.

As other authors have noted, each focusing device carries with it certain unique demands, and Table 4 shows that the speed-up in reading times on focused material after the presentation of contrastive alternatives has only been demonstrated with clefted foci. It is possible that this pattern would hold only of clefts, or would only hold of structures that shared some property with clefts; this would account for why it was not observed in the present Experiment 1.

Support for this possibility, that the effect of contrastive alternatives interacts with the differences among focus constructions, comes from the theoretical semantics literature (Jackendoff, 1972; Rooth, 1985), which distinguishes between *free foci* and *bound foci* (also called "associated" foci, a term we avoid to forestall confusion with semantic association). Bound foci are those that are signaled by a particular syntactic construction, such as (pseudo)clefts or focus particles (e.g., *only, even*), while free foci are those that are merely mandated by context, such as by a preceding question, not by any expression in their immediate sentence. Clefts and focus particles generally require comprehenders to calculate construction-specific inferences or presuppositions, some of which are only optional for free foci. Moreover, bound foci are also more predictable in incremental processing, because many of the devices that bind them must precede them in the linguistic signal. The locations of the foci of clefts are predictable with a high degree of certainty before those foci have themselves been fixated, because it is always a word after *was* (or *is*) that is focused. Similarly, a focus bound by *only* always follows this particle, although this can be at a small distance. In contrast, a comprehender can only predict the positions of free foci that are relatively later in the sentences in which they occur, because it is only after some linguistic material has elapsed that this material could have sufficiently narrowed the possible continuations that would be congruent with the preceding focusing structure.

This point is illustrated by the example sentences in (13), in which the question in (13a) provides some restriction as to what will be focused (i.e., the part of the response that answers the question), but a comprehender does not know, when they first turn to reading that

Table 4
Overview of previous investigations of focus in reading. *Inhibition* versus *Facilitation* columns indicate whether a study reported slower or faster reading times on focused material, respectively. *Construction* indicates which focusing structure was employed. *Alternatives* indicates whether alternative expressions to target words were presented. *Newness* indicates the new versus given status of the focus condition versus the condition(s) to which it was compared.

	Inhibition	Facilitation	Construction	Alternatives	Newness	
					Focus	Baseline
Birch and Rayner (2010)	×	✓	clefts	present	new	new
Morris and Folk (1998)	×	✓	clefts	present	new	new
Ward and Sturt (2007)	×	×	wh-phrase	absent	new	new
Birch and Rayner (1997), Exp. 1	✓	×	clefts	absent	new	new
Lowder and Gordon (2015)	✓	×	(pseudo)clefts	absent	new	new
Birch and Rayner (1997), Exp. 2	✓	×	q/a pairs	absent	new	given
Benatar and Clifton (2014), Exp. 1, 2	✓	×	q/a pairs	absent	new	given
Benatar and Clifton (2014), Exp. 3	✓	×	indefinites	absent	new	given

answer, whether the response will be more like (13b), where the answer and so the focus is later in the sentence and by that point more predictable, or more like (13c), where the answer and focus is unsignalled in the sentence before it occurs.

(13) a. What did Sarah want for dessert?
b. Sarah wanted **cake** for dessert. late focus
c. **Cake**, I think it was. early focus

It is thus the case that different focusing constructions could have different processing profiles in the presence or absence of earlier contrastive alternatives because of differences in how readily a reader can anticipate where a focus will occur. In line with Lowder and Gordon's proposal of deeper encoding and greater integration of focused material, it may be that readers are better able to allocate their resources toward an upcoming focus ahead of time when reading a bound focus than a free focus. Combining such an account with an alternatives-based understanding of focus, it may be that contextually mentioned contrastive alternatives can more greatly facilitate the processing of bound than free foci. Since incorporating the meaning of foci involves reasoning about alternatives, the presence of contrastive alternatives in the context, as well as the presence of a focusing device that cues the location of the upcoming focus, may allow comprehenders to initiate some of this reasoning process earlier on, i.e., upon encountering the focusing device, thereby facilitating comprehension of the subsequent focus itself.

This distinction between free and bound focus is also reflected in results from priming studies, which suggest that contrastive alternatives only become differentially pre-activated due to the presence of focus particles like *only* or *also*, but not due to the presence of question/answer focus (Braun, Asano, & Dehé, 2018). Converging evidence for the idea that comprehenders already start reasoning about contrastive alternatives when they first encounter a focus-sensitive particle also comes from visual world studies in which comprehenders' looks start converging to a depiction of a focused target upon encountering a focus particle before that focus (Kim, Gunlogson, Tanenhaus, & Runner, 2015).

It is in bound focus constructions that readers can take advantage of the pre-activation of contrastive alternatives to the foci they will need to comprehend. In the case of free foci, however, comprehenders must identify that material has been focused by preceding context anew in each situation; there are not such clear and consistent cues as dedicated lexical items like *only* or *it was a...* to signal focus.

In light of this similarity between focus particles and clefts, both these constructions were tested, respectively, in the presence or absence of alternatives in Experiment 2 and Experiment 3, which were designed to reveal the interaction of contrastive alternatives with bound focus constructions. The results of these studies suggest that for both clefts and particles like *only*, explicit mention of a contrastive alternative facilitates processing of the subsequent focus: Maze response times on foci were found to be shorter in the presence of contrastive alternatives than in their absence. This finding therefore also suggests a crucial difference between free and bound focus along the lines suggested above, where explicitly mentioned contrastive alternatives facilitate the processing of bound foci, but not free foci.

Experiment 2

The only difference between Experiment 1 and Experiment 2 is that in the latter, the focus particle *only* was added to the target sentences. This particle unambiguously bound the target word and served two purposes. First, it provided a cross-experiment comparison between free and bound focus. Because the stimuli of Experiment 1 and Experiment 2 employed different focusing constructions, they differed in how clearly the position of their foci were signaled by these constructions. Second, the focus particle put the target word into narrow focus in all of the conditions of Experiment 2. It therefore allowed assessment of the effect of providing alternatives to target words in preceding contexts when focus status is held constant by a bound focus construction.

Method

Participants. 58 native speakers of English were recruited via Prolific. All participants were compensated at a $12 hourly rate. Completion of the experiment usually took 50 min including the practice phase. Data from 48 participants were included in the analysis; 10 participants were excluded because they failed to complete more than 70% of the Maze sentences.

Materials. The stimuli of Experiment 2 were the same as in Experiment 1, except for the presence of the particle *only* immediately before the target word in each target sentence. An example of an item is shown in (14) below.

(14) *Speaker A:* This company often makes bad decisions, but...
a. Did they hire a **lawyer** last fall, or an accountant? (NF) alt, given
b. Did they hire a **lawyer** last fall? (NF) no alt, given
c. Did they hire an accountant last fall? (NF) alt, new
d. What did they announce last time? (NF) no alt, new

Speaker B: I think they announced they hired **only** a **lawyer** last fall, but I'm not sure.

As in Experiment 1, a newness manipulation determined whether the target word was new or given by the time participants read it. Unlike Experiment 1, the target word of every condition in Experiment 2 was narrow focused. In (14), *lawyer* is bound by *only*, which puts it in narrow focus. This position of *only* immediately before the target word prevented it from being interpreted as binding any other word or phrase in the sentence. In order to facilitate comparison of these conditions with the ones from Experiment 1, the label "NF" is shown next to all the conditions of Experiment 2 in (14).

In order to interpret the meaning of *only* in a target sentence, comprehenders require contextually relevant alternatives to the target word, because the meaning of *only* is that nothing other than its bound focus is true in that context. As noted above, the preceding context questions used in both Experiment 1 and Experiment 2 manipulated the presence versus absence of an alternative to the target word. In Experiment 2, where all target words are focused by binding with

Table 5
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
(NF) alt, given	776.95 (17.49)	676.20 (10.40)	799.06 (11.58)	758.62 (10.83)	735.89 (17.40)
(NF) no alt, given	793.35 (16.33)	683.25 (9.66)	791.93 (14.13)	782.73 (13.58)	747.24 (13.06)
(NF) alt, new	811.27 (16.95)	711.83 (12.45)	901.01 (13.32)	793.31 (12.19)	713.45 (11.16)
(NF) no alt, new	909.15 (17.25)	772.62 (16.40)	968.70 (17.32)	876.82 (15.27)	835.61 (18.91)

Table 6
Parameter values for fixed effects in mixed linear regression model of LogRTs in Experiment 2.

	Estimate	Std. Error	t value
(Intercept)	6.64594	0.02132	311.722
Alternative	−0.02849	0.02232	−1.276
Newness	0.10723	0.02257	4.752
Alternative:Newness	0.09016	0.02820	3.197

Table 7
Parameter values for fixed effects in mixed linear regression model of untransformed RTs in Experiment 2.

	Estimate	Std. Error	t value
(Intercept)	803.153	19.503	41.180
Alternative	−3.699	24.699	−0.150
Newness	94.717	22.865	4.142
Alternative:Newness	86.692	34.821	2.490

only, this manipulation thus assessed the effect of explicitly provided alternatives in processing a focus that strongly supported an inference about alternatives, unlike the weaker inferences of the question/answer pairs in Experiment 1. Since an alternative question like that in (14a) presupposes that the mentioned alternatives are the only possible hires, *accountant* formed a salient alternative to the target *lawyer*. Similarly, the polar question in (14c) explicitly mentioned an alternative to the target word, *accountant*, but did not mention the target word itself. Thus, (14a) and (14c) are labeled "alt", while (14b) and (14d) are labeled "no alt". All materials of Experiment 2 were assessed in an acceptability judgment study, the results of which can be found in "Appendix A".

Procedure. As in Experiment 1, the target sentences in Experiment 2 were implemented in the Maze task, in which response times were measured as the time it took for participants to choose between the actual continuation word and a foil.

Since the materials for Experiment 1 and Experiment 2 were the same except for the word *only* in the target sentence, the foils generated for Experiment 1 were used to create the foils for Experiment 2. To do so, the target sentences of Experiment 2 were used as the input to the AutoMaze algorithm to generate the appropriate foils for the word *only* in each item. Then, these foils for *only* were inserted into the foils that were already generated for Experiment 1. In this way, the differences between Experiment 1 and Experiment 2 were kept as minimal as possible, to provide maximal comparability between the two experiments. Fillers, practice items, comprehension questions, and presentation lists were the same as in Experiment 1.

Analysis. The analysis was the same as that of Experiment 1, except that fixed effects were newness and the presence versus absence of alternatives, again dummy-coded. The presence of an alternative to the a focused target word was treated as the reference level of this factor.

Results

Mean response times for the target word and its surrounding regions in all conditions are given in Table 5. They are plotted with 95% confidence intervals in Fig. 2.

The mean comprehension question accuracy was 88%, and the mean completion rate of the maze target sentences of Experiment 2 was 87%.

Fixed effects estimates for the model fitted to log-transformed responses are reported in Table 6; those for the model fitted to untransformed responses are reported in Table 7 (See "Appendix D" for results on spillover regions).

As in Experiment 1, the significant positive estimates for newness indicate that new targets were responded to more slowly than targets that were mentioned in the preceding question. A significant interaction between newness and the presence of alternatives indicated that the difference in response times between the two new conditions was larger than the difference between the two given conditions. Pairwise comparisons on log-transformed response times confirm that the effect of presence of alternatives does not reach significance in the given conditions (t = .31), while the difference between the new conditions with and without previously mentioned alternatives was significant (t = −3.60). Responses were thus slower in conditions without an alternative compared to those with an alternative only when the target was also new. On the word following the target word, models on log-transformed response times revealed a significant main effect of newness (t = 2.77), but did effect did not reach significant on raw RTs (t = 2.26). Models on log-transformed RTs revealed a significant interaction between newness and the presence of alternatives (t = 2.54), but again this effect did not reach significance in models run on untransformed RTs (t = 2.18). Pairwise comparisons on log RTs revealed that the effect of alternatives was only significant in the new conditions (t = 4.99).

Discussion

Experiment 2 replicated the effect of newness found in Experiment 1: responses were slower when a target word was new compared to when it was given. This finding also conceptually replicated the results of Benatar and Clifton's eye movement studies in which new information focus was found to cause significant slowdowns.

Experiment 2 also found limited evidence that preceding contextual information modulates the reading of foci. Narrow foci that were preceded by a contextually-mentioned alternative expression to the target word were read faster than narrow foci that were not preceded by an alternative, but only when the target itself was new. This may be because previously encountering the exact expression in focus or previously encountering an alternative to the focus can both aid in comprehending the focus itself. If these are the reasons why there was no difference between the two given conditions in Experiment 2, then this would point to an important difference between free and bound focus, because it would suggest that the presence of alternatives in the preceding context only aided comprehension of bound foci (as in Experiment 2), not free foci (Experiment 1).

However, it was also true that, across the two new conditions, the alternative-mentioned condition contained more repetition of words before the target word than the no-alternative-mentioned condition. This could have facilitated processing of the alternative-mentioned condition throughout the sentence, as suggested by the generally lower reading times before and after the target word in this condition relative to the new one without an alternative. Without a difference between the given conditions, it is impossible to adjudicate between a contrastive alternatives-based versus simple repetition-based explanation for the faster response times to new alternative-mentioned conditions here in this study alone. However, the results of Experiment 3 for another kind of bound focus, clefts, suggest that the lack of difference

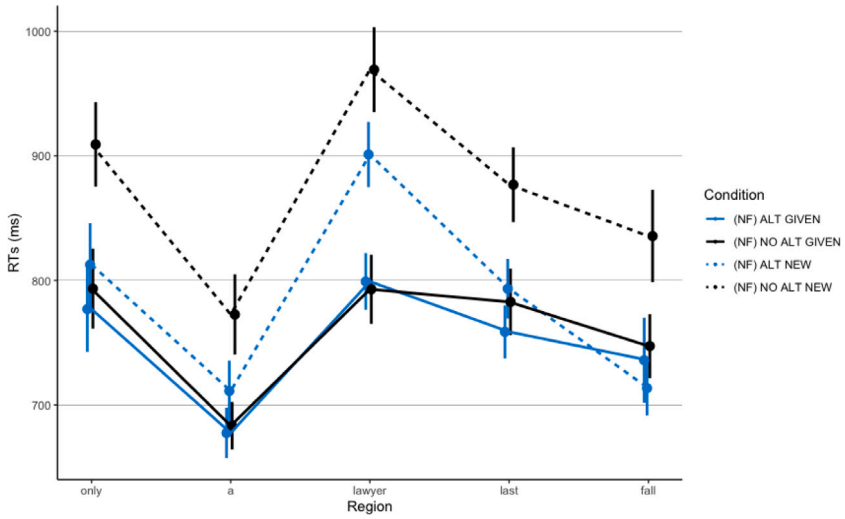

Fig. 2. Experiment 2: mean RT in each region in each condition. Error bars represent the 95% confidence interval.

between the given conditions in Experiment 2 is, itself, a floor effect, because the cleft structure reverses this difference in word repetition across conditions.

Experiment 3

The difference between Experiment 2 and Experiment 3 is that in the latter, an *it*-cleft was used to focus target words, instead of the focus particle *only*.

Method

Participants. 53 native speakers of English were recruited via Prolific. Data from 48 participants were included in the analysis; 5 participants were excluded because they failed to complete more than 70% of the Maze sentences.

Materials. The items of Experiment 3 consisted of modified versions of those of Experiment 2. An example of an item is in (15), below.

(15) *Speaker A:* This company often makes bad decisions, but...

a. Did they hire a **lawyer** last fall, or an accountant? (NF) alt, given

b. Did they hire a **lawyer** last fall? (NF) no alt, given

c. Did they hire an accountant last fall? (NF) alt, new

d. What did they announce last time? (NF) no alt, new

Speaker B: I think they announced it was a **lawyer** that they hired, but I'm not sure.

As in Experiment 2, the preceding context questions of Experiment 3 manipulated whether an alternative to the expression in focus was either mentioned or not (alt vs. no alt) and whether the focus itself was previously mentioned or not (given vs. new). Like the focus particle *only*, the cleft structure (*it was a...*) caused target words in all conditions of Experiment 3 to be unambiguously narrow bound foci.

Besides replacing *only* with a cleft, another difference in the target sentences between Experiment 3 and Experiment 2 was that the verbs of which the target words were direct objects were moved to immediately after target words, as in *lawyer that they hired* in (15). For this reason, Experiment 3 no longer confounded givenness with the simple repetition of the words immediately before the target word; if anything, it was the new condition without alternatives mentioned in (15d) that contained the most repetition across context and target sentences before the target word. This reversed the pattern of which conditions contained the most repetition from the one in Experiment 2. The final difference between the materials for these two experiments was that the phrase that previously functioned as a short spillover region (*last fall* in Experiment 1 and Experiment 2) was removed from the target sentence in Experiment 3 to make the target sentence slightly shorter and more natural.

All materials of Experiment 3 were first assessed in an acceptability judgment study; the results of this can be found in "Appendix A". Fillers, practice items, and comprehension questions were the same as in the previous two experiments.

Procedure. As in Experiment 1 and Experiment 2, target sentences were implemented in the Maze task. Maze foils for Experiment 3 were independently generated using the AutoMaze algorithm, with the result that the foils in this experiment were not directly based on those generated for Experiment 1 and Experiment 2. This was necessary, because the target sentences in Experiment 3 are of a different structure from the target sentences in Experiment 1 and Experiment 2. For this same reason, a direct comparison between response times obtained in these experiments and those from Experiment 3 would not have been possible regardless of the way in which the foils were generated.

Analysis. The analysis was the same as that of Experiment 2.

Results

Mean response times for the target word and its surrounding regions in all conditions are presented in Table 8. They are plotted with 95% confidence intervals in Fig. 3.

The mean comprehension question accuracy was 86%, and the mean completion rate of the maze target sentences of Experiment 3 was 83%.

Fixed effects estimates for the model fitted to log-transformed response times are reported in Table 9; those for the model fitted to untransformed response times are reported in Table 10 (See "Appendix D" for results on spillover regions). Significantly positive estimates for newness again indicated a slowdown on new targets compared to targets that were mentioned in the previous question. Unlike Experiment 2, however, the models fitted to Experiment 3 revealed a small but significant effect of the presence of alternatives, indicating that foci were read faster in the presence of a contextual alternative than in the absence of one. Finally, a significant interaction between newness and the presence of alternatives was also found, suggesting that the difference in response times between the two new conditions was larger than the difference between the two given conditions. As in Experiment 2, pairwise comparisons on log-transformed response times revealed that the effect of the presence of alternatives only reaches

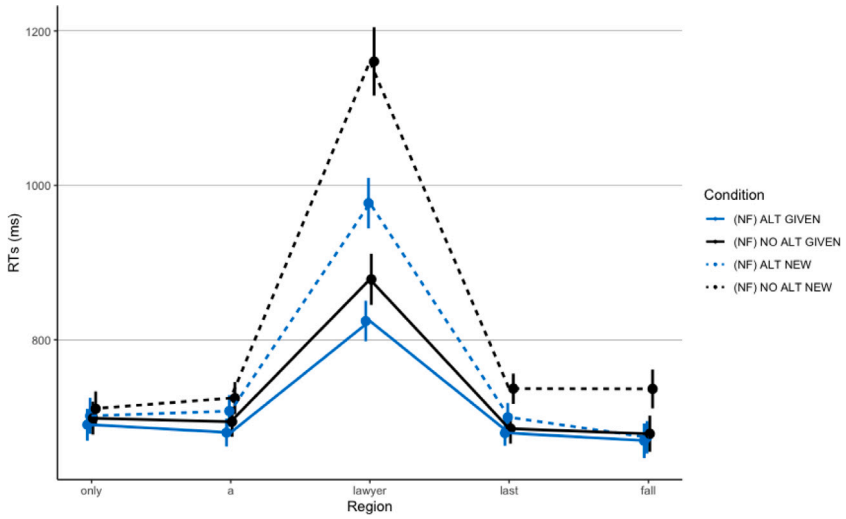

Fig. 3. Experiment 3: mean RT in each region in each condition. Error bars represent the 95% confidence interval.

Table 8
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
(NF) alt, given	691.65 (10.52)	681.21 (9.19)	823.68 (9.19)	680.58 (8.42)	670.26 (11.46)
(NF) no alt, given	697.15 (10.78)	693.09 (9.84)	877.81 (16.87)	683.79 (9.84)	678.52 (11.91)
(NF) alt, new	701.00 (11.80)	707.76 (10.44)	977.80 (16.66)	698.56 (9.43)	673.30 (10.68)
(NF) no alt, new	710.78 (11.44)	724.57 (10.59)	1160.42 (22.56)	736.85 (9.96)	736.54 (12.81)

Table 9
Parameter values for fixed effects in mixed linear regression model of LogRTs in Experiment 3.

	Estimate	Std. Error	t value
(Intercept)	6.66	0.034	197.58
Alternative	0.05	0.023	2.03
Newness	0.17	0.021	8.19
Alternative:Newness	0.11	0.040	2.81

Table 10
Parameter values for fixed effects in mixed linear regression model of untransformed RTs in Experiment 3.

	Estimate	Std. Error	t value
(Intercept)	826.99	31.93	25.90
Alternative	55.48	24.13	2.30
Newness	157.85	23.82	6.63
Alternative:Newness	128.39	44.78	2.87

significance in the new conditions ($t = -5.60$), but not in the given conditions ($t = 2.03$). On the words following and preceding the target word, no significant effects were observed in both untransformed and log-transformed RTs.

Discussion

Experiment 3 replicated the newness effects reported above for both Experiment 1 and Experiment 2: In all experiments, responses were slower to new foci than given foci. Experiment 3 also replicated Experiment 2 in finding that the slowdown for new foci was smaller when a context mentioned an alternative expression to the target word. This effect from Experiment 2 was replicated in Experiment 3, even though it was the new conditions without alternatives that contained the most repetition across contexts and targets, that is, that would have been read the fastest if the difference between given conditions in Experiment 1 had been due to repetition. This suggests that, for bound foci generally, either previously encountering the expression in focus or previously encountering an alternative to the focus facilitates the comprehension of the focus itself. Together with Experiment 2, these results support the hypothesis that the presence of alternatives in the context aids the comprehension of a subsequent bound focus.

But neither Experiment 2 nor Experiment 3 identifies *how* explicitly mentioned alternatives have this attenuating effect in on-line focus processing. Contextually relevant alternatives will most often be semantically associated with the foci with which they contrast. Could it thus be that the benefit for focus processing of contextually mentioning an alternative reported here is due to semantic associate priming of foci from their preceding alternatives? Understanding the real-time comprehension of focus requires understanding whether contrastive alternatives have any explanatory status in our theory of human language processing, or whether semantic associate priming can subsume the effects found in Experiment 2 and Experiment 3.[1]

Within the current studies, there is reason to hypothesize that contrastiveness cannot be reduced to semantic priming (convergent evidence for this from other studies will be introduced in the General Discussion). In Experiment 1, the same alternatives as those in Experiments 2 and 3 were used in preceding context questions to put target words in narrow focus. Instead of resulting in a significant speed up on this target region, the presence of these alternatives slowed down reading on this word. This finding already suggests that even if semantic priming drove the effects in Experiment 2 and 3, it did not affect reading times in the same way in Experiment 1. For this reason, in order to appeal to semantic priming to explain the effects of Experiments 2 and 3, one would also have to explain how the effect of semantic priming could be modulated by the type of focus construction (either free or bound focus) employed in these experiments. The predictability of bound foci may be a part of the explanation for why they are read less slowly in the presence of alternatives. But this greater predictability alone, absent any consideration of contrastive alternatives, does not account for the fact that the presence of alternatives seems to facilitate the comprehension of bound foci, because this predictability does

[1] We thank anonymous reviewers for emphasizing the importance of this point.

not encompass the reason that alternatives are needed in the first place. It is the meaning of clefts and the particle *only* that requires the comprehender to consider contrastive alternatives, because those constructions require the comprehender to draw inferences about what is *not* happening. This requirement would not be satisfied by semantic associates of a focused word that did not contrast with that word on a salient dimension or were not replaceable.

A clear prediction of this account is that the focus slowdown should be attenuated more by the presence of an alternative (whether this alternative is semantically associated to the focus or not), than a mere associate of a focused word. This motivated Experiment 4.

Experiment 4

To show that it is the presence of contrastive alternatives in a context, and not just semantic priming, that aids the comprehension of subsequent bound foci, the context sentences of Experiment 4 manipulated the presence of expressions that were semantically associated to upcoming foci independently of the presence of alternatives to those foci. If the facilitatory effect of alternatives found in Experiment 2 and Experiment 3 is solely due to semantic priming, then attenuation effects should parallel the degree of semantic association between contextually-mentioned words of any kind (whether alternatives or not) and foci. In the presence of alternatives that are not closely semantically related to the focus (*non-associated alternatives*), the benefit from alternatives should be weak or non-existent. Moreover, in such a scenario, semantically related words that nevertheless are unlikely to serve as contrastive alternatives to the focus (*associated non-alternatives*) would be expected to give rise to similar facilitatory effects as semantically associated alternatives. This must be understood as the likelihood of being an alternative, and not absolute possibility, because, for example, all imageable nouns could be alternatives in answer to the question *What did you see?*, but the intuition is still retained that the President of the United States and a pink armchair are unlikely to be alternatives to each other in many other scenarios.

Experiment 4 therefore investigates the extent to which both non-associated alternatives and associated non-alternatives aid the comprehension of a subsequent focus, by crossing contextual mention of alternatives with contextual mention of semantically associated expressions. If the benefit from the presence of contrastive alternatives in reading foci is not reducible to semantic associate priming, then alternatives that are not associated with foci would be expected to attenuate the focus slowdown more than associates that are not alternatives of those foci.

Method

Materials. The items of Experiment 4 consisted of modified versions of those of Experiment 3. Preceding context questions manipulated whether a relatively likely alternative to the expression in focus was either mentioned or not (alt vs. no alt) and whether an associated expression to the focus was mentioned or not (assoc vs. non-assoc). An example of an item is in (16), below.

In the associated alternative condition, the context question contained an alternative (*painter*) that was strongly associated with the focus (*sculptor*), as in (16a). Here, *painter* and *sculptor* are related expressions, but both expressions contrast with each other along a single dimension. The associated non-alternative context question mentioned an associated expression that would not usually serve as a relevant alternative to the focus (*statue*), as in (16c). This is because in order to be a contrastive alternative to a focus, an expression needs to be replaceable with that focus, and there are fewer situations in which *statue* and *sculptor* are expected to be replaceable with each other than *painter* and *sculptor* (see *Alternatives in the processing of focus*, above). In this particular target sentence, *statue* ultimately cannot replace *painter*, because statues cannot give talks. For the purpose of incremental reading in the Maze, where participants do not encounter the material that rules out *statue* until after the target word in the sentence, the manipulation depends on the fact that word *statue* is more generally unlikely be an alternative to *sculptor*, because the first is inanimate but the second is animate. The non-associated alternative condition mentioned an alternative that was only weakly associated (*lawyer*) with the focus, as in (16b). Even though this expression may not immediately come to mind when reading a word like *sculptor*, the word *lawyer* can still serve as a contrastive alternative because both are animate, and are therefore likely to both take similar roles in the events in which they participate and participate in similar events. Finally, the non-associated non-alternative condition mentioned neither a possible alternative nor a semantically associated expression, as in (16d).

(16) *Speaker A:* I can't really remember what talks are happening at the conference today.

a.	Will the last speaker be a painter?	alt assoc
b.	Will the last speaker be a lawyer?	alt nonassoc
c.	What did you say about a statue?	no alt assoc
d.	What did you see on the program?	no alt nonassoc

Speaker B: I think I saw that it was a **sculptor** who will be giving the last talk of the day, but I'm not sure.

As in Experiment 3, a cleft was used to unambiguously put the target word (*sculptor*) in narrow focus in all conditions. In all conditions, the focused target expression itself was new.

The level of semantic association between the expressions in focus and their alternatives or associated non-alternatives was based on their Latent Semantic Analysis scores (Landauer, 1999). For each experimental item, semantic association between the focus and both the associated alternative and the associated non-alternative was at least .3, while the semantic association between the focus and the non-associated alternative was at most .13. Moreover, the difference in LSA score between the associated non-alternative and the associated alternative was larger than −0.1. Out of the 48 items that were written, 24 were based on quadruplets from (Husband & Ferreira, 2015), where only those items were selected that satisfied the criteria above. A list of the key words of all 48 items is given in "Appendix C".

Practice items and comprehension questions were similar to the previous three studies. 64 additional filler sentences were included, 48 of which also involved either a cleft or a focus particle and 16 of which involved only broad focus.

Procedure. Maze foils for Experiment 4 were generated using the AutoMaze algorithm. Again, context sentences were presented normally on a separate screen prior to the target sentence which was presented using the Maze task.

Participants. 52 native speakers of English were recruited via Prolific. Data from 48 participants were included in the analysis; 4 participants were excluded because they failed to complete more than 68% of the Maze sentences.

Analysis. The analysis was the same as that of Experiment 3, except that fixed effects were the presence versus absence of an associated expression and the presence versus absence of alternatives, again dummy-coded. The presence of an alternative/associated expression was treated as the reference level of these factors.

Results

Mean response times for the target word and its surrounding regions in all conditions are presented in Table 11. They are plotted with 95% confidence intervals in Fig. 4.

The mean comprehension question accuracy was 87%, and the mean completion rate of the maze target sentences of Experiment 4 was 85%.

Table 11
Experiment 4: 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
(NF) alt, assoc	708.90 (11.52)	713.56 (12.01)	1004.01 (19.76)	712.43 (19.35)	750.47 (25.72)
(NF) alt, non-assoc	695.02 (11.20)	718.57 (13.63)	1299.07 (27.80)	717.78 (21.94)	724.22 (22.91)
(NF) no alt, assoc	699.40 (10.56)	725.78 (12.31)	1151.17 (23.26)	752.25 (26.81)	837.98 (29.74)
(NF) no alt, non-assoc	702.22 (12.03)	755.04 (13.67)	1378.40 (30.08)	755.86 (26.82)	850.51 (36.47)

Table 12
Parameter values for fixed effects in mixed linear regression model of LogRTs in Experiment 4.

	Estimate	Std. Error	t value
(Intercept)	2.971274	0.009348	317.847
Association	0.056674	0.009509	5.960
Alternative	0.102899	0.012722	8.088
Alt:Assoc	−0.029763	0.014339	−2.076

Table 13
Parameter values for fixed effects in mixed linear regression model of untransformed RTs in Experiment 4.

	Estimate	Std. Error	t value
(Intercept)	1002.80	28.30	35.439
Association	149.50	32.02	4.669
Alternative	294.85	44.43	6.636
Alt:Assoc	−69.62	51.70	−1.347

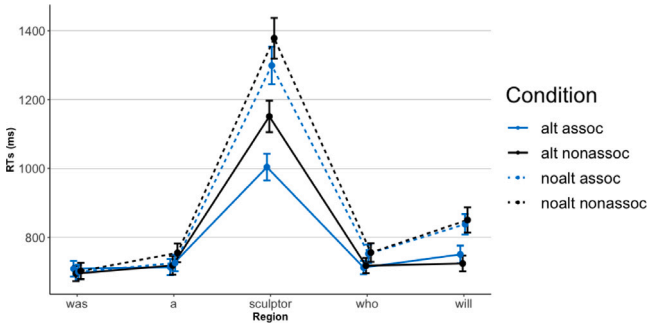

Fig. 4. Experiment 4: mean RT in each region in each condition. Error bars represent the 95% confidence interval.

Fixed effects estimates for the model fitted to log-transformed response times are reported in Table 12; those for the model fitted to untransformed response times are reported in Table 13. Like Experiment 3, the models fitted to Experiment 4 revealed a significant effect of the presence of alternatives, indicating that foci were read faster in the presence of a contextual alternative than in the absence of one (t = 8.08). Moreover, models also revealed a significant effect of association, indicating that foci were read faster in the presence of a semantically associated word than in the absence of one (t = 5.97). Pairwise comparisons on log-transformed response times moreover revealed that foci were read faster in the non-associated alternative conditions than in the associated non-alternative condition (t = −3.69). No significant effects were observed on the region immediately preceding and following the target, but models fitted to raw and log RTs on the region two words after the target revealed a significant main effect of the presence of alternatives, suggesting that this region was read faster in the presence of alternatives than in their absence (t = 4.79).

Discussion

Experiment 4 replicated the effect of the presence of alternatives found in Experiment 2 and Experiment 3. In the presence of alternatives in the context, responses were faster than in their absence. Although the presence of semantically associated expressions generally facilitated the reading of a subsequent focus, the presence of alternatives in the context reduced response times independently of semantic association: responses on targets were faster in the presence of alternatives even when these alternatives were not associated with the focus. Moreover, the facilitatory effect of alternatives in the context (whether associated or not) was stronger than the facilitatory effect of expressions that were semantically associated to the focus but could not serve as an alternative to the focus. This suggests that the facilitatory effect of alternatives in Experiment 2 and Experiment 3, too, cannot solely be due to semantic priming.

General discussion

Experiment 1 demonstrated that focus effects in reading are not reducible to the newness/givenness distinction, and Experiments 2–4 demonstrated that, instead, the appropriate understanding of focus for language processing research must reference which parts of a sentence are most relevantly contrasted with alternatives. The results of Experiment 1 supported this conclusion, because longer response times were found on narrow focused words than words that were part of broad focused phrases, even after controlling for newness versus givenness. This manipulation of the size of foci was achieved by presenting contrastive alternatives in context questions before target sentences which did not, themselves, contain an explicit focusing construction. At the same time, Experiments 2–3 demonstrated that the earlier mention of alternative expressions can somewhat attenuate the cost of processing new material when that new material is also a bound focus, i.e., when it is explicitly cued and gives rise to a strong inference. Experiment 4 showed that this effect of explicit mention of alternatives in the context cannot solely be due to the fact that those alternatives semantically prime focused targets.

Overall, these results accord with the findings of Benatar and Clifton (2014), Birch and Rayner (1997), and Lowder and Gordon (2015), who all argued for a general processing cost of focus based on observed longer reading times in both early and late eye movement measures (with significant slowdowns reported in e.g., first fixation, gaze duration, regression path duration and total reading times). But this general focus slowdown cannot be reduced to the newness/givenness distinction, even in the absence of extra inferences or contextual support. Instead, our findings are more in accord with Lowder and Gordon's interpretation of focus processing costs as deeper encoding or more effortful integration of focus. We suggest that part of this more effortful integration could involve either the computation or retrieval of contrastive alternatives, but we must leave investigation of the exact nature of this for future work.

The present findings are also potentially compatible with the speed-ups on focused material reported by Birch and Rayner (2010) and Morris and Folk (1998), who report shorter reading times in first fixations, gaze durations, as well as total reading times. Both Birch and Rayner and Morris and Folk made use of materials in which target words were new, alternatives to the expression in focus were explicitly mentioned, and target words were focused by clefts, just as in the present Experiment 3 and Experiment 4. While Experiment 3 and 4 only found that new foci which followed alternatives were read *less slowly* than new foci which did not follow alternatives, without evidence of a speed-up, the present studies' baseline conditions were unlike Morris and Folk's and Birch and Rayner's. Looking more closely at Morris and Folk's materials may provide a clue for an alternative explanation about

why such a speed-up may have arisen. An example item is repeated in (17) below, in which an alternative, *waiter*, always preceded the target word *accountant*.

(17) a. While the waiter watched, it was the **accountant** who balanced the ledger a second time. focus
b. It was the waiter who watched while the **accountant** balanced the ledger a second time. defocus

It may be that this earlier alternative expression, *waiter*, was ultimately also understood as focused by the readers of both of these sentences, because it was understood in clear contrast with *accountant*. In other words, the target word *accountant* in the defocus condition might have received contrastive focus as well, because it seemed like a relevant alternative to *waiter*. In that case, the comparison made in this study would have been one between an ultimately contrastive focus without a preceding alternative as in (17b), because the waiter was ruled out of potentially balancing the ledger by the time the accountant was encountered, and a focus inside a cleft with a preceding alternative as in (17a), where the latter type of focus gave rise to shorter reading times than the first. This could have been due to the focused target word in (17a) being more clearly demarcated as focused by its preceding cleft than the ultimately contrastively focused target word in (17b) requiring more inference on the comprehender's part. If the speed-up in reading the focused target words in Morris and Folk's study was due in large part to a combination of the presentation of alternatives before the target word and the clarity of focus marking provided by cleft constructions, then this explanation would extend to Birch and Rayner's (2010) faster reading times on focused words as well.

If inherent ambiguity in the location and size of a *free* focus makes it more costly to process than a focus bound by a cleft or *only*, which are overtly signaled, then the different patterns of processing times on foci after alternatives in the present Experiment 1 versus Experiment 2 and 3 could be understood as the construction-specific demands of focus processing. Morris and Folk's and Birch and Rayner's studies may, in fact, be better understood as more similar to comparing the broad focused, no-alternatives-mentioned, new condition of Experiment 1 as a baseline against the narrow focused, alternatives-mentioned, new condition of Experiment 3. At this point, it is not clear whether the facilitated reading of bound foci after explicitly mentioned alternatives that we observe here is due to the particular properties of the syntactic expressions (clefts, *only)* they were linked to, or whether it is due to the general fact that they are bound at all (and hence, we do not know whether foci bound by other particles, such as *too* or *even*, would show the same effect).

One possible piece of evidence in support of the suggestion that free foci are generally more costly to process comes from self-paced reading studies reported by Fraundorf et al. (2013), who also showed slowdowns on foci that occurred even after the explicit mention of contrastive alternatives. Unlike any of the studies discussed here thusfar, in Fraundorf et al.'s materials, foci were marked using font emphasis. No focusing device, whether contextual or syntactic, signaled the presence of the upcoming focus in advance of the emphasized word. Fraundorf et al.'s studies may therefore have yielded a slowdown even following contextually-mentioned alternatives, because their conditions had in common with free foci the property that comprehenders were not able to confidently anticipate a focus before they encountered it.

The studies reported here thus provide support for Benatar and Clifton's and Lowder and Gordon's suggestion that different focus constructions may all be processed slightly differently. The attenuation of a newness slowdown when alternatives to foci were explicitly mentioned was only observed for material that was focused by either the particle *only* or a cleft, that is, for bound foci. In Experiment 1, the narrow foci necessarily occurred in a context in which an explicit alternative was mentioned, but these new narrow-focused target words were not read faster than target words that were part of a new broad-focused phrase. Alternative expressions seem to be most useful when the focus structure of a sentence is clearly signaled.

If comprehending foci also requires some understanding of the most relevant contrastive alternatives, then it would be expected that the contextual mention of alternatives can aid in the reading of foci. We thus propose that psycholinguistic theories adopt the concept of contrastive alternatives to explain the comprehension of focus. If bound and free focus share an interpretive dependence on alternatives, a view the formal semantics literature has adopted, contrastive alternatives may be involved even in the processing of free foci, though future work will have to determine in what way. At a more algorithmic level, alternative expressions may provide some semantic associate priming benefit to upcoming foci, and the process of fully comprehending a focus may encompass, first, the activation of semantically associated expressions, followed, second, by the narrowing of those associated expressions into only the set that would be contrastive in the current context, as has been suggested by Husband and Ferreira (2015), inter alia. But it may also be the case that contrastive alternatives are either more directly computed or more directly retrieved via their contrastiveness, because comprehenders already have *a priori* expectations about the salient dimensions along which concepts are likely to be contrasted with each other when they first encounter a focused word. Either of these mechanistic understandings of focus processing would accord well with both the studies that have found activation of alternative sets from focused words (Braun & Tagliapietra, 2010; Fraundorf et al., 2013, 2010; Gotzner et al., 2016; Husband & Ferreira, 2015) and a growing body of reading studies that have demonstrated that comprehenders use the content of focused expressions to anticipate the upcoming mention of contrastive alternatives (Ferreira & Lowder, 2016; Filik, Paterson, & Liversedge, 2009; Lowder, Ryan, Opie, & Kaminsky, 2021).

For the present, this paper puts forward evidence that previous reading results can be explained by adopting the appropriate (computational level) understanding of focus for psycholinguistic theories. This is the same as the understanding of focus in formal linguistics — the only property that unifies all focus constructions is the requirement that contrastive alternatives be considered in order to understand the meaning of focused expressions. Such an alternatives-based conceptualization of focus for language processing is compatible with results from eye movements, the Maze task, self-paced reading, semantic priming, and event-related potentials, which all converge on the conclusion that the comprehenders automatically consider alternatives to focused expressions during the course of real time language processing.

CRediT authorship contribution statement

Morwenna Hoeks: Conceptualization, Methodology, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Maziar Toosarvandani:** Conceptualization, Resources, Writing – original draft, Writing – review & editing, Funding acquisition, Supervision. **Amanda Rysling:** Conceptualization, Resources, Writing – original draft, Writing – review & editing, Funding acquisition, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All materials, data and analysis code of this and subsequent experiments are made available via the Open Science Framework and can be accessed via https://osf.io/k6tbw/?view_only=71d86431090046929d56f1ba94dcc38b. This study's design and its analysis were not pre-registered.

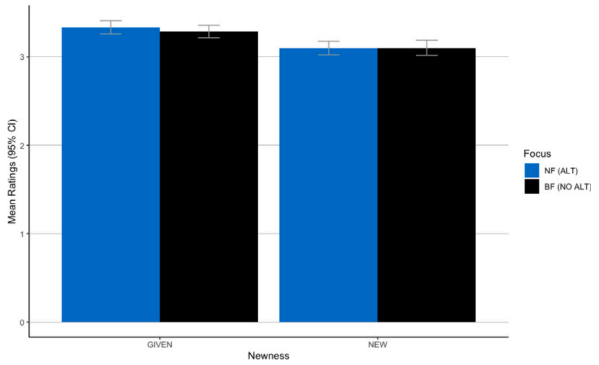

Fig. 5. Experiment A.1: mean rating in each condition. Error bars represent the 95% confidence interval.

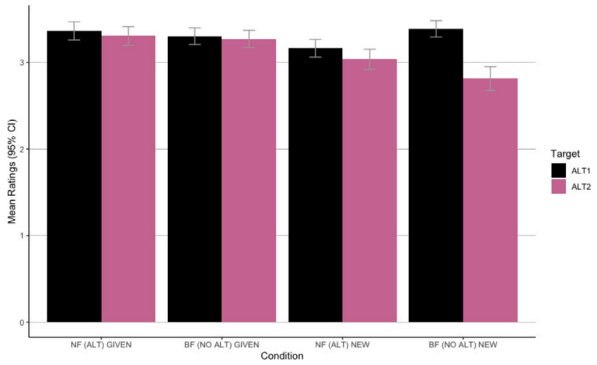

Fig. 6. Experiment A.1: mean rating by alternative in each condition. Error bars represent the 95% confidence interval.

Acknowledgments

We are grateful to Pranav Anand, Adrian Brasoveanu, Charles Clifton, Scott Fraundorf, Lyn Frazier, Jesse Harris, Dean McHugh, Kelsey Sasaki, Shayne Sloggett, Adrian Staub, Nick Van Handel, and Matt Wagers for their feedback, as well as audiences at the UC Santa Cruz Graduate Research Symposium and CUNY 2021 for their comments and questions.

Appendix A. Offline acceptability ratings

The offline acceptability judgment studies discussed in this section aimed to establish the extent to which the materials used in the Maze task online reading studies were considered natural by native speakers of English. To that end, Experiments A.1-3 use the same stimulus and filler materials as Experiments 1-3. Since reduced acceptability ratings have been shown repeatedly to provide an indication of a significant processing cost, these offline studies also provided preliminary and convergent evidence for potential focus costs.

Participants were from the same population as Experiments 1-3 and recruited in the same way. Sentences were presented using the Ibex Farm platform for web-based experiments (Drummond, 2013).

In each trial, participants read a full dialogue on a single screen and were asked to judge the naturalness of the full discourse on a 4-point Likert scale. The practice items provided guided feedback to make sure participants were familiar with the use of the scale.

All of the studies reported here were analyzed with mixed effects ordinal regression models fitted to the rating data using the `clmm` function of the ordinal package in R (Christensen, 2019; R Core Team, 2021). All fixed and random effects structures parallel those used for the Maze studies, unless otherwise noted.

A.1. Experiment A.1

An example item of Experiment A.1 is given in (18) below.

(18) *Speaker A:* This company often makes bad decisions, but...

a. Did they hire a **lawyer** last fall, or an accountant? NF, given
b. Did they hire a **lawyer** last fall? BF, given
c. Did they hire an accountant last fall? NF, new
d. What did they announce last time? BF, new

Speaker B: I think they announced they hired a **lawyer** last fall, but I'm not sure.

Mean acceptability ratings per condition are shown in Figure Fig. 5 and Table 14.

Table 14
Experiment A.1: Mean rating and standard error of the mean by condition and by target and presence of alternatives.

Condition			Target Identity		
NF (alt), given	3.34	(0.043)	alt1	3.36	(0.053)
			alt2	3.30	(0.055)
BF (no alt), given	3.29	(0.035)	alt1	3.30	(0.048)
			alt2	3.27	(0.050)
NF (alt), new	3.10	(0.039)	alt1	3.16	(0.051)
			alt2	3.03	(0.059)
BF (no alt), new	3.10	(0.038)	alt1	3.28	(0.048)
			alt2	2.81	(0.70)

This acceptability rating study also aimed to establish whether, in the NF given condition as in (18a), the eventual target word was considered a natural alternative expression to the alternative mentioned the preceding question and vice versa. If the target and the alternative expression were indeed proper alternatives to each other, it would be expected that it would not matter which one was mentioned in the question and which one was mentioned in the target sentence. In Experiment A.1, both the intended question/answer pairs and the question/answer pair in which the position of the target and the alternative expression were switched were tested.

The identity of target and alternative expression was treated as a between-subjects manipulation: one group of participants (n = 48) were presented with the set of items that were be used in our reading studies, while a second group of participants (n = 48) were presented with the version of all the items in which the target and the alternatives were switched (see Fig. 6).

Results of the mixed effects ordinal regression models are shown in Table 15. A t-value of 2 was considered to be the critical value for significance. The broad focus and given conditions were treated as baselines throughout.

The model that included the between-subjects fixed effect did not find a significant main effect for target identity ($z = -0.322$). However, the triple interaction between target identity, newness and focus turned out to be significant ($z = -4.100, p < 0.001$). Pairwise comparisons revealed that the only significant difference between the two target forms occurred in the BF (no alt), new condition, in which items with alt2 as the target received lower ratings than items with alt1 as the target ($z = -4.286, p < .01$ after applying Bonferroni-corrections for multiple comparions).

Table 15
Parameter values for fixed effects in mixed ordinal regression model of acceptability judgments in Experiment A.1.

	Estimate	Std. Error	z value
New	−0.8658	0.1515	−5.716
Focus	−0.3615	0.1707	−2.118
New:Focus	0.5338	0.2539	2.102

Table 16
Experiment A.2: mean rating and standard error of the mean by condition and by target and presence of alternatives.

Condition			Target Identity		
(NF) alt, given	3.30	(0.033)	alt1	3.27	(0.047)
			alt2	3.34	(0.048)
(NF) no alt, given	2.77	(0.039)	alt1	2.73	(0.056)
			alt2	2.82	(0.053)
(NF) alt, new	3.19	(0.034)	alt1	3.12	(0.051)
			alt2	3.26	(0.046)
(NF) no alt, new	3.04	(0.040)	alt1	3.19	(0.052)
			alt2	2.89	(0.600)

Table 17
Parameter values for fixed effects in mixed ordinal regression model of acceptability judgments in Experiment A.2.

	Estimate	Std. Error	z value
Newness	−0.3194	0.1460	−2.188
Alternative	−1.4143	0.1661	−8.516
Newness:Alternative	0.9407	0.2306	4.079

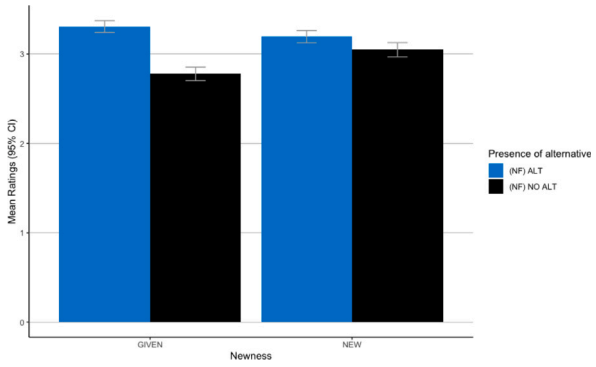

Fig. 7. Experiment A.2: mean rating in each condition. Error bars represent the 95% confidence interval.

A.2. Experiment A.2

An example item of Experiment A.2 is repeated in (19) below.

(19) *Speaker A:* This company often makes bad decisions, but...
- a. Did they hire a **lawyer** last fall, or an accountant? (NF) alt, given
- b. Did they hire a **lawyer** last fall? (NF) no alt, given
- c. Did they hire an accountant last fall? (NF) alt, new
- d. What did they announce last time? (NF) no alt, new

Speaker B: I think they announced they hired **only** a **lawyer** last fall, but I'm not sure.

Mean acceptability ratings per condition are shown in Fig. 7 and Table 16. Results of the mixed effects ordinal regression models are shown in Table 17.

The same between-subjects manipulation of target identity was used as in Experiment A.1 to investigate the effect of the specific lexical material making up the target and the alternative expressions (see Fig. 8). In the model including the between-subjects manipulation of target identity, the main effect of target identity did not reach significance ($z = 0.65$). However, this model revealed a significant three-way interaction between target identity, presence of an alternative and newness ($z = -2.55, p < 0.05$). Again, this indicates that acceptability judgments for items with alt2 as the target were only significantly lower than items with alt1 as the target in the (NF) no alt, new condition ($z = -2.989, p < 0.05$ after Bonferroni correction for multiple comparisons).

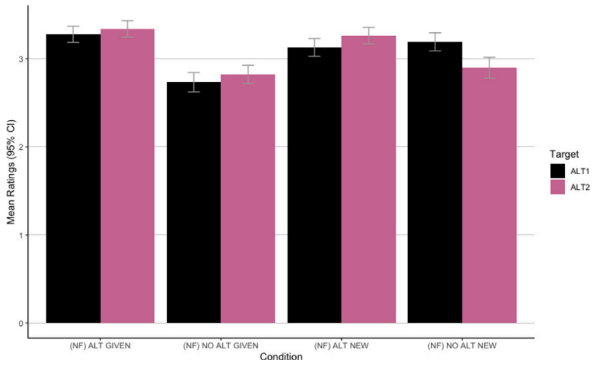

Fig. 8. Experiment A.2: mean rating by alternative in each condition. Error bars represent the 95% confidence interval.

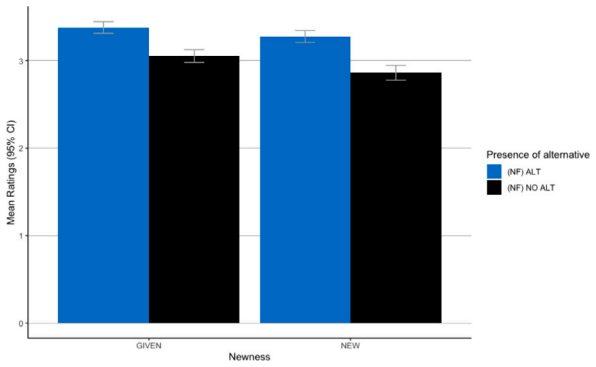

Fig. 9. Experiment A.3: mean rating in each condition. Error bars represent the 95% confidence interval.

Table 18
Experiment A.3: mean rating and standard error of the mean by condition.

Condition		
(NF) alt, given	3.38	(0.03)
(NF) no alt, given	3.05	(0.03)
(NF) alt, new	3.27	(0.03)
(NF) no alt, new	2.86	(0.04)

A.3. Experiment A.3

Mean acceptability ratings per condition of Experiment A.3 are summarized in Fig. 9 and Table 18. The data analysis was again analogous to that of Experiment A.2, except that it did not include a between-subjects fixed effects for target identity. Model outputs are given in Table 19.

An example item of Experiment A.3 is given in (20) below.

(20) *Speaker A:* This company often makes bad decisions, but...
- a. Did they hire a **lawyer** last fall, or an accountant? (NF) alt, given
- b. Did they hire a **lawyer** last fall? (NF) no alt, given
- c. Did they hire an accountant last fall? (NF) alt, new
- d. What did they announce last time? (NF) no alt, new

Speaker B: I think they announced it was a **lawyer** that they hired, but I'm not sure.

Appendix B. Materials Experiment 1

Materials for Experiment A.1 and Experiment 1 are given below. These materials were then adapted to create materials for the other experiments.

Table 19
Parameter values for fixed effects in mixed ordinal regression model of acceptability judgments in Experiment A.3.

	Estimate	Std. Error	z value
Newness	−0.3790	0.1485	−2.552
Alternative	−0.9406	0.1503	−6.257
Newness:Alternative	−0.2012	0.2471	−0.815

(1) **Context:** Abbie is a very picky eater.
- a. Did she want chocolate cake for dessert, or apple pie? NF (alt), given
- b. Did she want chocolate cake for dessert? BF (no alt), given
- c. Did she want apple pie for dessert? NF (alt), new
- d. Do you remember what she said? BF (no alt), new

Target: I think Abbie said she wanted chocolate cake for dessert, but I'm not sure.

(2) **Context:** Ben is feeling very sick and we're trying to figure out why.
- a. Did he eat pasta at the restaurant, or pizza? NF (alt), given
- b. Did he eat pasta at the restaurant? BF (no alt), given
- c. Did he eat pizza at the restaurant? NF (alt), new
- d. What do you remember about yesterday? BF (no alt), new

Target: I think I saw him eating pasta at the restaurant, but it could have been something else.

(3) **Context:** We need a few computers for the lab.
- a. Did Charlie buy a desktop at the store, or a laptop? NF (alt), given
- b. Did Charlie buy a desktop at the store? BF (no alt), given
- c. Did Charlie buy a laptop at the store? NF (alt), new
- d. What did Charlie tell you again? BF (no alt), new

Target: I think Charlie told me he bought a desktop at the store, although I could be wrong.

(4) **Context:** Dave had to get rid of a lot of his stuff.
- a. Did he sell his washing machine when he moved out, or his dryer? NF (alt), given
- b. Did he sell his washing machine when he moved out? BF (no alt), given
- c. Did he sell his dryer when he moved out? NF (alt), new
- d. What did he say about it? BF (no alt), new

Target: I believe he said he sold his washing machine when he moved out, but he didn't tell his roommates.

(5) **Context:** I wonder how Erik is doing these days.
- a. Does he have regrets from his previous marriage, or fond memories? NF (alt), given
- b. Does he have regrets from his previous marriage? BF (no alt), given
- c. Does he have fond memories from his previous marriage? NF (alt), new
- d. What did he say the other day? BF (no alt), new

Target: I believe he said he has regrets from his previous marriage, but I'm not sure.

(6) **Context:** I'm looking for someone who can drop this off at work.
- a. Are you going to the store today, or to the office? NF (alt), given
- b. Are you going to the store today? BF (no alt), given
- c. Are you going to the office today? NF (alt), new
- d. What did you decide to do? BF (no alt), new

Target: I decided that I am going to the store today, but I might change my mind.

(7) **Context:** I'm just trying to figure out the logistics for this weekend.
- a. Are you dropping people off at the train station tomorrow afternoon, or at the bus stop? NF (alt), given
- b. Are you dropping people off at the train station tomorrow afternoon? BF (no alt), given
- c. Are you dropping people off at the bus stop tomorrow afternoon? NF (alt), new
- d. What do you think? BF (no alt), new

Target: I think I will be dropping people off at the train station tomorrow afternoon, but I can pick you up wherever.

(8) **Context:** I don't know what I should get.
- a. Are you drinking beer tonight, or wine? NF (alt), given
- b. Are you drinking beer tonight? BF (no alt), given
- c. Are you drinking wine tonight? NF (alt), new
- d. What do you think? BF (no alt), new

Target: I think I will be drinking beer tonight, but I don't know about the others.

(9) **Context:** I wonder how the reimbursement process works.
- a. Would it be better to pay with cash tomorrow, or with card? NF (alt), given
- b. Would it be better to pay with cash tomorrow? BF (no alt), given
- c. Would it be better to pay with card tomorrow? NF (alt), new
- d. What did Andrew say? BF (no alt), new

Target: Andrew said it would be better to pay with cash tomorrow, although it doesn't really matter.

(10) **Context:** Do you remember,
- a. did Faye order rice with her meal, or fries? NF (alt), given
- b. did Faye order rice with her meal? BF (no alt), given
- c. did Faye order fries with her meal? NF (alt), new
- d. what did Faye say just now? BF (no alt), new

Target: I believe she said she ordered rice with her meal, but we should ask her when she's back.

(11) **Context:** Greg offered to help me move my stuff next weekend, but
- a. does he drive a car these days, or a van? NF (alt), given
- b. does he drive a car these days? BF (no alt), given
- c. does he drive a van these days? NF (alt), new
- d. what did he say exactly? BF (no alt), new

Target: I believe he said he drives a car these days, but I would give him a call.

(12) **Context:** I'm thinking of buying Hana a birthday present.
- a. Has she been a fan of fantasy since her teenage years, or of science fiction? NF (alt), given
- b. Has she been a fan of fantasy since her teenage years? BF (no alt), given
- c. Has she been a fan of science fiction since her teenage years? NF (alt), new
- d. What did she say again? BF (no alt), new

Target: I think she said she has been a fan of fantasy since her teenage years, but I'm not sure.

(13) **Context:** I'm not sure what to get at the supermarket.

a. Does Jonathan like vanilla as an ice cream flavor, or strawberry? NF (alt), given
b. Does Jonathan like vanilla as an ice cream flavor? BF (no alt), given
c. Does Jonathan like strawberry as an ice cream flavor? NF (alt), new
d. What did Jonathan say before he left? BF (no alt), new

Target: I remember that he said he likes vanilla as an ice cream flavor, although I could be wrong.

(14) **Context:** I might have left my stuff at Kate's place after the event yesterday.

a. Did she find a jacket last night, or a sweater? NF (alt), given
b. Did she find a jacket last night? BF (no alt), given
c. Did she find a sweater last night? NF (alt), new
d. What did she say again? BF (no alt), new

Target: I think she said she found a jacket last night, but I would give her a call.

(15) **Context:** I'm trying to find out about the dietary restrictions of our guests.

a. Has Logan been allergic to peanuts ever since she was little, or to seafood? NF (alt), given
b. Has Logan been allergic to peanuts ever since she was little? BF (no alt), given
c. Has Logan been allergic to seafood ever since she was little? NF (alt), new
d. What did Logan say last time? BF (no alt), new

Target: I believe she said she has been allergic to peanuts ever since she was little, but I will double check.

(16) **Context:** We have to update your immunization record before we can proceed.

a. Were you vaccinated for tetanus recently, or for chicken pox? NF (alt), given
b. Were you vaccinated for tetanus recently? BF (no alt), given
c. Were you vaccinated for chicken pox recently? NF (alt), new
d. What did your doctor say? BF (no alt), new

Target: I think my doctor said I was vaccinated for tetanus recently, although I could be wrong.

(17) **Context:** I'm just wondering who made such a mess on this table.

a. Did Maria read a newspaper this morning, or a magazine? NF (alt), given
b. Did Maria read a newspaper this morning? BF (no alt), given
c. Did Maria read a magazine this morning? NF (alt), new
d. What did Maria say? BF (no alt), new

Target: I think she said she was reading a newspaper this morning, but I'm not sure.

(18) **Context:** I'm looking for some recommendations.

a. Does Tony like to listen to music while driving to work, or to a podcast? NF (alt), given
b. Does Tony like to listen to music while driving to work? BF (no alt), given
c. Does Tony like to listen to a podcast while driving to work? NF (alt), new
d. What did Tony say again? BF (no alt), new

Target: He said he usually likes to listen to music while driving to work, but he has horrible taste.

(19) **Context:** I'm trying to gauge his background knowledge.

a. Did he study biology in high school, or chemistry? NF (alt), given
b. Did he study biology in high school? BF (no alt), given
c. Did he study chemistry in high school? NF (alt), new
d. What did he tell you? BF (no alt), new

Target: I remember that he said he studied biology in high school, but you should ask him yourself.

(20) **Context:** Oliver really was an annoying kid.

a. Did he always make fun of his mother when he was younger, or of his sister? NF (alt), given
b. Did he always make fun of his mother when he was younger? BF (no alt), given
c. Did he always make fun of his sister when he was younger? NF (alt), new
d. What did his dad say again? BF (no alt), new

Target: I think his dad said he always made fun of his mother when he was younger, but it wasn't too bad.

(21) **Context:** I wonder how your mom got the information.

a. Did she talk to a nurse at the hospital, or to a doctor? NF (alt), given
b. Did she talk to a nurse at the hospital? BF (no alt), given
c. Did she talk to a doctor at the hospital? NF (alt), new
d. What did she tell you? BF (no alt), new

Target: I believe she said she talked to a nurse at the hospital, but I might be mistaken.

(22) **Context:** I'm not sure what to bring tomorrow night.

a. Are you making a main dish for the dinner party, or a dessert? NF (alt), given
b. Are you making a main dish for the dinner party? BF (no alt), given
c. Are you making a dessert for the dinner party? NF (alt), new
d. What did you decide? BF (no alt), new

Target: I think I decided to make a main dish for the dinner party, but I'm not really a good cook.

(23) **Context:** What is your plan for tomorrow?

a. Is your dad coming over for lunch tomorrow, or for dinner? NF (alt), given
b. Is your dad coming over for lunch tomorrow? BF (no alt), given
c. Is your dad coming over for dinner tomorrow? NF (alt), new
d. What did your dad say? BF (no alt), new

Target: I think he said he is coming over for lunch tomorrow, but I will check.

(24) **Context:** I was thinking of buying some wool for Liz.

a. Is she knitting a scarf for her granddaughter, or socks? NF (alt), given
b. Is she knitting a scarf for her granddaughter? BF (no alt), given
c. Is she knitting socks for her granddaughter? NF (alt), new
d. What did she say yesterday? BF (no alt), new

Target: I think she said she is knitting a scarf for her granddaughter, but I will ask her again.

(25) **Context:** This road has been closed for quite a while now.

a. Are they building a bridge here, or a tunnel? NF (alt), given
b. Are they building a bridge here? BF (no alt), given
c. Are they building a tunnel here? NF (alt), new
d. What do you know about the situation? BF (no alt), new

Target: I think they are building a bridge here, but they will be done very soon.

(26) **Context:** I wonder if Rachel already knows about the recent divorce in her family.

a. Did she call her aunt last week, or her uncle? NF (alt), given
b. Did she call her aunt last week? BF (no alt), given
c. Did she call her uncle last week? NF (alt), new
d. What did she tell you last night? BF (no alt), new

Target: I think she said she called her aunt last week, but I don't think she knows anything.

(27) **Context:** I'm not sure what is appropriate in this case.

a. Are you giving them money for their wedding, or a giftcard? NF (alt), given
b. Are you giving them money for their wedding? BF (no alt), given
c. Are you giving them a giftcard for their wedding? NF (alt), new
d. What do you think? BF (no alt), new

Target: I think I am giving them money for their wedding, but I might change my mind.

(28) **Context:** There was an accident on the highway.

a. Does Stephanie take the bus to school every day, or the train? NF (alt), given
b. Does Stephanie take the bus to school every day? BF (no alt), given
c. Does Stephanie take the train to school every day? NF (alt), new
d. What did Stephanie's mother say? BF (no alt), new

Target: Her mom said Stephanie takes the bus to school every day, but I'm not sure.

(29) **Context:** I haven't heard anything yet.

a. Did Dan receive a letter last month, or an email?NF (alt), given
b. Did Dan receive a letter last month? BF (no alt), given
c. Did Dan receive an email last month? NF (alt), new
d. What did Dan tell you? BF (no alt), new

Target: He told me he received a letter last month, but you should just give them a call.

(30) **Context:** I'm not sure when we should have our new furniture delivered.

a. Did you paint the walls this week, or the ceiling?NF (alt), given
b. Did you paint the walls this week? BF (no alt), given
c. Did you paint the ceiling this week? NF (alt), new
d. What did you decide? BF (no alt), new

Target: I decided to paint the walls this week, and I hope to be done with the first floor next week.

(31) **Context:** I must be going deaf!

a. Did you hear the door bell just now, or the microwave? NF (alt), given
b. Did you hear the door bell just now? BF (no alt), given
c. Did you hear the microwave just now? NF (alt), new
d. What did you say? BF (no alt), new

Target: I said I heard the door bell just now, but I might be wrong.

(32) **Context:** I'm updating the roster.

a. Did Tom choose to write a paper for this class, or to take the exam? NF (alt), given
b. Did Tom choose to write a paper for this class? BF (no alt), given
c. Did Tom choose to take the exam for this class? NF (alt), new
d. What did Tom say? BF (no alt), new

Target: I think Tom said he chose to write a paper for this class, but he could change his mind.

(33) **Context:** We should find a place to stay for next weekend.

a. Is Caroline renting a house in the city, or an apartment? NF (alt), given
b. Is Caroline renting a house in the city? BF (no alt), given
c. Is Caroline renting an apartment in the city? NF (alt), new
d. Do you remember what Caroline said? BF (no alt), new

Target: I remember Caroline said she is renting a house in the city, but we should ask her again.

(34) **Context:** We're almost done with the side dishes, but

a. did Vera cut up cucumbers for the salad, or tomatoes? NF (alt), given
b. did Vera cut up cucumbers for the salad? BF (no alt), given
c. did Vera cut up tomatoes for the salad? NF (alt), new
d. what did Vera say? BF (no alt), new

Target: I think Vera said she cut up cucumbers for the salad, although it doesn't really matter.

(35) **Context:** Wendy is not allowed to watch everything.

a. Did she watch a sitcom yesterday, or a documentary? NF (alt), given
b. Did she watch a sitcom yesterday? BF (no alt), given
c. Did she watch a documentary yesterday? NF (alt), new
d. What did she tell you? BF (no alt), new

Target: I believe she said she watched a sitcom yesterday, but I'm not sure.

(36) **Context:** Something is different here!

a. Did Saul move the table to the other side of the room, or the sofa? NF (alt), given
b. Did Saul move the table to the other side of the room? BF (no alt), given
c. Did Saul move the sofa to the other side of the room? NF (alt), new
d. What did Saul say? BF (no alt), new

Target: He said he moved the table to the other side of the room, although I'm not sure if I like it.

(37) **Context:** This company often makes the wrong decisions.

a. Did they hire a lawyer last fall, or an accountant? NF (alt), given
b. Did they hire a lawyer last fall? BF (no alt), given
c. Did they hire an accountant last fall? NF (alt), new
d. What did they announce this time? BF (no alt), new

Target: I think they announced that they hired a lawyer last fall, but I might be wrong.

(38) **Context:** What are you doing for the holidays?

a. Are you celebrating new year's with family this year, or with friends? NF (alt), given
b. Are you celebrating new year's with family this year? BF (no alt), given
c. Are you celebrating new year's with friends this year? NF (alt), new
d. What did you decide? BF (no alt), new

Target: I decided I will celebrate new year's with family this year, but I might change my mind.

(39) **Context:** I'm making the same recipe as Zara did last time.

a. Did she use basil for the sauce, or parsley? NF (alt), given
b. Did she use basil for the sauce? BF (no alt), given
c. Did she use parsley for the sauce? NF (alt), new
d. What did she say? BF (no alt), new

Target: She said she used basil for the sauce, although I could be wrong.

(40) **Context:** The police are trying to find out how the burglar got in.

a. Did Amanda close the door when it got cold, or the window? NF (alt), given
b. Did Amanda close the door when it got cold? BF (no alt), given
c. Did Amanda close the window when it got cold? NF (alt), new
d. What did Amanda tell them? BF (no alt), new

Target: I think she said she closed the door when it got cold, but she didn't lock it.

(41) **Context:** I'm making Bill's schedule right now.

a. Is he teaching on Tuesdays this quarter, or on Wednesdays? NF (alt), given
b. Is he teaching on Tuesdays this quarter? BF (no alt), given
c. Is he teaching on Wednesdays this quarter? NF (alt), new
d. What did he tell you? BF (no alt), new

Target: I believe he said he will be teaching on Tuesdays this quarter, but I'm not sure.

(42) **Context:** I don't know what the weather will be like.

a. Should I wear shorts today, or jeans? NF (alt), given
b. Should I wear shorts today? BF (no alt), given
c. Should I wear jeans today? NF (alt), new
d. What do you think? BF (no alt), new

Target: I think you should wear shorts today, but you should decide for yourself.

(43) **Context:** I wonder if we have all the ingredients already.

a. Do you still need milk for this recipe, or eggs? NF (alt), given
b. Do you still need milk for this recipe? BF (no alt), given
c. Do you still need eggs for this recipe? NF (alt), new
d. What do you think? BF (no alt), new

Target: I think we still need milk for this recipe, but I will check the fridge.

(44) **Context:** I'm trying to decide if I should make a reservation.

a. Would you like to sit in the back during the show, or in the front? NF (alt), given
b. Would you like to sit in the back during the show? BF (no alt), given
c. Would you like to sit in the front during the show? NF (alt), new
d. What do you think? BF (no alt), new

Target: I think I would like to sit in the back during the show, but you don't have to make a reservation.

(45) **Context:** What was going on?

a. Was Jack looking for his wallet in the car, or for his keys? NF (alt), given
b. Was Jack looking for his wallet in the car? BF (no alt), given
c. Was Jack looking for his keys in the car? NF (alt), new
d. What did Jack tell you? BF (no alt), new

Target: I think he said he was looking for his wallet in the car, but he didn't find anything.

(46) **Context:** We're trying to give away the leftovers.

a. Did Claire bring the roasted vegetables to the potluck, or the fruit salad? NF (alt), given
b. Did Claire bring the roasted vegetables to the potluck? BF (no alt), given
c. Did Claire bring the fruit salad to the potluck? NF (alt), new
d. What did Claire say? BF (no alt), new

Target: I believe she said she brought the roasted vegetables to the potluck, but we should ask her again.

(47) **Context:** Yesterday the jewelry store was held up.

a. Did the thief steal a bracelet from the store, or a necklace? NF (alt), given
b. Did the thief steal a bracelet from the store? BF (no alt), given
c. Did the thief steal a necklace from the store? NF (alt), new
d. What did your hear about it? BF (no alt), new

Target: I heard that they stole a bracelet from the store, and it wasn't a very expensive one.

(48) **Context:** We already did a lot of chores today!

a. Did Dean do the dishes this morning, or the laundry? NF (alt), given
b. Did Dean do the dishes this morning? BF (no alt), given
c. Did Dean do the laundry this morning? NF (alt), new
d. What did Dean tell you? BF (no alt), new

Target: He said he did the dishes this morning, but I'm not sure.

Appendix C. Foci, alternatives and primes for Experiment 4

See Table 20.

Appendix D. Results of spillover regions

D.1. Experiment 1

See Tables 21 and 22.

D.2. Experiment 2

See Tables 23 and 24.

Table 20

Item	Focus (target)	Assoc. alt	LSA targ.	Nonassoc. alt	LSA targ.	Assoc non-alt	LSA targ.
1	Swan	Duck	0.43	Fish	0.14	Nest	0.49
2	Puppy	Kitten	0.43	Dinosaur	0.04	Furry	0.44
3	Sleet	Snow	0.50	Leaves	0.09	Frozen	0.67
4	Garden	Lawn	0.41	Sidewalk	0.39	Hoe	0.56
5	Ponies	Horses	0.79	Ducks	0.06	Riding	0.70
6	Newspapers	Magazines	0.75	Cigarettes	0.06	Reporter	0.80
7	Toad	Turtle	0.59	Bee	0.10	Pond	0.49
8	Oranges	Lemons	0.67	Candle	−0.06	Sour	0.67
9	Chair	Table	0.61	Flowerpot	0.03	Dinner	0.57
10	Tv	Radio	0.70	Bike	0.01	Viewing	0.79
11	Chemistry	Biology	0.68	Sports	0.04	Transformative	0.60
12	Church	Cathedral	0.43	Post office	0.06	Priest	0.39
13	Eel	Anemone	0.37	Rock	0.14	Slimy	0.41
14	Wool	Cotton	0.67	Metal	0.02	Dyed	0.57
15	Windows	Doors	0.66	Tape	0.08	Open	0.73
16	Tulips	Roses	0.58	Card	−0.01	Pink	0.48
17	Rice	Noodle	0.43	Parchment paper	0.00	Fried	0.43
18	Muffins	Cake	0.52	Swords	0.02	Birthday	0.50
19	Hurricane	Earthquake	0.31	Kidnapping	0.08	Clouds	0.41
20	Airplane	Helicopter	0.62	Bus	0.13	Fly	0.68
21	Tomatoes	Cucumbers	0.58	Beers	0.02	Cooking	0.50
22	Jeans	Scarf	0.45	Book	0.11	Skinny	0.35
23	Sculptor	Painter	0.58	Lawyer	0.09	Statue	0.52
24	Stove	Oven	0.57	Garage	0.11	Pan	0.51
25	Flour	Milk	0.56	Sponges	−0.01	Baking	0.59
26	Sweater	Jacket	0.56	Puzzle	0.07	Wool	0.49
27	Umbrella	Raincoat	0.38	Sunscreen	0.01	Raining	0.4
28	Moon	Sun	0.28	Locket	0.03	Bright	0.39
29	Napkin	Fork	0.43	Flowers	0.04	Lap	0.36
30	Pipe	Cigarette	0.30	Incense	0.12	Lungs	0.33
31	Leash	Collar	0.42	Nail clippers	0.13	Leather	0.39
32	Pillows	Blankets	0.54	Food	0.03	Couch	0.54
33	Doctors	Nurses	0.52	Carpenters	0.02	Clinic	0.57
34	Pears	Apples	0.51	Salmon	0.05	Ripe	0.52
35	Soap	Shampoo	0.50	Vitamins	0.02	Dermatologists	0.55
36	Parsley	Thyme	0.49	Candy	0.12	Soup chef	0.53
37	Cherries	Strawberries	0.48	Toilet paper	0.04	Cakes	0.49
38	Sink	Faucet	0.48	Alarm system	0.04	Water	0.59
39	Kayaks	Canoes	0.47	Camper vans	0.06	Island	0.43
40	Eyeshadow	Lipstick	0.46	Keys	0.09	Saleslady	0.43
41	Goats	Sheep	0.46	Fish	0.04	Grass	0.47
42	Pizza	Pasta	0.44	Water	0.02	Pastry	0.49
43	Gloves	Scarves	0.44	Teapots	0.1	Bag 0	.52
44	Pines	Palms	0.43	Tulips	0.04	Trees	0.68
45	Nails	Screws	0.43	Lightbulb	−0.03	Sanding	0.39
46	Necklace	Tiara	0.32	Computer	0.02	Posh	0.28
47	Bacteria	Parasite	0.52	Hernia	0.03	Microscope	0.42
48	Apartments	Houses	0.40	Bridge	0.06	Construction	0.33

Table 21
Parameter values for fixed effects in mixed linear regression model of LogRTs on spillover regions in Experiment 1.

	Critical +1			Critical +2		
	Estimate	Std. Error	t value	Estimate	Std. Error	t value
(Intcpt)	2.8502	0.0103	275.472	2.828	0.010	278.086
Focus	0.0034	0.0064	0.532	0.002	0.006	0.373
Newness	0.0642	0.0102	**6.237**	0.036	0.008	**4.064**
Focus:New	−0.0387	0.0108	**−3.579**	−0.031	0.010	**−2.945**

Table 22
Parameter values for fixed effects in mixed linear regression model of raw RTs in Experiment 1.

	Critical +1			Critical +2		
	Estimate	Std. Error	t value	Estimate	Std. Error	t value
(Intcpt)	667.43	17.57	37.982	701.71	17.61	39.859
Focus	−21.22	12.90	−1.645	9.29	13.66	0.680
Newness	89.22	19.62	**4.548**	72.88	18.87	**3.862**
Focus:New	−56.72	20.71	**−2.738**	−65.32	22.70	**−2.878**

D.3. Experiment 3

See Tables 25 and 26.

D.4. Experiment 4

See Tables 27 and 28.

Table 23
Parameter values for fixed effects in mixed linear regression model of LogRTs on spillover regions in Experiment 2.

	Critical +1			Critical +2		
	Estimate	Std. Error	t value	Estimate	Std. Error	t value
(Intcpt)	2.8584	0.0127	224.737	2.8365	0.0118	239.129
Focus	0.0088	0.0066	1.321	0.0099	0.0073	1.354
Newness	0.0200	0.0072	**2.764**	−0.0027	0.0071	−0.387
Focus:New	0.0288	0.0113	**2.536**	0.0435	0.0106	**4.096**

Table 24
Parameter values for fixed effects in mixed linear regression model of raw RTs in Experiment 2.

	Critical +1			Critical +2		
	Estimate	Std. Error	t value	Estimate	Std. Error	t value
(Intcpt)	754.48	24.99	30.195	722.664	23.796	30.369
Focus	24.57	15.81	1.554	18.065	15.875	1.138
Newness	37.14	16.39	2.266	−9.888	16.296	−0.607
Focus:New	61.66	28.18	2.188	99.523	28.040	**3.549**

Table 25
Parameter values for fixed effects in mixed linear regression model of LogRTs on spillover regions in Experiment 1.

	Critical +1			Critical +2		
	Estimate	Std. Error	t value	Estimate	Std. Error	t value
(Intcpt)	2.8158	0.0112	249.463	2.8034	0.0140	199.983
Focus	0.0012	0.0068	0.188	0.0059	0.0066	0.890
Newness	0.0107	0.0074	1.447	0.0054	0.0064	0.839
Focus:New	0.0220	0.0087	2.514	0.0270	0.0094	2.857

Table 26
Parameter values for fixed effects in mixed linear regression model of raw RTs in Experiment 1.

	Critical +1			Critical +2		
	Estimate	Std. Error	t value	Estimate	Std. Error	t value
(Intcpt)	680.821	19.877	34.251	671.880	27.606	24.338
Focus	4.286	13.742	0.312	7.624	13.745	0.555
Newness	19.486	14.400	1.353	3.321	13.694	0.243
Focus:New	34.373	17.256	1.992	53.188	21.656	2.456

Table 27
Parameter values for fixed effects in mixed linear regression model of LogRTs on spillover regions in Experiment 4.

	Critical +1			Critical +2		
	Estimate	Std. Error	t value	Estimate	Std. Error	t value
(Intcpt)	711.363	12.302	57.824	2.851783	0.009	299.359
Association	8.701	17.369	0.501	−0.012721	0.008	−1.586
Alternative	39.313	19.846	1.981	0.046003	0.008	**5.517**
Assoc:Alt	−1.888	23.331	−0.081	0.014242	0.011	1.211

Table 28
Parameter values for fixed effects in mixed linear regression model of raw RTs on spillover regions in Experiment 4.

	Critical +1			Critical +2		
	Estimate	Std. Error	t value	Estimate	Std. Error	t value
(Intcpt)	711.363	12.302	57.824	748.97	19.85	37.722
Association	8.701	17.369	0.501	−22.47	17.77	−1.264
Alternative	39.313	19.846	1.981	92.48	19.30	**4.792**
Assoc:Alt	−1.888	23.331	−0.081	30.73	27.27	1.127

References

Baayen, R. Harald, Davidson, Douglas J., & Bates, Douglas M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, *59*(4), 390–412.

Bates, Douglas, Mächler, Martin, Bolker, Ben, & Walker, Steve (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Benatar, Ashley, & Clifton, Charles (2014). Newness, givenness and discourse updating: evidence from eye movements. *Journal of Memory and Language*, *71*(1), 1–16.

Birch, Stacy L., & Garnsey, Susan M. (1995). The effect of focus on memory for words in sentences. *Journal of Memory and Language*, *34*(2), 232–267.

Birch, Stacy L., & Rayner, Keith (1997). Linguistic focus affects eye movements during reading. *Memory & Cognition*, *25*(5), 653–660.

Birch, Stacy L., & Rayner, Keith (2010). Effects of syntactic prominence on eye movements during reading. *Memory & Cognition*, *38*(6), 740–752.

Boyce, Veronica, Futrell, Richard, & Levy, Roger P. (2020). Maze made easy: better and easier measurement of incremental processing difficulty. *Journal of Memory and Language*, *111*, 104082.

Braun, Bettina, Asano, Yuki, & Dehé, Nicole (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, Bettina, & Tagliapietra, Lara (2010). The role of contrastive intonation contours in the retrieval of contextual alternatives. *Language and Cognitive Processes*, *25*(7-9), 1024–1043.

Christensen, R. H. B. (2019). Ordinal—regression models for ordinal data. R package version 2019.12-10. https://CRAN.R-project.org/package=ordinal.

Drummond, Alex (2013). Ibex farm. *Online server: http://spellout. net/ibexfarm*.

Ferreira, Fernanda, & Lowder, Matthew W. (2016). Prediction, information structure, and good-enough language processing. *65*, In *Psychology of learning and motivation* (pp. 217–247). Elsevier.

Filik, Ruth, Paterson, Kevin B., & Liversedge, Simon P. (2009). The influence of only and even on online semantic interpretation. *Psychonomic Bulletin & Review*, *16*(4), 678–683.

Forster, Kenneth I., Guerrera, Christine, & Elliot, Lisa (2009). The maze task: measuring forced incremental sentence processing time. *Behavior research methods*, *41*(1), 163–171.

Fraundorf, Scott H., Benjamin, Aaron S., & Watson, Duane 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, Scott H., Watson, Duane G., & Benjamin, Aaron S. (2010). Recognition memory reveals just how contrastive contrastive accenting really is. *Journal of memory and language*, *63*(3), 367–386.

Gernsbacher, Morton Ann, & Jescheniak, Jörg D. (1995). Cataphoric devices in spoken discourse. *Cognitive psychology*, *29*(1), 24–58.

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

Groenendijk, Jeroen, & Stokhof, Martin (1984). Studies on the semantics of questions and the pragmatics of answers. University of Amsterdam.

Harris, Jesse A., & Carlson, Katy (2014). Focus preferences for focus-sensitive particles (and why). Poster presented at The 27th Annual CUNY Human Sentence Processing Conference. The Ohio State University. March 13-14, 2014..

Harris, Jesse A., & Carlson, Katy (2017). Association with focus for focus-sensitive particles: differences between only and even in silent reading. Talk presented at The 91st Annual Meeting of the Linguistic Society of America. Austin, TX. February 5-8, 2017..

Hintikka, Jaakko (1976). The semantics of questions and the questions of semantics: case studies in the interrelations of logic, semantics, and syntax.

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

Jackendoff, Ray S. (1972). Semantic interpretation in generative grammar. Cambridge, MA: MIT Press.

Kim, Christina S, Gunlogson, Christine, Tanenhaus, Michael K, & Runner, Jeffrey T (2015). Context-driven expectations about focus alternatives. *Cognition*, *139*, 28–49.

Landauer, Thomas K. (1999). Latent semantic analysis: a theory of the psychology of language and mind. *Discourse Processes*, *25*, 259–284.

Lowder, Matthew W., & Gordon, Peter C. (2015). Focus takes time: structural effects on reading. *Psychonomic Bulletin & Review*, *22*(6), 1733–1738.

Lowder, Matthew W, Ryan, Gwynna, Opie, Jaclyn, & Kaminsky, Emily (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.

McKoon, Gail, Ratcliff, Roger, Ward, Gregory, & Sproat, Richard (1993). Syntactic prominence effects on discourse processes. *Journal of Memory and Language*, *32*(5), 593–607.

Michel, Jean-Baptiste, Shen, Yuan Kui, Aiden, Aviva Presser, Veres, Adrian, Gray, Matthew K, Pickett, Joseph P, et al. (2011). Quantitative analysis of culture using millions of digitized books. *Science*, *331*(6014), 176–182.

Morris, Robin K., & Folk, Jocelyn R. (1998). Focus as a contextual priming mechanism in reading. *Memory & Cognition*, *26*(6), 1313–1322.

R Core Team (2021). R: A language and environment for statistical computing.

Rooth, Mats (1985). Association with focus. University of Massachusetts, Amherst.

Rooth, Mats (1992). A theory of focus interpretation. *Natural language semantics*, *1*(1), 75–116.

van Rooy, Robert (2003). Negative polarity items in questions: strength as relevance. *Journal of Semantics*, *20*, 239–273.

Schwarzschild, Roger (1999). GIVENness, AVOIDF and other constraints on the placement of accent. *Natural Language Semantics*, *7*, 141–177.

Singer, Murray (1976). Thematic structure and the integration of linguistic information. *Journal of Verbal Learning and Verbal Behavior*, *15*(5), 549–558.

Witzel, Naoko, Witzel, Jeffrey, & Forster, Kenneth (2012). Comparisons of online reading paradigms: eye tracking, moving-window, and maze. *Journal of Psycholinguistic Research*, *41*(2), 105–128.