Focus Alternatives and Contextual Domain Restriction: a Visual World Eye-tracking Study on the Interpretation of 'Only'

Christina Kim
ckim@bcs.rochester.edu

Christine Gunlogson
gunlog@ling.rochester.edu

Michael Tanenhaus
mtan@bcs.rochester.edu

Jeffrey Runner
runner@ling.rochester.edu

Dept. of Brain & Cognitive Sciences
Dept. of Linguistics
University of Rochester

Abstract

The interpretation of sentences with focus-sensitive elements like 'only' depends on context to restrict the domain of relevant alternatives for evaluating the focused expression. But what kinds of contextually available information do listeners actually use to restrict interpretive domains? Three visual world eye-tracking experiments show that listeners use at least previous mention (Experiment 1), real-world knowledge about specific scenarios (Experiment 2), and conceptual similarity to recently mentioned items (Experiment 3).

1 Introduction

The semantic contribution of the focus particle 'only' in sentences like (1-a)-(2-a) has two components, under standard assumptions: (i) the proposition expressed by the sentence without 'only'[1]—e.g., Matt's acing the exam in (1-a); and (ii)

---

[1]This proposition is often referred to as the *prejacent*; there has been a lot of debate about its status, which we do not address here.

the claim that no alternative to the focus value associated with 'only' makes the sentence true. The focus value and its alternatives are understood to be drawn from some appropriately restricted domain, as suggested in (1-b)-(2-b).

(1)   a.   Only Matt got a perfect score on the exam.
       b.   students in some class

(2)   a.   I only had a crush on Jared Leto.
       b.   cast of mid-90s teen TV drama *'My So-called Life'*

A likely context for (1-a), for instance, is a discussion of a specific class in a specific term. In other cases, like (2), the alternatives that the sentence is interpreted against seem to vary easily with the particular discourse context the sentence appears in.

Our concern in this paper is how this contextual narrowing of alternatives takes place. The work of spelling out the role of context generally falls to pragmatics, as von Fintel (1998) suggests in connection with similar issues of domain restriction for generalized quantifiers:

> The idea is to (temporarily) restrict the domain of evaluation for the whole sentence or even the whole discourse. The pragmatics will help in choosing a suitable universe for the evaluation of a particular sentence, but the semantics can just operate abstracting away from any such choice of a universe.

Rooth (1996) similarly characterizes the domain variable posited for interpretation of focus as pragmatically determined. How exactly the pragmatics accomplishes the task of suitably restricting the domain remains largely unarticulated. Our approach in the present study is to investigate experimentally potential sources of relevant contextual information by considering their effects on processing of sentences with 'only'.

We examine three factors, starting with preceding mention, cited by Rooth (1996) as one pragmatic factor affecting interpretation of 'only' sentences. With reference to (3) (Rooth's 24), he observes that "the domain of quantification is understood as consisting of just three propositions, rather than the full set of propositions of the form 'John introduced y to Sue' ".

(3)   John brought Tom, Bill, and Harry to the party, but he only introduced Bill$_F$ to Sue. *(Rooth 1996, example 24)*

That is, the domain is restricted to the set of propositions featuring the individuals just mentioned. In Experiment 1, we manipulate the factor of previous mention, as in Rooth's example.

Experiment 2 varies, in addition to linguistic mention, 'how much' context there is—that is, how much the nature of the scene described by the context-setting

sentences constrains likely alternatives. To illustrate this, consider a shopper described as being at a farmers market vs. one who is at a shopping mall. Potential purchases for the first shopper are most likely confined to produce and other food items, whereas the mall shopper could be buying just about anything. Relative to the shopping mall, the farmers market context is more restrictive and hence more informative, in a sense, about the kinds of things available for purchase.

Finally, Experiment 3 introduces the factor of conceptual similarity with previously mentioned items. In principle, the pair of sentences (4-a)-(4-b) can be interpreted with respect to the alternatives in (4-c), but we most easily construe this as meaning 'strawberries, but not other types of fruit' (4-d).

(4)   a.   Jill likes apples and nectarines.
             b.   Abby only likes [strawberries].
             c.   {strawberries, apples, nectarines, grapes, peas, socks, fountain pens,...}
             d.   {strawberries, apples, nectarines, grapes,..}

The remainder of this paper is structured as follows. First, in Section 1, we review some relevant previous psycholinguistic work on domain restriction. Section 2 introduces the Visual World paradigm in general, and describes specifically how eye movements can be used to probe comprehenders' expectations about focus alternatives. Sections 3-5 present three eye-tracking experiments examining effects on focus alternatives of previous mention, informativity of the context, and conceptual similarity, respectively. Section 6 concludes with directions for future research.

2   Using eye-tracking to investigate domain restriction

Our methodology involves monitoring of participants' eye movements in a 'visual world' paradigm. In a typical visual world eye-tracking study, participants move or click on objects in a visual display as they are listening to a sentence that indicates what item in the display is the target. Eye-movements have been shown to be closely time-locked to salient linguistic events in auditorily presented stimuli (Tanenhaus et al., 1995), and therefore provide a means to track listeners' expectations about upcoming linguistic input given the visual context and what they have heard so far. By manipulating the availability of different information types available in the visual or linguistic context, one can ask to what extent each of these potential information sources helps the listener restrict the referential domain to the point that the single intended referent can be picked out.

Previous experimental work has shown that language comprehenders rapidly integrate multiple sources of information for the purpose of referential disambiguation. Tanenhaus et al. (1995) showed that reference resolution can be guided by

what we know about the meanings of definite descriptions, in conjunction with properties of the visual context. Participants' eye movements were tracked as they followed instructions to manipulate items in a display. For example, they would hear *'Put the apple on the napkin in the box'*, while viewing a display containing one apple on a napkin, an empty napkin, an unrelated item, and a box. They found that whether the PP *'on the napkin'* was interpreted as a modifier or as a goal depended on properties of the visual display. When the display contained only one referent that matched the description *'apple'*, at the point when participants had heard *'the apple'*, they had all the information they needed to pick out the intended unique referent in the scene. As a result, *'on the napkin'* was not construed as a modifier but as a goal: participants looked at the empty napkin and sometimes even started to put the apple on the empty napkin. However in a display containing two apples, after hearing *'Put the apple'*, listeners interpreted *'on the napkin'* as a restrictive modifier picking out one of the two apples, not a goal.

These findings demonstrate that reference resolution is an incremental process that is sensitive to the visual context—in fact small changes to the visual context can bias comprehenders in favor of one parse over another. Moment to moment biases are reflected in participants' anticipatory eye movements as they are interpreting a sentence in a particular visual context. Subsequent studies have established language comprehenders' sensitivity to a variety of information sources during online processing: selectional properties of lexical items (Altmann & Kamide, 1999), the presence of contrast (Sedivy et al., 1999), information about the preceding linguistic discourse (Chambers et al., 2002), and knowledge about possible eventualities in the world (Chambers et al., 2004).

The current study take the same methodological approach to investigating what factors determine what is included in the set of focus alternatives that a sentence like (5) is interpreted with respect to.

(5)   Jane only has some candy.

Under our standard assumptions, (5) conveys that Jane has some candy and that she has nothing other than candy. What is included in this 'nothing else'? Since the eventual target word (*'candy'*) must be included among the focus values, having an expectation about what that word will be amounts to having stronger or weaker expectations about what will be a possible alternative.

For each trial, we record continuously what item the participant is fixating in a display like Fig. 1, as they are hearing the target sentence. After averaging this information over many trials (for a number of subjects), we can look at the proportion of fixations to a particular display item (for example, fixations to the target item) over time. Once we have the proportion of fixations to the target, the cohort competitor,[2] and the distractors, we can look at a particular time interval

---

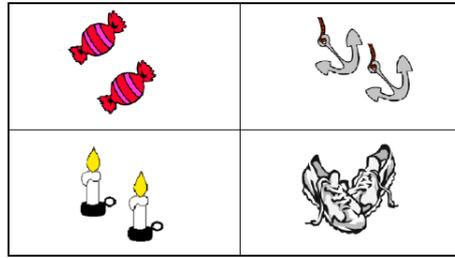[2]The cohort competitor shares initial phonology with the target word; see Section 3.1.

Figure 1: Example display. *target* = candy; *cohort* = candles; *unrelated* = anchors, sneakers.

and ask whether there is a difference between the proportion of looks to each display item.

3 Experiment 1: Focus alternatives are constrained by previous mention

Even out of the blue, one might expect (5) to be interpreted with respect to just the *relevant* alternatives (6-a).

(6) a. {candy, cupcakes, apples, sandwiches, gum, dry erase markers, refrigerators, pickup trucks ...}
     b. {candy, cupcakes, apples, sandwiches, gum, ...}

(7) Mark has some candy and some apples.

But in the context of a sentence like (7), the mentioned subset of the focus alternatives seem much more salient (6-b): the mentioned alternatives are somehow 'preferred'. Is the set of alternatives considered in interpreting a sentence like (5) constrained by the set of things just mentioned in the discourse?

3.1 Design, Procedure

The same basic paradigm is used (with variations) in all four experiments. The pre-recorded stimuli each consist of one or more *context sentences*, the last one of which includes references to particular types of objects such as boots or candy. The *target sentence* follows, as exemplified in (8).

(8) a. [*Context*] Mark has some candy and some apples.
     b. [*Target*] Jane (only) has some . . .
         (i)   candy
         (ii) candles

What Jane is described as having varies by experimental condition, e.g., the mentioned *candy* above vs. unmentioned *pencils*. The presence vs. absence of *only* is systematically varied as well. The task required of subjects is simply to click on the item(s) identified in the target sentence.

We manipulated (i) whether the target word was mentioned in the context sentence (Mention), and (ii) whether 'only' appears in the target sentence. Examples of the four resulting conditions are in Table 1; target sentences are to be interpreted with respect to the four-item display in Fig. 1.

On each trial, participants heard a pair of sentences like (7) (context sentence) and (5) (target sentence). At the onset of the target sentence, four pictures appeared (Fig. 1), one in each quadrant of the computer screen. Participants were instructed to click on the items in the target sentence (i.e. the things Jane had). 28 University of Rochester students who were native speakers of American English participated in the experiment.

Context	*No Mention*	*Mention*
Mark has...	...some gloves and some pencils.	...some candy and some pencils.
Target	*No Only*	*Only*
Jane...	...has some candy.	...only has some candy.

Table 1: Experiment 1 design and example stimuli.

In experimental trials, two of the four pictures were members of the same phonological cohort (*'candy'*-*'candles'*). In the absence of biasing factors, participants will begin to shift fixation to words that match the acoustic input about 200 ms after the onset of the word (Allopenna et al., 1998). Therefore we expect participants to look equiprobably at the target item and the cohort competitor at the point in the target sentence when they've heard just the beginning of the direct object (*'can...'*). As the unfolding auditory input disambiguates the target referent (*'...ndy'*)—the point of disambiguation—the proportion of fixations to the target item should rise as fixations to the competitor drop off. This means that if looks to the target item increase earlier than the point of disambiguation, there is a bias toward the target item due to some other property of the stimulus.

If recent mention of a particular type of object makes that kind of object more salient, as seems plausible, then we expect earlier identification of the target *candy* just in the Mention conditions. Since the target item in critical trials is always either an item mentioned in the Context sentence, or a phonological competitor of a mentioned item, using previous mention as a cue would effectively allow participants to identify the target early, despite the fact that the initial syllables of the target and competitor are identical.

Whether the presence of 'only' by itself can be expected to facilitate identification of the target is not clear. A more interesting question is whether 'only'

interacts with the mention factor. If the presence of 'only' strengthens the mention effect, we will see fastest identification of the target in Mention Only conditions—faster than can be expected on the basis of Mention NoOnly and NoMention Only conditions.

3.2   Results

In order to examine the time course of fixations, we calculated the proportion of fixations to the target at every 33 ms time slice, aggregating trials for each condition first within a participant and then across participants. Fig. 2.a shows proportion of fixation curves plotted as a function of time. The average time to convergence on the target referent (where target looks reliably exceed looks to the competitor) for each condition is shown in Fig. 2.b. For example, the 'Mention-Only' curve in Fig. 2.a corresponds to the average proportion of looks to the quadrant containing the target referent, *candy*, in Fig. 1, as listeners hear the sentence *'Jane only has some candy'*. In this condition, the target word will have been mentioned in the preceding context sentence. The corresponding bar in Fig. 2.b (rightmost) represents the average time for fixations to converge on *candy*—that is, diverge from fixations to the competitor, *candles*.

There were main effects of Mention ($F1(1, 24) = 46.8$, $p < .0001$) and Only ($F1(1, 24) = 6.2$, $p < .05$), as well as a Mention-Only interaction ($F1(1, 24) = 14.8$, $p < .0005$). On No Mention trials, listeners were able to disambiguate the target referent from the phonological competitor only after hearing the entire word, on average 560 ms after the onset of the target word (left-hand bars in Fig. 2.b). Thus in the absence of Mention, listeners had no preference for *candy* over *candles*. There was no advantage for the Only condition over the No Only condition ($t = 1.4$, $p = .15$).

The Mention-No Only trials (right-hand bars, Fig. 2.b) showed an effect of Mention independent of any effect of Only: fixations converged on the target referent 404 ms after target word onset. Thus when a previously mentioned item appeared as part of the visual context, listeners had a preference for the mentioned item. When 'only' was present, fixations converged on the target referent 139 ms after target word onset, well before the input disambiguated the target and the cohort competitor. In the 200-400 ms post-target onset interval, fixations to the target in Mention-Only trials exceeded those in Mention-No Only trials ($t = 10.4$, $p < .001$), while No Mention trials did not differ as a function of Only ($t = .9$, $p = .35$). Thus, after hearing only the initial part of the target word, listeners strongly expected the possible referents to be constrained by the set of just mentioned items. When this expectation is violated, as in the No Mention-Only condition, the point of disambiguation is late (in fact, later in absolute terms than in the No Mention-No Only condition).

These results suggest that upon hearing 'only', listeners have a strong expectation that the upcoming focus will be a recently mentioned item. We might
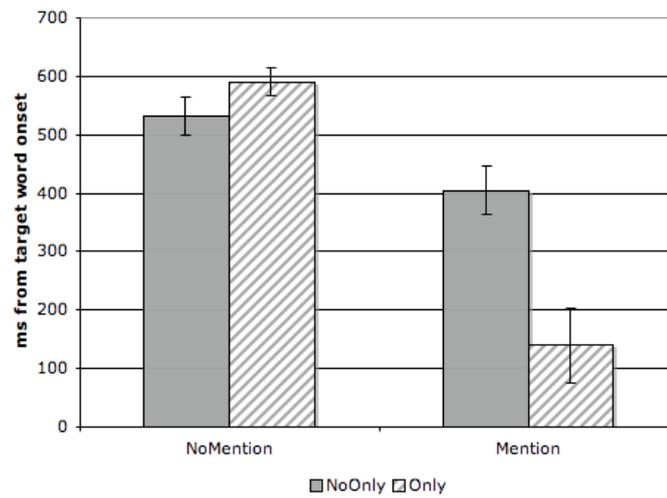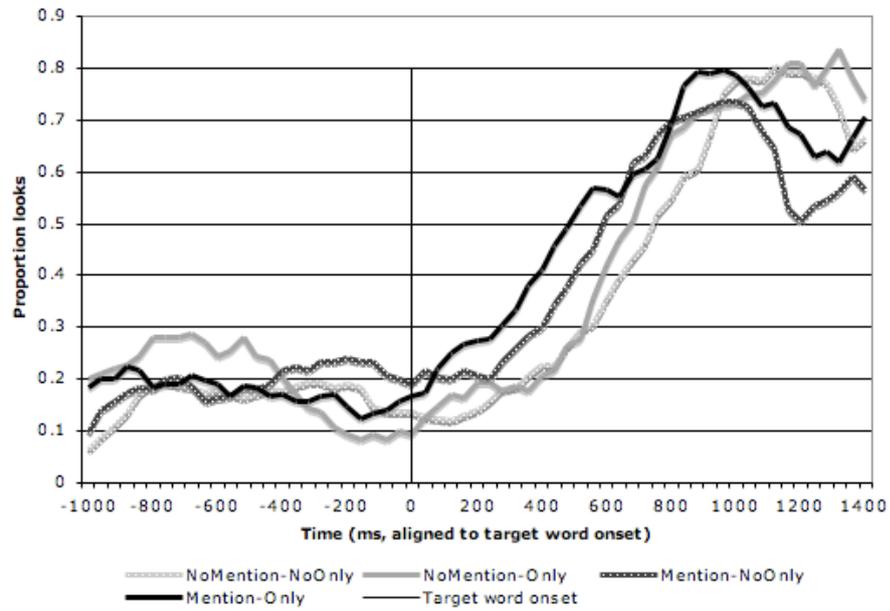
a.

b.

Figure 2: Experiment 1: a. Proportion target fixations over time, b. Mean point of disambiguation (error bars are Standard Error).

think of 'only' as functioning as a cue that increases listeners' sensitivity to aspects of the preceding discourse context.

4 Experiment 2: Informativity contributes to restricting alternatives

Presumably the manipulation of Mention in Experiment 1 has the effect of making some set of things salient in the context. We might then expect to observe the same restrictive effect just by enriching the information in the context (i.e. making the context more 'restrictive'). Experiment 2 tests this hypothesis, asking whether having richer information in the context contributes to restricting focus alternatives in sentences like (9). Compare (10-a) and (10-b).

(9)   Peter only wants to buy [some magazines].

(10)   a.   Jill and Peter are at the drugstore.
         b.   Jill and Peter are at the newsstand.

Intuitively, (10-b) provides more information, since our knowledge about the world tells us that the range of items that can be purchased is relatively narrow compared to a drugstore, where a wider set of items can be purchased. In addition to repeating the experimental conditions from Experiment 1 (Mention x Only), Experiment 2 varied the informativity of the discourse context.

4.1   Design, Procedure

Experiment 2 crossed three factors: Context Informativity (Informative, Underinformative), Mention, and Only; the resulting eight conditions are given in Table 2. The corresponding visual display is in Fig. 3.

Context 1	*Underinformative*	*Informative*
Jill and Peter are...	...at the drugstore.	...at the newsstand.
Context 2	*No Mention*	*Mention*
Jill is getting...	...some comic books and some cigarettes.	...some magazines and some cigarettes.
Target	*No Only*	*Only*
Peter is...	...getting some magazines.	...only getting some magazines.

Table 2: Experiment 2 design and example stimuli.

The procedure was as in Experiment 1, except that two context sentences (Context 1 and 2 in Table 2) preceded the target sentences. As participants heard the target sentence, they were shown a visual display like Fig. 3, with a target item

Figure 3: Experiment 2 example display.

(*magazines*), a cohort competitor (*magnets*), and two unrelated distractor items (*scissors, lamps*). Notice these are all items consistent with the Underinformative Context (here, 'drugstore'), while only the target item is compatible with the Informative Context ('newsstand'). 24 native English speakers participated in the experiment.

If more informative contexts function in the same way as mention, they could restrict the domain of interpretation specifically when 'only' is present. We might then expect faster convergence on the target item only in Only conditions (on top of the Only-Mention effect from Experiment 1). On the other hand, we might find that enriching the context has a restrictive effect on subsequent interpretation, but in a general way that isnt specific to the presence of 'only'. In that case, we would expect across-the-board faster convergence on the target item in Informative conditions, irrespective of the presence of 'only'.

4.2 Results

Fig. 4 shows the average time to convergence on the target referent (Underinformative conditions on the left, Informative conditions on the right).

4.2.1 Underinformative contexts

There were main effects of Informativity ($F(1, 20) = 34.0$, $p < .0001$), Mention ($F(1, 20) = 11.5$, $p < .001$), and Only ($F(1, 20) = 9.8$, $p < .005$), and no interactions.

Underinformative contexts patterned much like Experiment 1. This is expected: the most underinformative thing to say is nothing at all, and in this case the four Underinformative conditions reduce to the conditions in Experiment 1. The target referent was disambiguated latest in the No Mention-No Only and Mention-No Only conditions, earlier in the No Mention-Only condition, and earliest in the Mention-Only condition (Fig. 4, left-hand bars). As in Experiment 1, 'only' seems to increase sensitivity to information in the preceding linguistic
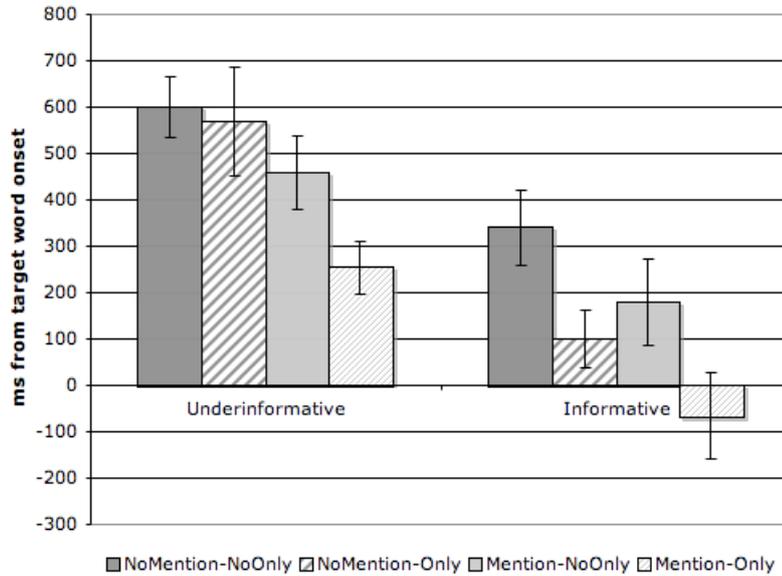
Figure 4: Experiment 2. Mean point of disambiguation.

context, creating a bias in favor of discourse-old items.

4.2.2 Informative contexts

First, there was a general restrictive effect of context informativity: Informative context conditions had on average a 335 ms earlier convergence on the target referent relative to the corresponding Underinformative context conditions.

In addition, the benefit due to informativity was strengthened in the presence of 'only': there was a 399 ms advantage due to Informative context in Only conditions, compared to a 271 ms advantage for No Only conditions. In Mention-Only trials, target fixations start rising well before the onset of the target word, soon after the onset of 'only'. The largest advantage occured in the Mention-Only condition, where listeners were able to disambiguate the target referent after hearing 'only', but well before the onset of the target word.

5 Experiment 3: Generating expectations about likely alternatives

In Experiment 3, we asked whether conceptually similar alternatives are preferred over conceptually unrelated ones: after hearing *'Jane likes apples and nectarines'*, a continuation like *'Mark only likes oranges'* seems more expected than one like *'Mark only likes pickup trucks'*. If this contrast is real, we might be able to use it

to ask a question about the nature of the expectations comprehenders have about the members of the alternative set.

What might listeners be doing to produce the results of Experiments 1-2? At least two explanations seem possible. First, maybe given the items in the visual display, listeners are ruling out certain referents as unlikely (based on the previous discourse context, etc.). This could explain the pattern of results we observe in both experiments. But another possibility is that listeners use the information from the discourse context to start generating hypotheses about what items are likely to be in the alternative set. If listeners are actively generating candidate alternatives, they might do this on the basis of something like conceptual similarity; this would predict earlier target disambiguation for same-category over different-category items, even without previous mention.

5.1   Design, Procedure

The structure of Experiment 3 is virtually identical to Experiment 1. Participants heard sequences consisting of a context sentence and a target sentence (Table 3; the corresponding visual display is shown in Fig. 5.

Context	*Mention*	*Novel-Same category*	*Novel-Different cat.*
Mark has...	...some apples and some oranges.	...some pears and some oranges.	...some boots and some sandals.
Target	Jane only has some apples.		

Table 3: Experiment 3 design and example stimuli.

At the onset of the target sentence, a display containing a target item (*apples*), a cohort competitor (*anchors*), and two unrelated distractors (*candycanes, speakers*) appeared. Based on Experiments 1-2, we expect a Mention preference. The question of interest is whether there is an advantage for Same-category over Different-category Novel items. 16 native English speakers participated in the experiment.

5.2   Results

Average points of disambiguation are in Fig. 6. Fixations converged on the target referent earlier in the Mention condition than in the Novel conditions ($t = 3.7$, $p < .0001$), consistent with the results of Experiments 1-2.

Interestingly, within Novel conditions, the target was disambiguated earlier when the target word was in the same category as recently mentioned items (*'some pears and some oranges ... some apples'*) than when it was in a different category (*'some boots and some sandals ... some apples'*) ($t = 2.4$, $p < .05$). This advantage cannot be due to explicit mention, since Same and Different category conditions
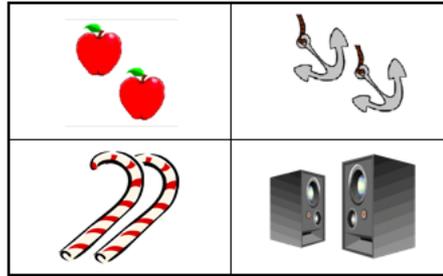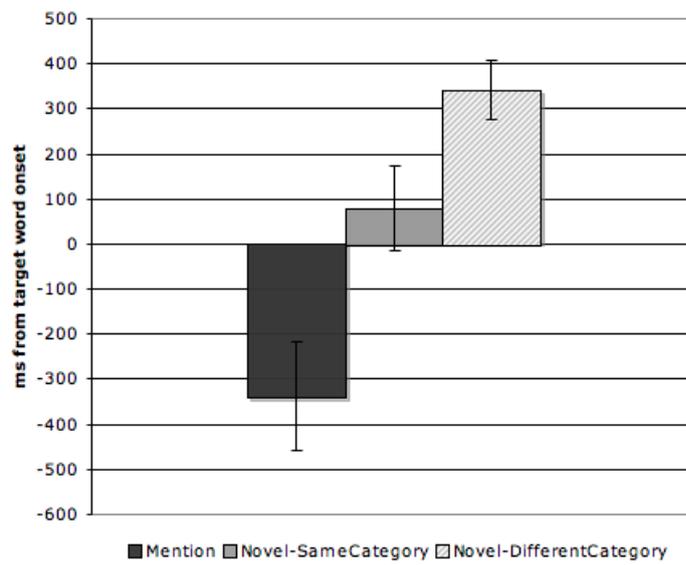
Figure 5: Experiment 3 example display.

Figure 6: Experiment 3: Mean point of disambiguation.

both contained novel target words. Instead, it suggests that previous mention of *'pears'* and *'oranges'* activates not only the meanings of those particular words and their corresponding conceptual representations, but also the conceptual category they are members of; this in turn makes other category members (like *'apples'*) more salient as possible members of the alternative set.

6   General discussion

In the current study, we address the question of how alternative sets are established for the purpose of interpreting sentences containing focus operators, looking specifically at sentences with adverbial 'only'. We use comprehenders' eye movements in a visual scene as a measure of their changing expectations about possible referents; in critical cases, the presence of 'only' earlier in the sentence served as a cue to attend to aspects of the linguistic context.

In three eye-tracking experiments, we show that recent mention (Experiment 1), the informativity of the linguistic context (Experiment 2), and conceptual similarity (Experiment 3) are among the factors that contribute to the restriction of focus alternatives in the context of 'only'. These factors speed recognition of targets for sentences without 'only' as well, suggesting they have a general role in comprehension. Their enhanced effect in the presence of 'only' is striking, raising the possibility that 'only' has a general function of directing attention to contextual cues about the relevant domain for interpretation. The results of Experiment 3 further suggest that listeners' expectations about likely alternatives underly the contrasts observed in Experiments 1-2: given the linguistic context, comprehenders immediately begin generating hypotheses about likely focus alternatives.

These findings for 'only' raise interesting questions about the behavior of other focus operators. Future work comparing 'only' with other alternative-sensitive operators like 'also' will help pull apart the specific contributions of these lexical items from general aspects of focus interpretation. In particular, our conclusions about 'only' lead us to specific predictions about how the behavior of 'also' will diverge from 'only', allowing us to substantiate the hypothesis that comprehenders actively generate candidate alternatives. Even more generally, we have been treating focus alternatives as analogous to quantifier domains, but whether the same factors influence domain restriction is an empirical question. We anticipate addressing this question by comparing 'only' with quantifiers like 'every' or quantificational adverbs like 'always', which share with 'only' the general problem of domain restriction, but also differ along other dimensions (for instance, the presuppositions carried by an 'only' sentence versus an 'every' sentence) that may influence the types of information comprehenders take into consideration.

A very general problem to be addressed from the point of understanding language comprehension has to do with cue combination; that is, how do prosody, discourse parallelism, discourse old-new status, and other potentially relevant fac-

tors combine with each other? Once we can adequately characterize how different kinds of information interact in various instances of contextual domain restriction, we will be in a position to ask how the linguistic properties of particular lexical items predict what contextual information they will draw on, given general facts about how different information types are integrated during interpretation.

Acknowledgements

References

Allopenna, P, Magnuson, J and Tanenhaus, M (1998) "Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models", *Journal of Memory and Language* **38**, 419–439.

Altmann, G and Kamide, Y (1999) "Incremental interpretation at verbs: Restricting the domain of subsequent reference", *Cognition* **73**, 191–238.

Chambers, C, Tanenhaus, M, Eberhard, K, Filip, H and Carlson, G (2002) "Circumscribing referential domains in real-time sentence comprehension", *Journal of Memory and Language* **37**, 37–49.

Chambers, C, Tanenhaus, M and Magnuson, J (2004) "Action-based affordances and syntactic ambiguity resolution", *Journal of Experimental Psychology: Learning Memory and Cognition* **30**, 687–696.

von Fintel, K (1998) "Quantifiers and If-Clauses", *The Philosophical Quarterly* **48**, 209–214.

Rooth, M (1996) "Focus", In S Lappin, ed. *The Handbook of Contemporary Semantic Theory*, 271–297.

Sedivy, J, Tanenhaus, M, Chambers, C and Carlson, G (1999) "Achieving incremental interpretation through contextual representation", *Cognition* **71**, 109–147.

Tanenhaus, M, Spivey-Knowlton, M, Eberhard, K and Sedivy, J (1995) "Integration of visual and linguistic information in spoken language comprehension", *Science* **268**, 1632–1634.