Scalar reference, contrast and discourse: Separating effects of linguistic discourse from availability of the referent

Lynsey Wolter,
Department of English University of Wisconsin, Eau Claire

Kristen Skovbroten Gorman, and
Department of Brain and Cognitive Sciences University of Rochester

Michael K. Tanenhaus
Department of Brain and Cognitive Sciences University of Rochester

Abstract

Listeners expect that a definite noun phrase with a pre-nominal scalar adjective (e.g., *the big …)* will refer to an entity that is part of a set of objects contrasting on the scalar dimension, e.g., size (Sedivy, Tanenhaus, Chambers & Carlson, 1999). Two visual world experiments demonstrate that uttering a referring expression with a scalar adjective makes all members of the relevant contrast set more salient in the discourse model, facilitating subsequent reference to other members of that contrast set. Moreover, this discourse effect is caused primarily by linguistic mention of a scalar adjective and not by the listener's prior visual or perceptual experience. These experiments demonstrate that language processing is sensitive to which information was introduced by linguistic mention, and that the visual world paradigm can be use to tease apart the separate contributions of visual and linguistic information to reference resolution.

Listeners rapidly integrate visual and linguistic information in real-time language comprehension, with linguistic input influencing visual processing (e.g., Spivey, Tyler, Eberhard & Tanenhaus, 2001), and visual information influencing language processing (Cooper, 1974; McGurk & MacDonald, 1976; Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995). For example, the number of objects in a display determines whether a temporarily ambiguous prepositional phrase, such as *in the box*, in the instruction, *Put the towel in the box*, is interpreted as a goal–where the towel should be put–or as a modifier– which towel is the intended theme of the putting event (Farmer, Anderson & Spivey, 2007; Tanenhaus et al., 1995; Spivey, Tanenhaus, Eberhard & Sedivy, 2002; Trueswell, Sekerina, Hill & Logrip, 1999; Novick, Thompson-Schill & Trueswell, 2008). When listeners interpret instructions such as *put the cube in the can*, the affordances of the available real-world objects (e.g. an object's size in relation to potential goal locations) affect which containers are included in the initial referential domain as the listener hears *in the can* (Chambers, Tanenhaus, Eberhard, Filip & Carlson, 2002). Affordances such as whether or an object can be poured (e.g., a solid or liquid egg) also affect whether an ambiguous prepositional phrase,

in the bowl, is parsed as a goal or a modifier in an instruction such as, *pour the egg in the bowl...* (Chambers, Tanenhaus & Magnuson, 2004).

These results can be interpreted as support for a unified conceptual representation that integrates linguistic and non-linguistic information in real time language comprehension, without differentiating the source of the input, for example, whether information is introduced linguistically or is made available by general knowledge or specific non-linguistic context. A number of other recent results also suggest a conceptual representation that does not distingish linguistic from non-linguistic sources of information. For example, Elman, Hare, McRae and their colleagues have argued that effects previously attributed to verb-based thematic roles are better explained by conceptual event-based representations (Bicknell, Elman, Hare, McRae & Kutas, 2010; Matsuki et al., in press). Kaiser (2009) has recently demonstrated that effects of causal discourse coherence relations on pronoun interpretation also occur when a sentence is primed by a non-linguistic event that implies causality. For example, a priming sentence that describes two events which are causally linked (e.g., The patient pressed the red emergency button near the bed and a nurse quickly ran into the room) and a visual event that implies causality (e.g., a short video showing an object moving after it is bumped by another object) have similar effects on the interpretations of an ambiguous pronoun in a target sentence with unrelated content.

In a striking series of studies, Altmann and his colleagues find that people will look to a blank location where an object, e.g., a wine glass, would have been moved given a linguistic description of an event that involves moving the glass, more often than they will look to the actual location of the glass. Based on these results, Altmann and Kamide (2007) have argued that in situations where language is interpreted in the context of a visual display, and eye movements are the dependent measure, linguistic and visual inputs are both used to construct event-based conceptual representations. Feature overlap between the conceptual representations evoked by language use and visual perception determines the relative salience of objects in a scene, thus mediating the attentional shifts that are reflected in eye movements. A number of studies (Altmann & Kamide 2009; Altmann & Mirković 2009; Knoeferle & Crocker 2006; Mayberry, Crocker & Knoeferle 2009) have investigated similar hypotheses, focusing on an event-based conceptual representation that is constructed from both linguistic and visual inputs, without differentiating the source of the information. Finally, it is worth noting that the strong form of the embodiment hypotheses in which sentence meaning is constructed through perceptual simulation (for review, see Glenberg, 2007) would suggest that the mental representation of linguistically presented information is not tied to its modality.

Linguistic theories of discourse and information structure provide a different perspective, assuming that *linguistic mention* is important for some types of linguistic structures, for example certain classes of elliptical expressions and referring expressions. Models of the semantics of discourse often assume that listeners construct a discourse model of the events and entities that are introduced linguistically in the conversation (Groenendijk, Stokhof & Veltman, 1997; Heim, 1982; Kadmon, 2001; Kamp, 1981). Similarly, linguistic models of information structure distinguish between information that is given by virtue of linguistic mention and information that is unmentioned or new to a discourse (e.g. Gundel, Hedberg, & Zacharski, 1993; Ariel, 1990; Gordon, Grosz, & Gilliom, 1993). It remains unclear, however, whether effects that are discussed in terms of discourse models or information structure could also be explained by a conceptual representation that integrates linguistic and non-linguistic information without encoding the source of the information. For example, models that seek to explain the felicity conditions for different referring expressions (e.g. whether it is most felicitous for a speaker to use a pronoun, or a definite or indefinite noun

phrase) often appeal to constructs such as cognitive status (Gundel et al., 1993), salience (Ariel, 1990), or attentional state (Gordon et al., 1993), which are not intrinsically linguistic.

Nonetheless, there are reasons to believe that linguistic mention might affect on-line sentence processing in a way that is distinct from general accessibility or salience. The distinction between visual and linguistic information from the context is potentially relevant to various ellipsis constructions (e.g., Arregui, Clifton, Frazier & Moulton 2006; Frazier & Clifton, 2005, 2006; Hankamer & Sag, 1976; Johnson, 2001; Kehler & Ward, 1999, 2004; Kim, Kobele, Runner & Hale, in press; Kim & Runner 2009; Murphy, 1990; Snider & Runner 2010; Tanenhaus & Carlson, 1990), de-accenting phenomena (Dahan, Tanenhaus & Chambers, 2002; Ladd, 1980, 1996; Schwarzschild, 1999), and referring expressions such as third person pronouns versus demonstrative pronouns (Brown-Schmidt, Byron & Tanenhaus, 2005). Moreover, Gunlogson (2003) has argued that the source of information plays a crucial role in determining the felicity conditions for different types of utterances. Perhaps most directly, Kaiser and Trueswell (2004) examined anticipatory eye movements to Finnish OVS constructions, where the post-verbal NP typically refers to a discourse-new entity. Discourse context increased anticipatory looks to unmentioned potential referents. All of the entities were present in the display so the given/new distinction was established by linguistic mention, thus indicating that Finnish post-verbal NPs are sensitive to information with a specific source – what has been mentioned. However, even this study does not provide conclusive evidence against a conceptual representation that does not differentiate sources of information, because linguistic mention could have increased the salience of the previously mentioned, visually co-present potential referent relative to the unmentioned potential referent by drawing attention to it. The same argument holds for studies that demonstrate different effects of given-new status with fluent and disfluent utterances. despite the fact that all of the objects are present in the display (Arnold, Tanenhaus, Altmann & Fagnano, 2003).

The goal of the present studies is to use the context sensitivity associated with pre-nominal scalar adjectives (e.g., *big, small, tall, short*) to evaluate whether when attention is controlled for, a particular context-sensitive expression is preferentially interpreted as referring to information that is introduced linguistically. Before we turn to the details of our experiments, we briefly discuss the relevant properties of how scalar adjectives are processed in real-time language comprehension.

Sedivy, Tanenhaus, Chambers and Carlson (1999) established the *basic contrast effect*. In the critical trials in Sedivy et al.'s experiments, participants saw either a display with a contrast, e.g., a tall glass, a short glass, and a pitcher, or a display without a contrast in which, for example, an unrelated object such as a pencil replaced the short glass. They then heard an instruction such as *Pick up the tall glass*. In the no-contrast condition, participants did not shift their gaze to the glass until several hundred ms after the onset of the noun. In the one-contrast condition, however, participants began to shift their gaze to the glass shortly after the onset of the adjective *tall,* even though the pitcher was taller than the glass. These results suggest that scalar adjectives induce a *contrastive implication*: listeners hearing a referring expression beginning with *the tall/short* expect the referent to belong to a set of two or more objects that can be described with the same noun and that contrast in height. The *basic contrast effect* associated with scalar and certain other pre-nominal adjectives has been replicated several times and has proved to be a useful tool for investigating a growing number of questions about language comprehension (Grodner & Sedivy, 2011; Hanna & Tanenhaus, 2004; Hanna, Tanenhaus & Trueswell, 2003; Heller, Grodner & Tanenhaus, 2008; Nadig & Sedivy, 2002; Sedivy, 2003) and language production (Brown-Schmidt & Konopka, 2008; Brown-Schmidt & Tanenhaus, 2006; Wardlow-Lane, Groisman & V. Ferreira, 2006, Wardlow-Lane & Ferreira, 2008).

The basic contrast effect depends on the rapid integration of information from the visual scene – which objects are present and which ones contrast along various dimensions – and information from the lexical semantics of the scalar adjective, which is interpreted relative to a *comparison class* of similar entities that contrast with respect to a particular scale (Kennedy 1999; Kennedy & McNally 2005). In a visual world study, use of a scalar adjective to refer to an entity introduces the comparison class via linguistic mention, but without explicitly mentioning the other members of the comparison set. This property makes it particularly useful for determining whether or not being introduced, even indirectly, via linguistic mention is encoded in a listener's representation in a way that is distinct from the status of entities introduced by non-linguistic means.

Consider the sequences of instructions in (1) uttered in the presence of a visual scene containing a big and small candle, a big and a small tie, and several shapes (including a heart). (1) Put the big candle above the heart; now put the small… When the listener is processing the referring expression that begins with *the small*, what expectations does he or she have about the referent? In example (1), the previous mention of *the big candle* presumably evokes the contrast set consisting of both candles and therefore might facilitate subsequent reference to the (unmentioned) small candle. However, facilitation might not be due to linguistic mention *per se*. It could arise because attention was drawn to the small candle or because participants expect to move another object of the same type as they previously moved. In a conceptual model that does not encode information about source, facilitation might occur because the listener must consider the contrast set of candles while processing *the big candle* in context (1). Therefore the set of candles – including the unmentioned small candle – should be more available as potential referents for the subsequent referring expression beginning *the small*. But, crucially, we would expect to find the same facilitation when the set of candles is made salient by other means. For example, consider context (2) below. Assume that the instructions in (2) are uttered in the same visual scene as (1), and in addition, the big candle is in a yellow square and the big tie is in a blue square. (2) Put the candle that's in the yellow square above the heart; now put the small… The experience of looking at the candles and/or manipulating the big candle while processing the first sentence of context (2) should make the big candle and the visually similar small candle more salient, again facilitating subsequent reference to either of the candles. If the relative salience of objects in the display is mediated by a uniform conceptual representation, it shouldn't matter *how* the candles are made more salient – linguistic information such as the previous mention of *the big candle* and nonlinguistic information such as having spent more time looking at the candles than other objects in the display should result in the same effect.

In the experiments reported in this article, we demonstrate that linguistic mention of a referring expression with a scalar adjective, e.g. *the big candle*, facilitates subsequent reference to other members of the contrast set to which the big candle belongs. Most importantly, we evaluate whether this effect is primarily due to linguistic mention or whether it can attributed to nonlinguistic factors that typically co-occur with the utterance of an expression like *the big candle*, such as the participant's prior experience of physically manipulating one of the objects in the display or of having directed visual attention to the members of a particular contrast set.

Experiment 1

The first goal of Experiment 1 is to determine whether in a display with two contrast pairs (e.g., two candles that differ in size and two ties that differ in size), linguistic mention of a contrast member using a pre-nominal scalar adjective (e.g., *Put the big candle…*) will cause the listener to expect a subsequent scalar adjective (e.g., *Now put the small…*) to refer to the

other member of the previously evoked contrast set. To our knowledge, the cross-sentential effect of mentioning a scalar adjective has not been investigated previously.

If linguistic mention of a referring expression such as *the big candle* does facilitate subsequent reference to other members of the same contrast set, this bias is consistent with at least two hypotheses. The *linguistic mention hypothesis* is that information introduced by linguistic mention is marked as such, and the facilitation of reference to the small candle is triggered by prior linguistic mention of the phrase *the big candle*. The *undifferentiated representation hypothesis* is that information from visual and linguistic sources is integrated into a single conceptual representation without maintaining the source of that information, and reference to the small candle can therefore be facilitated equally well by information from linguistic sources (mentioning *the big candle*) or from non-linguistic sources (e.g. causing the listener to look at or move the set of candles). Therefore, the second goal of Experiment 1 was to tease apart these hypotheses by separating effects of linguistic mention from effects of manipulation and/or visual attention. We compared instructions like those in (3), in which the participant moves an object described with a scalar adjective, to those in which a participant hears a referring expression that contains a scalar adjective but does not move the referent of that expression, as in (4). We also considered cases in which the participant moves an object that is a member of a contrast set, but is not described using a scalar adjective, as in (5), where *the object in the yellow square* is the big candle. These conditions were included in order to separate effects of linguistic mention from effects of attention due to visually-guided reaching: effects of linguistic mention would be expected to affect the processing of *the small candle* in both (3) and (4), whereas an effect of attention due to visually-guiding reaching would affect the processing of *the small candle* in (3) and (5). Our control was instructions like (6), in which no contrast members are mentioned or manipulated during the first instruction. Comparing these situations allowed us to disentangle the linguistic mention of a scalar adjective from the attentional effects of manipulating an object in a scene. In Experiment 1A, discussed in the next section, we consider the potential effect of visual attention (as opposed to the attention caused by moving an object) in more detail.

```
(3) Put the big candle above the heart;
    now put the small {candle/tie} below the heart.

(4) Put the circle below the big candle;
    now put the small {candle/tie} to the left of the diamond.

(5) Put the object in the yellow square above the triangle;
    now put the small {candle/tie} below the heart.

(6) Put the diamond above the heart;
    now put the small {candle/tie} to the right of the diamond.
```

The linguistic mention hypothesis predicts a bias to the small candle as the listener is processing *small* in (3) and (4), but not (5), where a contrast set member has been moved but not mentioned, or (6), where contrast set members have neither been mentioned nor moved. By contrast, the undifferentiated representation hypothesis predicts a bias to the small candle as the listener is processing *small* in (5) as well as (3-4).

METHODS

Participants—Nineteen participants were recruited from the University of Rochester undergraduate community and paid for their participation. All participants in this and the

following experiments were native speakers of North American English, with normal or corrected-to-normal hearing and vision. One participant was excluded from analysis due to equipment failure, and two participants were excluded because they made frequent errors in following the instructions; the remaining sixteen participants were at least 95% accurate.

Materials—During each trial, participants viewed a computer screen displaying a five-by-five square grid similar to one of the grids shown in Figure 1. The grid contained four gray shapes whose starting locations remained constant, and two pairs of objects that contrasted in size. The contrast pairs were either vertically aligned or horizontally aligned. While the four shapes were always the same, different objects were used in the contrast pairs for each trial. Finally, four cells of the grid had colored backgrounds, as indicated by the shaded cells in Figure 1. During most trials, the colored cells were in the corners of the grid, as shown on the Figure 1, Panel A. During some trials, two objects were displayed in colored cells, as shown in Figure 1, Panel B. When an object appeared in a colored cell, it was always referred to using its location rather than its identity, e.g. *the object in the green square*. In critical trials, the nouns describing the pairs of objects, which matched in frequency, did not share a phonological onset. No objects had a /b/ or /s/ onset, to avoid phonological cohort effects with the scalar adjectives *big* and *small*. The shapes and contrasting objects could be moved around the grid by the participant; the colored backgrounds of cells could not. (Figure 1.)

Experimental trials—During each trial, the participant heard a sequence of two sentences instructing the participant to move objects in the display. Example pairs of instructions for Experiments 1 and 1A are shown in Table 1. We discuss Experiment 1 first and then Experiment 1A. The experiment had a 2 (Mention) × 2 (Move) × 2 (2nd Target Type) design. The Mention manipulation allowed us to address our first experimental question, namely, does mention of a scalar adjective affect the interpretation of subsequent scalar adjectives? In Mention conditions, the first sentence included a noun phrase with a scalar adjective and basic-level noun, such as *the big candle*, while in No-Mention conditions, the first sentence did not include a scalar adjective. The Move manipulation addressed our second experimental question by teasing apart the participant's experience of moving an object in the display from the participant's experience of processing a scalar adjective. In Move conditions, the target of the first sentence – i.e. the object moved by the participant – was a member of a contrast pair. (Note that in the No-Mention, Move conditions illustrated in Table 1, *the object in the yellow square* refers to the big candle.) In No-Move conditions, the target of the first sentence was a shape rather than a member of a contrast pair. Finally, the 2nd Target Type manipulation balanced our design and allowed us to compare Same-Set conditions, in which noun phrases in both sentences referred to members of the same contrast pair, with Different-Set conditions, in which noun phrases in the two sentences referred to members of different contrast pairs. In No-Mention, No-Move conditions, the distinction between Same-Set and Different-Set trials was arbitrary.

The initial locations of the target objects, as well as the target locations for each instruction, were evenly distributed. We constructed four trials for each display/instruction type, for a total of 32 test trials. Eight lists were created using a Latin square design.

Filler trials—The test trials were intermixed with 72 filler trials designed to prevent the participants from developing expectations based on contingencies between the instructions and the display. For example, in some filler trials the participant moved the same object twice, or moved two objects of the same size. In other filler trials, some objects were displayed in colored cells, but the target of the instruction was an object not in a colored cell. As a result, the contingencies were balanced so that, at the beginning of each sentence, every object and shape in the display was equally likely to be the target.

Recording procedure—The instructions were recorded by a male native speaker of English, with training in phonetics, using a Marantz PMD 671 solid-state audio recorder, reading from a script consisting of pairs of sentences. The first sentence in the pair ended with a semi-colon and the speaker was asked to produce a "neutral" intonation pattern ending with a continuation rise. The second sentence in each pair never repeated adjectives or nouns (other than *square*) from the first sentence, and the speaker was asked to produce a "neutral" intonation pattern. As a result, there were small, non-contrastive pitch accents on each of the content words in the second sentence. After recording, the sentences were re-ordered to construct the actual pairs of sentences used in the experiment. For example, sentences from the excerpt of the script shown in (7) corresponded to the actual trial shown in (8).

```
(7) Put the small taxi above the triangle; now put the big hanger to the
left of the circle.
    Put the small egg above the heart; now put the big taxi to the right of
the diamond.
```

(8) Put the small taxi above the triangle; now put the big taxi to the right of the diamond.

This recording procedure produced instructions whose intonation did not disambiguate the target object. Finally, in instructions such as (9), the first noun and post-nominal modifier were uttered in a single intonation phrase, so that this instruction would not give rise to a garden-path interpretation in which *in the yellow square* was misinterpreted as the goal of the movement.

(9) Put the object in the yellow square above the triangle;

Procedure—Throughout the experiment, the participant wore an EyeLink II head-mounted eye tracker, which provided point-of-gaze screen coordinates every 4ms. Participants were told to follow pre-recorded instructions by clicking and dragging objects on the computer screen. The experiment began after the eye tracker was calibrated using a standard nine-point calibration procedure. Before each trial, a blank screen with a black cross in the middle was displayed. The participant clicked on the cross to begin the trial. The display for the trial was then shown, and after a two-second delay, the first instruction played. The second instruction played as soon as the participant correctly followed the first instruction by dragging an object to a new position with the mouse. The entire experiment lasted about twenty minutes.

Results—Eye movements were analyzed using EyeLink software to parse the eye-position record into three categories: saccades, fixations, and blinks. A look to a pictured object was defined from the onset of the saccade to the end of the fixation on that object. Twelve of the 512 critical trials were excluded because the participant did not initially click on the correct target object during one of the instructions.

We first describe the pattern of results using proportion-of-fixation plots (Allopenna, Magnuson & Tanenhaus, 1998). Unless otherwise noted, our descriptions are supported by the statistical analyses, which follow. For all figures, we only plot proportions of fixations to the targets and competitors, aligned to the onset of the adjective in the second instruction (time 0). Vertical lines indicate the average duration of each region: 220 ms for the pre-adjective and adjective regions and 290ms for the noun region. In the analyses, the regions are offset from each adjective using 200 ms as an estimate of the time needed to program and launch an eye-movement (Hallet, 1986). As described earlier, the target is the referent of the direct object in the instruction (e.g. the small candle in the Same-Set conditions and the

small tie in the Different-Set conditions) and the competitor is the member of the other contrast set which is the same size as the target (e.g. the small tie in the Same-Set conditions and the small candle in the Different-Set conditions).

Figures 2 and 3 plot proportions of fixations over time during the second instruction for the four Mention conditions, i.e. the conditions in which the first instruction contained a reference to a member of a contrast set described with a scalar adjective.

Figure 2 suggests an apparent, but as it turns out, statistically non-significant, baseline preference for the mentioned contrast set in the Mention/Move/Same-Set condition. We return to this point shortly. In both Figures 2 and 3, there is a clear bias towards the small candle (i.e. the small member of the previously mentioned contrast set) during the adjective *small*, regardless of whether it was the eventual target, as in the Same-Set condition, or the competitor, as in the Different-Set condition. Another way of observing this bias in the graphs is to note that looks toward the target diverge from looks toward the same-size competitor noticeably later in the Different-Set conditions (where the target was the small tie) than in the Same-Set conditions.

In contrast, Figure 4 shows that manipulating an object which is described without a scalar adjective does not give rise to any systematic bias towards other objects in the display. The results from the conditions shown in Figure 4 are comparable to those shown in Figure 5, in which no contrast set members are moved or mentioned during the first instruction, and again there is no systematic bias towards an object in the display during the second instruction.

We calculated target ratios during each region for each trial for each subject by dividing the proportion of fixations to the target by the proportion of fixations to the target and the competitor (Salverda, Dahan, Tanenhaus, Crosswhite, Masharov & McDonough, 2007); recall that the "competitor" was the object in the display that was the same size as the target. We defined three regions based on words in the instruction. The pre-adjective region was equal to the size of the adjective region and ended 200 ms after the actual onset of the adjective. The adjective region spanned from 200 ms after the actual onset of the adjective in the second instruction to 200 ms after offset of the adjective and the noun region spanned from 200ms after the offset of adjective to 200 ms after the offset of the noun. Regions were calculated individually for each trial based on the actual length of the words in the trial. Data points were collected every 4ms and each time point at which a participant was fixated on a given object was counted as a "look." Conditions where the participant did not fixate on the target or the competitor during the region on any of the trials were counted as 0.5 to reflect an equal likelihood of looking at either of the objects. To avoid issues inherent to proportional data, we performed our analyses on both the participant and item ratios and on ratios that were quasi-logit transformed (Jaeger, 2008). There were no differences in significance levels between the two analyses, so for ease of exposition, we will report the more intuitive untransformed ratios.

We calculated target ratios for the Pre-Adjective, Adjective and Noun Regions and conducted a 3 (Region) $\times$ 2 (Mention) $\times$ 2 (Move) $\times$ 2 (2nd Target Type) ANOVA on these ratios. There was a reliable effect of Region ($F1(2,14)=23.206$, MSE=0.065, $p<0.001$; $F2(2,30)=49.712$, MSE=0.047, $p<0.001$; $minF'(2,27)=15.821$, $p<0.001$) as expected, such that the target ratios were larger in the Noun region ($m=0.753$, SE=0.032, CI=±0.068) than in the Adjective region ($m=0.519$, SE=0.031, CI=±0.066) and the Pre-adjective region ($m=0.507$, SE=0.025, CI=±0.054). There was also a reliable main effect of 2nd Target Type ($F1(1,15)=7.904$, MSE=0.205, $p<0.02$; $F2(1,31)=26.429$, MSE=0.101, $p<.001$; $minF'$ (1,24)=6.084, $p<0.03$), which reflects a bias towards the second member of a contrast set

evoked in the first instruction (i.e. the small candle in our example). For purposes of exposition, we will henceforth refer to the unmentioned member of the previously evoked contrast set as "the candle." The target ratio was greater in the Same-Set conditions (m=0.658, SE=0.029, CI=±0.062) than in the Different-Set conditions (m=0.528, SE=0.035, CI=±0.076).

The time-course differences illustrated in the Figures are reflected in a Region by Mention by 2nd Target Type interaction ($F1(2,14)$=4.488, MSE=0.046, p<0.04; $F2(2,30)$=4.912, MSE=0.040, $p<0.02$; $minF'(2,37)$=2.345, p=0.11). In order to look at this interaction more closely, we performed a series of sub-analyses. First, we analyzed each region individually by conducting 2 (Mention) × 2 (Move) × 2 (2nd Target Type) ANOVAs. There were no significant effects in Pre-Adjective region (F's <1.0), indicating that there were not any significant baseline differences. Although there were no significant baseline effects during the Pre-Adjective region, there are numerically more looks to the target in the Same Set condition, which raises a concern that some effects or interactions with Target Type might be due to baseline differences that began in the pre-adjective region. However, contingent analysis in which we excluded trials where the participant was looking at the target or the competitor at the beginning of the Adjective region showed the same pattern as the full analyses (F's <1.0).

Crucially, the bias towards the candle was reflected in a Mention by 2nd Target Type interaction in both the Adjective region ($F1(1,15)$=25.438, MSE=0.060, p<0.001; $F2(1,31)$=13.850, MSE=0.079, p<0.001; $minF'(1,46)$=8.96, p<0.005) and Noun region ($F1(1,15)$=29.130, MSE=0.049, p<0.001; $F2(1,31)$=10.978, MSE=0.110, p<0.003; $minF'(1,45)$=7.97, p<0.01).

Separate sub-analyses of the Mention and No-Mention conditions confirmed that the bias towards the candle was only present in the Mention conditions, where the contrast set was actually mentioned during the first instruction. A 3 (Region) × 2 (Move) × 2 (2nd Target Type) ANOVA on just the Mention conditions revealed a reliable main effect of 2nd Target Type ($F1(1,15)$=26.342, MSE=0.157, p<0.001; $F2(1,31)$=62.477, MSE=0.075, p<0.001; $minF'(1,28)$=18.529, p<0.001), such that the target ratios were larger in the Same-Set conditions (m=0.732, SE=0.026, CI=±0.055) than in the Different-Set conditions (m=0.438, SE=0.043, CI=±0.092). In contrast, for the No-Mention conditions, the effect of 2nd Target Type did not approach significance (F's <1.0).

Discussion—The first goal of this experiment was to determine whether the contrast set evoked by mentioning *the big candle* affects the subsequent interpretation of referring expressions with scalar adjectives. There clearly was a bias, as predicted by the linguistic mention hypothesis, such that *the small candle* (i.e. the other member of the evoked contrast set) was more available as a referent than the small member of an unmentioned contrast set.

The second goal of the experiment was to determine whether the cross-sentential effect of evoking a contrast set is caused by the linguistic mention of a referring expression that contains a scalar adjective, as opposed to non-linguistic factors that typically co-occur with such a referring expression. Since the bias we observed in Experiment 1 occurred even when the previously mentioned contrast set member was not moved by the participant (see again the Mention, No-Move conditions in Figure 3), we can conclude that the participant's previous experience in manipulating an object in the display is not the cause of the effect. Furthermore, the effect disappears when the previously mentioned contrast set member is described by location (see again the No-Mention, Move conditions in Figure 4), suggesting that the previous mention of a scalar adjective is crucial to the effect.

Experiment 1A

At least one alternate explanation of the results of Experiment 1 remains viable that does not require a theoretical model in which information introduced by linguistic mention is marked as such: perhaps the bias towards the candle is induced by greater frequency of looks to the mentioned contrast set during the first instruction. Indeed, Sedivy et al. (1999) found that after a scalar reference (e.g., *the big candle*), listeners are more likely to look at the entity evoked by the mentioned contrast (the small candle) than another small object in the display. Before concluding that the bias towards the candle during the second instruction reflects the special status of the evoked contrast set in the discourse model, it is important to rule out an explanation based solely on the participant's previous looks. This alternate explanation would lend itself well to a theoretical model in which processing is mediated by a conceptual representation that does not encode linguistic mention.

In Experiment 1A we examine the frequency-of-looks explanation by using a post-nominal relative clause to disambiguate the referent of the first instruction (e.g. the big candle referred to using *the candle that's near the red square...*). Following Eberhard et al. (1995), we expect participants to look more at both candles than at other objects in the display during this first instruction. If the "mention" bias is due to visual attention having been directed at the small candle during the first instruction, then we should observe a bias towards the candle compared to another small object, which is also part of a potential contrast set (e.g. big and small tie), even though it has not been evoked by prior mention using a scalar adjective. The conditions that we will present as Experiment 1A were designed with this goal in mind. These conditions were embedded in Experiment 2, which is described in more detail later. For ease of exposition, however, we postpone discussing the full design until we introduce Experiment 2.

METHODS

Participants—Sixteen participants were recruited from the University of Rochester undergraduate community and paid for their participation.

Materials—As in Experiment 1, participants were shown a square grid on a computer display. In this experiment, each display consisted of a 4 by 4 square grid such as the one shown in Figure 6. Three of the four central cells had colored backgrounds, and the fourth central cell contained a picture of a trash can. The location of the colored cells and the trash remained constant throughout the experiment. The trash can is relevant only to Experiment 2, discussed in the next section.

In each trial, the display contained four objects, initially displayed in the corners. All of these objects could be moved by the participant. In all critical trials, two of the objects in the display were identical except that they contrasted in size (as shown in Figure 6). The contrast pairs were either vertically or horizontally aligned, and no objects were used in more than one trial.

During each trial, participants heard a sequence of two instructions. These instructions were recorded by a female native speaker of English, using the procedure described for Experiment 1. In each pair of instructions, the first sentence was produced with a "neutral" intonation pattern with a small non-contrastive pitch accent on each content word, and ended with a continuation rise. The second sentence was also produced with a "neutral" intonation pattern with a small non-contrastive pitch accent on each content word.

Two types of critical trials were constructed. In each critical trial, the first instruction asked the participant to move an object, which was described using a basic-level noun and locative

post-nominal modifier. The destinations for these movements were the colored squares in the center of the display. The second instruction asked the participant to move an object of a different size from the first target. In the Same-Set condition, participants moved two objects from the same contrast pair, while in the Different-Set condition, participants moved objects from different contrast pairs. Sample instructions for the critical trials are illustrated in Table 1 above. While processing the first instruction of these conditions and looking at the display shown in Figure 6, participants should spend more time looking at the candles than the ties. Therefore, if the cross-sentential bias observed in Experiment 1 was caused by visual attention, the same bias will be observed when participants hear *small* in the second instruction. On the other hand, if the bias observed in Experiment 1 was caused by the previous linguistic mention of a scalar adjective, then no bias is expected during *small*.

There were four experimental trials for each display/instruction type, for a total of eight experimental trials. The nouns describing the objects in the displays were matched in frequency, did not share phonological onsets, and did not begin with /b/ or /s/. The initial locations of the objects, as well as the locations the objects were moved to, were evenly distributed.

There were forty filler trials designed to prevent the participants from recognizing contingencies on the eight critical trials. As in Experiment 1, from the point of view of the participant, at the beginning of each sentence, all four objects were equally likely to be the target, and the target was equally likely to be described by location (e.g. *the candle that's near the yellow square*) or by size (e.g. *the big tie*). Two lists were created so that critical items were used in both conditions.

Procedure—The procedure was the same as Experiment 1 except that participants were told that they would be asked to move objects either into colored squares or into the trash. Participants were told that once an object was in the trash, it could not be retrieved.

Results—Eye movements were analyzed using the same method as for Experiment 1. One of the 128 critical trials was excluded because the participant did not initially click on the correct target object. The logic of this experiment depends on inducing more visual attention to one set of objects than the other during the first instruction. To verify that this was indeed the case, we calculated the proportion of fixations to each of the objects during a region that spanned from 200ms after the noun onset of the first instruction (the earliest point at which listeners would be more likely to look at the candles than the ties) to the end of that part of the trial. We compared the proportion of fixations for the objects that would be the targets and competitors in the second instructions. One of these objects, which for convenience we will call the "attended object," could be described with the same noun as the target of the first instruction (e.g., the small candle). During the second instruction, the attended object was the target of Same-Set trials and the competitor of Different-Set trials. The competitor of Same-Set trials and the target of Different-Set trials during the second instruction was the same size as the attended object.

As expected, from 200ms after the first noun onset of the first instruction to the end of the trial, participants looked more at the attended object (M=0.091, $S.D.$=0.05) than at the other same-size object (M=0.040, $S.D.$=0.03). A paired two-sample t-test confirmed that this difference was reliable (t=4.488, p<0.001). The participant's experience while listening to and carrying out the first instruction thus creates a situation in which the participant's visual attention has been more focused on just one of the objects that is consistent with the scalar adjective in the second instruction (*small*). This differs from the pattern of looks that has been documented across numerous studies whenever there are two contrast sets and one of the contrast sets is mentioned with a prenominal scalar adjective at the beginning of the trial.

Under these conditions, participants typically do not shift gaze to any of the potential referents until they hear the noun (Grodner & Sedivy, 2011; Hanna & Tanenhaus, 2004; Hanna, et al., 2003; Heller, et al., 2008; Nadig & Sedivy, 2002; Sedivy, 2003).

Figure 7 shows the proportions of fixations to the targets and competitors over time during the second instruction. As before, these proportions are aligned to the onset of the adjective in the second instruction (time 0). The female speaker used for Experiment 1A & Experiment 2 had a slightly slower speech rate than the speaker used for Experiment 1, so the average length of the Pre-Adjective and Adjective region was 290ms and the average length of the Noun region was 350ms.

Recall that the mention effect in Experiment 1 was reflected by more looks to members of the evoked contrast set, resulting in a clear bias towards the target for the Same-Set conditions compared to the Different-Set conditions. There is not a hint of this pattern in Experiment 1A. Figure 7 shows that participants fixated on the target of the second instruction equally quickly in both conditions. There was a slight baseline advantage for the competitor in the Same-Set condition (a tendency in the opposite direction of the results of Experiment 1), suggesting that the previous reference to a contrast member by location did not make the mentioned contrast set more salient in the discourse model. Neither the main effect of, nor any interactions with Target Type approached significance (F's <1.0).

Discussion—Taken together, Experiments 1 and 1A establish a cross-sentential effect of scalar adjectives: mentioning a referring expression with a scalar adjective makes an entire contrast set more salient, facilitating subsequent reference to other members of that contrast set. Experiment 1 showed that the bias towards a previously mentioned contrast set was not caused by the participant's experience of manipulating one object in a display; Experiment 1A showed that the bias was not caused by the frequency of looks to different objects in the display prior to the referring expression with a scalar adjective in the second instruction. This leads us to conclude that the bias that we observed is caused by the linguistic mention of a referring expression containing a scalar adjective. These results are clearly consistent with a theoretical model in which information evoked by linguistic mention is differentiated by information that arises from another source. The results are more difficult to reconcile with a theoretical model in which information is rapidly integrated into an undifferentiated conceptual representation that does not track the source of information.

Experiment 2

Interpreting the results of Experiments 1 and 1A as an effect that depends primarily on previous linguistic discourse leads to a strong prediction: previous mention of a referring expression with a scalar adjective should increase the salience of other members of the contrast set even if the referent of the first NP is no longer visible, thus eliminating a visually salient contrast set. Experiment 2 was designed to test this prediction by dissociating the prior mention of a contrast set from from its visual presence. In the critical trials in this experiment, a previously mentioned member of a contrast set removed from the display,,creating a situation in which only an unmentioned contrast set is fully visually co-present.

Experiment 2 introduces conditions in which the referent of a noun phrase that evokes a contrast set is subsequently removed from the display by placing it in the trash. For example, in (10a), with the display in Figure 6 above, *the big candle* evokes a contrast set consisting of candles that contrast in size. In (10b), a continuation of the same discourse, the big candle is no longer visible; the question is whether the contrast set of candles nevertheless has a special status in the discourse model. If the linguistic mention hypothesis is correct, the

processing of *small* in (10b) should be influenced by the previous mention of the noun phrase *the big candle,* even though the big candle itself is neither visible nor available to be moved. That is, when the scalar adjective is the second instruction is being processed, the only visually co-present contrast set involves the unmentioned entity; the singleton is a contrast member solely because it was evoked by linguistic mention.

```
(10) a. Put the big candle in the trash; …
     b. …now put the small…
```

We used the trash manipulation to evaluate the linguistic mention hypothesis in two ways. The first and strongest comparison was between trials that arrived at the same display from different initial displays. In some trials, the initial display contained two contrast pairs (e.g. ties, candles), as in Figure 8, Panel A. In other trials, the initial display contained one contrast pair (the ties) and two singleton objects (small candle, big package), as in Figure 8, Panel B. Participants were first instructed to discard an object (Put the big candle in the trash/Put the package in the trash), thus creating a display that looked like Figure 8, Panel C.

Notice that in both of these trial types, the display at the beginning of the second instruction contains an unmentioned contrast set (the ties) and a singleton object (the small tie). In conditions which arrive at this display from the initial display in Figure 8, Panel B, there is just one contrast set throughout the trial, and we therefore expect a large *basic contrast effect* at the adjective for an instruction beginning *Now put the small...*, with a strong bias towards the small tie. In conditions which arrive at the display in Figure 8, Panel C from an initial display with two contrast sets, one might expect to see the same *basic contrast effect* – after all, there is only one contrast set visible in the display. But if, as we predict, prior mention of the big candle creates a contrast set in the discourse model, then we should see a bias towards the small candle. In effect, this would be a reverse basic contrast effect: a bias towards a singleton in a display with only one visible contrast set (the two ties).

The second way we evaluated the linguistic mention hypothesis involved trials beginning with the display in Panel 8A, in which the target of the first instruction was a contrast member described with a scalar adjective. We compared trials in which a contrast member was placed in the trash (and therefore removed from the display) and trials in which a contrast member was moved to a colored square in the center of the display. The latter trials replicated the Mention, Move conditions of Experiment 1. Here we predict a similar effect in both trial types, regardless of whether the mentioned object was moved within the display (as it was in Experiment 1) or moved into the trash.

METHOD

Participants—Sixteen participants were recruited from the University of Rochester undergraduate community and paid for their participation.

Materials—This experiment was conducted in conjunction with Experiment 1A; trials from both experiments were interleaved. Sample displays used in Experiment 2 are illustrated in Figure 8, and sample instructions are shown in Table 2.

The experiment had eight conditions that differed on three dimensions: the number of contrast sets in the initial display (One-Contrast Display vs. Two-Contrast Display), the type of action performed by the participant after the first instruction (Center vs. Trash), and the second target type (Same-Set/Singleton vs. Different-Set/Contrast). In One-Contrast Display conditions, the initial display contained one contrast set, one large singleton object and one small singleton object, as shown in Figure 8, Panel B. In Two-Contrast Display conditions, the initial display contained two contrast sets, as shown in Figure 8, Panel A. In Trash

conditions, the participant's first action in the trial was to put an object into the trash, after which that object disappeared and a sound effect was played to suggest the trash was being emptied. In Center conditions, the participant's first action in the trial was to put an object into one of the colored cells in the center of the display, where it remained visible and could be moved later in the trial. Finally, we manipulated the second target type as in Experiments 1 & 1A. For the Same-Set/Singleton conditions, the second target is the unmentioned singleton item in the One-Contrast Display conditions and a member of the previously mentioned contrast set in the Two-Contrast Display conditions (e.g. the small candle). For the Different-Set/Contrast conditions, the second target is a member of the contrast set in the One-Contrast Display conditions and a member of the unmentioned contrast set in the Two-Contrast Display conditions (e.g. the small tie). We constructed four critical trials in each condition for a total of 32 critical trials.

In addition, we constructed 52 filler trials filler trials designed to remove contingencies from the critical trials. As in the previous experiments, from the participant's point of view at the beginning of each instruction, each object was equally likely to be the target. Different sets of critical items were used for the One-Contrast Display conditions and Two-Contrast Display conditions. For each set, four lists were created using a Latin square rotation through all of the conditions with that display type.

Procedure—The procedure was the same as Experiment 1 except that participants were told that they would be asked to move objects either into colored squares into the trash. Participants were told that once an object was in the trash, it could not be retrieved. To reinforce this, when participants moved an object to the trash, a sound effect played that suggested the trash was being emptied.

Results—Eye movements were analyzed in the same way as for Experiments 1 and 1A. Two of the 128 critical trials were excluded because the participant initially did not click on the correct target object. As in Experiments 1 and 1A, we will first present proportion of fixation plots, followed by statistical analyses.

We begin by considering the One-Contrast Display conditions and comparing them to the Two-Contrast Trash conditions. As can be seen in Figures 9 and 10, we replicated Sedivy et al.'s (1999) basic contrast effect in the One-Contrast Display conditions. Upon hearing the scalar adjective in the second instruction, participants expected the target to belong to a contrast set. In the Center conditions shown in Figure 9, where the first instruction had participants move a singleton to a colored square, participants have an early bias towards the member of the contrast set, regardless of whether this is the eventual target. That is, participants' fixations to the target more quickly diverge in the Different-Set/Contrast condition, where the target is the small candle, than in the Same-Set/Singleton condition, where the target is the small tie, so there is a large garden-path-like effect in the Same-Set/Singleton condition when the member of the contrast set is not the target. In the Trash conditions shown in Figure 10, where participants first moved one of the items into the trash, the same pattern of results is present and in fact, even more dramatic. This is most likely because there are only three moveable items during the second instruction.

Recall that the same display for the second instruction in the One-Contrast Trash conditions also arises for the second instruction in the Two-Contrast Trash conditions. Comparing these conditions provides a strong test of the linguistic mention hypothesis, which predicts that the basic contrast effect seen in the One-Contrast Trash conditions will reverse when the singleton in the display for the second instruction was created by discarding its contrast during the first instruction (Two-Contrast Trash conditions). The prediction from the linguistic mention hypothesis is clearly supported, as can be seen in Figure 11, which plots

the proportion of fixations to the target from these conditions on the same graph. In the One-Contrast conditions (shown with solid shapes in Figure 11), the target is clearly identified earlier in the Different-Set/Contrast condition, in which the target is a member of the only visible contrast set (i.e. the small tie), than in the Same-Set/Singleton condition. The opposite pattern is observed in the Two-Contrast conditions (shown with open shapes in Figure 11), in which the target is instead identified slightly earlier in the Same-Set/Singleton condition than in the Different-Set/Contrast condition. An ANOVA comparing target ratios for One-Contrast and Two-Contrast Trash conditions in the Adjective and Noun Regions reveals an interaction between Display Type and 2nd Target Type ($F1(1,15)=5.103$, MSE=0.109, $p<0.04$).

We now turn to comparisons of the Two-Contrast conditions. Recall that in the full design, Two-Contrast conditions (in which the initial display contained two contrast sets) were either Trash conditions, in which participants first discarded one object from the display, or Center conditions, in which participants first moved one contrast member to one of the central colored cells. The linguistic mention hypothesis predicts a bias towards the small candle in all Two-Contrast conditions, which should be reflected in earlier looks to the target when it was a member of a mentioned contrast set. Figures 12 and 13 are consistent with this prediction. In the Center conditions shown in Figure 12, we replicated the bias observed in Experiment 1: participants' looks to the target diverge from the competitor earlier in the Same-Set condition than the Different-Set condition, suggesting that the participants were expecting the target to refer to the same contrast set in the second instruction. A similar pattern of results holds for the Trash conditions, which are presented in Figure 13.

Although the data pattern for these conditions seems to shows a clear bias towards the mentioned set, when we conducted a 3 (Region) × 2 (First Move) × 2 (2nd Target Type) ANOVA on the Two-Contrast Display Conditions, there was only a reliable effect of Region ($F1(2,14)=20.917$, MSE=0.075, $p<0.001$; $F2(2,14)=30.116$, MSE=0.045, $p<0.001$; $minF'$ $(2,27)=12.34$, $p<0.001$), such that the target ratios were larger in the Noun Region ($m=0.775$, $SE=0.035$, $CI=\pm0.074$) than in the Adjective Region ($m=0.524$, $SE=0.042$, $CI=\pm0.090$) and Pre-Adjective Region ($m=0.514$, $SE=0.049$, $CI=\pm0.104$). Unlike in Experiment 1 where the mention effect resulted in interactions with Target type and Region as well as significant effects on Target on sub-analyses in the adjective and Noun regions, there were no effects of target-type.

A more detailed inspection of the figures reveals why there is discrepancy between the apparent data patterns in the figures and the statistical analyses. Note that the timing of the effects does not map clearly onto the regions defined by the adjective and the nouns. The bias bridges the adjective and noun regions: it begins late in the adjective region and ends in the middle of the noun region. In order to evaluate whether the apparent bias is statistically reliable, we created 100 ms regions beginning 100 ms after the actual onset of the noun (i.e., 100 ms before we would expect to see looks in response to information from the noun). For each region, we calculated target ratios for each subject and item in each condition. We then performed a one-way t-test for each region testing for whether the target ratio was above 50%, indicating that the target had diverged from the competitor. If there is a bias due to previous mention of a scalar adjective, then the Same-Set conditions should diverge earlier than the Different-Set conditions.

For the Center conditions, the target ratio in the Same-Set condition was above 50% beginning in the region 300-400ms after the noun onset ($t=2.509$, $p<0.03$), whereas in the Different-Set condition, the ratio wasn't above 50% until the region 400-500ms after the noun onset ($t=4.804$, $p<0.001$). In the Trash conditions, both the Same-Set and Different-Set target ratios are above 50% in the region 300-400ms after the noun onset (Same-Set $t=4.804$,

$p<0.001$; Different-Set $t=2.405$, $p<0.03$); however, the Same-Set condition is marginally significant in the prior region, 200-300 ms after the noun onset ($t=2.036$, $p=0.06$). This indicates that although it is a weaker effect than for the Center conditions, the bias seen in Experiment 1 is still present even when both members of a contrast set are no longer present in the display. In sum, then, there is a bias towards the mentioned contrast set for the Two-Contrast conditions, although this bias is weaker than it was in Experiment 1 for reasons that are unclear to us.

Discussion—Experiment 2 replicated the results of Experiment 1, showing that linguistic mention of a scalar adjective facilitates subsequent reference to members of the same contrast set. We extended the results further to show that this effect is observed even when the set is no longer entirely present in the visual domain. Crucially, we were also able to observe effects of the linguistic context that were separate from the visual context. In the Trash conditions, the One-Contrast Display and the Two-Contrast Display conditions were visually identical but participants' fixations were different as a result of the previous linguistic context. These results are clearly consistent with a theoretical model that tracks which information was introduced by linguistic vs. non-linguistic means, and again pose difficulties for a model in which processing depends on an undifferentiated conceptual representation.

General Discussion

Previous work on the processing of scalar adjectives has established that listeners assume that a noun phrase containing a scalar adjective will refer to a member of a contrast set. The two experiments presented here further demonstrate that uttering a referring expression with a scalar adjective makes all members of the relevant contrast set more salient in the discourse model, facilitating subsequent reference to other members of that contrast set. Most importantly, our results suggest that this discourse effect is caused by the linguistic mention of a scalar adjective. The effect is not caused merely by attending to objects that have been moved, because the cross-sentential effect of scalar adjectives is observed even when the listener did not actually move the first object described with a scalar adjective. Likewise, the effect does not arise merely as a result of looking more at some objects in the display than others, because the effect is not observed when listeners have spent more time looking at objects that were described using a noun and a locative modifier. Finally, the results of Experiment 2 showed that the bias towards members of a contrast set evoked by mention of a scalar adjective persists even when the visual domain does not contain both contrasting objects.

We do, however, want to acknowledge a possibility suggested by one of the reviewers, who argued that the size dimension was highlighted only in conditions in which a scalar adjective was mentioned. For example, in processing an expression such as *the candle in the yellow square*, participants' visual attention was drawn to both of the candles, but not explicitly to the size dimension. Could our results be explained by positing a conceptual model that tags different dimensions of contrast, but does not differentiate sources of information? This is a question that suggests an avenue for future research. One could draw attention to size non-linguistically, perhaps by using a display that would initially have two identical objects. One object would then expand and the other would shrink. This would presumably draw attention to a size difference non-linguistically. The linguistic mention hypothesis predicts that mentioning a scalar adjective would have an effect above and beyond any such perceptual manipulation–a prediction that will be important to evaluate in future research.

In spite of these new questions that arise about the processing of specific types of linguistic and visual information, in our view, the experiments discussed here showed strong evidence

that it is possible and necessary to tease apart the independent contributions of visual and linguistic sources of contextual information. The primary theoretical import of this result is to suggest that our model of sentence processing in context must be capable of distinguishing the source of contextual information. While it is of course generally agreed that language comprehension in context relies on the rapid integration of linguistic and visual information, theoretical models differ with respect to how this integration happens. A model in which linguistic and visual information is rapidly integrated into an undifferentiated conceptual representation will have difficulty explaining the effect we have found, in which linguistic information has a different effect than visual attention or physical interaction with a scene, and this effect persists over a sentential boundary. Our effect can be captured much more straightforwardly in a model that tags information introduced via linguistic mention.

Finally, we want to emphasize the methodological contribution made by these experiments, especially the "trash" paradigm. Distinguishing among types of representations poses a difficult challenge for psycholinguistic methods. For example, in studies with text, all information is introduced linguistically. Thus, one cannot easily tease apart effects of linguistic mention from conceptual availability of the mentioned entity. In standard visual world studies, the entities referred to linguistically are either visually co-present in a display or, in the blank screen paradigm (Altmann 2004), were visually co-present with the language and the effects of language are mediated by eye-movements to visual display. Again, then it is difficult to tease apart linguistic representation from the effects linguistic representation has on conceptual/perceptual representations. We believe that the methodology introduced here will allow cognitive scientists to ask new questions about roles of linguistic and visual sources of information in sentence processing, leading us towards a more nuanced model of how these sources of information are represented and integrated during sentence processing.

One of the questions that arises most naturally from our results is what determines the boundaries of linguistic context – how much previously mentioned material is relevant to on-line processing. In addition to simply asking whether this information fades away over time, we might wish to investigate the role of discourse relations, as indicated by discourse markers or by prosody, in determining which previous utterances influence on-line processing to which degree. It is also important to emphasize that the present studies focused only on pre-nominal scalar adjectives. It remains an open question whether other nominal modifiers give sets of objects a privileged status in the discourse model. For example, will the same effects occur with other pre-nominal adjectives? In addition, our results suggest that pre- and post-nominal modifiers behave differently; a more detailed investigation of this difference would help us to explore the interaction of syntactic structure with the establishment of sets of objects in a discourse model. Because languages differ with respect to which nominal modifiers occur before or after the noun, cross-linguistic investigation of these issues may also be of interest. Most generally, extensions of the trash paradigm should prove useful for investigating a range of context-dependent expressions that have been argued to be selectively sensitive to linguistic mention.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

Allopenna PD, Magnuson JS, Tanenhaus MK. Tracking the time course of spoken word recognition: evidence for continuous mapping models. Journal of Memory and Language. 1998; 38:419–439.

Altmann GTM. Language-mediated eye movements in the absence of a visual world: the 'blank screen paradigm.'. Cognition. 2004; 93:79–87.

Altmann GTM, Kamide Y. The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. Journal of Memory and Language. 2007; 57:502–518.

Altmann GTM, Kamide Y. Discourse-mediation of the mapping between language and the visual world: Eye movements and mental representation. Cognition. 2009; 111:55–71. [PubMed: 19193366]

Altmann GTM, Mirković J. Incrementality and prediction in human sentence processing. Cognitive Science. 2009; 33:583–609. [PubMed: 20396405]

Ariel, M. Accessing noun-phrase antecedents. Routledge; London: 1990.

Arnold JA, Tanenhaus MK, Altmann RJ, Fagnano M. The old and, theee, uh, new: Disfluency and reference resolution. Psychological Science. 2004; 9:578–582. [PubMed: 15327627]

Arregui A, Clifton C, Frazier L, Moulton K. Processing elided VPs with flawed antecedents. Journal of Memory & Language. 2006; 55:232–246. [PubMed: 17710192]

Bicknell K, Elman JL, Hare M, McRae K, Kutas M. Rapid effects of event knowledge in processing verbal arguments. Journal of Memory & Language. 2010; 63:489–505. [PubMed: 21076629]

Brown-Schmidt S, Byron D, Tanenhaus MK. Beyond salience: interpretation of personal and demonstrative pronouns. Journal of Memory and Language. 2005; 53:292–313.

Brown-Schmidt S, Konopka AE. Little houses and casas pequeñas: message formulation and syntactic form in unscripted speech with speakers of English and Spanish. Cognition. 2008; 109:274–280. [PubMed: 18842259]

Brown-Schmidt S, Tanenhaus MK. Watching the eyes when talking about size: An investigation of message formulation and utterance planning. Journal of Memory and Language. 2006; 54:592–609.

Chambers CG, Tanenhaus MK, Eberhard KM, Filip H, Carlson GN. Circumscribing referential domains in real-time sentence comprehension. Journal of Memory and Language. 2002; 47:30–49.

Chambers CG, Tanenhaus MK, Magnuson JS. Action-based affordances and syntactic ambiguity resolution. Journal of Experimental Psychology: Learning, Memory & Cognition. 2004; 30:687–696.

Cooper RM. The control of eye fixation by the meaning of spoken language. Cognitive Psychology. 1974; 6:84–107.

Dahan D, Tanenhaus MK, Chambers CG. Accent and reference resolution in spoken-language comprehension. Journal of Memory and Language. 2002; 47:292–314.

Eberhard KM, Spivey-Knowlton MJ, Sedivy JC, Tanenhaus MK. Eye movements as a window into real-time spoken language comprehension in natural contexts. Journal of Psycholinguistic Research. 1995; 24:409–436. [PubMed: 8531168]

Farmer TA, Anderson SE, Spivey MJ. Gradiency and visual context in syntactic garden-paths. Journal of Memory and Language. 2007; 57:570–595. [PubMed: 18037980]

Frazier L, Clifton CE. The syntax-discourse divide: Processing ellipsis. Syntax. 2005; 8:121–174. [PubMed: 17356681]

Frazier L, Clifton CE. Ellipsis and discourse coherence. Linguistics and Philosophy. 2006; 29:315–46. [PubMed: 16896367]

Glenberg, AM. Language and action: creating sensible combinations of ideas. In: Gaskell, G., editor. The Oxford handbook of psycholinguistics. Oxford University Press; Oxford, UK: 2007. p. 361-370.

Gordon PC, Grosz BJ, Gilliom LA. Pronouns, names, and the centering of attention in discourse. Cognitive Science. 1993; 17:311–347.

Grodner, D.; Sedivy, JC. The effect of speaker-specific information on pragmatic inferences. In: Pearlmutter, N.; Gibson, E., editors. The processing and acquisition of reference. MIT Press; Cambridge, MA: 2011. p. 239-272.

Groenendijk, J.; Stokhof, M.; Veltman, F. Coreference and modality. In: Lappin, S., editor. The Handbook of Contemporary Semantic Theory. Blackwell; Oxford: 1997.

Gundel JK, Hedberg N, Zacharski R. Cognitive status and the form of referring expressions in discourse. Language. 1993; 69:274–307.

Gunlogson, C. True to form: Rising and falling declaratives as questions in English. Routledge; New York: 2003.

Hallett, PE. Eye movements. In: Boff, KR.; Kaufman, L.; Thomas, JP., editors. Handbook of perception and human performance. Wiley; New York: 1986. p. 10.1-10.112.

Hankamer J, Sag I. Deep and surface anaphora. Linguistic Inquiry. 1976; 7:391–426.

Hanna JE, Tanenhaus MK. Pragmatic effects on reference resolution in a collaborative task: evidence from eye movements. Cognitive Science. 2004; 28:105–115.

Hanna JE, Tanenhaus MK, Trueswell JC. The effects of common ground and perspective on domains of referential interpretation. Journal of Memory and Language. 2003; 49:43–61.

Heim, I. Doctoral dissertation. University of Massachusetts; Amherst: 1982. The semantics of definite and indefinite noun phrases.

Heller D, Grodner D, Tanenhaus MK. The role of perspective in identifying domains of reference. Cognition. 2008; 108(3):831–836. [PubMed: 18586232]

Jaeger TF. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. Journal of Memory and Language. 2008; 59:434–446. [PubMed: 19884961]

Johnson, K. What VP ellipsis can do, what it can't, but not why. In: Baltin, M.; Collins, C., editors. The handbook of contemporary syntactic theory. Blackwell Publishers; Oxford: 2001. p. 439-479.

Kaiser E, Trueswell JC. The role of discourse context in the processing of a flexible word-order language. Cognition. 2004; 94(2):113–147. [PubMed: 15582623]

Kehler, A.; Ward, G. On the semantics and pragmatics of '*identifier so*.'. In: Turner, K., editor. The semantics/pragmatics interface from different points of view. Elsevier; Amsterdam: 1999. p. 233-256.

Kehler, A.; Ward, G. Constraints on ellipsis and event reference. In: Horn, LR.; Ward, G., editors. The handbook of pragmatics. Blackwell Publishers; Oxford: 2004. p. 383-403.

Kadmon, N. Formal pragmatics. Blackwell; Oxford: 2001.

Kaiser, E. Effects of anaphoric dependencies and semantic representations on pronoun interpretation. In: Devi, SL.; Branco, A.; Mitkov, R., editors. Anaphora processing and applications. Springer; Heidelberg: 2009. p. 121-130.

Kamp, H. A theory of truth and discourse representation. In: Groenendijk, J.; Janssen, T.; Stokhof, M., editors. Formal methods in the study of language. Mathematical Centre; Amsterdam: 1981.

Kennedy, C. Projecting the adjective: the syntax and semantics of gradability and comparison. Garland; New York: 1999.

Kennedy C, McNally L. Scale structure, degree modification, and the semantics of gradable predicates. Language. 2005; 81:345–381.

Kim C, Kobele G, Runner JT, Hale J. The acceptability cline in VP ellipsis. Syntax. (in press).

Kim, C.; Runner, JT. Semantics and linguistic theory 19 proceedings. The Ohio State University; 2009. Discourse structure and parallelism in VP ellipsis.

Knoeferle P, Crocker MW. The coordinated interplay of scene, utterance, and world knowledge: Evidence from eye tracking. Cognitive Science. 2006; 30:481–529. [PubMed: 21702823]

Ladd, DR. The structure of intonational meaning: Evidence from English. Indiana University Press; Bloomington: 1980.

Ladd, DR. Intonational Phonology. Cambridge University Press; Cambridge: 1996.

Matsuki K, Chow T, Hare M, Elman JL, Scheepers C, McRae K. Event-based plausibility immediately influences on-line language comprehension. Journal of Experimental Psychology: Learning, Memory, & Cognition. (in press).

Mayberry MR, Crocker MW, Knoeferle P. Learning to attend: A connectionist model of situated language comprehension. Cognitive Science. 2009; 33:449–496. [PubMed: 21585477]

McGurk H, MacDonald JW. Hearing lips and seeing voices. Nature. 1976; 264:746–748. [PubMed: 1012311]

Murphy GL. The interpretation of verb phrase anaphora: Influences of task and syntactic context. The Quarterly Journal of Experimental Psychology Section A. 1990; 42(4):675–692.

Nadig A, Sedivy J. Evidence of perspective-taking constraints in children's on-line reference resolution. Psychological Science. 2002; 13:329–336. [PubMed: 12137135]

Novick JM, Thompson-Schill S, Trueswell JC. Putting lexical constraints in context into the visual-world paradigm. Cognition. 2008; 107:850–903. [PubMed: 18279848]

Salverda AP, Dahan D, Tanenhaus MK, Crosswhite K, Masharov M, McDonough J. Effects of prosodically modulated sub-phonetic variation on lexical competition. Cognition. 2007; 105(2): 466–476. [PubMed: 17141751]

Schwarzschild R. GIVENness, Avoid F and other constraints on the placement of focus. Natural Language Semantics. 1999; 7(2):141–177.

Sedivy JC. Pragmatic Versus Form-Based Accounts of Referential Contrast: Evidence for Effects of Informativity Expectations. Journal of Psycholinguistic Research. 2003; 32(1):3–23. [PubMed: 12647560]

Sedivy JC, Tanenhaus MK, Chambers CG, Carlson GN. Achieving incremental interpretation through contextual representation: Evidence from the processing of adjectives. Cognition. 1999; 71:109–147. [PubMed: 10444906]

Snider, N.; Runner, JT. Structural parallelism aids ellipsis and anaphor resolution: Evidence from eye movements to semantic and phonological neighbors. Paper presented at Architectures and Mechanisms of Language Processing; York, England. 2010.

Spivey MJ, Tanenhaus MK, Eberhard KM, Sedivy JC. Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. Cognitive Psychology. 2002; 45:447–481. [PubMed: 12480476]

Spivey MJ, Tyler MJ, Eberhard KM, Tanenhaus MK. Linguistically mediated visual search. Psychological Science. 2001; 12:282–286. [PubMed: 11476093]

Tanenhaus MK, Carlson GN. Comprehension of deep and surface verb phrase anaphors. Language and Cognitive Processes. 1990; 5:257–80.

Tanenhaus MK, Spivey-Knowlton MJ, Eberhard KM, Sedivy JC. Integration of visual and linguistic information in spoken language comprehension. Science. 1995; 268:1632–1634. [PubMed: 7777863]

Trueswell JC, Sekerina I, Hill NM, Logrip ML. The kindergarten-path effect: studying on-line sentence processing in young children. Cognition. 1999; 73:89–134. [PubMed: 10580160]

Wardlow-Lane L, Ferreira VS. Speaker-external vs speaker-internal forces on utterance form. Do cognitive demands override threats to referential success? Journal of Experimental Psychology: Learning, Memory & Cognition. 2008; 34:1466–1481.

Wardlow-Lane L, Groisman M, Ferreira VS. Don't talk about pink elephants! Speaker's control over leaking private information during language production. Psychological Science. 2006; 17(4):273–277. [PubMed: 16623681]

Highlights

- We study the cross-sentential effect of uttering a scalar adjective.

- A scalar adjective facilitates subsequent reference to the same contrast set.

- This effect is caused by linguistic mention, not visual or perceptual experience.

- This suggests that linguistic mention plays a special role in language processing.

- Our paradigm can tease apart contributions of visual and linguistic information.
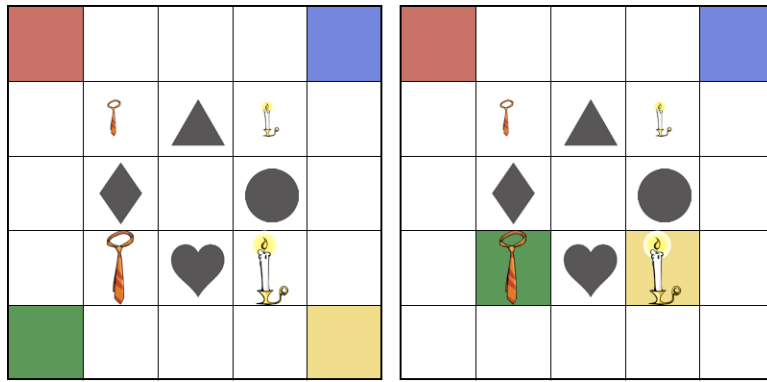
Figure 1.
Sample displays for Experiment 1. The display on the left was used for all conditions except the No-Mention, Manipulation conditions where the display on the right was used instead.
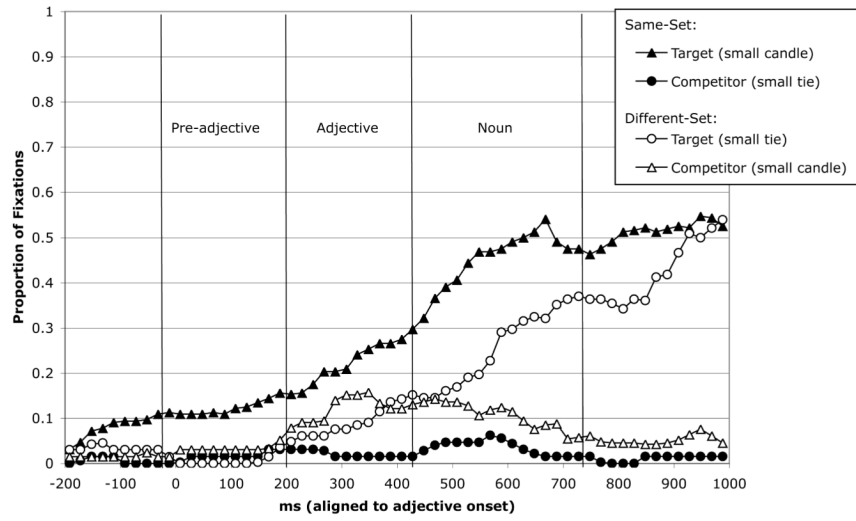
Figure 2.
Proportion of fixations to target vs. competitor during the second instruction: Mention, Move conditions.
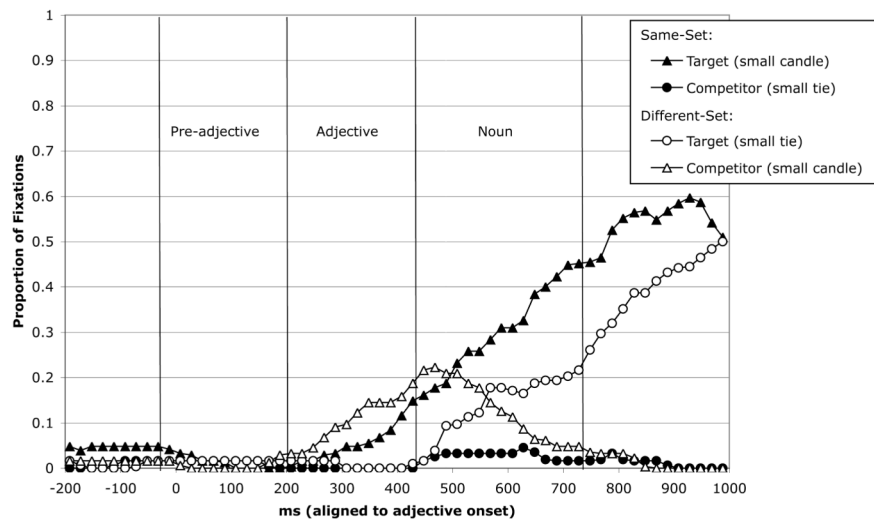
Figure 3.
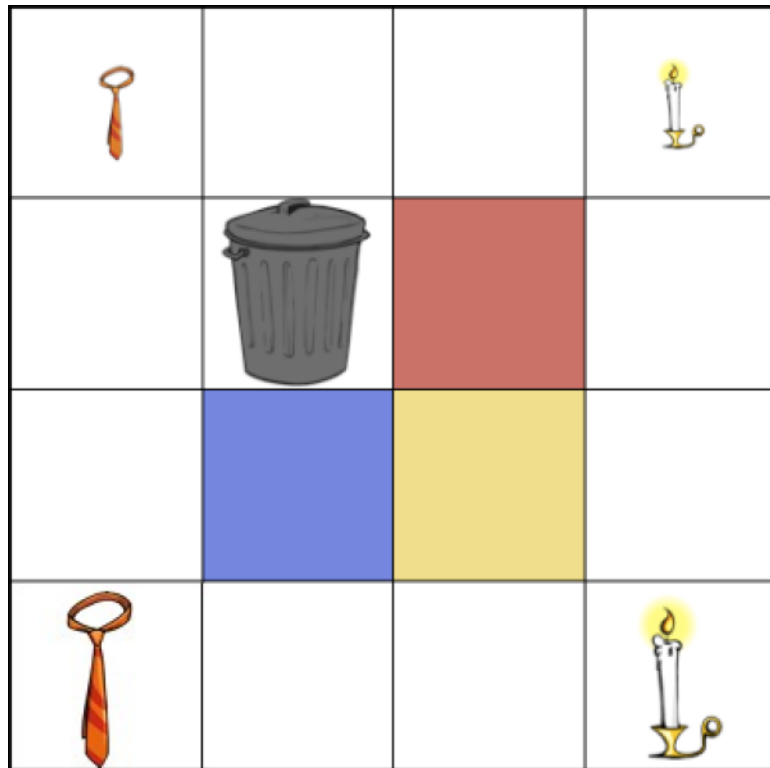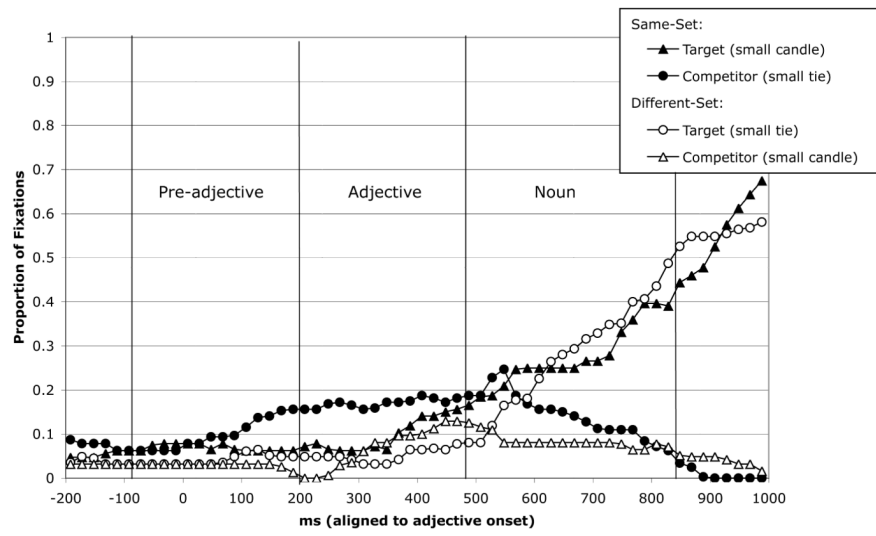Proportion of fixations to target vs. competitor during the second instruction: Mention, No-Move conditions.
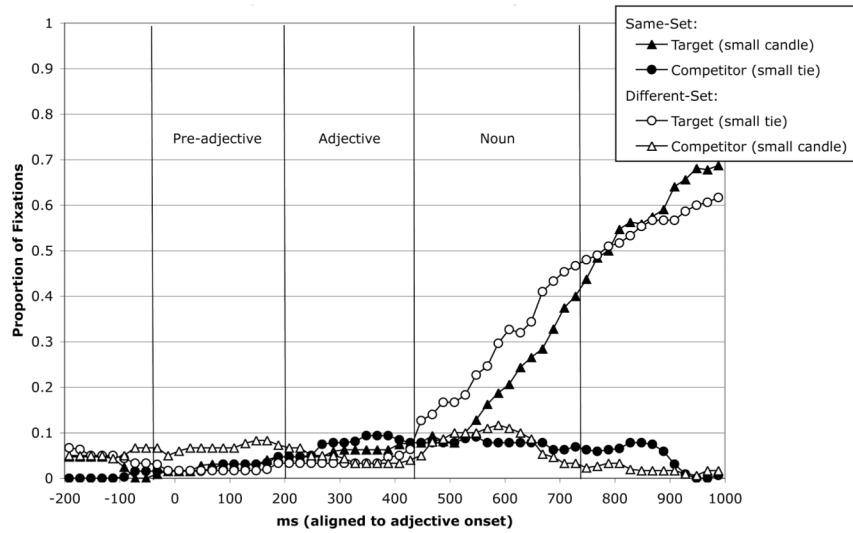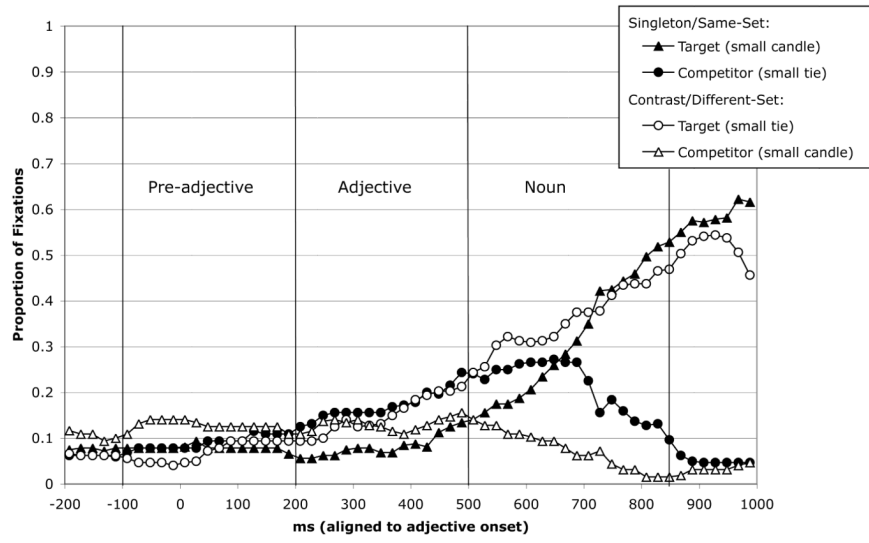
Figure 4.
Proportion of fixations to target vs. competitor during the second instruction: No-Mention, Move conditions.
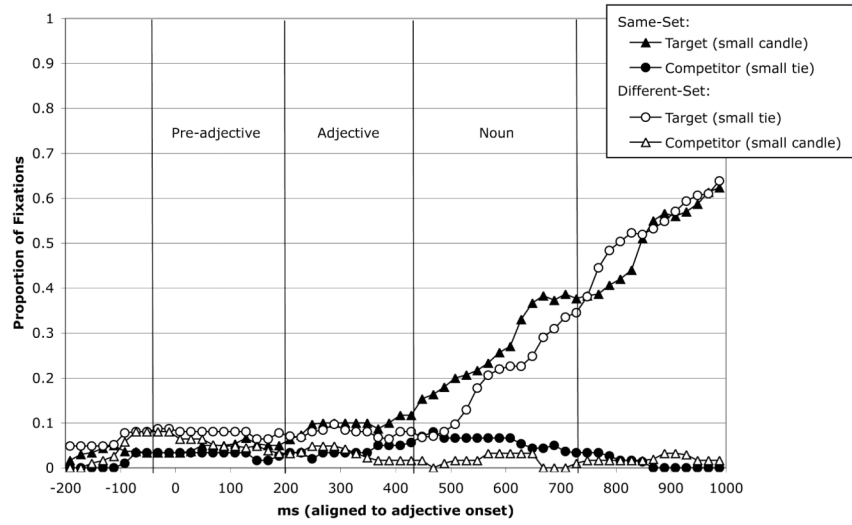
Figure 5.
Proportion of fixations to target vs. competitor during the second instruction: No-Mention, No-Move conditions.

Figure 6.
Display for Experiment 1A.

Figure 7.
Proportion of fixations to target vs. competitor during the second instruction: Experiment 1A.
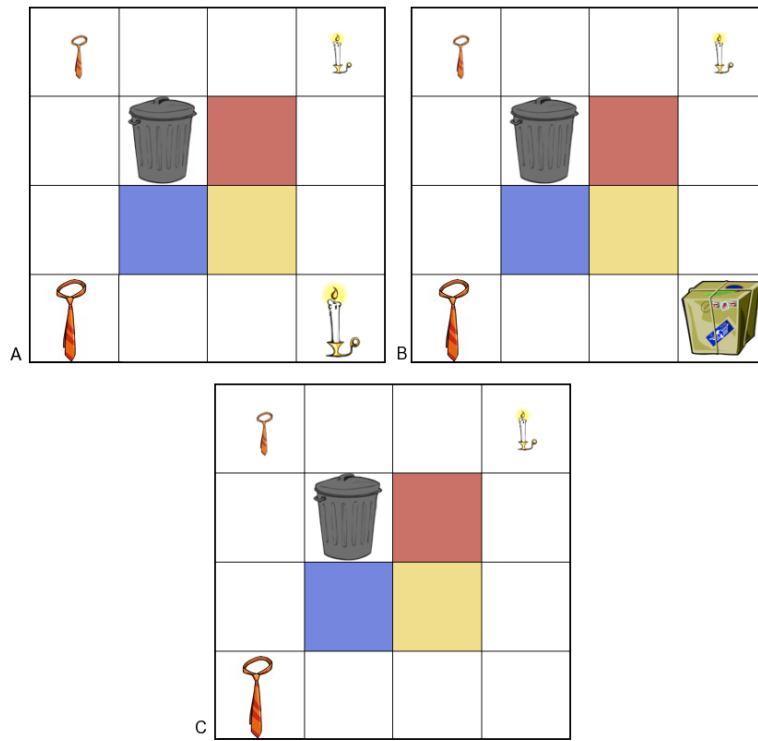
Figure 8.
Sample displays for Experiment 2 (A) Two-Contrast display at the first instruction. (B) One-Contrast display at the first instruction. (C) One-contrast and Two-Contrast displays at second instruction in Trash conditions.

Figure 9.
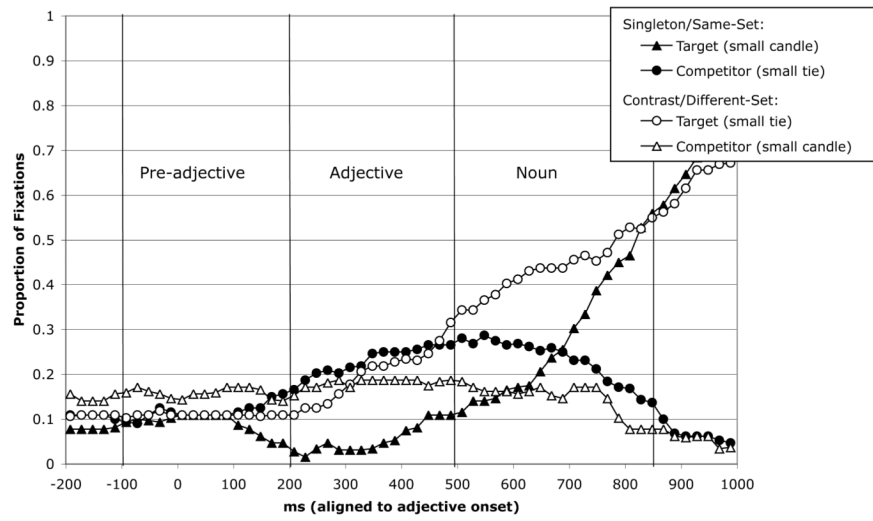Proportion of fixations to target vs. competitor during the second instruction: One-Contrast, Center conditions.

Figure 10.
Proportion of fixations to target vs. competitor during the second instruction: One-Contrast, Trash conditions.
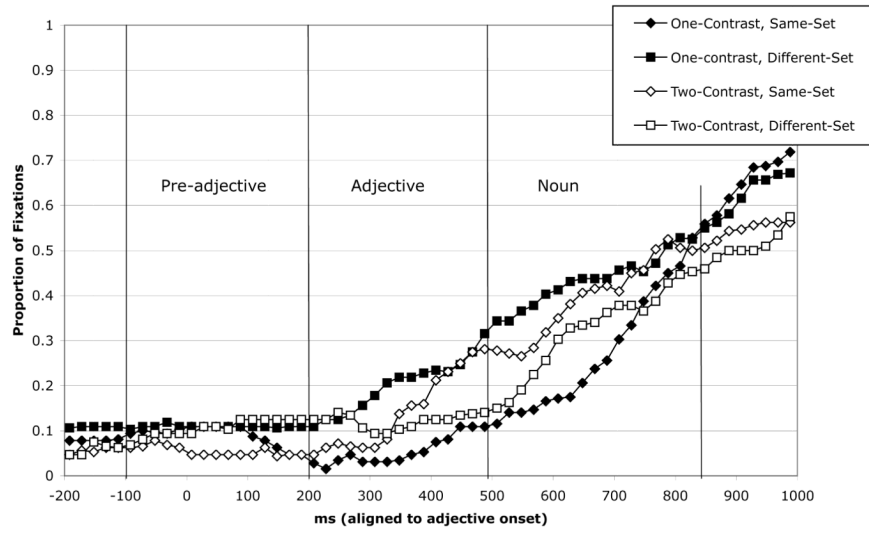
Figure 11.
Proportions of fixations to the target during the second instruction in the four Trash conditions.
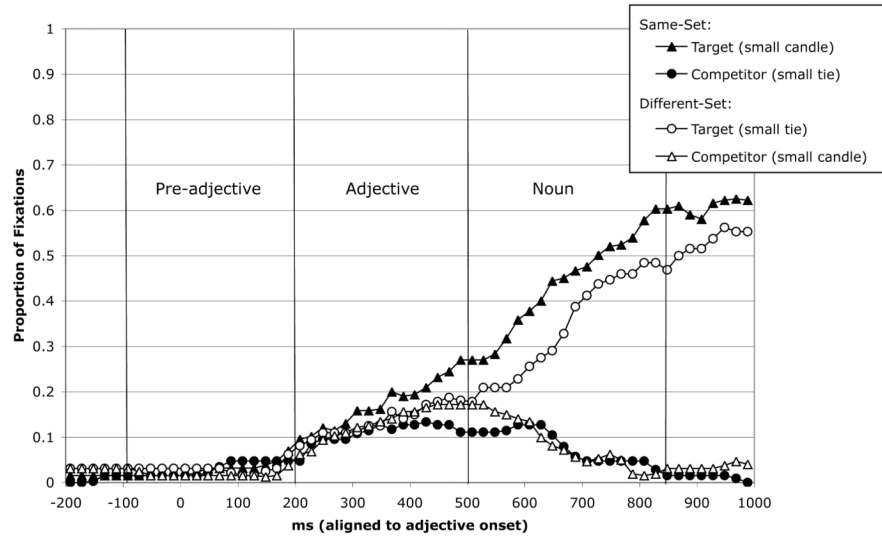
Figure 12.
Proportion of fixations to target vs. competitor during the second instruction: Two-Contrast, Center conditions.
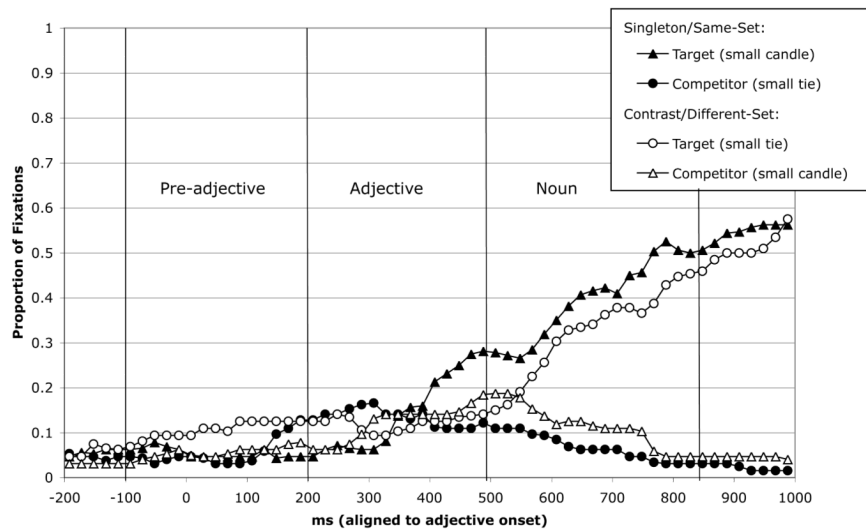
Figure 13.
Proportion of fixations to target vs. competitor during the second instruction: Two-Contrast, Trash conditions.

TABLE 1

Sample instructions for Experiments 1 & 1A

Condition			First Instruction	Second Instruction
Mention	Move	Same-Set	*Put the big candle above the heart;*	*Now put the small candle below the heart.*
		Different-Set		*Now put the small tie below the heart.*
	No-Move	Same-Set	*Put the circle below the big candle;*	*Now put the small candle below the heart.*
		Different-Set		*Now put the small tie below the heart.*
No-Mention	Move	Same-Set	*Put the object in the yellow square above the triangle;*	*Now put the small candle below the heart.*
		Different-Set		*Now put the small tie below the heart.*
	No-Move	Same-Set	*Put the diamond above the heart;*	*Now put the small candle below the heart.*
		Different-Set		*Now put the small tie below the heart.*
Experiment 1A		Same-Set	*Put the candle that's near the red square into the blue square;*	*Now put the small candle below the heart.*
		Different-Set		*Now put the small tie below the heart.*

TABLE 2

Sample instructions for Experiment 2

Condition			First Instruction	Second Instruction
One-Contrast Display	Center	Same-Set/Singleton	*Put the package into the blue square;*	*Now put the small candle below the heart.*
		Different-Set/Contrast		*Now put the small tie below the heart.*
	Trash	Same-Set/Singleton	*Put the package into the trash;*	*Now put the small candle below the heart.*
		Different-Set/Contrast		*Now put the small tie below the heart.*
Two-Contrast Display	Center	Same-Set/Singleton	*Put the big candle into the blue square;*	*Now put the small candle below the heart.*
		Different-Set/Contrast		*Now put the small tie below the heart.*
	Trash	Same-Set/Singleton	*Put the big candle into the trash;*	*Now put the small candle below the heart.*
		Different-Set/Contrast		*Now put the small tie below the heart.*