How to compute a focus: Evidence from incremental processing

Morwenna Hoeks, Maziar Toosarvandani & Amanda Rysling*

Abstract. The global interpretation of a focus marked sentence with a particle like *only* arises due to the interplay of several formal components: F-marking, the semantics of the particle, the nature of contrastive alternatives, and a dependence on context. In this paper, we argue that on-line reading measures can be used to probe which of these components are computed when. In order to isolate which components give rise to reading slowdowns typically observed on foci (Birch & Rayner 1997, Benatar & Clifton 2014, Lowder & Gordon 2015, Hoeks et al. 2023), a Maze reading study tested whether such slowdowns still arise on second-occurrence foci (SOF)—foci whose inferences have already been computed in prior discourse and are therefore entirely predictable—and on foci whose size and location can only be determined via a previously introduced contrast. Results indeed showed slowdowns on such foci, suggesting that these cannot solely be attributed to readers computing focal inferences anew, nor to comprehenders initiating their reasoning about the relevant alternatives.

Keywords. focus; alternatives; Maze task; reading; second-occurrence focus; predictability

1. Introduction. This paper is about the interpretation of strings like (1), which contain the focus-sensitive particle *only* and give rise to different *exhaustivity inferences*, depending on the size of its associated focus.

(1) Sarah only read a book about BATS.
  a. Sarah only read a book about $[\text{BATS}]_F$. ⇝ She did not read a book about anything else.
  b. Sarah only read $[\text{a book about BATS}]_F$. ⇝ She did not read anything else.
  c. Sarah only $[\text{read a book about BATS}]_F$. ⇝ She did not do anything else.

We can think of the interpretation of such a string in at least two ways. The first is in terms of the inferences that it gives rise to from a global perspective, considering the entire sentence as a whole, as is done in theoretical semantics; the second is in terms of the string's word-by-word interpretation, and how those inferences are deduced by a comprehender in real-time.

The goal of this paper is to connect these two perspectives in a more direct way than has been done, by using incremental reading measures to probe, at a relatively fine-grained level, what aspects of a focus meaning is computed when. We present the results of a reading study, and argue that by investigating slowdowns that arise during reading of material in carefully controlled contexts, the incremental interpretation of focus can be linked up with the formal representations that describe its global meaning.

These formal representations have several distinct components. Therefore, what it means to understand the meaning of any of the sentences in (1a–c) is to know, minimally, the information

*Authors: Morwenna Hoeks, University of Osnabrück (morwenna.hoeks@uni-osnabrueck.de), Maziar Toosarvandani, University of California Santa Cruz (mtoosarv@ucsc.edu) & Amanda Rysling, University of California Santa Cruz (rysling@ucsc.edu)

corresponding to each of these components. Taking seriously the fact that the offline inferences triggered by focus arise due to an interplay between different pieces of formal machinery also suggests that the real-time interpretation of foci may be decomposable into several distinct subprocesses. Below, we outline what these formal components are, before discussing how the incremental computation of their corresponding subprocesses can be probed using on-line measures.

At the same time, we will use these subcomponents to structure a discussion of the preceding psycholinguistic literature, which has shown—using a variety of tasks and methods—that the presence of focus marking leads to a number of distinct behavioral effects. By conceptualizing the processing of focus in terms of these formal subcomponents, we can associate these behavioral effects with distinct subprocesses in real-time comprehension.

While this will engender many questions, we will concentrate on the interaction between these subprocesses and the preceding discourse context. We will suggest that investigating the processing of *second-occurrence focus (SOF)*—a focus for which some of the subcomponents have already been computed in prior discourse—uniquely sheds light on this issue, because it allows us to isolate those subcomponents responsible for structural predictions about focus from those subcomponents that give rise to the focus-related inferences and integration with the preceding context.

In a Maze reading study, we test reading of sentences with an SOF in discourse contexts that specify the alternative sets before the interpretation of those foci. Results from this experiment suggest that effects of focus still arise in reading measures when some aspects of a focus' meaning have already been computed in prior discourse, and when accidental properties of focus are accounted for.

2. Decomposing the computation of focus inferences. Since natural language is not used in isolation, we consider (2), which adds a preceding discourse for the example in (1a).

(2) Lily read a book about whales and penguins, but Sarah only read a book about [BATS]$_F$.
$\rightsquigarrow$ Sarah didn't read a book about whales or penguins.

Theoretical accounts of focus marking, designed to capture the exhaustivity inference arising in (2), typically involve formal representations that include at least four distinct components:

(i) These accounts, first, include some abstract representation of focus, or *F-marking* as we will refer to it here, which is often understood as a feature that is part of the syntactic structure of a sentence (Jackendoff 1972). In (2), for example, F-marking is placed solely on the constituent *bats* (indicated with a subscript F), and is ultimately responsible for both the phonological and interpretational effects of focus marking.
(ii) The exhaustivity inferences arise because the presence of F-marking evokes a set of alternatives that contrast with the focus (Jacobs 1983, Rooth 1985, 1992) on *bats*, as in (2); the contrastive alternatives would be computed, roughly, by replacing *bats* within that sentence with other expressions that contrast with it.
(iii) These evoked alternatives affect the basic meaning of a sentence as they interact with focus-sensitive expressions inside that sentence. For *only*, the evoked focus alternatives are negated, leading to an exhaustivity inference.
(iv) Finally, representations of the discourse context are also assumed to play a role in the interpretation of focus marked sentences: In Rooth's (1992) Alternative Semantics, for instance,

this effect is captured by requiring the alternatives evoked by a focus to be a subset of a set of alternatives made salient by the context. This contextual restriction leads, in the case of (2) specifically, to the inference that *whales* and *penguins* are among the contextually-relevant alternatives to the focus, and therefore among the things that Sarah did not read a book about.

In terms of a human comprehender interpreting the above sentence word-by-word in real time, there is a sense in which the pieces of information in (i-iv) may not be computed all at once. It could be, for instance, that in some cases the location of F-marking is determined prior to reasoning about the evoked alternatives, while in other cases the specification of an alternative set in the context may guide a reader's assumptions about the exact location of an upcoming focus.

This latter scenario may in fact be what happens when (2) is interpreted incrementally. As shown in (1a-c), the second clause of this sentence is in principle compatible with multiple focus structures. Because a focal accent is only placed on the last word of this sentence, this string is ambiguous with respect to size of the F-marked constituent even in listening where comprehenders have full access to the prosodic signature. But for (2), the comprehender may still conclude that only *bats* is F-marked, because the preceding context sets up a contrast between *whales*, *penguins* and *bats*, signalling that the relevant alternative set consists of alternatives to those expressions (and not alternatives to larger constituents). Here, information about the alternative set may thus be deduced from context before the F-marked phrase itself is encountered.

Assuming that interpreting focus may be decomposed into several non-contemporaneous subprocesses, as exemplified above, may also suggest that different aspects of incremental focus comprehension have their own distinct behavioral signatures, or that the exact nature of the effects observed on foci can depend on the information that is available before a focus is encountered. Behaviorally, it has been shown that foci are more deeply encoded in memory than non-foci (Singer 1976, McKoon et al. 1993, Birch & Garnsey 1995, Gernsbacher & Jescheniak 1995), and that focus marking leads to more accurate responses in change- and error-detection tasks (Bredart & Modolo 1988), as well as longer reading (Birch & Rayner 1997, Benatar & Clifton 2014, Lowder & Gordon 2015) and response times (Hoeks et al. 2023). The comprehension of focused material has therefore been argued to require more processing resources than that of non-focused material. But in light of the above, the exact reasons such effects arise are not fully understood, as it still remains unclear what specific aspect(s) of focus comprehension are driving these observed effects.

Hoeks et al. (2023), for instance, have suggested that the reading slowdowns observed on foci can at least in part be explained by comprehenders having to infer what the contextually-relevant alternative set is. Contrasting the way in which target sentences like (3) are read in multiple different contexts, they showed that slowdowns on focused material (*apple pie*) are diminished in contexts that specify the relevant alternatives to a focus, as in (3b), compared to cases where no alternatives to the focus were mentioned in the preceding context, as in (3a).

(3) a. **Context A:** Did Sarah want apple pie for dessert? NO ALT
b. **Context B:** Did Sarah want apple pie for dessert, or chocolate cake? ALT

Target: Sarah said it was [apple pie]$_F$ that she wanted for dessert

In other words, readers spend more time interpreting a focus in the absence of such contextually specified alternatives, suggesting that perhaps part of the additional time it takes comprehenders to read a focus is spent on inferring what the contextually relevant alternatives may be—component (iv) above. However, since their studies also found suggestive evidence that focus slowdowns do not entirely disappear in the presence of explicit contextual alternatives, it may be that other aspects of interpreting focus, like deducing the location of F-marking (i) or incorporating the alternatives with the sentence meaning (iii), also play a role. The observed reading slowdowns on foci are likely complex and may be driven by several aspects of focus comprehension at once.

Besides the additional inferences that need to be computed for focus, other (non-mutually exclusive) explanations for the observed focus slowdowns have been pointed out elsewhere in the psycholinguistic literature on focus. It has been suggested, for instance, that focus causes readers to slow down because foci are generally less predictable than non-foci, as they often constitute new information (Benatar & Clifton 2014). In general, readers may spend more time on material that is less predictable from their context, including some foci. But although discourse-new foci were indeed found to be read more slowly than foci that had already been mentioned in the preceding context, Hoeks et al. (2023) also found that focus slowdowns still arose for foci that were themselves discourse-given, like the ones in (3). This suggests that these slowdowns can only in part be attributed to differences in newness/givenness.

It may still be, however, that focus slows down reading because foci are unpredictable in some other way. For instance, even though *apple pie* had already been mentioned in both contexts preceding (3), this was also the phrase that was explicitly asked about in those preceding questions. The answer—i.e., what it is that Sarah wanted for dessert—had not been established in the common ground up to that point by either of those contexts. The focused phrase may therefore have been less predictable than other material in this sense, not because it had not been mentioned before, but perhaps because that phrase resolves a thus far open Question-under-Discussion (QUD).

It has also been proposed that (some aspect of) the interpretation of focus marked material may generally be prioritized by the comprehension system (Cutler & Fodor 1979, Morris & Folk 1998). It may be that readers slow down on focused material, not only because it takes additional effort and time to compute the relevant inferences, but also because comprehenders strategically allocate more attention and processing resources to interpret them. Perhaps material put in focus generally carries more important information, exactly because such material often answers an open QUD, as in (3). Such an account could explain the observation that foci are better remembered and attended to as well (Singer 1976, Cutler & Fodor 1979, Bredart & Modolo 1988, McKoon et al. 1993, Birch & Garnsey 1995, Gernsbacher & Jescheniak 1995).

Finally, it may be that the fact that foci typically receive contrastive accents also plays a role in these focus slowdowns. It may be, for instance, that the human perceptual system is particularly attuned to allocating attention to focal accents and that this affects the general processing of language. Such an explanation could also account for the reading slowdowns on focus, because it may be that the implicit prosodic structure that is assigned during silent reading could be used to differentially allocate resources (Breen 2014)—resulting in longer reading times on material that is predicted to be implicitly accented as well (see e.g., Lowder & Gordon 2015 for an account that would be consistent with this hypothesis).

In short, the behavioral effects observed on most foci could be attributed to any or all of these potential aspects of focus comprehension: Such effects may be due to readers not expecting a focus, or to the content of the focus somehow being new or unpredictable, or providing novel information in some other way; they may be due to the fact that a reasoning process about the contrastive alternatives to the focus is set in motion, or to the fact that a context in which such alternatives are relevant has to be accommodated; they could also be due to the fact that an (implicit) accent was present on those foci, and/or to the fact that more processing resources are being allocated to the focus, once this focus has been identified. In the basic scenario that has been tested thus far in the psycholinguistic literature, focus marking is typically unambiguously signaled by a particle, cleft or by the auditory signal, sometimes even in the absence of any preceding context. In those cases, it is hard to distinguish between all the possible processes related to the interpretation of foci, because they may all be triggered at once, the moment the focus is recognized as such.

The goal of the present experiment is to eliminate some of these explanations. What will be crucial in motivating the design of the experiment is the observation that some aspects of focus processing may be set in motion before the focus is encountered, as we saw for (2). Unlike previous work, the present experiment involved sentences that would be ambiguous with respect to the focus structure they contained in isolation, but for which context disambiguated the location of F-marking. This allows us to disentangle effects of focus from effects of focus prosody because the location of F-marking can be manipulated via context while keeping the prosody of that sentence constant. It also allows us to test how the presence of contrastive expressions in the preceding context affects reading of a subsequent focus. Specifically, we hypothesize that if such preceding contexts guide readers in assigning focus marking further downstream, reasoning about the alternative sets must already be set in motion as the reader is first incrementally proceeding through the sentence, meaning that, on the focus itself, no preceding context would have to be accommodated in order to interpret it. This predicts the resulting focus slowdowns to either disappear altogether (showing that focus slowdowns are solely due to the novel computation of such alternatives), or to be diminished (in which case remaining slowdowns must index a cost associated with other components of focus processing). Finally, we also compare novel foci with foci that are fully predictable within their discourse context, in order to determine to what extent focus slowdowns are driven by their relative unpredictability. We further motivate the design of this experiment next.

3. Using second-occurrence focus to decompose focus slowdowns. This experiment used sentences with second-occurrence foci (SOF), as in (4b), to probe the incremental processing of foci.

(4) a. Sarah read a book about whales and penguins, and Bob only read a book about [BATS]$_{F1}$
b. No, LILY$_{F2}$ only read a book about [bats]$_{F1}$

These sentences have several properties which make them well-suited for current purposes. First discussed by Partee (1999), they have since risen to fame mainly because only the focus on *Lily* is marked by a pitch accent, while the focus on *bats* that is associated with the particle *only* is not. It must still be that the associate of *only* is underlyingly F-marked, because it must evoke a set of alternatives that is operated on by this particle. After all, (4b) still entails that Lily did not read books about alternative topics besides bats. Most theoretical accounts of such foci explain this mismatch between interpretation and prosodic realization by making reference to the fact that these foci have typically been mentioned as regular foci before their second occurrence as a focus, as (4a) illustrates (Selkirk 2008, Rooth 2010, Beaver & Velleman 2011). These two properties of

SOF (non-accenting and givenness) allow us to measure effects of focus marking that cannot be due to material being either new or marked with a pitch accent.

In fact, unlike the foci tested by Hoeks et al. (2023), SOF do not provide novel answers to a preceding question, nor do they provide novel information in any other way. The utterance that (4b) responds to already states that someone only read a book about bats, and it therefore does not just make the entire phrase *only read a book about bats* given, it also makes it already entailed by the discourse context. SOF examples like these are therefore a good test case to study whether previously observed slowdowns on foci are due to them being unpredictable given the preceding discourse, or whether they are due to other aspects of focus comprehension.

As outlined above, perhaps readers slow down on foci not only because they are updating their representation of the discourse context with incoming information provided by the focus itself, but also because they have to compute what the contextually relevant alternatives to a focus are. Crucially, for SOF it is not just the case that the focus itself is entailed, but also that the evoked alternatives, as well as the the relevant focal inferences, have already been computed during the interpretation of the previous sentence. Already in (4a) it is the case that the particle *only* operates over the set of alternatives, negating alternative statements about types of things that could have been read about instead of bats. If the behavioral effects observed on foci are due to the novel computation of such inferences, we would expect such effects to be diminished for SOF. Comparing reading times on new foci with those on SOF gives us an indication of how much such slowdowns are driven by the novel computation of focal inferences as well.

Moreover, these sentences crucially contain multiple foci which each serve to signal a distinct contrast, and are each associated with their own operator. Such sentences thus require there to be multiple distinct sets of salient contextual alternatives. Complex cases like these could potentially reveal how comprehenders deal with multiple such alternative sets. Again, a comparison between new foci and SOF, in particular, may tell us whether comprehenders are able to encode and maintain multiple alternative sets during their incremental assignment of focus structure, or whether it is only the alternative sets that have not been computed before that drive the observed slowdowns.

Because SOF are fully predictable themselves, do not carry a focal accent, and do not introduce any novel inferences, it may very well be that slowdowns on these foci entirely disappear. Since we test the reading strings for which preceding contexts disambiguate F-marking, the alternatives that are evoked by these foci must also be fully known at the point in time where these foci are encountered. Even if focus slowdowns are due to the general prioritization of focused material, it is conceivable that such slowdowns disappear for SOF, because such foci may not constitute important information. In fact, the comprehension of SOF may be de-prioritized, as they occur inside the background of another focus that may receive higher priority instead. Crucially, however, results of this experiment in fact found a consistent slowdown, even for SOF, thus ruling out any explanation for the general focus slowdown in terms of newness, unpredictability, calculation of novel inferences, or accenting. Next, we describe the design of this experiment in more detail.

3.1. METHOD. Previous work has established that focus marking causes readers to slow down, and the goal of the experiment presented here was to further investigate the exact source of such focus slowdowns. The design of the experiment did this in several ways: because the string of words of the target sentence, on its own, would be ambiguous with respect to the assignment of

F-marking, this experiment first tested to what extent salient sets of alternatives in the preceding context can guide comprehenders in their assignment of focus on subsequent material. Moreover, by comparing reading times on second-occurrence foci with that of new foci, the experiment tested whether such slowdowns still arise even if the material that is being read is already entailed by the context and is entirely recoverable and predictable from it. Finally, it also aimed to test whether focus slowdowns still occur on material that would not be accented if pronounced out loud.

3.1.1. MATERIALS. Every item constituted a dialogue between two speakers, *Speaker A* and *Speaker B*, where the utterance by *Speaker B* was considered the target sentence and the utterance of *Speaker A* served as the context against which the target sentence was interpreted. The target sentence, uttered by *Speaker B*, always contained the focus particle *only* which was placed inside that sentence in a position that was ambiguous in terms of the associate it takes. The preceding context sentence always consisted of two clauses in which two alternatives were contrasted with each other, and the Size of the focus in the target sentence was therefore manipulated by the size of these alternatives, such that in the NARROW conditions only single nouns were contrasted while in the WIDE conditions complex noun phrases consisting of two nouns were contrasted with each other. Orthogonal to this manipulation of focus Size, the utterance of Speaker A also manipulated focus Type since it determined whether the focused phrase uttered by Speaker B was either NEW or second-occurrence (SOF). An example item in all four conditions is shown in (5).

(5) a. **Speaker A:** Abby read a book about penguins and whales,
and Bob read a book about [gorillas]$_F$
Speaker B: And Lily only read a book about [bats]$_F$ NARROW NEW

b. **Speaker A:** Abby read an article about penguins and a report on whales,
and Bob read [an article about gorillas]$_F$
Speaker B: And Lily only read [a book about bats]$_F$ WIDE NEW

c. **Speaker A:** Abby read a book about penguins and whales,
and Bob only read a book about [bats]$_F$
Speaker B: No, [Lily]$_F$ only read a book about [bats]$_F$ NARROW SOF

d. **Speaker A:** Abby read an article about penguins and a report on whales,
and Bob only read [a book about bats]$_F$
Speaker B: No, [Lily]$_F$ only read [a book about bats]$_F$ WIDE SOF

In all conditions, the first object noun in the target sentence (*book*) constituted the critical region of interest, because—if focus structure in the context affected focus structure on the target sentence—it is this region that would either be focused (in the WIDE conditions) or not (in the NARROW conditions). Since the material inside the target sentence was held constant across conditions, response time differences in this region between the WIDE and NARROW conditions could only indicate a difference in the projected focus structure of this sentence. If an effect of size can be observed on this first NP, this means that readers indeed must be able to use previously specified contrasts to guide their assignment of focus marking. If a slowdown can be found in the WIDE SOF relative to the NARROW SOF conditions on this region, this would indicate a slowdown due to focus marking, and would thus indicate that such slowdowns cannot solely be explained in terms of their unpredictability, accenting, or the accommodation of alternatives in the context.

In total, 48 items like (5) were constructed, each with the four conditions illustrated above. Another 64 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.

3.1.2. PROCEDURE & PARTICIPANTS. Target sentences were presented using the Maze task, while context sentences were presented normally. 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 et al. (2020). An example of the AutoMaze output for one target sentence is given in (6) below. On the second line, the distractor word is presented below its corresponding word of the target sentence.

(6) No, Lily only read a book about bats, but I might be misremembering it.
x-x-x Came fine call ew been trump must, jack hill glass laws hypothyroidism am.

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.

To ensure careful reading of the context, all experimental trials were followed by a comprehension question that probed various properties of the context preceding the target sentence. For instance, the example item in (5) was followed by the comprehension question in (7).

(7) Did Speaker A mention books about sharks?

When participants chose the wrong Maze word (the foil), they automatically exited the Maze trial and were sent immediately to the comprehension question. 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 Maze-sentence and a comprehension question. After the short practice phase, the experimental items were presented along with the fillers in a pseudo-random order.

55 native speakers of English were recruited via Prolific and compensated at a $12 hourly rate. Data from 48 participants were included in the analysis; 7 participants were excluded because they failed to complete more than 70% of the Maze sentences.

3.2. RESULTS. The mean comprehension question accuracy was 79%, and the mean completion rate of the Maze target sentences of Experiment 1 was 89%. Mean response times per condition for each region are given in Table 1 and plotted with 95% confidence intervals in Figure 1.

Data were analyzed using R, version 3.6.3 (R Core Team 2021). Bayesian (generalized) linear mixed-effect models were fit using Stan, as implemented in the brms package, version 2.18.0 (Bürkner 2017), with the default priors. Separate models were fit to log-transformed response

Condition	Subject	Particle	Verb	NP1	NP2
NARROW NEW	834.10 (23.74)	803.18 (14.11)	696.93 (11.48)	789.48 (13.22)	954.33 (15.80)
WIDE NEW	810.50 (17.55)	844.61 (17.10)	728.04 (16.33)	936.04 (17.26)	939.57 (20.26)
NARROW SOF	971.72 (31.47)	753.85 (14.45)	704.20 (12.08)	774.20 (14.89)	862.86 (16.41)
WIDE SOF	947.68 (25.50)	771.36 (14.67)	714.24 (12.70)	859.30 (17.59)	886.34 (16.95)

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.

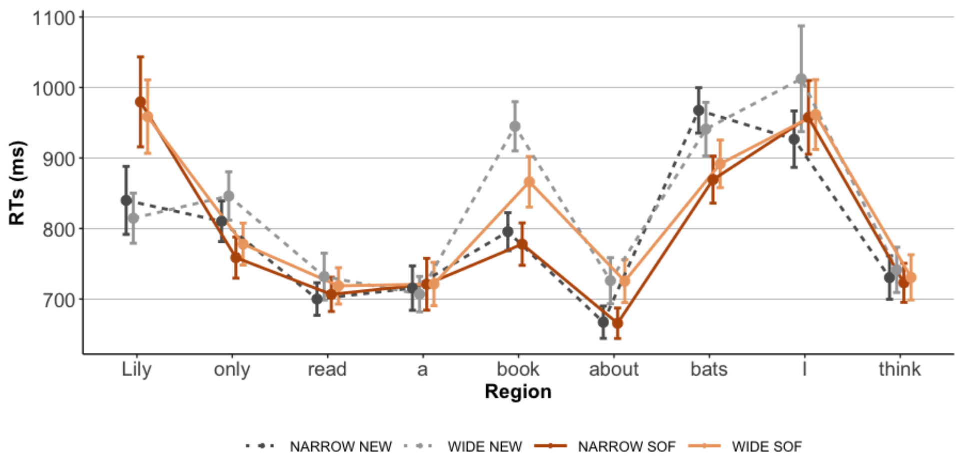

Figure 1: Experiment 1: mean RT in each region in each condition. Error bars represent the 95% confidence interval.

times and untransformed response times as dependent measures. Models included population-level effects of focus and newness (coded 0.5, -0.5), with NARROW focus and NEW conditions treated as reference levels, and random slopes and intercepts for both subjects and items (Baayen et al. 2008). For each model, we ran four chains, each with 5000 steps (warmup = 1000 steps). Rhat statistics in all models approached 1.00 and no warnings emerged. Table 2 and 3 present the posterior estimates of the fixed effects as well as 95% credible intervals for both models of Experiment 1. Below, reliable effects on each region will be discussed. No reliable effects were found at either the verb region, or the spillover regions of NP2.

On the subject (*Lily*), positive estimates for focus Type indicate that subjects in the SOF conditions were read reliably slower than subjects in the NEW conditions. Contrastive focus on the SOF subjects may thus have induced a slowdown relative to NEW conditions in which subjects were also new but were not contrastively focused. On (*only*), positive estimates for Type indicate that this particle was read faster in the SOF conditions than in the NEW focus conditions. This may be expected as this particle was new in the NEW but not in the SOF conditions.

On the first NP (*book*), positive estimates for focus Size indicate that this noun was read faster in the WIDE conditions than in the NARROW conditions, and positive estimates for focus Type indicate that SOF foci were read faster than NEW foci. There was also a reliable interaction between focus Size and Type, though the focus Size effect was reliable among both the SOF (β =71.89,

	Est.	Error	95%CrI	Rhat
Intcpt	2.90	.01	[2.87,2.92]	1.00
Size	0.03	.01	[0.02,0.04]	1.00
Type	0.05	.01	[0.03,0.07]	1.00
Si:Ty	0.04	.01	[0.01,0.06]	1.00

Table 2: Posterior estimates for the population-level effects of logRTs on NP1.

	Est.	Error	95%CrI	Rhat
Intcpt	884.00	38.94	[807, 962]	1.00
Size	51.90	21.26	[9.50,94.1]	1.00
Type	120.00	28.62	[64.1,175]	1.00
Si:Ty	96.22	44.93	[7.94,186]	1.00

Table 3: Posterior estimates for the population-level effects of raw RTs on NP1.

95%CrI=[3.51, 140.59]) and the NEW conditions (β=168.11, 95%CrI=[94.73,242.04]). Positive estimates for Size on the NP1 spillover region again indicate that this phrase was read faster in the NARROW than in the WIDE conditions. However, the estimated credible interval for the main effect of focus Type as well as the interaction overlapped with zero so these will not be considered reliable.

Finally, on the second NP (*bats*), positive estimates for focus Type again indicate that NP2 was read slower in the SOF than in the NEW conditions, but estimates for Size and the interaction were not reliable. This may be unsurprising because the second NP was focused in both WIDE and NARROW conditions, and therefore the only difference between the conditions was their newness.

4. Discussion. By carefully manipulating the context preceding a focus, the experiment presented here was designed to disentangle various explanations for the reading slowdowns typically observed on foci, with the goal to better understand what aspects of a focus' meaning are computed when. To test whether focus slowdowns still arise on second-occurrence foci and in cases where their alternative sets were already fully determined in the prior context, the experiment crossed the type of focus (NEW or SOF) with the size of that focus inside a target sentence.

Reading times were indeed found to be affected by the manipulation in focus Size: Slowdowns were found on NPs that were put in focus in the WIDE focus condition, relative to the NARROW focus condition in which that NP was not F-marked. Slowdowns were thus found at the left edges of wide foci, indicating that the source of these slowdowns cannot lie in the anticipation of a focal accent, which would only be assigned at their right edge, if assigned at all. Since the size of these foci could only be determined via a previously specified contrast, these results also confirmed that contextual alternatives can guide readers' projection of F-marking onto subsequent material.

The results of this experiment revealed an effect of focus Size even in the SOF conditions. This slowdown allows us to draw several conclusions about the processing of focus: It first indicates that focus slowdowns cannot be explained solely in terms of predictability, because these SOF were fully recoverable from their context and did not present any new information at the point where they were encountered. It also indicates that slowdowns can be found even for foci that had already occurred *as* foci, and whose corresponding inferences had already been computed. The presence of such slowdowns, although diminished compared to focus slowdowns in the NEW conditions, suggests that these focus slowdowns cannot solely be attributed to a cost associated with computing focal inferences anew, nor can they be explained by comprehenders starting to set a process in motion that involves reasoning about these alternatives. After all, if readers had not already encoded some information about these alternatives prior to the occurrence of the SOF, they would not have been able to tell whether these phrases themselves were F-marked at all.

One possibility is that the SOF slowdown on NP1 does not indicate the assignment of focus marking, but instead indicates that readers slowed down because they expected new material in those regions even though those regions were in fact always given. But even in that case, some explanation would be needed for the Size effect: Since the only difference between the NARROW SOF and WIDE SOF conditions was the size of the alternatives that were mentioned in the preceding context, all else being equal, reading times in these conditions could only have been affected by this contextual manipulation. Even if readers slowed down because the target sentence was in some way different than expected, such expectations would still have to have been formed based on the alternatives mentioned in the context. Any explanation of this effect would need to make reference to this contextual manipulation of alternatives, which would ultimately entail that some representation involving these alternatives guides readers' downstream behavior.

There are, as far as we can tell, only three possible explanations for the SOF slowdowns that remain. First, it may be readers slow down on such foci because computing their inferences is a relatively automatic process triggered by the presence of a particle signaling an upcoming focus—i.e., a process triggered even if the current sentence is entirely parallel to a preceding sentence for which such inferences have just been computed. Alternatively, it could be that such slowdowns indicate additional care and attention for the processing of foci, which would have to hold, again, even for foci that have just been interpreted. Note that such prioritization itself therefore cannot be explained in terms of a general prioritization of material that is new or provides a novel answer to an (implicit) question. Finally, it may be that these SOF slowdowns arose because there is some level of representation at which events and their participants are being linked together, and that the contrastive focus on a subject like *Lily* results in an over-writing process where subsequent material describing a specific type of event is actively being linked to that new individual, after it has been de-linked from the previously encoded agent of that event (*Bob*).

However, whatever the exact characterization of this cost may be, it is clear from these results that the cost itself must be sensitive to the preceding discourse that sets up a particular alternative set: In all three cases, readers must still be sensitive to the particular size of the contrast involved. This suggests that comprehenders must somehow encode a representation of that discourse context which goes beyond the linear organization of the preceding sentences, e.g., involving the kind of representation adopted in a QUD-based framework—one which would represent, for the examples involved in our experiment, an overarching QUD asking either *who read what* or *who read a book about what topic*. Finally, what these results moreover suggests, at a more general level, is that the interpretation of focus has consequences for the comprehension system that cannot be reduced to some accidental property of focus, like accenting, newness, or predictability. They suggest, instead, that the very presence of focus marking may be driving these effects, even if that focus does not contain any new information or introduce any novel inferences.

References

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

390–412. https://doi.org/10.1016/j.jml.2007.12.005.

Beaver, David & Dan Velleman. 2011. The communicative significance of primary and secondary accents. *Lingua* 121(11). 1671–1692.

Benatar, Ashley & Charles Clifton. 2014. Newness, givenness and discourse updating: Evidence from eye movements. *Journal of Memory and Language* 71(1). 1–16. https://doi.org/10.1016/j.jml.2013.10.003.

Birch, Stacy L. & Susan M. Garnsey. 1995. The effect of focus on memory for words in sentences. *Journal of Memory and Language* 34(2). 232–267. https://doi.org/10.1006/jmla.1995.1011.

Birch, Stacy L. & Keith Rayner. 1997. Linguistic focus affects eye movements during reading. *Memory & Cognition* 25(5). 653–660. https://doi.org/10.3758/BF03211306.

Boyce, Veronica, Richard Futrell & Roger P. Levy. 2020. Maze made easy: Better and easier measurement of incremental processing difficulty. *Journal of Memory and Language* 111. 104082. https://doi.org/10.1016/j.jml.2019.104082.

Bredart, Serge & Karin Modolo. 1988. Moses strikes again: Focalization effect on a semantic illusion. *Acta Psychologica* 67(2). 135–144. https://doi.org/10.1016/0001-6918(88)90009-1.

Breen, Mara. 2014. Empirical investigations of the role of implicit prosody in sentence processing. *Language and Linguistics Compass* 8(2). 37–50. https://doi.org/10.1111/lnc3.12061.

Bürkner, Paul-Christian. 2017. brms: An r package for bayesian multilevel models using stan. *Journal of statistical software* 80. 1–28. 10.18637/jss.v080.i01.

Cutler, Anne & Jerry A Fodor. 1979. Semantic focus and sentence comprehension. *Cognition* 7(1). 49–59. https://doi.org/10.1016/0010-0277(79)90010-6.

Gernsbacher, Morton Ann & Jörg D. Jescheniak. 1995. Cataphoric devices in spoken discourse. *Cognitive psychology* 29(1). 24–58. https://doi.org/10.1006/cogp.1995.1011.

Hoeks, Morwenna, Maziar Toosarvandani & Amanda Rysling. 2023. Processing of linguistic focus depends on contrastive alternatives. *Journal of Memory and Language* 132. 104444. https://doi.org/10.1016/j.jml.2023.104444.

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. Cambr., MA: MIT Press.

Jacobs, Joachim. 1983. *Fokus und Skalen: Zur Syntax und Semantik der Gradpartikeln im heutigen Deutsch*. Tübingen: Niemeyer.

Lowder, Matthew & Peter Gordon. 2015. Focus takes time: structural effects on reading. *Psychonomic Bulletin & Review* 22(6). 1733–1738. https://doi.org/10.3758/s13423-015-0843-2.

McKoon, Gail, Roger Ratcliff, Gregory Ward & Richard Sproat. 1993. Syntactic prominence effects on discourse processes. *Journal of Memory and Language* 32(5). 593–607. https://doi.org/10.1006/jmla.1993.1004.

Morris, Robin K. & Jocelyn R. Folk. 1998. Focus as a contextual priming mechanism in reading. *Memory & Cognition* 26(6). 1313–1322. https://doi.org/10.3758/BF03201203.

Partee, Barbara H. 1999. Focus, quantification, and semantics-pragmatics issues. In: Peter Bosch & Rob van der Sandt (eds.), *Focus: Linguistic, cognitive, and computational perspectives*, 213–231. Cambridge: Cambridge University Press.

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

Rooth, Mats. 1985. *Association with focus*. Amherst, MA: UMass Amherst dissertation.

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

https://doi.org/10.1007/BF02342617.
Rooth, Mats. 2010. Second Occurrence Focus and Relativized Stress F. In: Caroline Fery & Malte Zimmermann (eds.), *Information Stucture: Theoretical, Typological, and Experimental Approaches*, 15–35. Cambridge: Cambridge University Press.
Selkirk, Elisabeth. 2008. Contrastive focus, givenness, and the unmarked status of "discourse-new". *Acta Linguistica Hungarica* 55(3–4). 1–16.
Singer, Murray. 1976. Thematic structure and the integration of linguistic information. *Journal of Verbal Learning and Verbal Behavior* 15(5). 549–558.