The Use of Lexical and Referential Cues in Children's Online Interpretation of Adjectives

The use of referential context in **children's** online interpretation of adjectives

Yi Ting Huang[1] & Jesse Snedeker[2]

1. The University of North Carolina at Chapel Hill
2. Harvard University

Abstract

Adult language comprehension is guided by both linguistic and contextual information. In contrast, recent work on syntactic parsing suggests that children typically fail to use contextual cues, like the number of potential referents for a noun phrase, during syntactic ambiguity resolution. We examine whether this cue influences children's interpretation of noun phrases with scalar adjectives (*big* and *tall*) which are typically produced only when there are two objects of the same kind in the discourse. Five-year-olds heard instructions like "Point to the big (small) coin," while their eye-movements were measured to displays containing 1 or 2 coins. They looked to the Target more rapidly in trials containing 2 referents, demonstrating that contextual cues rapidly influence children's real-time language comprehension.

Keywords: language processing; referential contrast; comprehension, scalar adjectives

1. Introduction

Our moment-to-moment interpretation of language depends not only by the words that we hear or read but also on the situation in which they occur. Take for example the fragment in (1).

(1)   I'll eat the pastry with…

If you heard this snippet while waiting in line at a Dunkin' Donuts, you would probably expect the speaker to complete the sentence with a modifier like *the sprinkles* or *chocolate icing*. With so many pastries around, a more specific description is called for. In contrast, if the same comment was made by a friend who had just been served dessert, you might expect it to end with an instrument like *my dinner fork* or *chopsticks*. With only one pastry in sight, a modifier would simply be redundant (Crain & Steedman, 1985; Altmann & Steedman, 1988).

As adults, we seamlessly integrate the linguistic information in the utterance with the extra-linguistic context in which it occurs. Our ability to coordinate multiple types of information during comprehension has led many to characterize the mature language system as rapid, incremental, and opportunistic in its use of information (MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell & Tanenhaus, 1994; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; Sedivy, Tanenhaus, Chambers, & Carlson, 1999). However the evidence to date suggests that children may have difficulties coordinating linguistic and contextual cues during language comprehension. Critically, when confronted with syntactic ambiguities like (1), children draw on lexical information to guide analysis but fail to use information about the number of potential referents to decide whether an ambiguous phrase is likely to be a modifier (Trueswell, Sekerina, Hill, & Logrip, 1999; Hurewitz, Brown-Schmidt, Thorpe, Gleitman, & Trueswell, 2000; Snedeker & Trueswell, 2004; Choi & Trueswell, 2006). While this disparity between adults and children is striking, the scope of the phenomenon is unclear. Are children's failures to use

referential information limited to syntactic parsing? Or are they unable to use context to make any predictions about noun phrase modification during real-time comprehension?

In this paper, we address these questions by examining another aspect of language comprehension which is sensitive to the set of possible referents: the interpretation of scalar adjectives. Adjectives, like the post-nominal modifiers described above, are typically used in situations where an unadorned noun would be insufficient. Thus they are more felicitous when there are at least two objects in the discourse context which are members of the same basic-level kind. This connection between adjectives and the referential context can support predictive inferences during language comprehension (Sedivy et al., 1999). Specifically, encountering an adjective (*tall*) should allow the listener to predict that the noun that follows will be one which would not have a unique referent in the context (two glasses, one tall and one short) – since if it were unique (one glass), the use of the adjective would not be necessary to pick out that object.

In the remainder of the Introduction, we will do three things. First, we briefly review recent studies on children's use of referential context during parsing and introduce three hypotheses that could account for their behavior. Next we review the linguistic properties of scalar adjectives, children's knowledge of these terms, and prior studies on the ability of adults to predict the upcoming noun on the basis of adjective use. Finally, we describe the present experiments and how they address these questions.

1.1. Children's failure to use referential context during parsing

Several decades of research in reading and spoken language processing has demonstrated that adults are able to use referential information to guide syntactic parsing (e.g., Altmann & Steedman, 1988; Tanenhaus et al., 1995; van Berkum, Brown & Hagoort, 1999; inter alia). When comprehenders encounter a potentially ambiguous phrase that follows a definite noun,

they are more likely to interpret it as a post-nominal modifier when the noun in question has more than one potential referent in the discourse context.

Young children, on the other hand, fail to use this information. For example, Trueswell and colleagues, following Tanenhaus et al. (1995), gave children and adults spoken instructions to move objects about on a table while their eye movements were recorded. The critical trials contained a temporary PP-attachment ambiguity, in which the verb's argument preferences strongly supported an initial analysis of "on the napkin" as a destination, see (2) below.

(2)   Put the frog on the napkin in the box.

In contexts with just one frog, adults initially looked over to the incorrect destination (the empty napkin), suggesting that they were misanalyzing the first prepositional phrase as a destination. But when two frogs were provided (one of which was on a napkin), the participants were able to immediately use the referential context to avoid this garden path, resulting in eye movements similar to unambiguous controls (e.g., "Put the frog that's on the napkin…").

In contrast, five-year-olds were unaffected by the referential context. In both the 1-referent and 2-referent contexts, children frequently looked at the incorrect destination, suggesting that they attached the prepositional phrase to the verb and interpreted it as a destination, regardless of the number of frogs. On over half of the trials, children actually performed an action that involved the incorrect destination, suggesting that they never revised this initial misanalysis. This robust failure to use information from the scene in the interpretation of prepositional attachments has been replicated under a variety of conditions and extended to other languages (Hurewitz et al., 2000; Snedeker & Trueswell, 2004; Choi & Trueswell, 2005).

Subsequent studies have ruled out several potential explanations for this pattern. First, children are clearly capable of interpreting ambiguous phrases as post-nominal modifiers. Like

adults, young children show robust preferences for the modifier analysis when the verb in the utterance supports this interpretation (i.e. *choose*) but they continue to show no effect of referential context (Snedeker & Trueswell, 2004). Second, their failure is not due to a global inability to use non-lexical information during parsing. Four- and five-year-olds are able to use the prosody of an utterance to interpret global attachment ambiguity (Snedeker & Yuan, 2008). Finally, children's failure to use referential context during parsing does not appear to stem from ignorance of the how post-nominal modification can be used to avoid referential ambiguity. Hurewitz and colleagues (2000) demonstrated that children will readily produce a post-nominal modifier when the context requires it. Five-year-olds who were asked to identify an animal in a 2-referent context (e.g., "Which frog visited Mrs. Squid's house?") readily produced situationally appropriate restrictive modifiers (e.g., "The frog on the napkin"). But just moments later, these same children misinterpreted parallel modifiers when asked to "Put the frog on the napkin in the box." This strongly suggests that even when children successfully encode the presence of two referents and the contrast between them, they fail to use this information to predict the likelihood of post-nominal modification during comprehension.

How can we account for this failure? Constraint-Based Lexicalist approaches offer one potential explanation (MacDonald et al., 1994; Trueswell & Tanenhaus, 1994). This account was initially advanced to explain comprehension in adults but has recently been extended to capture the developmental process in children (Trueswell & Gleitman, 2004). It posits that the language processing system has two critical characteristics: (1) an architecture that represents input at many different levels (prosodic, syntactic, semantic, phonological, discourse) and (2) a statistical mechanism that is highly attuned to the grammatical regularities of individual lexical items. Processing at each level of representation makes use of constraints from the other levels as well

as from stored lexical information to resolve ambiguity and make predictions about material that

has yet to come. While in principle, comprehension is sensitive to multiple sources of

information from the earliest stages of acquisition, in practice, the links between the various

linguistic representations must be acquired through experience and may change over the course

of development. The rapidity with which these links form will crucially depend upon the strength

of the correlation between the two phenomena. This provides a clear role for experience: even

after the acquisition of the representational systems is largely complete, children must still learn

how different levels of representation constrain one another during interpretation.

How well does the number of referents predict the use of a post-nominal modifier?

Surprisingly, evidence from referential communication tasks suggests there is only a weak

correlation between the two (Brown-Schmidt, Campana & Tanenhaus, 2002). Brown-Schmidt

and colleagues found that adults quite often utter bare definite noun phrases (e.g., "Pick up the

square") in the presence of multiple potential referents (e.g., many squares). However, their

listeners had little difficulty understanding these "ambiguous" noun phrases because the

participants' current goals and prior discourse made the referents clear. This suggests that the

actual learning problem is quite complex.   If the input does not support a direct mapping

between the number of objects in a situation and the presence of noun phrase modification, then

the child is confronted with the much trickier problem of tracking the relative salience of

different referents in the discourse model. This task would require knowledge of the speakers'

goals and an ability to rapidly update the discourse model with each new utterance and may

explain why children persistently fail to use referential context to guide parsing. Critically, the

Constraint-Based Lexicalist theory suggests that use of non-lexical information should emerge

early in cases where the cue reliably predicts the intended meaning of the utterance (Trueswell &

Gleitman, 2004). This highlights the need to explore children's ability to use contrast in contexts where it might be a more reliable predictor of language use.

1.2. Test case: Interpretation of scalar adjectives

In our current experiments, we turn to a case where the used in a modifier is highly correlated with the number of referents in the scene (Sedivy, 2003). Scalar adjectives like *big* and *tall* describe a class of terms which are typically interpreted relative to the nouns that they describe (Kennedy, 1999; Bierwisch, 1987). At the lexical level, their semantics specify a scale along which entities can be compared (e.g., size or height) and the relevant pole on that scale (e.g., greater or smaller along this dimension). However to extend and interpret these adjectives, language users must set a flexible standard of comparison to determine what values along that scale are sufficient to count as examples of the adjective. For example, to evaluate a statement like "The Hummer is a big car," one must know information not only about the size of a Hummer but also the size of cars in general. This information allows one to determine that a Hummer is a *big* relative to Honda Civics and Ford Escorts. However, the appropriate scale can shift dramatically when the same referent is evaluated relative to a different category. Thus the Hummer would not a *big vehicle* relative to trains, cruise ships, and Boeing 747s. Similarly, even within a single modified category like *big cars*, the standard of comparison can change with respect to a contextually defined comparison class (cars in Europe vs. cars in Texas).

Prior developmental studies have demonstrated that knowledge of the comparison class affects off-line judgments of adjective/noun combinations in children as young as three- and four-years of age (Barner & Snedeker, 2008; Syrett, Kennedy, & Lidz, in press). Similarly, psycholinguistic research has shown that adults quickly detect the presence of a comparison class and use this to make predictions about the upcoming referents of scalar adjectives (Sedivy et al.,

1999). Adult listeners were quicker to comprehend utterances containing *tall glass* in the presence of another contrasting referent of the same category (a short glass) but slower when then item was replaced with an unrelated object.

But how do children interpret adjectives during real-time comprehension? Previous studies have found that when three-year-olds were asked for the *blue car*, they abandoned their looks to a different colored competitor (red car or red house) shortly after adjective onset (Thorpe, Baumgartner, & Fernald, 2007). Similarly, five- to seven-year-olds who heard adjective noun combinations like *red cat* were quicker to shift their gaze to the correct referent when it was the only red item in the display (Sedivy, Demuth, Chunyo, & Freeman, 2000).   However, while these findings clearly demonstrate that children incrementally interpret adjectives (rather than waiting for the noun), they do not directly address the role of contrast sets since they examine simpler, non-scalar terms which are not typically interpreted contrastively (Sedivy, 2003).

In fact, when focusing specifically on scalar adjectives, previous research suggests that children's comprehension may be insensitive to the contrastive function of these terms. Nadig, Sedivy, Joshi, and Bortfeld (2003) extended the paradigm developed by Sedivy et al. (1999) to five- and six-year-olds and found that unlike adults, children's interpretation of *big car* was not facilitated by the presence of a same category contrast (a small car). Moreover, reference resolution in these 2-referent trials was about 600 ms slower compared to 1-referent trials in which the contrast item was replaced with an unrelated object (a baseball). While this difference demonstrates that children are sensitive to the number of referents in the displays, the direction of this disparity is opposite from what would be predicted by the referential implications of these terms. A closer inspection of the data reveals that children's reference restriction was primarily driven by the onset of the noun (*car*) rather than the adjective (*big*). This led to more

competition between potential referents in displays featuring two cars compared to those featuring only one. A similar pattern has been found in the on-line interpretation of non-scalar adjectives in young two-year-olds (Thorpe et al., 2007) and suggests that scalar adjectives may be subject to similar difficulties throughout the early school years. Accordingly, Nadig et al. (2003) concluded that "children may not yet have the processing capacity to successfully incorporate referential context" (pg. 577).

Nevertheless, the authors also noted that referential contrast did appear to facilitate interpretation of scalar adjectives in one important way: children in 2-referent trials made fewer spurious looks to a competitor that was of the same size as the correct target but was not paired with a contrast object (a big turtle). This suggests that children's real-time interpretations of scalar adjectives may be sensitive to the number of referents in the scene but this sensitivity was concealed by other aspects of the previous study. In the following section, we describe some of these differences and their possible effects on children's ability to use contrast from the scene.

1.3. The present study

In the following experiments, we employ a task that is similar in structure to those used by Sedivy et al. (1999) and Nadig et al. (2003) but we modify the materials in several key respects. Adults and five-year-olds were given instructions like "Point to the big coin" and their eye-movements were measured to visual display containing four items which varied in size and category membership (see Figure 1). These displays always featured a Target object that matched the adjective/noun combination (a big coin) and a Contrast object that differed in size. In the 2-referent trials, this item came from the same category as the Target (a small coin) while in the 1-referent trials, it came from a different category (a small button). The displays also

featured a `Competitor` that matched the Target in size but not by category (a big stamp) and an unrelated object that matched the Target in neither size nor category (a small marshmallow).

Critically, unlike previous studies, these displays always featured objects that were consistent with their real-world size/height (e.g., coins, buttons, stamps, marshmallows). The stimuli in the Nadig et al. (2003) experiment were "familiar household objects," however many of these objects were miniature models of items that are ordinarily quite large. The relative scale of these objects may have interfered with children's interpretation of the adjectives. For example, the presence of a same category contrast may be irrelevant in evaluating whether a 6" car is in fact a *big car* since the toy vehicle is by design much smaller than normal-sized cars. By using objects that fell within the range of sizes typical of real-world referents, we ensure that interpretation of the adjective is not be complicated by the question of scale.

We also tightly controlled for two additional aspects of the displays. First, we made certain that Contrast objects always differed in size from the Target in both 1-referent and 2-referent trials (small button and small coin). This ensures that any difference that emerges between these conditions can be specifically attributed to the category membership of the Contrast and from not other extraneous features. Second, we arranged the objects in a way that increased the likelihood of encoding the size difference between the Target and Contrast. Thus the Contrast was always placed to the left/right of the Target while the Competitor was always placed above/below it.

Finally, in Nadig et al. (2003) it was not clear that children's interpretations were incrementally sensitive to adjective meaning since looks to the Target only differed from the other objects after the onset of the critical noun. In our current study, we assess the processing the lexical semantics by determining when the meaning of the scalar adjective rules out referents that are incompatible with the specified pole. When children hear *big*, we would expect their

fixations to the big objects to increase, quickly exceeding their fixations to the small objects. Similarly, when children hear *small*, we would expect their fixations to the small objects to increase and exceed their fixations to the big objects.

To assess children's use of referential contrast, we examined whether the presence of a within-category contrast item facilitates interpretation of the adjective. We compared looks to the Target in the 2-referent trials versus those in the 1-referent trials. If children's interpretation of adjectives is sensitive to referential contrast, then we would predict facilitation of reference resolution in the 2-referent trials. If, however, interpretation depends solely on the numbers of items in the scene (or the number having the property encoded in the adjective), then we would expect no differences between the two conditions. Finally, if children primarily rely on the meanings of nouns to establish reference then we might expect greater looks to the Target for the 1-referent trials than the 2-referent trials following the onset of the noun.

In Experiment 1, we first examine the use of lexical semantics and referential information in adult interpretation of scalar adjectives. The goals of this experiment were three-fold. First, we wanted to replicate the contrast effects seen in the previous studies by Sedivy and her colleagues (1999, 2003). Second, we wanted to situate these contrast effects with respect to the use of lexical semantic information about the scalar adjectives. Finally, our experimental design provided tight controls for several features of the display that were somewhat different from those used in the previous studies. Thus additional data from adult participants was necessary to establish the expected pattern of performance in this task. In Experiment 2, we turn our attention to five-year-olds and explore how lexical semantics and referential contrast influences children's interpretation of scalar adjectives.

2. Experiment 1

2.1. Methods

2.1.1. Subjects

Thirty-two undergraduates at Harvard University participated in this study and received either course credit or $5 for their participation. All participants were native English speakers.

2.1.2. Procedure

Participants sat in front of an inclined podium divided into four quadrants (upper-left, upper-right, lower-left, and lower-right), each containing a shelf where objects could be placed. A camera at the center of the display focused on the participant's face and recorded the direction of their gaze while they were performing the task. A second camera, located behind the participant, recorded both their actions and the location of the items in the display. For every trial, the experimenter took out four objects from a bag and placed them each on a shelf in a pre-specified order. This was followed by a pre-recorded utterance which instructed participants to point to one of the objects. Once the participant pointed to an object, the trial ended, the objects were removed from the display, and the next trial began.

2.1.3. Materials

Scalar adjectives were selected from the size (*big, small*) and height scale (*tall, short*). Each item was rotated through the four conditions of a 2 x 2 design. The first factor, polarity, indicated whether the Target item was from the negative pole (*small, short*) versus the positive pole (*big, tall*) of the scale. The second factor, contrast, indicated whether the Contrast item belonged to the same basic-level category (2-referent trials) versus a different one (1-referent trials). On each trial, participants heard prerecorded commands like (3).

(3)   Point to the big coin.

These sentences were recorded by a female actor and each digital waveform was examined to ensure that the sentences had a consistently natural and unmarked prosody.

The objects featured in the visual displays consisted of 16 sets of four household objects. Within each set, objects were selected to match the relative scale of other members of the set. To ensure that these items were good exemplars of the adjective/noun combination, we conducted an object ratings task to see how participants perceived their size/height. A separate group of 36 participants were asked to rate how an object (e.g., big coin) compared to typical members of its category (e.g., coins) along a particular dimension (e.g., size). Participants were asked to make their judgments on a 1 to 7 scale where 1 indicated "much smaller/shorter than usual" and 7 indicated "much bigger/taller than usual." In order to avoid any direct comparison across objects from the same category, participants only saw one member of each kind (e.g., either a big coin or small coin). We found that big/tall items were rated significantly higher ($M = 5.3$, $SD = 1.3$) than small/short items ($M = 1.6$, $SD = 0.7$) along this scale, $p < .001$.

In addition to the two critical factors mentioned above, we also counterbalanced the Target and Competitor items so that all Targets were used as Competitors and vice versa. This counterbalancing was achieved by creating eight versions of each base item that were used to create eight presentation lists such that each list contained two items in each of the eight cells (four items in each of the critical condition) and each base item appeared just once in every list.

Finally, we created eight filler trials that featured similar display configurations as the 2-referent trials (big ball vs. small ball) but instead asked for a non-contrast item ("Pick up the big tomato"). This was critical since the effect of contrast sets on interpretation is assessed by comparing trials in which the Target appears with a contrast item from the same category with ones in which it does not. Thus it was possible that if the remaining two distractor items were

never members of a contrast set, then participants could learn that whenever they see two items of the same kind, one of those two items will always be the Target. This could facilitate Target identification on 2-referent trials relative to 1-referent trials but it would not reveal whether participants were sensitive to the informational implications of modification or whether they were simply sensitive to a specific contingency in our stimuli. To ensure that early eye-movements to the Target were not merely a reflection of this type of strategy, we included an equal number of filler trials in which a contrast set is present but another item is the Target.

2.1.4. Coding

Trained research assistants watched videotapes of the participants' actions and coded the object that was selected on each trial. Across both experiments, we only included trials where participants correctly selected the Target in subsequent analyses of eye movements. However in Experiment 1, no trials were excluded on this basis. Approximately 0.5% of test trials were excluded from further analyses because of experimenter error.

Eye movements were coded by a research assistant, who was blind to the location of each object, using frame-by-frame viewing of the participant's face on a Sony digital VCR (Snedeker & Trueswell, 2004). Each recorded trial began from the onset of the instruction and ended with completion of the corresponding action. Each change in direction of gaze was coded as towards one of the quadrants, at the center, or missing due to looks away from the display or blinking. These missing frames accounted for approximately 2% of all coded frames and were excluded from analysis. Twenty-five percent of the trials were checked by second coder who confirmed the direction of fixation for 94.6% of the coded frames. Any disagreements between the two coders were resolved by a third coder.

2.2. Results

Our analyses focused on two different time scales. For each analysis, we first identified differences across conditions by conducting an ANOVA over three broad time windows:

1. Baseline region: This 667 ms period began at the onset of the instruction and ended just before the onset of the adjective ("Point to the").

2. Adjective region: This 433 ms period began at the onset of the adjective ("big") and ended just before the onset of the noun.

3. Noun region: This 667 ms period began at the onset of the final noun ("coin") and ended at the offset of the command.

Each period was shifted 200 ms after the relevant marker in the speech stream to account for the time it would take to program saccadic eye-movements (Matin, Shao, & Boff, 1993). Across these broad time windows, any differences in looks to the objects between our conditions of interest were followed up by a second analysis that explored these divergences in greater detailed. These fine-grained analyses compared looks to the objects across conditions during 100ms time windows after the onset of the adjective.

2.2.1. Use of polarity information

Our first analysis examined whether adults use polarity information from the lexical semantics of the adjective to rule out objects from the opposite end of the height or size scale. In order to directly compare the objects from the two polarity conditions, we recoded participants' fixations in terms of looks to the *big/tall* objects or the *small/short* objects. Thus on a positive polarity trial, the Target and Competitor (e.g., big coin and big stamp) would be categorized as *big/tall* objects while the Contrast and Unrelated object (e.g., small coin and small marshmallow) would be categorized as *small/short* objects. On a negative polarity trial, this coding would be reversed: the Target and Competitor would now be categorized as *small/short* objects while the

Contrast and Unrelated object would be categorized as *big/tall* objects. Our dependent measure was total looking time to the *big/tall* objects as a proportion of looking time to all four objects. These scores were analyzed with an ANOVA with polarity (negative vs. positive) and time window (Baseline vs. Adjective vs. Noun region) as within-subject variables and list as a between-subjects variable.

Figure 2 illustrates that during the Baseline region, looks to the *big/tall* objects remained around chance across both positive (52%) and negative polarity trials (55%). However during the Adjective region, adults hearing positive adjectives quickly shifted their fixations to the *big/tall* objects (68%) while hearing negative polarity adjectives abandoned these objects (44%). Finally, during the Noun region, looks to the *big/tall* object continued to diverge in the positive (84%) and negative polarity trials (19%). This led to a significant interaction between polarity and time window, $F(2, 48) = 99.42$, $p < .001$.

To explore the timing of these polarity effects, we calculated the proportion of fixations to the *big/tall* objects for 100 ms intervals beginning from the onset of the adjective and continuing for 2000 ms later. Each time window was defined by the period from the labeled time point to the frame prior to the onset of the next interval and corresponded to the real-time onset of speech information. Using a series of one-way ANOVA's, we found that adults reliably differentiated the referents based on lexical polarity approximately 300 ms after adjective onset, $F(1, 24) = 24.53$, $p < .001$. Looks to the *big/tall* objects were greater in positive polarity trials (68%) than in the negative polarity trials (46%). The rapidity of this effect suggests that the polarity of scalar adjectives is processed and used in reference resolution at the earliest moments of language comprehension.

2.2.2. Use of referential contrast

17

Our second set of analyses examined whether referential contrast facilitated adults' interpretation of scalar adjectives. Here our dependent measure was the total looking time to the Target as a proportion of looking time to the Target over all four objects. These scores were submitted to an ANOVA with contrast (1- vs. 2-referent) and time window (Baseline vs. Adjective vs. Noun region) as within-subject variables and list as a between-subjects variable.

Figure 3 illustrates that that during the Baseline region, the proportion of looks to the Target initially remained around chance across both 1- and 2-referent trials (21% vs. 22%). Following the onset of the adjective, looks to the Target increased for both trial types (33% vs. 29%) but did not substantially diverge. During the Noun region, however, adults made more looks to the Target in the 2-referent trials compared to the 1-referent trials (65% vs. 57%). This led to a significant interaction between contrast and time window, $F(2, 48) = 3.51$, $p < .05$.

Next we explored the critical difference between contrast conditions in greater detail using 100 ms time windows. There was an advantage for 2-referent trials which emerged around the onset of noun (or approximately 600 ms after the onset of the adjective). During this window, looks to the Target in the 2-referent trials exceeded those in the 1-referent trials (47% vs. 37%; $F(1, 24) = 5.33$, $p < .05$) and continued to do so through the 900 ms time window (68% vs. 58%; $F(1, 24) = 7.13$, $p < .05$). This indicated that the presence of a within-category contrast item facilitated adults' real-time interpretation of scalar adjectives.

2.3. Discussion

In Experiment 1, we found that adults' rapidly used both the meaning of scalar adjectives and their referential implications to constrain the referent of the noun phrase. Within 300 ms of adjective onset, participants retrieved the lexical polarity of these terms and used it to rule out incompatible referents of the opposite size/height. We also replicated previous findings

Children's language processing 19

demonstrating sensitivity to referential contrast in the interpretation of scalar adjectives. Like Sedivy and her colleagues, we found that adults were faster to comprehend a modified noun like *big coin* in the presence of another contrasting member of the same category (Sedivy et al., 1999; Sedivy, 2003; Grodner & Sedivy, in press).

One curious feature of this contrast effect was that did not become reliable until after the onset of the modified noun. The fact that these contrast effects only emerge following the early use of polarity information suggests the possibility that contrast facilitates reference restriction after an initial period of semantic analysis. Furthermore, the lateness of this contrast effect in adults raises critical questions about whether sensitivity to referential contrast would be present early in development or whether processing limitations in children would prevent them from recruiting this cue during real-time comprehension (Nadig et al., 2003). It is possible that the greater complexity in the meanings of scalar adjectives relative to non-scalar adjectives may lead to delays not only in on-line processing but in acquisition as well.

In Experiment 2, we tested five-year-olds using the same materials and procedure to examine whether children use scalar adjectives to incrementally restrict reference and whether their interpretation is influenced by contextual cues. If children use the meaning of adjectives during real-time comprehension, then their looks to the *big/tall* objects should increase shortly after hearing adjectives with positive polarity but decrease after hearing ones with negative polarity. Critically, if children are sensitive to referential contrast, we would predict that their looks to the Target would increase more quickly in the 2-referent trials than in the 1-referent trials. If, however, children are unable to incorporate these contextual cues during interpretation, then there should be no difference in Target looks in 1- and 2-referent trials. Finally, it is possible that the complexity of the meanings of scalar adjectives would lead children to simply

wait until the onset of the noun before restricting reference. Thus consistent with prior findings from Nadig et al. (2003), children may generate fewer Target looks in 2-referent trials (where there are multiple objects from the same category) compared to 1-referent trials.

3. Experiment 2

3.1. Methods

3.1.1. Participants

Forty five-year-olds (ranging from 4;6 to 5;8, mean age 5;0) were recruited from the Arlington Children's Center in Arlington, Massachusetts and the McGlynn Elementary School in Medford, Massachusetts. This age group was targeted for two reasons. First, the studies of Trueswell and colleagues demonstrate that while eight-years-olds use referential contrast to resolve syntactic ambiguities, five-year-olds typically do not (Trueswell et al., 1999; Snedeker & Trueswell, 2004). Second, as we noted earlier, previous work demonstrates that five-year-olds are able to use non-scalar adjectives to incrementally restrict the reference of a noun phrase (Sedivy et al., 2000), but they fail to use the presence of contrast to facilitate interpretation of scalar adjectives (Nadig et al., 2003). All children were native English speakers.

3.1.2. Procedure and Materials

The procedure and materials was identical to Experiment 1.

3.1.3. Coding

The data was coded in the manner described in Experiment 1. Approximately 2.2% of trials were excluded from further analysis due to experimenter error while approximately 3.1% of trials were excluded because of a participant's incorrect action. Finally, missing frames due to blinks or looks away accounted for 5.4% of all coded frames and were also excluded from analysis. First and second coding had 93.8% inter-coder reliability.

3.2. Results

3.2.1. Use of polarity information

We examined children's use of lexical semantics and referential contrast using the same coarse- and fine-grained analyses employed in Experiment 1. To examine the use of polarity information, we again calculated the total looking time to the *big/tall* objects as a proportion of looking time to the *big/tall* and *small/short* objects. Figure 4 illustrates that during the Baseline region, looks to the *big/tall* objects remained around chance for both positive (48%) and negative polarity trials (46%). However during the Adjective region, children in the positive polarity trials quickly shifted their fixations to the *big/tall* objects (59%) while those in the negative polarity trials did not (54%). Finally, during the Noun region, looks to the *big/tall* object continued to diverge in the positive (81%) and negative polarity trials (34%). This lead to a significant interaction between polarity and time window, $F(2, 64) = 68.07$, $p < .001$. Our fine-grained analyses indicated that children reliably differentiated the referents of positive polarity terms (62%) from negative polarity terms (52%) approximately 400 ms after adjective onset, $F(1, 32) = 8.61$, $p < .01$. The rapidity of this effect suggests that the polarity of a scalar adjective is accessed as the word is being recognized and is quickly used to restrict reference.

3.2.2. Use of referential contrast

Next, we turned to children' sensitivity to referential contrast. We again calculated the total looking time to the Target as a proportion of looking time to the Target over all four objects. As Figure 5 illustrates, during the Baseline region, the proportion of looks to the Target remained around chance across both trial types (24% vs. 23%). In the Adjective region, however, there was a slight preference to look at the Target in the 2-referent trials relative to the 1-referent trials (30% vs. 24%). This preference disappeared following the onset of the noun

(51% vs. 48%). While there was no interaction between contrast and time (F(2, 64) = 0.46, p > .60), there was a reliable effect of contrast (F(1, 64) = 4.66, p < .05).

But a closer examination of Figure 5 also revealed that the earliest differences between the 1-referent and 2-referent conditions actually emerge in the -200 ms time window, before the adjective even begins. While this could reflect children's anticipation of an adjective (which might lead to greater attention to items in contrast sets), the absence of any similar effect in the adults suggests that this is unlikely. Instead, we suspect that it reflects a nonlinguistic preference or interest in items that match. To separate out effects of early perceptual preferences from effects in response to the adjective, we conducted an additional analysis of the first and second half trials. One possible explanation for this early contrast bias is that children realized over the course of the experiment that the presence of a contrast set is sometimes associated with the Target. While the inclusion of the filler trials decreases the predictive value of this approach, children may have nonetheless adopted such strategy after repeated presentations of the 2-referent trials. If this were the case, then we should expect no pre-adjective bias in the first half of the experiment when children have had fewer exposures to contrast sets and a more exaggerated bias in the second half of the experiment when they have had more exposures.

We again calculated the total looking time to the Target as a proportion of looking time to the Target over all four objects. The focus on performance by experiment halves inevitably led to greater subject variability in our analyses. In order to increase our ability to detect differences, we performed fine-grained analyses that collapsed across a single window across a time range rather than by individual 100 ms intervals. This method of analysis was adopted across all reported effects in this section.

Figure 6a illustrates that in the first half of the experiment, the proportion of looks to the Target remained around chance across both 1-referent and 2-referent trials throughout the Baseline region (23% vs. 24%). Analysis of 100 ms time windows revealed no period in which Target looks differed across the two trial types (all p's > .50). In contrast, Figure 6b illustrates that in the second half of the experiment, the proportion of looks to the Target in the 2-referent trials exceeded those in the 1-referent trials, particularly in the period just prior to the onset of the adjective. Fine-grained analyses confirmed a marginal contrast effect in the 100 ms window before adjective onset (26% vs. 16%; $F(1, 32) = 3.08$, $p < .10$). This early difference strongly suggests that the emergence of the pre-adjective contrast bias is linked to a predictive strategy developed over the course of the experiment.

Next we turned to regions following the onset of the adjective. In the first half of the experiment, Target looks in the 2-referent trials began to exceed those in the 1-referent trials in the Adjective region (30% vs. 25%) and this difference became more exaggerated in the Noun region (54% vs. 47%). Fine-grained analyses confirmed that there was a significant contrast effect from approximately 500 ms to 800 ms after adjective onset (31% vs. 41%; $F(1, 32) = 4.01$, $p < .05$). This time region overlaps with the contrast effects seen in adults and suggest a similar process guiding interpretation of adjectives in both populations. In contrast, effects in the second half of the experiment were less transparent. While there was a slight preference to look at the Target in the 2-referent trials relative to the 1-referent trials in the Adjective region (28% vs. 25%), this preference shifted in the opposite direction in the Noun region (46% vs. 43%). None of these differences reached statistical significance in fine-grained analyses (all p's > .30) and the overall pattern differs both from that of the first half trials and that of the adults in Experiment 1.

Altogether these results suggest that performance in the second half may reflect a more strategic process that is not specifically linked to the processing of the adjective.

3.3. Discussion

In Experiment 2, we found that five-year-olds patterned like adults in their interpretation of scalar adjectives. During the early moments of comprehension, children, like adults, were able to exploit the lexical polarity of scalar adjectives to distinguish between referents of different sizes. Critically, we also found that children were also sensitive to the presence of multiple members of the same category in the visual scene and used this information to facilitate resolution of the correct referent. Like adults, they were quicker to restrict reference in the 2-referent context compared to a 1-referent context. Thus our results contrast with the findings of Nadig and colleagues (2003). More broadly, this effect of number of referents demonstrates that children are sensitive to this aspect of the situational context and can make use of this information during real-time comprehension. The overall pattern of results across both experiments is consistent with an account of language comprehension which is predictive and interactive, drawing on information of multiple kinds from early in development.

4. General Discussion

This study explores the use of linguistic meaning and referential contrast in the real-time interpretation of scalar adjectives. We found that, comprehension in both adults and children was rapidly influenced by both these sources of information. These findings add to a growing literature demonstrating that children use multiple sources of information – including lexical meaning and usage (Trueswell et al., 1999; Snedeker & Trueswell, 2004), discourse constraints (Song & Fisher, 2005; Pyykkonen, Mathews, & Jarvikivi, 2007), and prosody (Snedeker & Yuan, 2008; Arnold, 2008) – to interpret language in real time.

Yet these findings are also somewhat surprising. Several studies on syntactic ambiguity resolution have found that children robustly fail to use the number of referents as a cue to parsing (Trueswell et al., 1999; Hurewitz et al., 2000; Snedeker & Trueswell, 2004). Why would children use referential contrast in one situation but not another? In the remainder of this section, we explore three potential explanations for the apparent discrepancy and discuss their implications for the development of language processing. We then turn our discussion to another literature that has focused on contextual effects in scalar adjectives and examine how our findings might inform the interpretation of prior results in this area.

4.1. Reconciling with prior findings

Recall that in the original Trueswell et al. (1999) study, participants were presented with temporarily ambiguous instructions like "Put the frog on the napkin in the box" where the modifier "on the napkin" could initially be interpreted as either the goal of the verb or the modifier of the noun. Critically, children's behavior indicated that they were only entertaining the goal interpretation even when the presence of multiple of referents in the visual scene (a frog on the napkin and a frog not on a napkin) supported the modifier analysis. In contrast, our study demonstrates that the presence of two referents from the same category (big coin and small coin) facilitates children's interpretations of scalar adjectives.

One possible reason for this divergence is that the two lines of work examine fundamentally different processes. The Trueswell et al. (1999) study, as well as much of the research since then, has focused on tasks which measured syntactic ambiguity resolution (Hurewitz, et al., 2000; Snedeker & Trueswell, 2004; Choi & Trueswell, 2006). However it is conceivable that while children fail to use contextual cues for syntactic parsing, they may be able to do so for lexical processes like predicting the noun. This could reflect an asymmetry in how

referential cues are integrated into different subsystems of language: Perhaps referential information is more systematically implicated in lexical processing or perhaps the coordination of multiple cues is more easily accomplished during lexical processing than it is during syntactic processing. This would predict that use of referential contrast should always emerge earlier in lexical processing than it does in syntactic processing.

Alternately, this difference may lie in the relative position of the modifier and the noun. In our task, the modifier occurred *before* the upcoming noun. In contrast, in the syntactic ambiguity tasks, the modifier occurred *after* the noun. This ordering may have critical implications for real-time comprehension. Children may have a strong bias to establish reference immediately after identifying a noun—regardless of whether they have sufficient evidence to do so. However, once this commitment is made, there is no referential indeterminacy and hence no relevant constraint on subsequent linguistic processes. Some support for this proposal comes from additional eye-movement analyses in the Trueswell study. In the 2- referent condition, children typically looked at one of the two referents shortly after hearing the direct-object noun ("the frog") and whichever frog they happened to look at generally became the preferred referent and was used to carry out the action. Thus by committing to an interpretation immediately after encountering the noun, the children may have resolved the referential ambiguity (for themselves) before they ever encountered the ambiguous prepositional phrase. Any subsequent integration of this phrase would call for a revision of reference assignment. In contrast, by moving the modifier to a position prior to the critical noun in our study, we may have created a context in which the presence of referential contrast can be used to facilitate the prediction of an up-coming referent rather than to revise a previous referential commitment. This hypothesis could be explored by looking at the effects of contrast on the interpretation of adjective/noun combinations across

languages with different word orders (e.g., in Spanish where the adjective typically appears after the noun, see Brown-Schmidt & Konopka, 2008; Weisleder & Fernald, 2009).

The final possibility is that that the number of referents is simply a more robust cue for adjective interpretation than it is for resolution of PP-attachment ambiguities, allowing children to acquire this constraint more rapidly. The research to date supports this hypothesis. As noted in the Introduction, the presence of referential contrast is a poor predictor of post-nominal modification: in a scene with multiple potential referents (many squares), speakers produce a bare definite NP ("the square") in nearly half of their utterances (Brown-Schmidt et al., 2002). In contrast, prior research has shown that there is a tight correlation between the number of referents in the scene and production of scalar adjectives (Sedivy, 2003; Gregory et al., 2003; Ferreira, Slevc, & Rogers, 2005). For example, Brown-Schmidt and Tanenhaus (2006) asked speakers to instruct listeners to select a picture like a *large triangle* among several items. In a portion of the trials, participants saw displays that included another same shaped item that in size (a small triangle). In the presence of this contrast, speakers produced the modifier 98% of the time while in its absence, they did so only about a quarter of the time.

4.2. Gricean inferences, context effects, and the interpretation of scalar adjectives

Contextual effects on the interpretation of scalar adjectives have also received considerable attention in adult psycholinguistic research (Sedivy et al., 1999; Sedivy, 2003; Gregory et al., 2003; Grodner & Sedivy, in press). However the mechanisms underlying these effects are not clearly understood. Sedivy (2003) suggested that the presence of a contrast item facilitates reference restriction by causing listeners to generate a rapid Gricean inference (Quantity Maxim: Grice, 1975). When listeners hear *tall*, they infer that it probably modifies a member of a contrastive set, otherwise adjectival modification would be over-informative since the item could

be uniquely identified from the noun alone. In the 2-referent context, this leads listeners to prefer the Target object over the Competitor. In the 1-referent context, neither item is compatible with this inference thus reference resolution awaits the noun. Recently, Grodner and Sedivy (in press) found that these context effects are also sensitive to a listener's perception of the speaker. When listeners were told that the speaker had "an impairment that causes social and language problems," they no longer showed facilitation in adjective interpretation in the 2-referent context. The authors suggest that listeners failed to calculate a Gricean inference when they perceived the speakers' utterances as irrational and uncooperative.

This construal of adjective contrast effects heightens the interest of studying children since this is a population notoriously poor at making these kinds of Gricean inferences (Smith, 1980; Noveck, 2001; Papafragou & Musolino, 2003; Chierchia, Crain, Guasti, Gualmini, & Meroni, 2001; Huang & Snedeker, in press). In fact, in contrast with our current findings, prior work demonstrates that children as old as seven- and nine-years of age consistently failed to make routine Gricean inferences for other scalar expressions. They instead prefer a literal interpretation in both on-line and off-line measures.

This difference raises questions concerning how scalar adjective interpretation might differ from other types of Gricean inferences. One possibility is that Gricean inferences are in fact a heterogeneous category with some expressions emerging earlier than others during development. While this alternative is logically conceivable, it would sweep aside the fact that children's difficulties with Gricean inferences span a variety of scalar terms including modals, quantifiers, and conjunctions (Smith, 1980; Noveck, 2001; Papafragou & Musolino, 2003; Chierchia et al., 2001; Huang & Snedeker, in press). It would fail to explain why these Gricean inference emerge so late in development while those for scalar adjectives appear so early.

28

A second possibility is that the use of contrast in adjective interpretation may not in fact be a Gricean inference at all. This would be consistent with standard linguistic analysis of these terms which include a contextual parameter that incorporates referential information within the lexical semantics. Thus rather than engaging in a post-semantic, pragmatic inference of the kind envisioned by Sedivy (2003), children may rapidly use referential contrast in their interpretation of scalar adjectives since the notion of a comparison class is part of the meanings of these words. However, while this explanation would remove the apparent discrepancy between current and prior findings, it would fail to account for other adult studies demonstrating similar contrast effects with non-scalar adjectives (Sedivy, 2003; Grodner & Sedivy, in press). Altogether the tension between these two lines of work suggests the need for a more detailed study of how these two kinds of pragmatic effects emerge over the course of development.

4.3. Conclusion

In conclusion, the findings of this study provide evidence for the use of referential contrast in children's real-time comprehension. Like adults, children's interpretations of scalar adjectives are rapidly influenced by multiple sources of information including the meanings of these terms and their referential implications. These findings suggest that the same fundamental features that characterize adult language comprehension are also present and operational in the child listener (Trueswell & Gleitman, 2004). They also demonstrate that the critical properties of language that affect sentence processing, such as the predictability of a cue or the position of a word/phrase, also influence the trajectory of acquisition in language development. This highlights an intrinsic relationship between the moment-to-moment processing during real-time language comprehension and the year-to-year changes over the course of language development.

References

Altmann, G. M., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, 30, 191-238.

Altmann, G. M., & Kamide, Y. (2004). Now you see it, now you don't: mediating the mapping between language and the visual world. In J. Henderson and F. Ferreira (eds.) *The integration of language, vision and action* (pp. 347-386). NY: Psychology Press.

Arnold, J. E. (2008). THE BACON not the bacon: How children and adults understand accented and unaccented noun phrases. *Cognition*, 108, 69-99.

Barner, D. & Snedeker, J. (2008). Compositionality and statistics in adjective acquisition: 4-year-olds interpret tall and short based on the size distributions of novel noun referents. *Child Development*, 79, 594-608.

Bierwisch, M. (1987). The Semantics of Gradation. In M. Bierwisch, E. Lang (eds.): *Dimensional Adjectives*. Berlin: SpringerVerlag.

Brown-Schmidt, S., Campana, E., & Tanenhaus, M. (2002). Reference resolution in the wild: Circumscription of referential domains by naive participants during an interactive problem solving task. In: J. Trueswell & M. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions*. Cambridge: MIT Press.

Brown-Schmidt, S. & Tanenhaus, M. K. (2006). Watching the eyes when talking about size: An investigation of message formulation and utterance planning. *Journal of Memory and Language*, 54, 592-609.

Brown-Schmidt, S. & Konopka, A. (2008). Little houses and casas pequeñas: Message formulation and syntactic form in unscripted speech with speakers of English and Spanish.

Cognition, 109, 274-280.

Chierchia, G., Crain, S., Guasti, M. T., Gualmini, A., & Meroni, L. (2001). The acquisition of disjunction: Evidence for a grammatical view of scalar implicatures. In A. H.-J. Do, L. Domingues, & A. Johansen, (Eds.), *Proceedings of the 25th Boston University Conference on Language Development* (pp. 157-168). Somerville, MA: Cascadilla Press.

Choi, Y. & Trueswell, J. (2005). Do Korean children hop frogs like English children? *Paper presented at the 33rd Annual Boston University Conference on Language Development*, Boston, MA.

Crain, S., & Steedman, M. (1985). On not being led up the garden path: The use of context by the psychological parser. In D. Dowty, L. Karttunnen, & A. Zwicky (Eds.), *Natural Language Parsing*. Cambridge, UK: Cambridge University Press.

Fernald, A., Pinto, J., Swingley, D., Weinberg, A., & McRoberts, G. (1998). Rapid gains in speed of verbal processing by infants in the second year. *Psychological Science*, 9, 228-231.

Ferreira, V. S., Slevc, L. R., & Rogers, E. S. (2005). How do speakers avoid ambiguous linguistic expressions? *Cognition,* 96, 263–284.

Gregory, M. L., Joshi, A., Grodner, D., & Sedivy, J. C. (2003). Adjectives and processing effort: So, uh, what are we doing during disfluencies? *Paper presented at the 16th Annual CUNY Conference on Human Sentence Processing*, Cambridge, MA.

Grice, H. P. (1975). Logic and Conversation. In P. Cole and J. L. Morgan (Eds.), *Syntax and Semantics, Vol. 3*. New York: Academic Press.

Grodner, D. & Sedivy, J. (in press). The effect of speaker-specific information on pragmatic inferences. In N. Pearlmutter & E. Gibson (eds), The Processing and Acquisition of Reference. MIT Press: Cambridge, MA.

Huang, Y. & Snedeker, J. (in press). Semantic meaning and pragmatic interpretation in five-year-olds: Evidence from real time spoken language comprehension. To appear in *Developmental Psychology.*

Hurewitz, F., Brown-Schmidt, S., Thorpe, K., Gleitman, L. & Trueswell, J. (2000). One frog, two frog, red frog, blue frog: Factors affecting children's syntactic choices in production and comprehension. *Journal of Psycholinguistic Research*, 29, 597-626.

Kennedy, C. (1999). Projecting the adjective: The syntax and semantics of gradability and comparison. New York: Garland Publishing.

MacDonald, M., Pearlmutter, N. & Seidenberg, M. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101, 676-703.

Matin, E., Shao, K. & Boff, K. (1993). Saccadic overhead: information processing time with and without saccades. *Perception & Psychophysics*, 53, 372-380.

Nadig, A., Sedivy, J., Joshi, A., & Bortfeld, H. (2003). The development of discourse constraints on the interpretation of adjectives. In B. Beachley, A. Brown, & F. Conlin (eds.), *Proceedings of the 27th annual Boston University Conference on Language Development* (pp. 568-579). Somerville, MA: Cascadilla Press

Noveck, I. A. (2001). When children are more logical than adults: experimental investigation of scalar implicatures. *Cognition*, 78, 165-188.

Papafragou, A. & Musolino, J. (2003). Scalar implicatures: experiments at the semantics-pragmatics interface. *Cognition*, 86, 253-282.

Pyykko̎nen, P., Matthews, D., & Ja̎rvikivi, J. (2007). Children's online comprehension of pronouns in spoken language: The role of verb semantics. Poster presented at the *20th annual CUNY conference on human sentence processing*, San Diego, CA.

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

Sedivy, J., Demuth, K., Chunyo, G. & Freeman, S. (2000). Incremental referentiality-based language processing in young children: Evidence from eye movement monitoring. In S. C. Howell, S. Fish, and T. Keith-Lucas (eds.) *Proceedings of the 24th Annual Boston University Conference on Language Development* (pp. 684-695). Somerville, MA: Cascadilla Press.

Sedivy, J. C. (2003). Pragmatic versus form-based accounts of referential contrast: Evidence for effect of informativity expectations. *Journal of Psycholinguistic Research*, 32, 3-23.

Smith, C. L. (1980). Quantifiers and question answering in young children. *Journal of Experimental Child Psychology*, 30, 191-205.

Snedeker, J. & Trueswell, J. (2004). The developing constraints on parsing decisions: The role of lexical biases and referential scenes in child and adult sentence processing. *Cognitive Psychology*, 49, 238-299.

Snedeker, J. & Yuan, S. (2008). Effects of prosodic and lexical constraints on parsing in young children (and adults). *Journal of Memory and Language*, 58, 574-608.

Song, H., & Fisher, C. (2005). Who's "she"? Discourse prominence influences preschoolers' comprehension of pronouns. *Journal of Memory and Language*, 52, 29–57.

Swingley, D. & Fernald, A. (2002). Recognition of words referring to present and absent objects by 24-month-olds. *Journal of Memory and Language*, 46, 39-56.

Syrett, K., Kennedy, C., & Lidz, J. (in press). Meaning and context in children's understanding of gradable adjectives.

Tanenhaus, M. K., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of

visual and linguistic information in spoken language comprehension. *Science*, 268, 1632.

Thorpe, K., Baumgartner, H. & Fernald, A. (2005). Children's developing ability to interpret adjective-noun combinations. In B. Beachley, A. Brown, & F. Conlin (eds.), *Proceedings of the 30th annual Boston University Conference on Language Development* (pp. 631-642). Somerville, MA: Cascadilla Press.

Trueswell, J. & Tanenhaus, M. (1994). Toward a lexicalist framework of constraint-based syntactic ambiguity resolution. In Clifton and Frazier (Eds), *Perspectives on sentence processing*. Hillsdale, NJ: Lawrence Erlbaum.

Trueswell, J., Sekerina, I., Hill, N. & Logrip, M. (1999). The kindergarten-path effect: studying on-line sentence processing in young children. *Cognition*, 73, 89-134.

Trueswell, J. & Gleitman, L (2004). Children's eye movements during listening: evidence for a constraint based theory of parsing and word learning. In J. Henderson & F. Ferreira (eds.). *Interface of Language, Vision, and Action: Eye Movements and the Visual World.* NY: Psychology Press.

Van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language*, 41, 147-182.

Weisleder, A. & Fernald, A. (2009). Real-time processing of postnominal adjectives by Latino children learning Spanish as a first language. In J. Chandlee, M. Franchini, S. Lord, & G. Rheine (eds.) *Proceedings of the 24th Annual Boston University Conference on Language Development* (pp. 611-621). Somerville, MA: Cascadilla Press.

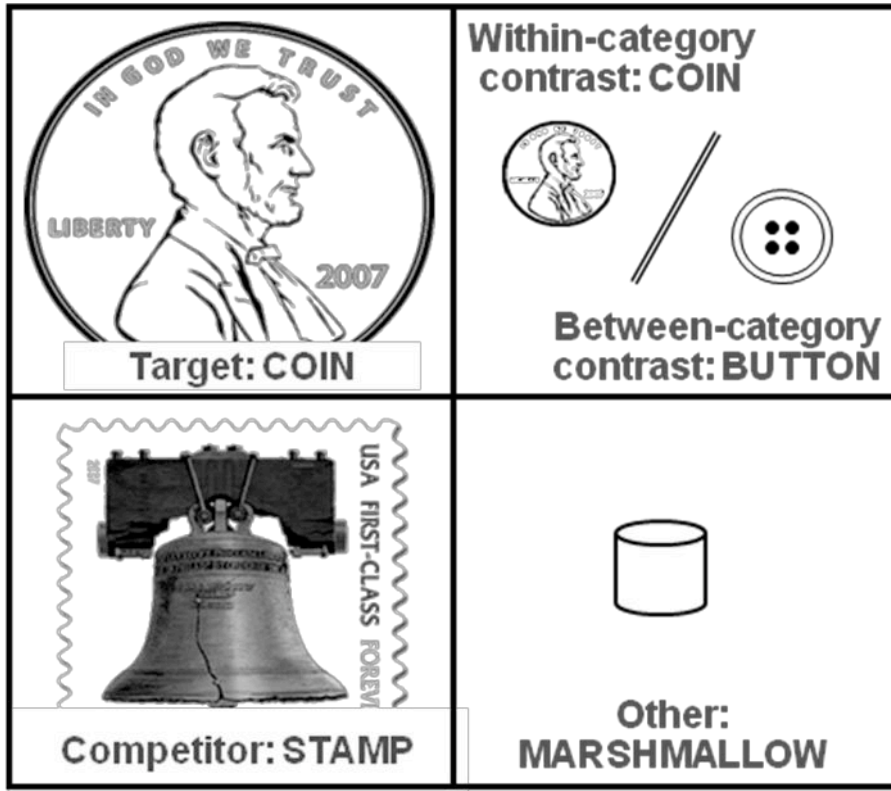Figure 1: Visual display for the "big coin" trial.
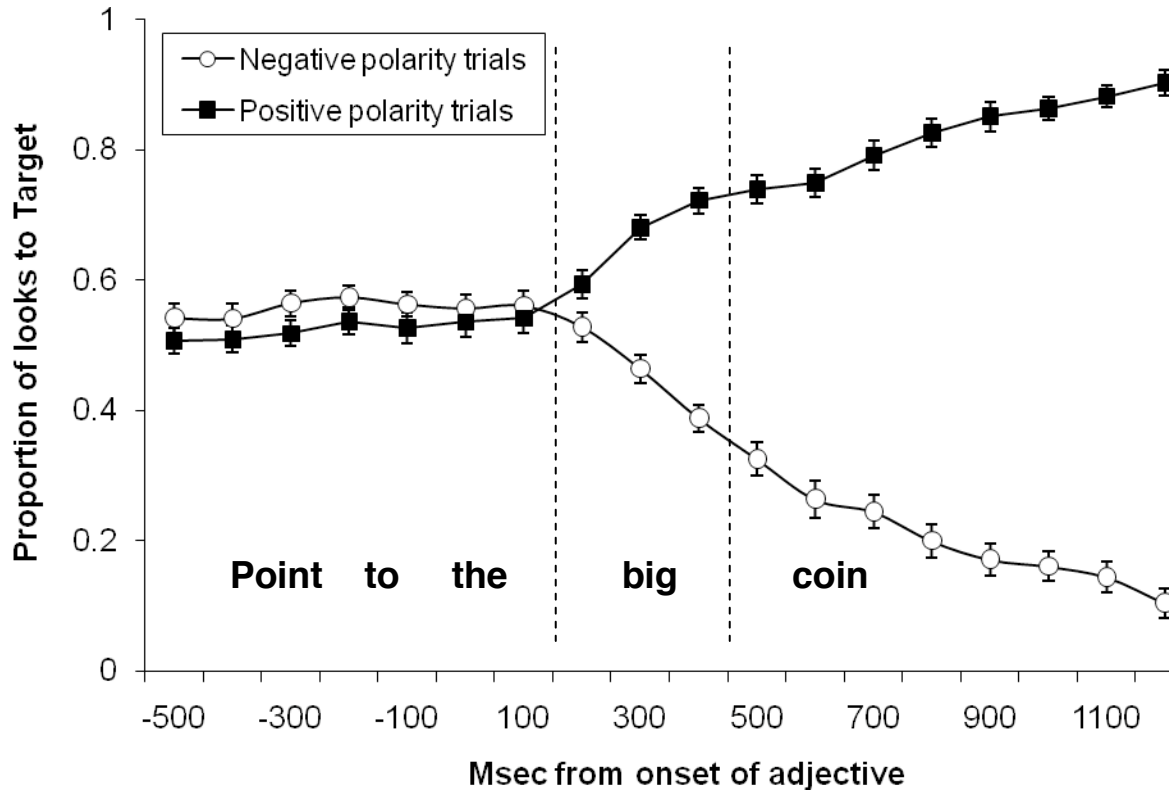
Figure 2: Processing of polarity in adults.

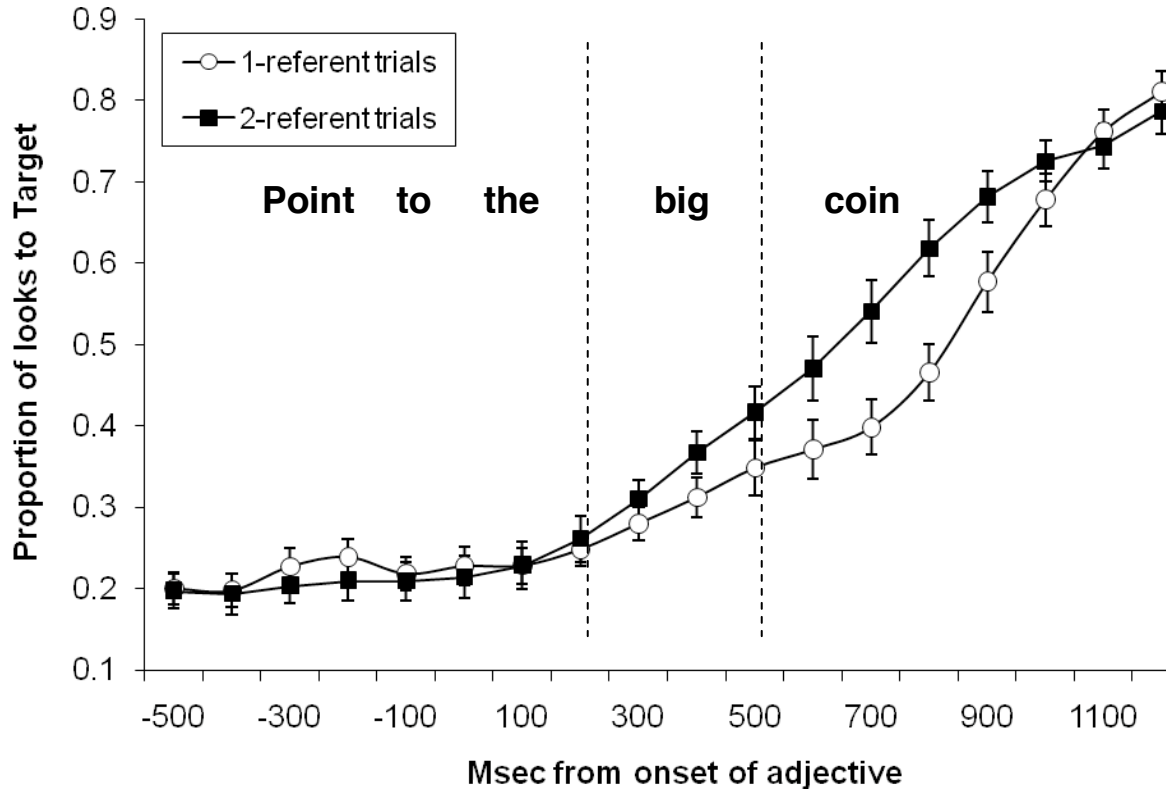Figure 3: Use of referential contrast in adults.
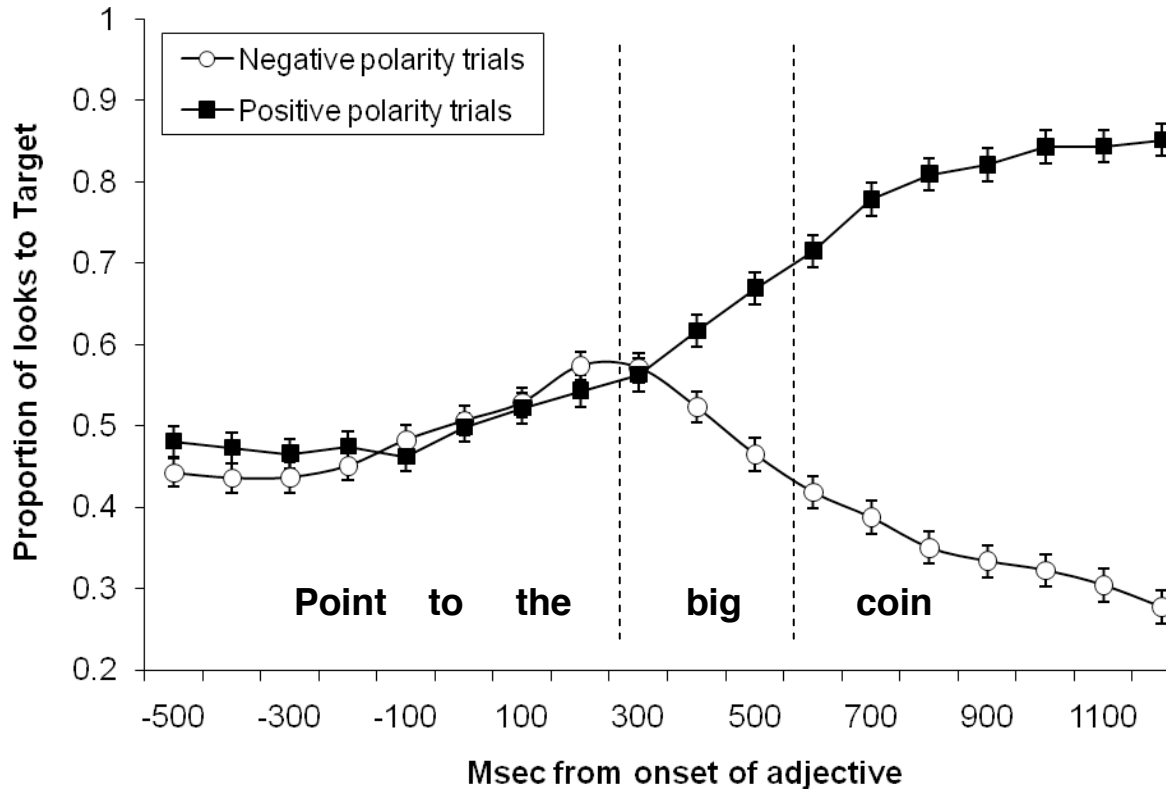
Figure 4: Processing of polarity in children.
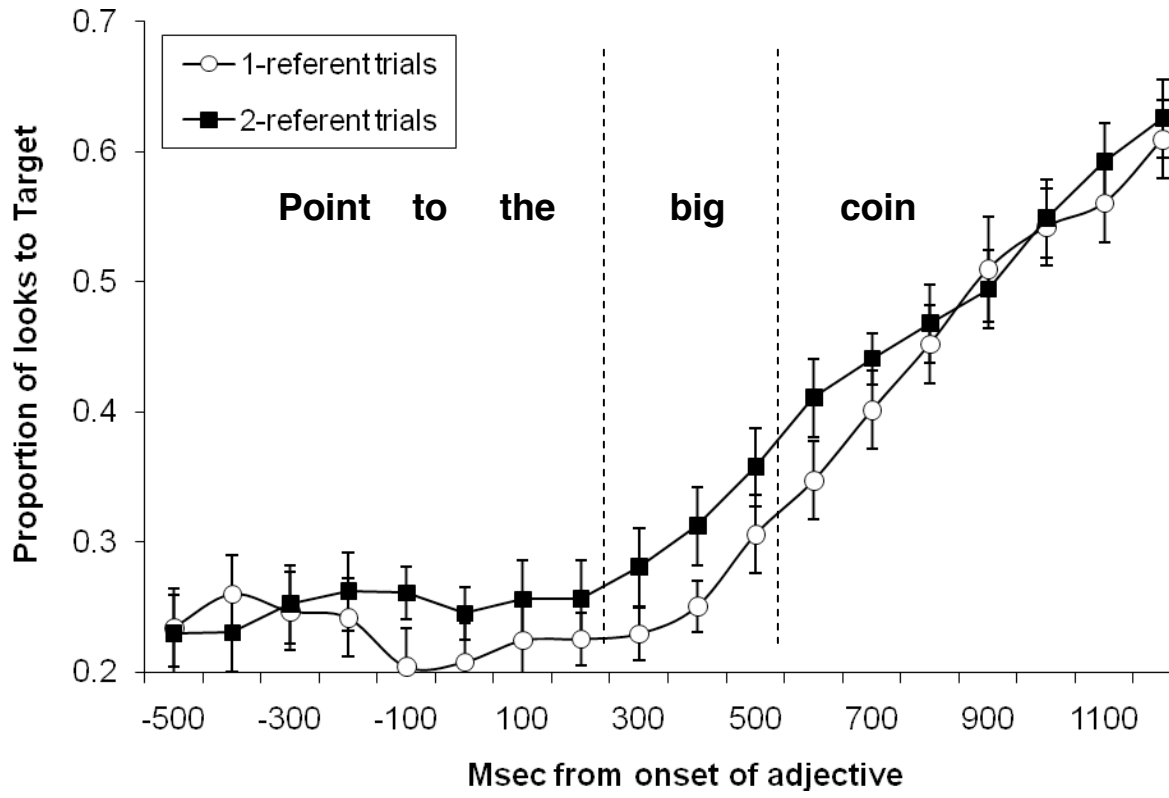
Figure 5: Use of referential contrast in children.

Figure 6: Looks to the Target in (a) first half trials and (b) second half trials.