# UC Merced
## Proceedings of the Annual Meeting of the Cognitive Science Society

**Title**

Using Intonationally-Marked Presuppositional Information in On-Line Language Processing: Evidence from Eye Movement s to a Visual Model

**Permalink**

https://escholarship.org/uc/item/0xx3v7wg

**Journal**

Proceedings of the Annual Meeting of the Cognitive Science Society, 17(0)

**Authors**

Sedivy, Julie C.
Tanenhaus, Michael K.
Spivey-Knowlton, Michael
et al.

**Publication Date**

1995

Peer reviewed

# Using Intonationally-Marked Presuppositional Information in On-Line Language Processing: Evidence from Eye Movements to a Visual Model

**Julie C. Sedivy**
Dept. of Linguistics
University of Rochester
Rochester, NY 14627
sedivy@psych.rochester.edu

**Michael K. Tanenhaus**
Dept. of Brain & Cognitive Sciences
University of Rochester
Rochester, NY 14627
mtan@psych.rochester.edu

**Kathleen Eberhard**
Dept. of Brain & Cognitive Sciences
University of Rochester
Rochester, NY 14627
eberhard@psych.rochester.edu

**Michael Spivey-Knowlton**
Dept. of Brain & Cognitive Sciences
University of Rochester
Rochester, NY 14627
spivey@psych.rochester.edu

**Gregory N. Carlson**
Dept. of Linguistics
University of Rochester
Rochester, NY 14627
carlson@ling.rochester.edu

## Abstract

This study evaluates the effect of presuppositional information associated with contrastive stress on on-line language processing. An eye-tracking methodology was used, in which eye movement latencies to real objects in a visual display are taken as a measure of on-line reference resolution. Results indicate that presupposed contrast sets are being computed on-line, and can be used to speed reference resolution by narrowing the referential domain of an utterance. In addition, presupposed contrast sets appear to play a role in managing attention in the processing of a discourse.

## Introduction

There is a growing awareness in the field of formal semantics that presuppositional phenomena, which have traditionally been relegated to an extralinguistic pragmatic component, are in fact necessary for a complete *semantic* characterization of a variety of linguistic constructions. The recent theoretical emphasis on the interplay between asserted and presupposed levels of meaning leads to some interesting questions regarding the interaction of these levels of meaning in the on-line processing of language. For instance, it becomes important to determine whether presuppositional information is readily accessible to the processing system, and if so, whether it is available on-line for immediate interpretation. Some recent psycholinguistic work indicates that specific presuppositions associated with the interpretation of definite noun phrases have immediate effects on the resolution of temporary syntactic ambiguities (e.g. Altmann & Steedman, 1988; Britt, 1994; Crain & Steedman, 1985; Spivey-Knowlton & Sedivy, in press).

This paper is primarily concerned with the processing role of presuppositional information in the interpretation of contrastive focus. Recent semantic analyses have converged on a representation of contrastive focus involving the integration of asserted and presupposed information (e.g. Rooth, 1992; Kratzer, 1991; Krifka, 1991). In particular, these theories have attempted to formalize the general observation that contrastive focus constructions are used not only to provide information about the material that is focused, but also to provide information about the relationship between the focused element and some presupposed set of contrasting entities. For instance, if a speaker were to assert "JOHN likes tofu", with contrastive stress on "John" (as indicated by capital letters), this utterance would likely be understood as drawing a comparison between John's attitudes towards tofu on the one hand, and those of a presupposed and contextually salient group of people on the other hand. This understanding arises from an interaction between asserted and presupposed levels of meaning. The dual representation of meaning is crucial to current semantic approaches to contrastive focus. This can be seen by observing the semantic effects of varying the placement of focus. For instance, a sentence such as "John only likes TOFU for breakfast" can be paraphrased as meaning that there are no foods that John likes for breakfast other than tofu. The lexical content of the sentence (and therefore its asserted propositional content) can be kept unchanged while varying the location of the focus (which reflects a change in the presuppositional content) as in the sentence "John only likes tofu for BREAKFAST" which means roughly that there are no meals other than breakfast for which John enjoys eating tofu. This change in the presuppositional information exerts a clear change in the semantics of the sentence. Thus, the former sentence is automatically false under conditions where John also likes to eat eggs for breakfast, whereas the latter can be true under the same conditions. In order to capture the truth conditional effects of focus, Rooth's (1992) analysis makes use of the pragmatic notion of a presupposed contrast set against which the asserted content must be

evaluated. The contrast set is derived essentially by replacing the focused element in the sentence with a variable. With respect to the sentence "John only likes TOFU for breakfast", the contrast set consists of entities that can serve as the variable for x in "John eats x for breakfast" (i.e. foods that can be eaten for breakfast). The contrast set serves as the input domain for the interpretation of the focus-sensitive word "only", yielding an interpretation which states that, of the members of the contrast set, tofu is the only entity that John likes to eat for breakfast. In the sentence "John only eats tofu for BREAKFAST", on the other hand, the contrast set consists of meals that might involve the eating of tofu.

This discussion of the analysis of focus demonstrates the central role of presuppositionally-based information in arriving at a semantic analysis of focus-related constructions. The centrality of presuppositional information suggests a more prominent role for presupposition than has frequently been assumed. In particular, it suggests that semantic interpretation may be representationally, and perhaps temporally dependent upon the processing of certain kinds of pragmatic information. In traditional processing models (e.g. Forster, 1979), it has been claimed that pragmatic information is processed at the latest stage of language processing. In this paper, we will present experimental evidence from eye movements to objects in a visual model which suggest that presuppositional information that is intonationally marked is used immediately in the mapping of referential phrases to real objects in a visual model.

## Experiment 1: Evidence for Contrast Sets in On-Line Processing

The study reported in this paper makes use of a new eye-tracking technique that is currently being adapted for use in experimental work with spoken language (Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, in press). 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. The subject sits in front of an array of objects presented on a table-top or upright board and manipulates objects in response to instructions from the experimenter. The objects in the visual display correspond to the discourse referents in the experimental instructions. The eye movements of the subject are monitored throughout the experiment, allowing for precise observation of the location of the subject's gaze and the point in time at which eye movements are launched in response to the experimenter's instructions.

A number of factors together make this methodology particularly well-suited for studying the on-line interpretation of referential phrases. First of all, the methodology relies on the subject being required to carry out an action that is closely linked with an experimental instruction. This ensures that the visual context is highly relevant for the interpretation of the target utterance. Second, this paradigm results in an eye movement record that is highly informative, as people typically look at an

object before reaching for it (Ballard, Hayhoe & Pelz, 1995). Thus, it is possible to get a very precise indication of which object in the display is being understood as the intended referent in the experimental instruction, and how long it takes the subject to process the referential phrase and launch an eye movement, (i.e. a saccade) to the understood referent.

In addition, this experimental paradigm exploits the highly incremental nature of language processing. It is clear from the eye movement records that subjects attempt to make early commitments to referents based on an integration of the linguistic input and the visual cues available in the display. Eye movements to objects are very closely time-locked to the point in the utterance where the referent becomes unique with respect to the alternative possibilities defined by the visual context together with the speech input. In a study reported in Tanenhaus et al. (in press), subjects were instructed to touch objects that were distinguished by marking, color, and shape. For instance, in a display consisting of a plain red square, a plain yellow rectangle, a plain blue square and a plain pink square, the referent in the instruction "Touch the plain red square" can be unambiguously identified at the word "red", as there is only one object in the display that is both plain and red. However, if the display contains more than one object that is both plain and red (e.g. a plain red square and a plain red rectangle), the point of disambiguation occurs only following the word "square". The eye movement record showed a strong sensitivity to the point of disambiguation, with the mean time to launch a saccade to the target object being 250 ms from the end of the word that disambiguated the referent. The weighing of possible alternatives compatible with the speech input at a certain point in time is seen even more dramatically in some more recent experimental manipulations using more complex displays and longer referential phrases (Eberhard, Tanenhaus, Spivey-Knowlton & Sedivy, 1995). Under these conditions, we observe eye movements before the point of disambiguation to objects whose visual properties are consistent with the linguistic description to date.

In the present study, we explored the contribution of presuppositions involved in the interpretation of contrastive focus to on-line processing of referential expressions. Unlike the referential presuppositions associated with definite expressions, presuppositional information related to contrastive focus is typically encoded intonationally, rather than lexically. This allows for the opportunity to examine the effects of prosodically marked information on language processing. In particular, this study focused on the potential of presuppositional contrast sets to provide constraining information allowing for early disambiguation of referential phrases. To see how this is possible, consider the example experimental display (1a) in Figure 1 below.

The point of disambiguation for this display in conjunction with the instruction "Touch the large red square" is at the word "red". However, when contrastive stress is used, as in the instruction "Touch the LARGE red square", a contrast set is presumably computed, which consists of the large and small red squares (by virtue of replacing the focused word "large" with a variable). If a contrast set consisting of objects differing only in size is computed

quickly enough, and is taken to be the relevant domain of reference, the phrase now becomes disambiguated prior to the word "red", rather than following it, as the relevant domain has only one large object. In this experiment, we looked for evidence of the on-line computation of contrast sets, and their effect on the incremental interpretation of referential expressions. We predicted that contrastive stress would speed the launching of saccades to a referent in a visual model in which the intended referent becomes unique at an earlier point in the speech stream due to the availability of a sufficiently restricted contrast set, as is the case in Figure 1a.. This facilitatory effect of contrastive stress should not be observed for a visual display in which contrastive stress does not allow for sufficient domain restriction to provide early disambiguation, as in Figure 1b, where the entire display serves as the referential domain, even when accompanied by an instruction involving contrastive stress on the size adjective.
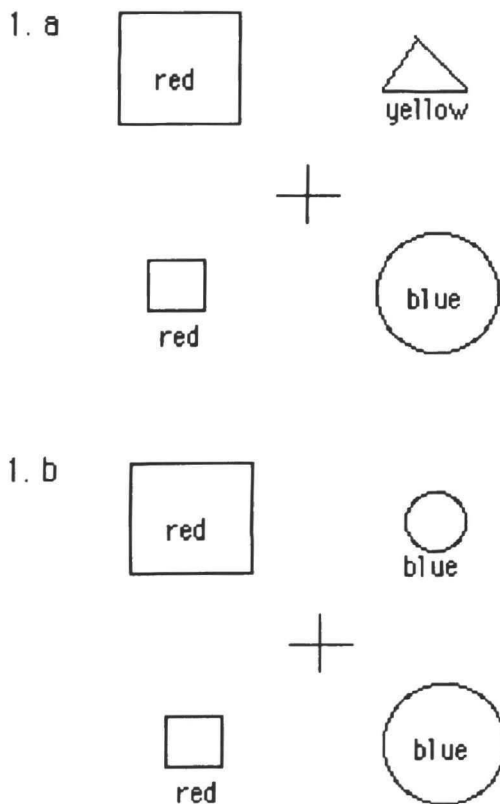


Figure 1: Two example visual displays serving as visual contexts for instructions in Exp.1

## Materials

The visual displays made use of arrangements of four cardboard shapes varying along the dimensions of size (large and small), color (red, yellow and blue), and shape (circles,

squares, and triangles). For nine of the fifteen subjects tested, each experimental display contained a pair of objects differing only in size, as in Figure 1 above. These displays were coupled with sets of experimental instructions spoken by the experimenter. There were three instructions accompanying each display, with only the first counting as a critical instruction. Each set of three instructions was preceded by the instruction "Look at the cross", in order to make sure that the subject fixated on the same place preceding each critical trial. Critical commands varied systematically as to whether or not they contained contrastive stress. Half of the instructions were uttered with contrastive stress on the size adjective (e.g. Touch the LARGE red square"), and the other half of the instructions were uttered with neutral intonation. There were six items in each of the two cells. We predicted that eye movement latencies would be shorter when the accompanying instruction was uttered with contrastive stress than when it received neutral intonation. For a subset of the subjects, we included experimental displays of two types: either the display contained one pair of objects differing only in size, as in Figure 1a (i.e. one potential contrast set). or it contained two pairs of minimally contrasting objects. (i.e. two potential contrast sets). Displays containing two potential contrast sets were included in order to rule out an alternative account of the same effect, in which the facilitatory effect of contrastive stress on the size adjective is due to an increase in the salience of the size dimension, and therefore an increase in the activation of the subset of large entities, of which the target referent was a member. For these displays, the increased salience account predicts that there should be equal facilitation due to stress for both display types, as the number of large objects is equal in both types. However, under the contrast set hypothesis, contrastive stress should not result in faster saccade latencies for displays such as Figure 1b, as both contrast sets must be considered as potential domains, which is equivalent to considering the entire display. The stress manipulation was crossed with the display type manipulation, yielding four cells with three items in each cell. We predicted an interaction between the presence or absence of contrastive stress and display type. All of the experimental displays were embedded with an equal number of filler displays and accompanying sets of instructions.

## Procedure

While the subject followed instructions to touch objects in the workspace, eye movement data was recorded using a lightweight Applied Scientific Laboratories (ASL) head-mounted video-based tracking system. The camera provides an infrared image of the eye at 60 Hz. Monocular eye position is determined by monitoring the locations of the center of the pupil and the cornea reflection. A scene camera is mounted on the side of the helmet, providing an image of the subject's field of view. Calibration is carefully monitored throughout each trial, and minor adjustments are occasionally made between trials.

A VCR record was made for each experimental trial, consisting of the instructions spoken by the experimenter into a microphone, as well as the subject's moment-by-moment gaze fixation superimposed over the scene camera image.

## Results

Latency data for trials in which the initial fixation was to the correct object was analyzed from the video tapes by identifying the beginnings and ends of critical words in the speech stream for each trial, and noting the time lapse between the critical speech points and the onset of a saccade to the intended object. Latencies were measured from the onset of the color adjective

Results of the experiment conformed to the predictions generated by the hypothesis that contrast sets are used on-line to constrain the interpretation of referential expressions. When the size adjective on the first instruction was uttered contrastively, latencies to launch a saccade to the correct referent were an average of 431 ms compared to 545 ms when the instruction had neutral intonation. This difference was reliable ($F(1,14)=6.17$, $p<.05$). In addition, an ANOVA on the subset of six subjects for whom we included displays containing two potential contrast sets showed a reliable interaction between display type and intonation type ($F(1,5)=18.04$, $p<.01$). When the display contained only one potential contrast set, launch latencies were 472 ms. with contrastive stress, compared to 635 ms. with neutral intonation, whereas for displays containing two potential contrast sets, the mean latencies were 588 ms. with contrastive stress and 555 ms. with neutral intonation.

## Discussion

Results of this experiment show that under conditions where the relevant contextual presuppositions required for interpreting contrastive focus are met, the presence of contrastive stress can facilitate the interpretation of referential phrases. The facilitatory effect of contrastive stress is observed only in visual displays in which computing a contrast set allows for a restriction of the referential domain. This data pattern rules out an account in which the effect of contrastive stress is limited to an increase in the salience of the focused property, without reference to a contrast set. These results provide experimental evidence for a central role of presupposition in interpreting language on-line.

## Experiment 2: Attentional function of contrast sets

As discussed above, the communicative goal of contrastive focus is to provide information about the relationship between the focused element and the set of contrasting entities, as well as to provide information about the focused element. In recognition of the heavy informational load associated with contrastive focus, a number of psycholinguistic studies have investigated the allotment of attentional resources to focused material. These studies have found that greater resources are allotted to focused material than unfocused material. For instance, Birch & Rayner (1994) demonstrate that more time is spent reading noun phrases that are clefted (in focus position) than identical noun phrases that occur in presuppositionally neutral sentences. Cutler & Fodor (1979) show that subjects are faster to respond to a target phoneme occurring at the onset of a focused word than an unfocused word. In addition, recall of focused material is superior to recall of unfocused material (Birch & Garnsey, in press; Singer, 1976).

However, to date, there have been no studies examining the attentional status of entities in the contrast set that are not explicitly referred to in the utterance. Given the crucial contribution of the contrast set to the interpretation of contrastive focus constructions, it is reasonable to hypothesize that the inclusion of an entity in the contrast set heightens its salience. We carried out a second experiment to investigate this hypothesis.

## Materials and Procedure

This experiment used the same eye-tracking technique that was used in Experiment 1. We reasoned that differences in the attentional salience of referents would be reflected in the ease of locating individual objects in the display, with highly salient objects taking less time to locate than less salient objects. We hypothesized that contrastive focus would serve to heighten the salience of contrast set members, resulting in shorter eye movement latencies.

As in the previous experiment, we used displays of four cardboard shapes arranged around a central fixation cross. All of the experimental displays in this experiment contained one minimal pair of shapes differing only in size, as in Figure 1a above.

Each target instruction was preceded by an instruction to look at one member of the minimal pair. The stress pattern of this instruction was manipulated such that half of the instructions were uttered with contrastive stress on the size adjective (e.g. "Touch the LARGE red square", with reference to Figure 1a above), and half of the trials were uttered with neutral intonation. The immediately following target instruction manipulated the identity of the target referent. In half of the trials, the target referent was the contrasting object (e.g. "Now touch the small red one"), and in the other half, it was an object that was outside of the contrast set, but that had the same value for the size dimension as the contrasting object (e.g. "Touch the small yellow triangle"). For both target referents, the point of disambiguation was at the word following the size adjective (i.e. "one" and "yellow" respectively). These two manipulations were crossed, to yield four cells, with four items in each cell. 16 filler trials were added, which were similar in structure to the experimental trials, and constructed such that, over the course of the experiment, the

second instruction in the set referred to each object in the display with equal probability, to discourage probability-based strategies. We predicted that contrastive stress on the first instruction would speed eye movement latencies for contrast set members. However, contrastive stress should have no effect on eye movement latencies for instructions referring to unrelated objects outside of the contrast set.

## Results and Discussion

Results thus far (based on six subjects) are preliminary, but suggestive. As predicted, the presence of contrastive stress preceding an instruction referring to the contrast member speeded eye movement latencies (216 ms after the onset of the disambiguating word for the contrastive stress condition versus 387 ms for the neutral stress condition). In addition, there was no facilitation of stress for instructions referring to an object that could not be construed as belonging to the contrast set (435 ms with contrastive stress versus 419 ms with neutral intonation). A statistical analysis indicates that the interaction between intonation type and referent type is marginally significant by subjects $(F(1,5)=6.3; p<0.06)$.

This data pattern lends support to the hypothesis that contrastive focus serves to raise the salience of objects that belong to the contrast set, despite the fact that these objects have not been explicitly referred to. The heightened salience is seen in the rapid resolution of subsequent reference to these objects.

## General Discussion

Much of the research in language comprehension over the last two decades has emphasized the incremental nature of language processing, showing evidence that linguistic information begins to be processed as it becomes available in the speech stream (e.g. Marslen-Wilson, 1975). The incremental property of the human language processor, however, introduces a specific problem for the processing system: that of temporary indeterminacy in the speech stream, such that linguistic information must frequently be processed before there is sufficient information in the input to exclude all but a single interpretation of the input. A great deal of the literature in language processing has focused on the problem of temporary structural indeterminacy, and the mechanisms by which the human processor resolves such indeterminacies (see Frazier, 1987; Tanenhaus & Trueswell, in press, for overviews). Two general accounts can be seen as representing the poles of this debate: serial models in which syntactic processing is autonomous and precedes the processing of semantic and pragmatic information (Forster, 1979, Frazier, 1987), and constraint-based models in which different sources of linguistic and nonlinguistic information interact to settle on one of a number of competing interpretations (MacDonald et al.,1994; Tanenhaus et al., in press b). The debate has revolved largely around evidence for or against syntactic parsing preferences which compute a first-pass default analysis based solely on syntactic information, and

impervious to information of a semantic or pragmatic nature. Implicit in much of this debate has been the assumption that the primary task of the human language processor is to recover the appropriate linguistic structure of the input.

Work in the head-mounted eyetracking paradigm, in which eye movements to a visual model are monitored, has demonstrated that another form of indeterminacy arises that is a result of the task of mapping referential expressions to objects in a model, rather than the result of a structural ambiguity (Eberhard et al., 1995; Tanenhaus et al., in press a, b). This mapping is also highly incremental in nature. Evaluation of the objects in the model against the properties denoted by modifiers in the referential phrase begins immediately, and does not depend upon the complete construction of the referential phrase as a syntactic unit. Indeed, it has been shown that information regarding the contextual model to which referential expressions must be mapped plays a crucial role in the resolution of structural ambiguities (e.g. Altmann & Steedman, 1988; Spivey-Knowlton, Trueswell & Tanenhaus, 1993; Tanenhaus et al., in press b).

The emphasis on the model-based aspect of language comprehension yields a rich opportunity for examining the resolution of indeterminacies that are referential, rather than structural, in nature. In the present study, we have begun to evaluate the role of presuppositional information in the resolution of this type of indeterminacy. The results of the two experiments reported here support the centrality of presuppositional information in language interpretation. Specifically, these studies provide evidence for the use of presupposed contrast sets in establishing reference to objects in a visual model. The results of Experiment 1 further show that contrast sets are computed rapidly enough to be useful in the highly incremental interpretation of referential phrases. The results of Experiment 2 suggest that, in addition to their relevance for resolving temporary indeterminacies of reference, contrast sets have an important function in managing attention in the interpretation of a discourse.

These data focus on the issue of referential, rather than structural indeterminacy. As such, they do not bear directly on the debate as to whether syntactic ambiguity resolution is serial or constraint-based, as the predictions of the serial account have hinged crucially on the presence of parsing preferences in temporarily ambiguous constructions. However, the results in Experiment 1 do suggest that complex presuppositional information is computed and used extremely rapidly in order to resolve referential indeterminacies. This certainly lends plausibility to accounts in which multiple constraints are rapidly integrated at the earliest stages in processing. In addition, these results are also congenial for recent trends in areas of semantics in which presuppositional information plays a central role in semantic interpretation. Specifically, the rapid interaction between lexical/semantic information pertaining to the properties of a referent and presupposed information regarding possible contrast sets lends plausibility to the interaction between levels of representation that is found in recent semantic work on contrastive focus.

## Acknowledgements

## References

Altmann, G & Steedman, M. (1988) Interaction with context during human sentence processing. *Cognition*, 30, 191-328.

Birch, T. & Rayner, K. (1994). Paper presented at the 1994 meeting of the Psychonomics Society, St. Louis, MO

Birch, T. & Garnsey, S. (In press). The effect of focus on memory for words in sentences. *Journal of Memory and Language*.

Ballard, D., Hayhoe, M & Pelz, J. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience*, 7, 68-82.

Britt, M. A. (1994). The interaction of referential ambiguity and argument structure in the parsing of prepositional phrases. *Journal of Memory and Language*, 33, 251-283.

Crain, S. & Steedman, M. (1985). On not being led up the garden path. In Dowty, Kartunnen & Zwicky (eds.), *Natural Language Parsing*. Cambridge U Press: Cambridge, MA.

Cutler, A. & Fodor, J.A. (1979). Semantic focus and sentence comprehension. *Cognition*, 7, 49-59.

Eberhard, K., Tanenhaus, M., Spivey-Knowlton, M., & Sedivy, J. (1995). Investigating the time-course of establishing reference: Evidence for rapid incremental processing. Paper presented at the 8th Annual CUNY Sentence Processing Conference, Tucson, AZ.

Forster, K. (1979). Levels of processing and the structure of the language processor. In Cooper & Walker (eds.), *Sentence Processing*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and Performance* (pp. 559-586). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kratzer, 1991. The representation of focus. In von Stechow & Wunderlich (eds.), *Semantik: Ein Internationales Handbuch der zeitgenossichen Forschung*. Berlin: Walter de Gruyter.

Krifka, M. (1991). A compositional semantics for multiple focus constructions. *Proceedings of Semantics and Linguistic Theory (SALT) 1, Cornell Working Papers*, 11.

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

Marslen-Wilson, W. (1975). Sentence percpetion as an interactive parallel process. *Science*, 189, 226-228.

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, 1 75-116.

Singer, M. (1976). Thematic structure and the integration of linguistic information. *Journal of Verbal Learning and Verbal Behavior*, 15, 549-558.

Spivey-Knowlton, M., Trueswell, J., Tanenhaus, M. (1993). Context effects in syntactic ambiguity resolution. *Canadian Journal of Epxerimental Psychology*, 47, 276-309.

Spivey-Knowlton, M. & Sedivy, J. (in press). Parsing attachment ambiguities with multiple constraints. *Cognition*.

Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K. & Sedivy, J. (in press a). Using eye movements to study spoken language comprehension. In T. Inoui & J, McClelland (eds.), *Attention & Performance XVI*.

Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K. & Sedivy, J. (in press b). Integration of visual and linguistic information in spoken language comprehension. *Science*.

Tanenhaus, M., & Trueswell, J. (In press). Sentence comprehension. In J Miller & P. Eimas (Eds.), Handbook of Perception and Cognition: Volume 11. New York: Academic Press.