

Systematic Uses of Context in the On-Line Interpretation of Adjectives

Julie C. Sedivy[†] (sedivy@ling.rochester.edu)

Craig G. Chambers[‡] (chambers@bcs.rochester.edu)

Michael K. Tanenhaus[‡] (mtan@bcs.rochester.edu)

Gregory N. Carlson[†] (carlson@ling.rochester.edu)

[†]Department of Linguistics / [‡]Department of Brain & Cognitive Sciences
University of Rochester, Rochester, NY 14627 USA

Abstract

This study evaluates the effects of a very specific aspect of discourse context, that of contrast among discourse entities. An eyetracking methodology monitoring eye movements to real objects in a visual display was used to investigate the interpretation of vague scalar adjectives. Two experiments were conducted in which we manipulated the typicality information associated with the modified noun and the presence or absence of a contrasting object in the visual context. The experiments provide general evidence for the systematicity of contextual contrast effects, and demonstrate their immediacy in on-line processing.

Introduction

Questions regarding the contextual conditions under which a sentence is typically uttered and understood have traditionally been set aside by a number of formal linguistic frameworks and by the sentence processing work which has built on these frameworks. This is due in part to the assumption that the relevant contextual information is largely unsystematic and unconstrained, and as a result, too computationally costly to be useful in on-line interpretation. However, these assumptions are currently being challenged on two sides. First, a growing body of work within formal linguistic theory has pointed to the widespread interaction of grammatical phenomena with contextually-defined information. Second, there has been growing consensus within the psycholinguistic community that much contextual information is readily accessible for on-line interpretation. In light of these developments, there is an increasing need for research to address in explicit detail the nature of contextually-relevant information and its interaction with core grammatical phenomena.

The goal of this paper is to discuss evidence for both the systematicity and accessibility of certain kinds of contextual information in the interpretation of adjectival meaning. Adjectives exemplify a particularly interesting phenomenon because they frequently resist simple rule-based

compositional semantic analyses. To illustrate, let us consider the simplest cases in which the adjective itself has a clear and relatively fixed core meaning (color adjectives are typical examples of this class). Consider, for example, the

expression **red bowl**. Here, it is easy to treat both the adjective and the noun as property-denoting expressions, that is, where the meanings of the words **red** and **bowl** independently correspond to a set of entities (i.e. the set of red entities and the set of entities that have the property of being bowls), and the meaning of the complex expression **red bowl** corresponds to the intersection of those two sets. This class of adjectives is frequently referred to in the semantics literature as *intersective*. However, the intersective analysis becomes impossible if the meaning of the adjective is not easily determined independently. It has been pointed out that the meaning of an adjective frequently depends on the noun it modifies. Contrast the meaning of the adjective **good** in the following sentences:

1. a) Tom is a good priest. He is caring and empathic.
- b) Tom is a good lawyer. He is ruthless and shrewd.

Here, the property that is picked out by the adjective is determined to a large extent by the head noun. These adjectives are often referred to as *subsective*, because the set of entities picked out by the adjective must constitute a subset of those picked out by the noun. Furthermore, in some cases, the meaning of the adjective cannot be determined independently of the general context of the utterance, as in the sentences below:

2. a) My two-year-old built a really tall snowman yesterday.
- b) The fraternity brothers built a really tall snowman yesterday. (from Kamp & Partee, 1995).

Thus, it is necessary to specify the ways in which contextual information contributes to fixing the meaning for adjectives that do not have a stable core meaning.

We take as a starting point for this investigation the observation that adjectives (as well as other noun modifiers such as relative clauses and prepositional phrases) have the discourse function of signaling a contrast between entities that share the property picked out by the head noun. That is, a modified expression such as **red bowl** is most naturally understood as presupposing a context which includes a red bowl, as well as a bowl of some other color. There is considerable psycholinguistic evidence for the importance of such contrastive information in on-line processing. A number of studies have shown that the tendency to understand modifiers as a means of distinguishing between similar entities has an impact on the on-line resolution of syntactic ambiguity (e.g. Altmann & Steedman, 1988; Crain & Steedman, 1985; Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995). Furthermore, it has been shown that people have rapid access to information regarding relevant contrasting entities in a visually-present context in the course of understanding phrases marked with contrastive stress (Sedivy, Tanenhaus, Eberhard, Spivey-Knowlton & Carlson, 1995).

We suggest that two somewhat different notions of contrast may be relevant in the interpretation of modifiers: 1) the contrast involved in explicitly distinguishing an entity, such as a tall glass, with another glass in the discourse context that is lower on the height dimension, and 2) an implicit notion of contrast, where **tall** serves to signal a contrast between the entity being referred to, and an implicit default representation for height that corresponds to a typical glass. The first of these notions relates to the properties of entities present in a particular discourse context, and the second relates to more stable, stored representations associated with the noun that is being modified. In this paper, we present the results of experiments designed to examine the interaction of information pertinent to the head with information from the general discourse context. experimental data

Experiment 1

We chose to look at the class of scalar adjectives such as **tall**, **thin**, **long**, etc. which are to a large degree relativized to the noun that they modify e.g. **tall man**, **tall building**, **tall glass**. We were thus able to manipulate the degree to which the adjective was felicitous given a particular instance of an object, as well as whether or not an explicitly contrasting object was present in the discourse context.

The study reported here makes use of an eye-tracking technique that has recently been adapted for use in experimental work with spoken language (Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995). This technique employs an eye-tracking device that is mounted onto a helmet and worn by the subject throughout the experiment, allowing for unrestricted head movement by the subject. Work within this experimental paradigm has shown that processing of spoken language is highly

incremental and strongly sensitive to the properties of the visual model.

We used a verification task in which subjects were asked to respond to questions such as "Is there a tall glass?" For half of the trials, the target object reflected a good fit with a description that involved a scalar adjective. For example, if the description was **tall glass**, the target was a typical example of a tall glass. Typicality was ascertained by means of a separate rating task, where subjects rated the target objects in isolation as being best expressed by an unmodified expression (e.g. glass), or a modified expression, with both poles of a particular scale as choices (e.g. **tall glass**, **short glass**) For the other half of the trials, the target was a poor token of a tall glass -- that is, it was rated as being best described by means of an unmodified bare noun (see Figure 1). Along with typicality, we also manipulated whether the visual display contained an explicit contrasting object -- that is, a glass that was rated as being consistent with the description **short glass**. There were a total of four objects in each display, with each display containing an object that could also be described by means of the particular scalar adjective used to describe the target (for instance, the display containing a glass also had a pitcher). We call this object the competitor. The competitor object was always further along the scale invoked by the adjective than the target in absolute terms, but was rated as being best described by means of an unmodified expression. In other words, the competitor **pitcher** was taller than the target **glass**, but not tall with respect to pitchers in general.

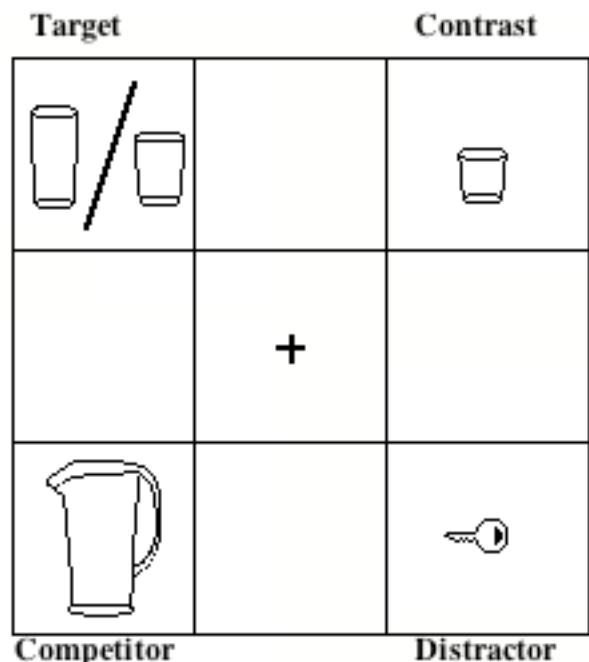


Figure 1: Experimental Display

Three measures were of interest: 1) the nature of the subjects' responses, that is, whether they accepted or rejected the modified expression as a description of any of the objects present, 2) the speed with which they provided their responses and 3) what objects they looked at in the course of making their decision.

Figure 2 shows the proportion of trials in which the subject accepted the modified description across the four conditions. As we might expect, for displays where the target is the only glass in the scene, we find that people are more reluctant to accept the description of "tall glass" for the poor token-- they answer "yes" only 58% of the time, whereas they accept the description all the time for the good token. What is striking, however, is how sensitive these judgments are to the presence of a contrasting glass. In these displays, the description of "tall glass" was accepted just as often for the poor tokens as for the good tokens (95% and 93% "yes" responses respectively). The interaction of typicality and contrast was robust ($F(1,11)=33.00$; $p<0.001$). The data indicate that the presence of a short glass neutralized the typicality effect for people's judgments, suggesting that their ultimate interpretation is strongly sensitive to the identity of the objects in the discourse context.

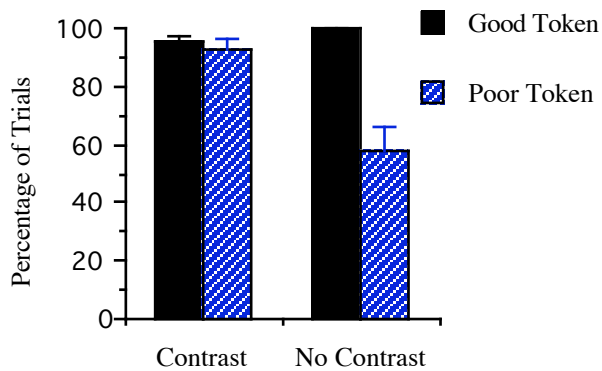


Figure 2: Percentage of "Yes" Responses, Experiment 1

Figure 3 shows the latencies of verbal responses when the subjects responded "yes". Here, we see the effects of the typicality information associated with the head noun most strongly. Overall, poor tokens elicited slower "yes" responses than good tokens ($F(1,11)=9.32$; $p<0.05$). For displays that did not contain an explicitly contrasting object, subjects took significantly longer to respond "yes" when the target was a poor token than when it was a good token. ($F(1,11)=6.22$; $p<0.05$). This difference was also present for the displays which did contain a contrasting object ($F(1,11)=5.78$; $p<0.05$). That is, when a short glass was present in the display, although subjects were no more likely to reject the modified description when the target was a poor exemplar, they showed higher uncertainty in the amount of time it took them to settle on a "yes" response.

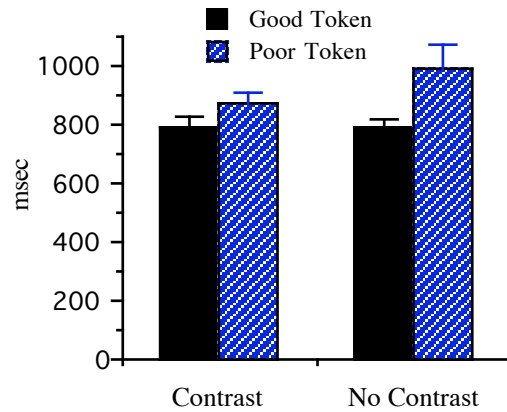


Figure 3: Latencies of "Yes" Responses, Experiment 1

The third measure of interest was the location of the subject's eye movements to objects in the display. These data are presented in Figure 4. There are two interesting phenomena here. First, if adjectives are being understood contrastively, we would expect subjects to look at the contrasting objects in an attempt to make an explicit comparison. We see a high proportion of trials that include a look to the contrasting object in displays that do contain such a contrast (47%). For displays that did not contain an explicit contrast, we calculated the looks to the object that was in the same location as the contrasting object in the other displays. Looks to this object were very rare (7%), with this difference easily achieving statistical reliability ($F(1,11)=75.77$; $p<0.001$). Second, recall that each display contained a competitor object -- that is, an object that could be described by the adjective used in the modified expression, such as the pitcher in the glass example discussed earlier. We calculated the proportion of trials for which the subject looked at this object in the course of making a decision: There were by far more looks to the competitor object when the display did not contain the contrasting short glass ($F(1,11)=9.58$; $p<0.05$). Thus, in the absence of a contrasting object, subjects are taking the other potentially tall object to be relevant, where in the displays that do contain a contrast, subjects are taking the contrasting object to be of greater importance than the competitor.

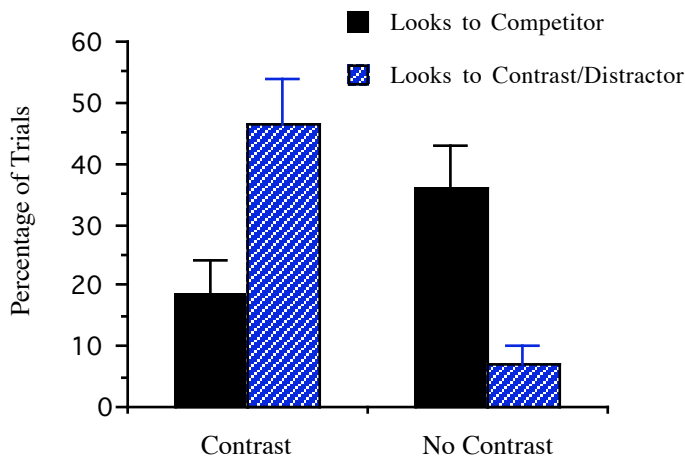


Figure 4: Looks to Objects Other than Target, Experiment 1

Interestingly, the time course of these false looks is different. The eye movements to the competitor object are launched on average 198 milliseconds after the onset of the head noun, whereas the eye movements to the contrasting object are launched on average 472 milliseconds after the same point ($F(1,5)=10.57$; $p<0.05$). The earliness of the looks to the competitor suggests that subjects are considering the competitor as relevant for the interpretation of the adjective (perhaps as a potential referent), whereas looks to the contrast are undertaken to compare the objects denoted by the head noun.

Overall, these data show strong evidence that subjects are sensitive to the presence of a contrasting object in making decisions about the appropriateness of modified expressions. This is particularly striking given that this is a task which invites a comparison with stored conceptual representations. The data from eye movements provides direct evidence that people are considering the contrasting object as relevant for interpretation.

Experiment 2

The previous experiment provides intriguing data with respect to looks to the competitor objects and the contrasting objects in the display. However, the looks to the target objects themselves were too noisy to yield clear results. We carried out a second experiment which was primarily designed to investigate how early subjects would be able to begin assigning an interpretation to the adjective, and whether this would differ across conditions. In particular, given the close relationship between the head noun and the identification of the specific value for the scalar adjective, we were interested to see whether providing a clear contrast in the discourse context would allow subjects to assign an early interpretation to the adjective. The measure we used

was the time it took for subjects to launch an eye movement to the target object. In the second task, instead of asking subjects to verify whether a particular object was present, we issued an instruction to pick up specific objects and move them to a different location on the board. A typical instruction would be: "Pick up the tall glass and put it below the pitcher." Notice that this task differs from the earlier one in one important way: whereas people were free in the earlier task to determine whether the modified description was appropriate for any of the objects in the display, this task carries a presupposition that the display does in fact contain an object of that description. In addition, subjects were permitted to look freely around while the display was set up, thereby reducing the amount of scanning that would be done during the instruction itself.

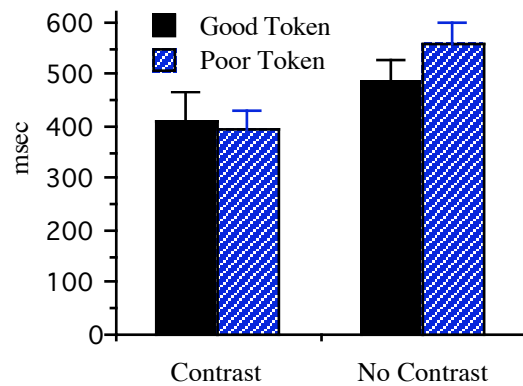


Figure 5: Eye movement latencies to target, Experiment 2

The data show that the presence of a contrasting glass does indeed speed eye movements to the target (see Figure 5). The main effect of contrast was statistically reliable ($F(1,11)=14.48$; $p<0.01$). This suggests that subjects are able to coordinate knowledge of the contrastive function of modifiers with evidence of contrast in the visual scene quickly enough to affect on-line reference resolution. What is of special interest is that there is no hint of a typicality effect for displays containing a contrast, indicating that in these displays, the typicality information associated with the head did not exert an influence in the interpretation of the modified expression. The effect of typicality in general seems very weak for this task -- for displays without a contrast, the numerical difference we see is nonreliable, as is the interaction between contrast and typicality.

Sensitivity to contrast is again seen in subjects' looks to objects other than the target, pictured in Figure 6. As in the earlier task, when the display lacks a clear contrast, subjects are particularly drawn to the competitor object. However, in the presence of a contrasting object, there are very few looks to the competitor, with the effect of contrast yielding a significant difference ($F(1,11)=23.16$; $p=0.001$). False looks to the contrasting object, however, are very frequent in displays containing a contrast (31% of trials), particularly when compared with false looks to the distractor object in

the same position as the contrasting object in the objects without a contrast (7%), again yielding a statistically reliable difference ($F(1,11)=23.37$; $p<0.001$). This pattern of eye movements the same as that found in the earlier experiment.

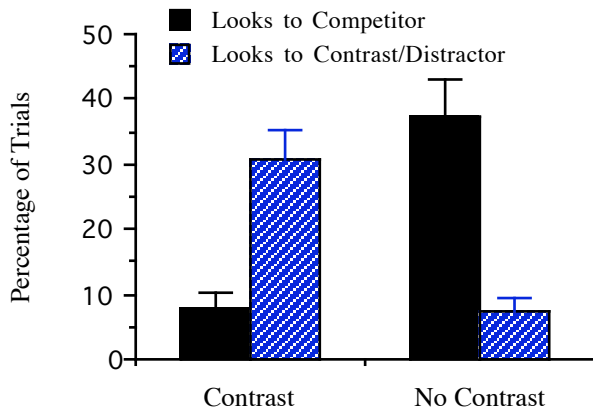


Figure 6: Looks to Objects other than target, Experiment 2

Also consistent with the first experiment is the time-course of the false looks. Recall that looks to the competitor generally occurred considerably earlier than looks to the contrast. This result is mirrored in the second experiment, with looks to the competitor occurring on average 269 milliseconds after the onset of the head noun, and looks to the contrast occurring 476 milliseconds after that point ($F(1,7)=22.93$; $p<0.01$).

Discussion

A number of phenomena are evident from the data thus far. First, we see evidence that typicality information relativized to the head plays a role in interpreting scalar adjectives. This is primarily evident in the verification task. We also see systematic effects of a very specific discourse function of adjectival modifiers, namely, the contrastive function. In fact, not only can discourse factors affect judgments, they virtually eliminate typicality effects in a task which is strongly restricted to the objects in the discourse model. As well as appearing systematic, the discourse-based function of modifiers is accessible quickly enough to be used to constrain the interpretation of scalar adjectives on-line.

More generally, these results have implications for mechanisms that combine simple expressions into more complex ones. The effects of typicality that are observed in the data reported here can be captured by appealing to lexical information and associated conceptual information evoked by the expression itself. However, we have seen

that the value for scalar adjectives can be shifted depending upon the properties of various entities in the discourse model. Most saliently, when there is an entity that can serve as an explicit contrast, there is a tendency to interpret the scalar adjective relative to this entity, rather than relative to some stored norm associated with the head noun. This suggests that the semantic representation of adjectives needs to be flexible enough to allow for such shifts in interpretation.

We are aware of at least one semantic analysis of scalar adjectives which builds in precisely the kind of flexibility that is indicated by our experimental data. Very informally stated, in Bierwisch's (1987) analysis of scalar adjectives, the meaning of scalar adjectives essentially corresponds to some underdetermined value on a certain scale that is further than some relevant comparison class. The comparison class can be determined either with respect to the average value on the scale for the set of entities denoted by the noun, or with respect to a contextually determined set of entities. Thus, the actual value or range of values corresponding to the scalar adjective is left undetermined in the lexical meaning of the adjective itself, with context playing a role in contributing to the fixing of this value via the identification of a set of contrasting entities. This analysis is reminiscent of a recent semantic analysis of focus (Rooth, 1992), where variables corresponding to contextually-defined contrast sets are used in the semantic interpretation of focus-sensitive phenomena.

As mentioned in the introductory section above, the importance of a notion of discourse contrast is evident in the experimental literature as well as the linguistic literature. The experiments reported here provide additional evidence for the centrality of the process whereby discourse entities are distinguished from and contrasted with other possible entities in the discourse model. Rather than concluding that context influences are unconstrained and unsystematic then, it would appear that it is possible to identify certain contextually-dependent processes as highly general and central to interpretation.

References

- Altmann, G. & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, 30, 191-238.
- Bierwisch, M. (1987). The semantics of gradation. In M. Bierwisch & E. Lang (Eds.), *Dimensional Adjectives* (pp. 71-262). Berlin: Springer-Verlag.
- Crain, S. & Steedman, M. (1985). On not being led up the garden path: the use of context by the psychological parser. In D.R. Dowty, L. Karttunen, & A. Zwicky (Eds.), *Natural Language parsing: Psychological, Computational and Theoretical Perspectives*. Cambridge, UK: Cambridge University Press.
- Kamp, H. & Partee, B. (1995). Prototype theory and compositionality. *Cognition*, 57, 129-191.
- Rooth, M. (1992). A theory of focus interpretation.

Natural Language Semantics, 1, 75-116.

- Sedivy, J.C., Tanenhaus, M.K., Eberhard, K.M., Spivey-Knowlton, M.J. & Carlson, G.N. (1995). Using intonationally-marked presuppositional information in on-line language processing: evidence from eye movements to a visual model. *Proceedings of the 17th Annual Conference of the Cognitive Science Society*. 275-380.
- Tanenhaus, M.K., Spivey-Knowlton, M. J., Eberhard, K.M. & Sedivy, J.C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, 1632-1634.