The Influence of Focus Alternative Sets on Memory for Text

WIND COWLES    MICHELLE PERDOMO
University of Florida
cowles@ufl.edu    mperdomo@ufl.edu

Abstract: This paper presents a psycholinguistic experiment testing the hypothesis that focus alternatives are activated during focus processing and remain associated with memory representations of focus referents. Research participants read short passages that used wh-questions and it-cleft structures to mark narrow focus and introduced small, two-entity contrast sets. Participants then judged whether a paraphrase of the sentence containing the focus-marked referent was accurate. Overall accuracy for this judgment was high (91%) when the paraphrase was true or when the focus referent was substituted with a referent that was not part of the contrast set. Accuracy was significantly lower (77.5%) when the focus was substituted with the alternative from the contrast set. These results support theoretical and processing proposals that focus interpretation involves processing of contrast alternatives, and further shows these alternatives may not be effectively suppressed in memory for focus.

1. Introduction
Several theoretical and processing accounts of informational focus have proposed that focus interpretation necessarily involves consideration of focus alternatives, that is – the activation of referents that could have been, but are not, given focus status (Cowles, Fionda, & Perdomo, 2012; Féry & Krifka, 2008; Fraundorf, Watson, & Benjamin, 2010; Husband & Ferreira, 2012; Rooth, 1992). In this paper we present an experiment that tests the hypothesis that these focus alternatives become part of the memory representation of focus and interfere with memory for focus referents by presenting readers with short passages in which focus status is conferred by using embedded wh-questions and it-clefts (e.g. Rochemont, 1998).

2. Background
Across different approaches to information structure there is a broad consensus that the focus in a sentence marks or reflects new or pragmatically unrecoverable information. Cross-linguistically, focus is associated with particular syntactic structures, pitch prominence and fuller forms of reference (e.g. Lambrecht, 1994; Selkirk, 1984). While there does not appear to be any one unique cue to focus status (Hedberg and Sosa 2007; Gundel and Fretheim 2001), answers to wh-questions (e.g. Rochemont, 1998) and it-clefts appear to provide reliable indicators of focus status. See examples (1) and (2), in which the focus constituent is underlined.
(1)  -What did the rabbits eat in your garden?
-They ate the lettuce.

(2)  It was the dog that chased the rabbit out of the garden.

In (1), the lettuce corresponds to the wh-word what in the preceding question, and in natural conversation would be likely to receive pitch prominence (they ate the lettuce). In this case, lettuce is the new, previously unknown information. In (2), the dog is in a clefted position, and is also considered to have focus status. In both of these cases, one may consider these focus referents against the possible alternatives – other things rabbits eat in gardens and other things that may chase rabbits (Féry & Krifka, 2008; Rooth, 1992).

In psycholinguistic research on focus, there was been an emphasis on understanding the consequences of focus status on sentence processing – does focus status cause greater attention to be placed on the focused constituent? Does it cause better memory for the focused referent? Not all of the data have been consistent; while there is evidence that focus does cause better encoding of focused information in memory, leading to better recall of focus information (Birch & Garnsey, 1995; Foraker & McElree, 2007; Fraundorf, Watson, & Benjamin, 2010; Sanford, Sanford, Molle, & Emmott, 2006), this may not arise from longer reading times that we would otherwise associate with more careful reading (Birch and Rayner 2010; Ward and Sturt 2007). In fact, Birch and Rayner (2010) suggest that focus status may actually facilitate early lexical access, causing faster reading times for focus. So, while focus status may lead to better memory representations compared to non-focus referents, it does not necessarily lead to longer processing times.

One area where focus has been shown to have a consistent impact is on depth of processing during reading and listening. Prior work has shown that at least under the conditions of laboratory testing, readers and listeners do not automatically fully or accurately comprehend the information conveyed, even in relatively simple sentences or passages (e.g. Cowles & Kim, 2011; Ferreira, 2003; Sanford & Sturt, 2002; Swets, Desmet, Clifton, & Ferreira, 2008); this has been referred to as shallow processing. Experimental evidence suggests that focus status may cause "deeper" processing, meaning that the information is more fully and/or accurately represented in memory. For example, Sturt et al. (2004) used a Text Change paradigm to test of effects of focus on depth of processing. In this task, one word in a short paragraph changes between two presentations of that paragraph. Sturt et al. (2004) found that when one word (the original) was replaced with a semantically related word (the replacement), readers were much less likely to notice than when the replacement word was semantically unrelated to the original. However, when the original word was given focus status (either via clefting or embedded question-answer pairs), readers were more likely to notice the change even for semantically related words. This, along with other evidence, has been used to argue that focus status causes a better memory representation of the focus.

Using the same Text Change task, Cowles et al. (2012) found that focus status did not always result in equally good change detection rates. In particular, for focused words, they found
lower change detection rates when the replacement word came from elsewhere in the passage, and the lowest rates of all when the replacement was a potential alternative to the focus word. An example of their materials is given in (3) below, with the replaced word (cellist) underlined, and replacement words for cellist given at the end.

(3) A soprano and a critic were waiting backstage after a performance. A cellist and a violinist were talking with a composer about the next show. The critic asked the soprano who she admired more, the cellist or the violinist. The soprano replied that it was the cellist that she liked. The critic agreed he was quite something.

Replacements: violinist (alternative), composer (mentioned), bassist (new)

Cowles et al. (2012) found that readers were most accurate at detecting the changed word when the replacement was new (e.g. bassist) and least accurate when it was the focus alternative (e.g. violinist). This finding is consistent with theoretical accounts of focus that propose that focus status necessarily indicates the presence of alternatives that are relevant for interpretation (Féry & Krifka, 2008; Rooth, 1992). For example, Féry and Krifka (2008) give the following definition for focus: “Focus indicates the presence of alternatives that are relevant for the interpretation of linguistic expressions.” (p.4). In this approach, the interpretation (and thus processing) of focus crucially involves not only the focus referent itself, but the alternatives to that focus referent, whether they are given explicitly (as in (4)) or implicitly (as in (5)).

(4) Some rabbits ate the lettuce in my garden, not my tomatoes or onions.

(5) Some rabbits ate the lettuce in my garden.

In both cases, interpreting the focus, lettuce, means taking into account that there are alternatives that could have been focus. These may be specified, but they need not be: in (5) the set of alternative is implicit, yet still part of the process of focus interpretation.

Thus, in this type of account of focus, the interpretation of focus crucially involves not only the focus referent itself, but also the set of alternatives to that referent. This raises the intriguing possibility that these focus alternatives are activated as part of focus processing, and there is some recent evidence in addition to Cowles et al. (2012) suggesting that this is indeed the case (Husband and Ferreira 2012; Fraundorf, Watson, and Benjamin 2010; Gotzner, Spalek, and Wartenburger 2013; Braun and Tagliapietra 2010). Further, if focus alternative sets are activated as part of sentence processing, then another question to ask is whether they are fully suppressed as part of focus selection during sentence processing. The results here are mixed, with results from Fraundorf et al. (2010) showing evidence that focus alternatives are specifically suppressed compared to non-focus alternatives, while the results from Cowles et al. (2012) suggest that the memory representation for focus, while more accurate overall compared to non-focus (e.g. Sturt et al., 2004), is none-the-less susceptible to interference from lingering representations from the alternative set.
The experiment presented in this paper was designed to test the hypothesis that readers would be less accurate in a task that requires correctly remembering and distinguishing a focus referent from a focus alternative referent. Specifically, we were interested in whether readers could accurately judge paraphrases of statements that involved focused referents. More broadly, this work also addresses the underlying architecture of focus interpretation by giving a window into whether alternatives are considered and what impact they have on memory for focus. We argue that while focus status does cause greater processing depth, our results show it also has specific drawbacks that processing other parts of the sentence may not – namely the need to effectively suppress competitive alternatives.

3. The Experiment
This study tests the impact of focus alternative set members on memory for a focused referent, and in particular whether processing focus alternatives interferes with memory for focus. Participants were asked to read short passages and after each one confirm whether a paraphrase of one of the passage's sentences was accurate. For critical trials, the passages were comprised of five precisely constructed sentences that introduced five discourse referents, created an explicit two-member focus alternative set and selected a focus referent from it. The paraphrase was either accurate (Same), or had one word substituted from the passage that was either the explicit focus alternative (Alternative) or was mentioned but not as a focus alternative (Mentioned). Accuracy and response times for the paraphrases were recorded. Our hypothesis was that focus alternatives are considered during focus processing, and thus create memory representations associated with focus that may linger and interfere with readers’ ability to recall the precise identity of the focus referent. Based on this, we predicted that responses to the Alternative Paraphrase would take the longest and be the least accurate, compared to the other two paraphrase types.

3.1. Participants
Twenty-nine undergraduates at the University of Florida participated in the experiment. All were recruited from the LIN-SLHS participant pool for linguistic and speech studies. Participants were native speakers of American English with no significant language exposure before the age of 6, had normal or corrected-to-normal vision and no reported learning or reading disabilities.

3.2. Design and Materials
Forty critical items were adapted from Cowles et al. (2012). Each of these items was comprised of five sentences that formed a coherent passage. All critical passages had the same basic structure. First, Sentence 1 introduced two discourse referents in a conjoined subject NP and provided a general place and activity context. Next, in Sentence 2, three further discourse referents were introduced, with two of them also introduced together in a conjoined subject NP. Sentence 3 contained an embedded wh-question in which one referent asks something of another referent, and gives two specific alternatives (contrast set). Sentence 4 provided the answer to the question, using an it-cleft to mark focus on the answer. Sentence 5 provided a neutral concluding statement that did not re-mention any of the referents that would be later probed in the paraphrase state-
Influence of Focus Alternative Sets on Memory

Each item had this passage and three possible paraphrases of Sentence 4: one that was correct (Same), one that instead used the alternative from the explicit contrast set (Alternative), and one that used a previously mentioned referent that was not part of the contrast set (Mentioned). An example of two items is given in (6a) and (6b) below.

(6a) A soprano and a critic were waiting backstage after a performance. A cellist and a violinist were nearby, talking with a composer about the next show. The critic asked the soprano who she admired more, the cellist or the violinist. The soprano replied that it was the cellist that she liked. The critic agreed he was quite something.
   
   Same Paraphrase: The soprano admired the cellist.
   Alternative Paraphrase: The soprano admired the violinist.
   Mentioned Paraphrase: The soprano admired the conductor.

(6b) A naturalist and a student were walking along a river trail. A bluebird and a finch landed briefly in the brush, and a hawk flew in trees above them. The student pointed and the naturalist smiled and asked him what he saw, the bluebird or the finch. The student said it was the bluebird that he saw. He was trying to get better at identifying birds.
   
   Same Paraphrase: The student said he saw the bluebird.
   Alternative Paraphrase: The student said he saw the finch.
   Mentioned Paraphrase: The student said he saw the hawk.

In addition to these 40 critical items, an additional 80 filler items were also taken and adapted from Cowles et al. (2012). The purpose of having a high number of filler items was to obscure the structure of the critical passages as much as possible, and to avoid participants developing strategies about where in the passage to expect the paraphrase to come from. For example, we did not want participants to place undue attention on Sentences 3 or 4, nor to anticipate that the paraphrase would be of Sentence 4 because of the particular structure of, e.g. Sentences 1 or 2. These filler items were thus divided into passage structures that were quite different from the critical items, and also structures that were similar to the critical items, but then paraphrased a different sentence from the passage. Crucially, many filler passages contained embedded it-clefts that were not subsequently paraphrased. The correct answer for the filler paraphrases was "no" in 44 of the items. An example of a two filler items is given in (7a) and (7b) below.

(7a) On a June Saturday afternoon, a family decided to go to the fair. The two youngest boys both wanted to ride the ferris wheel. The oldest child asked her mom if it was okay to visit the fun house with her cousin. The middle child announced that it was the carousel that she really wanted to ride. It was a fun day filled with rides and cotton candy.
   
   Paraphrase: The youngest boys wanted to ride the ferris wheel.
A girl walked into a bakery and began eyeing delicious desserts on display. There were scones, cinnamon buns and candied apples in the window. She wondered which ones were on sale for the day. It was the cinnamon buns that were on sale. The girl bought six and ate two on the way home.

Paraphrase: Fresh cookies were displayed in the window.

3.3. Procedure
Upon arrival, each participant gave informed consent by reading and signing a consent form that described the general procedure of the experiment. After completing a short background questionnaire that collected demographic and language experience information, participants were seated in a small, sound-attenuated room in front of a computer screen and button response box. A sheet with printed instructions was given to participants to read over, and the researcher reiterated these instructions verbally and answered any questions that participants had. After the researcher was confident that the participant understood the task, she left the room and the participant completed the experiment alone. The researcher remained close by and available in case any additional questions arose during the experiment trials.

For each trial, participants would press a button on the button box to reveal a passage, which would appear in the middle of the screen as a block of text in a black sans-serif font. After reading this passage, participants would press the same button to proceed to the next screen. On this screen, a 6-point scale was presented and participants were asked to rate how well the sentences in the passage "fit together" to produce a "coherent, natural-sounding passage". The end points of this scale were "Doesn't fit together at all" (1) and "Fits together well" (6). Each point in the scale corresponded to a specific button on the response box, and participants pushed the button that corresponded to their rating. After this, a sentence would appear in the middle of the screen. This was a paraphrase of one of the sentences in the passage, and participants were instructed to indicate whether the paraphrase was correct or incorrect by pressing one of two specific buttons that corresponded to these judgments.

All trials were pseudo-randomized using a Latin Square design to create three different lists. Each participant was assigned to one list, and each list contained each critical item once, with equal numbers of items in each paraphrase condition. This meant that each participant saw each critical passage exactly once, and saw equal numbers of each paraphrase type. The order of passages appeared to be random to participants but in fact was carefully controlled so that paraphrases from the same condition never appeared one after the other. The entire experiment took approximately one hour to complete.

3.4. Results
Three types of data were collected from participants: Rating data for the naturalness of the passages, Response Time data for how quickly participants indicated each paraphrase was correct or incorrect, and Accuracy data for whether the participants' response to the paraphrase was correct or not.

First, participants' ratings of the naturalness of the passages were analyzed to see whether items differed significantly in terms of how natural and coherent the passages sounded. The av-
Average rating across all items was 5.74 (+/-0.07, 95% Confidence Interval), on a 1-6 scale in which 6 corresponded to a natural-sounding passage in which the sentences fit together well. This suggests that overall the items were largely natural sounding to participants. Ratings were submitted to a one-way ANOVA with items as a random variable, and no significant difference was found ($F > 1$), showing that all of the items were judged similarly natural.

For the Response Time data, times less than 200 milliseconds and longer than 15 seconds were first omitted from analysis, as were response times for trials in which the participant gave the incorrect answer. These data points were removed because very short times (>200 milliseconds) are too fast to be the result of performing the task and very long times (>15 seconds) are more likely to be due to distraction on the part of the participant.Incorrect trials were omitted because in this case we cannot be sure why the trial was incorrect; that is, what processes the participant was engaged in that lead to an incorrect answer. These criteria removed 13% of the data from the Response Time analysis, which is in keeping with data from similar tasks. The proportion of accurate trials for each participant within each paraphrase type was also calculated, including all collected data. Average Response Time and Accuracy are reported in Table 1 below.

<table>
<thead>
<tr>
<th>Paraphrase Type</th>
<th>RT (95% CI)</th>
<th>Accuracy (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same</td>
<td>2859 (175)</td>
<td>91.2% (.03)</td>
</tr>
<tr>
<td>Alternative</td>
<td>2781 (155)</td>
<td>77.5% (.06)</td>
</tr>
<tr>
<td>Mentioned</td>
<td>2685 (159)</td>
<td>91.5% (.04)</td>
</tr>
</tbody>
</table>

For both Response Time and Accuracy measures, analyses were conducted using linear mixed effects regression (Baayen, Davidson, and Bates 2008) in R using lme4 and lmerTest packages to build models and evaluate (1) whether a model including Paraphrase Type as a fixed factor was a better fit to the data than a baseline model without Paraphrase Type, and (2) whether each of three conditions differed from the other conditions (pair-wise comparisons). Significance levels for model comparison were evaluated using a log-likelihood $X^2$ test, measuring the decrease in goodness of fit when Paraphrase Type was removed from the model. All analyses include crossed random intercepts for participants and items. Pair-wise comparisons between Paraphrase Type were evaluated using the diffmeans() function in lmerTest library.

For Response Times, the model including Paraphrase Type as a fixed factor did not significantly improve the goodness of fit over a baseline model that did not include Paraphrase Type ($X^2(2) = 2.55$, $p = .28$). This shows that the different paraphrases (same, alternative, mentioned) did not, in fact, cause differences in response times to judge whether the paraphrase was correct; participants responded equally quickly for all three conditions. However, for Accuracy, the model including Paraphrase Type as a fixed factor did significantly improve the goodness of fit over a baseline model that did not include it ($X^2(3) = 24.99$, $p < .0001$), establishing that the different paraphrase types did have an impact on accuracy. Looking within paraphrase type, both
the Same ($t(57.9) = 4.8, p < .001$) and Mentioned ($t(57.9) = 4.88, p < .001$) conditions significantly differed from the Alternative condition, each showing significantly greater accuracy. However, accuracy between the Same and Mentioned condition did not differ ($t < 1$). This established that while participants were not faster or slower to respond to the different paraphrase types, they were poorer at correctly rejecting the alternative to the focus when it was given as a paraphrase, compared to both correctly accepting the correct paraphrase and correctly rejecting a paraphrase with a previously mentioned (but not focus alternative) referent.

4. **Discussion**

These results support our hypothesis that focus alternatives are considered as part of focus processing and remain associated with focus referents in memory representations of discourse. These results show that overall readers were quite good at correctly judging whether a paraphrase of a focus-bearing sentence was correct, even when the focus word was substituted with a previously mentioned word. Considering these results in combination with the results comparing focus vs. non-focus change detection presented in Sturt et al., (2004), our results confirm the high degree of depth of processing found for focus in other studies. However, there was also a sharp decline in accuracy when the focus word was substituted with an explicitly given alternative set member. This is consistent with prior work that has shown that focus alternatives are processed as part of focus interpretation (e.g. Husband & Ferreira, 2012), with lingering effects of focus alternatives on memory for focus referents.

Importantly, these results also suggest that focus status for referents does not make them immune to interference from referents that may have been initially considered as part of focus processing. These results are consistent with prior results showing similar interference effects in a different task (Cowles et al., 2012), but appear, at least initially, to be in contrast to Fraundorf et al. (2010), who found evidence that focus alternatives are specifically and effectively suppressed in memory. There are a number of methodological differences between Fraundorf et al. and this experiment that could explain why we find different results. The most likely difference is that in the key experiment with differing results (Experiment 3), there was a 24-hour delay between when the participants heard the passages and when they were given statements about the passages to confirm or deny. It may be that focus alternative representations are effectively suppressed in the final representation of focus, but that they are not initially suppressed effectively and so interfere in tasks like ours that test for memory after a very short delay. Another difference that may be important is that Fraundorf et al. used an auditory presentation and marked focus using pitch accent, in contrast to our combination of *wh*-questions and *it*-clefts for focus marking with written materials. Thus, it could also be that this difference in modality (auditory vs. visual) caused different outcomes. Further work is needed to determine the time course of contrast alternative interference as well as whether there are systematic differences in processing focus in different modalities. Crucially, we do not believe that our results necessarily conflict with those of Fraundorf et al., and we agree with these authors’ conclusions that focus processing crucially involves activation and suppression of focus alternatives. But, we believe our results suggest that this suppression is not always immediately effective.
Influence of Focus Alternative Sets on Memory

Our results are most consistent with a processing account of focus in which focus alternatives are (re)activated during focus processing, and stored with the focus as part of its representation in a way that makes it more difficult to accurately separate the representations. In this type of account, focus is specifically prone to particular processing issues than are not true for non-focus referents: While non-focus referents may not be encoded as carefully in memory (e.g. Sanford et al., 2006), focus status does not guarantee a perfect memory representation either because the nature of its encoding makes it specifically vulnerable to interference from the alternatives that were considered as part of focus interpretation. Thus, while focus status may have certain benefits in terms of processing, e.g. drawing greater attention and causing faster lexical access, it comes with its own, particular issues for accurate memory representations.

It is important to note that in this experiment as well as in others investigating this same question (e.g. Cowles et al., 2012; Fraundorf et al., 2010), narrow focus with small, explicitly given contrast sets were used. Cowles et al. (2012) did have one condition without an explicitly given contrast set, and their results suggest that having an explicit contrast set strengthens the interference effect of focus alternatives. Thus, while results from prior and current work thus far are consistent with theoretical accounts of focus interpretation that invoke focus alternatives (e.g. Féry & Krifka, 2008; Rooth, 1992), more work is needed to establish whether larger and/or implicit contrast sets show the same effects on memory for focus. It may well be that effects of focus alternatives are limited to small and/or explicit alternative sets, or that effects of focus alternatives are modulated by these factors.

5. Conclusion and Future Directions
The results of this experiment provide evidence that focus alternatives in a narrowly-specified contrast set are associated with representations of focus referents in memory, and cause interference with accurate memory for focus referents. These results are consistent with theoretical accounts of focus that propose that focus alternatives are a necessary part of focus interpretation (e.g. Féry & Krifka, 2008; Rooth, 1992). Future studies are planned to investigate the time course of the impact of focus alternatives on memory for focus referents, and to test the hypothesis that interference effects will still be present in focus contexts with larger, implicit alternative sets.

References


Wind Cowles and Michelle Perdomo


