

Review of “A Model-Based Psycholinguistic Study of Semantic Contrast”
(Sedivy 1997)

Colin Phillips
University of Maryland

We all share the intuition from everyday experience that we interpret sentences incrementally and generally effortlessly, but there has been a big gap between this intuition and the state-of-the-art in linguistics and psycholinguistics. In psycholinguistics, semantic interpretation has been somewhat of a poor cousin, compared to the great deal of attention that has been devoted to syntactic processing. One reason for this may be the relative scarcity of detailed theories of what is involved in *incremental* compositional interpretation, in which interpretations are built following the surface word order of each sentence, rather than following the bottom-to-top hierarchical structure of the sentence.

Julie Sedivy’s dissertation provides a detailed investigation of one aspect of incremental semantic interpretation, contrast-sensitive focus. The results of Sedivy’s experiments are textbook material, in that they provide some of the best evidence to date on the speed of semantic interpretation, and at the same time involve findings which are clear and straightforward enough to be understood by an audience with very little technical background. The dissertation is very easy to read, and the results of the nine experiments are quite clear cut. At the same time, though, Sedivy’s results beg for incorporation into a more detailed theoretical model. With this in mind, I will first review what I take to be the most striking findings of the dissertation, and then I will go on to discuss the mechanisms that might be needed to account for these findings.

Findings

The main finding of the dissertation is that speakers are remarkably good at using information about contrast very rapidly while interpreting sentences. The most striking findings come from a series of experiments in which subjects’ eye-movements are tracked while they carry out simple instructions involving an array of objects on a board in front of them. For example, in the display shown in Figure 1 below, subjects might be told to “pick up the small yellow triangle”, or “pick up the large red square”. People typically look at an object before picking it up, and eye-movements to the target object occur *much* earlier than the picking-up actions that follow. Therefore, since vision researchers have independently determined how long it takes the brain to program an eye-movement, it becomes possible to use the eye-movements to get a remarkably precise idea of how long it takes for a listener to fix the reference of the target object.

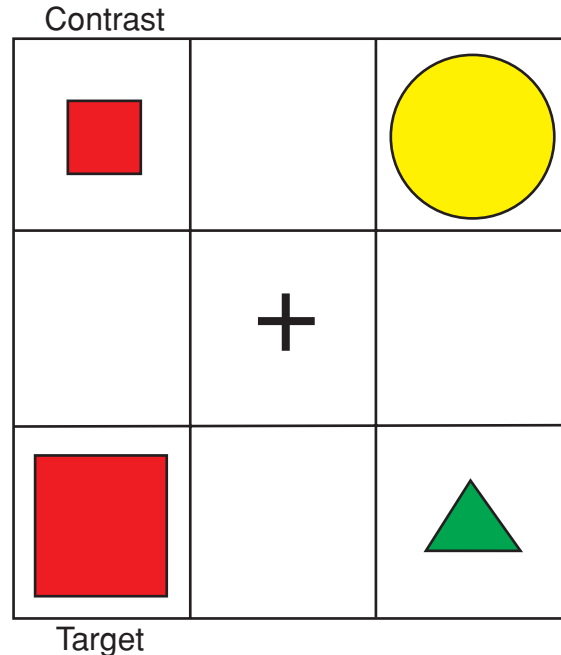


Figure 1: An example display accompanying instructions such as “Pick up the LARGE red square.”

It is important to note that this *head-mounted eye-tracking* technique which Sedivy uses is quite new to language processing research; it is a technique which Sedivy, Tanenhaus and colleagues at the University of Rochester pioneered in language research (cf. Tanenhaus et al. 1995). The vast majority of work on language processing has involved tasks in which the only input is presented in the form of written or spoken sentences. In these situations, the comprehender’s discourse model must be generated entirely based on the linguistic input, with no additional visual support. The head-mounted eye-tracking technique allows the experimenter to provide a *vastly* more vivid context than is possible in reading-based studies. The discourse model is placed directly in front of the comprehender’s eyes, rather than being left to his or her imagination.

The eye-tracking experiments provide two basic findings, with a number of additional experiments serving to reinforce the basic findings.

First, contrastive focus on an adjective can be used very quickly to uniquely identify an object. In the display shown above, there are two large objects – a red square and a blue circle. Therefore, the partial instruction “pick up the large ...” fails to uniquely identify the object to be picked up. On the other hand, if the instruction is changed to “pick up the LARGE ...”, and if the listener can understand the contrastive focus to mean that the target object is one that minimally contrasts with another object in the property *size*, then this may be enough to uniquely identify the large red square. Accordingly, Sedivy shows that her subjects looked to the target somewhat earlier when phonological focus allowed the object to be uniquely identified at the size adjective.

Second, a number of Sedivy’s experiments showed that even in the absence of phonological focus, people show a bias to interpret scalar adjectives contrastively wherever possible. This was shown using displays like the one shown in Figure 2 below, which contained a number of different household objects. A display like this might be

accompanied by an instruction like “pick up the tall glass”. In the display, the pitcher on the lower left is the tallest object, but the glass on the upper left is both tall by comparison to glasses in general, and tall by comparison to the shorter glass in the upper right-hand corner. After hearing the partial instruction “pick up the tall ...” subjects were most likely to choose the tall glass as the target object, rather than the (tall) pitcher. This is despite the fact that there was no phonological contrast and the glass was not uniquely identified by the linguistic input.

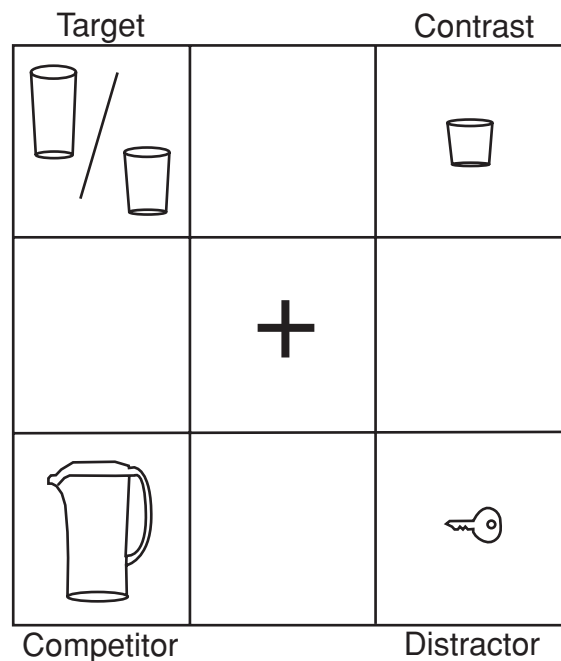


Figure 2: An example display accompanying instructions such as “Pick up the tall glass.” The target object, indicated in the top left corner was either a good token or a poor token of a description involving modification. In displays without a contrasting object, the glass in the upper right corner was replaced by an unrelated distractor item, such as a file folder.

In a pair of other experiments, using more standard measures of reading times and sentence completion, Sedivy shows that the presence of the focus-sensitive operator *only* can induce readers to parse ambiguous material following a subject NP as a restrictive modifier, rather than as a matrix VP, as readers are otherwise inclined to do. These studies are discussed further in Sedivy’s summary (above) and in my comments below.

The demonstration that people are able to act upon contrast set information so rapidly is quite impressive. This could easily have not been true, and contradicts the expectations of some who have doubted that people could achieve this, due the complex inferential processes that the interpretation of contrastive focus demands (e.g. Ferreira & Clifton 1989), or more generally that information about the world is too complex to recruit efficiently in the early stages of language processing (e.g. Katz & Fodor 1963). Therefore, I am inclined to agree with Sedivy that “The robust and pervasive effects of contrast across all of these varied experimental conditions challenge the frequent assumption in the sentence processing literature that information about the world is too cumbersome or idiosyncratic to be useful in the early stages of processing” (p.142). On

the other hand, I am less certain of another conclusion which Sedivy draws regarding *how* speakers manage to use this information so effectively: “The very early influence of contrast upon patterns of eye-movements to the visual display suggests that such information is integral to the processing of adjectival meaning, and does not reflect a set of complex inferences that are computed only after linguistic representations are constructed” (p.140). In order to evaluate this claim about the mechanisms involved, we need to look more carefully at ways in which semantic interpretations and contextual knowledge could interact in order to yield Sedivy’s results.

Mechanisms

Having established that contrast-relevant information about adjectives can be computed extremely quickly, even before the head noun of the NP is processed, we must ask how this feat is achieved. I can see at least two different possibilities. The dissertation is rather ambivalent when it comes to the specification of precise interpretive mechanisms.

The first possibility is that the processing of contrast-relevant information really does involve a cumbersome set of inferences, but that this is another of the many complicated things which humans are remarkably good at. We might then expect this interpretive flair to extend to many other areas beyond focus.

The second possibility is that the interpretation of focus is a highly grammaticized area of semantic knowledge, and that knowledge of the grammar of focus is organized in such a way that it is highly conducive to rapid on-line use. In this case, we need not expect that this helpful property of focus should extend to other aspects of interpretation, which could well be processed in a slow and cumbersome fashion.

We can see how either of these alternatives might work, if we spell out how Sedivy’s results might be accommodated in Rooth’s *alternative semantics* approach (Rooth 1992). In this approach, contrastive focus conveys a comparison between the *ordinary semantic value* of an expression and its *focus semantic value*. The focus semantic value is related to the ordinary semantic value by simply replacing the focused material with a variable. For example, the sentence “WALLACE likes cheese” conveys the information that Wallace likes cheese, and conveys the information that this is to be contrasted with other expressions of the form “x likes cheese”. This much is specified by the semantics. The range of possible values for the variable x is determined by the discourse context (e.g. by a question like “Does Wallace or Wendolene like cheese?”). This part of the interpretation is therefore left to pragmatics.

We can straightforwardly apply this approach to imperative sentences like “pick up the LARGE red square”. This instruction corresponds to the interpretation in (1). The instruction to pick up the large red square is contrasted with instructions to pick up other red squares which are not large.

- (1) ordinary semantic value: pick up ιx , such that large (x) & red (x) & square (x)
focus semantic value: pick up x, such that P (x) & red (x) & square (x)

However, in Sedivy’s experiments the critical eye-movements occurred at a point when the instructions were not yet completed, and were therefore responses to *partial*

instructions, such as “pick up the LARGE...”. This requires that the semantics for contrastive focus be extended to provide interpretations for partial sentences. This raises interesting questions. Under one possible application of alternative semantics to incomplete instructions, we get the interpretation in (2).

- (2) ordinary semantic value: pick up ιx , such that large (x)
focus semantic value: pick up x , such that P (x)

This interpretation yields an instruction to pick up a large object, contrasted with an instruction to pick up an object with some other property. This is not enough, because it does not require the large object to match the non-large object(s) in any other respects, and therefore fails to predict Sedivy’s finding that in a display like Figure 1, subjects look to the large red square more than to the large yellow circle. In the interpretation for the full sentence in (1), the requirement that the target object be contrastive on the focused dimension and non-contrastive on all other dimensions is contributed by the remainder of the noun phrase. How do subjects manage to perform as if they have heard the entire NP when they have only heard the determiner and the contrastively focused adjective? I can see at least two possibilities.

The first possibility is that the additional constraint is provided *semantically*: the interpretation in (3) is just like (2), except that it also adds a free variable Q_1 (ranging over properties) to both the ordinary semantic value and the focus semantic value of the partial instruction. By virtue of the fact that the value of Q_1 must co-vary in the OSV and the FSV, the instruction now successfully identifies the object that is *contrastively* large.

- (3) ordinary semantic value: pick up ιx , such that large (x) & $Q_1(x)$
focus semantic value: pick up x , such that P (x) & $Q_1(x)$

The free variable could be contributed by one of two sources. First, it could be contributed by the interpretation of a predicted noun contributed by the syntactic/semantic parser. The parser may need to build a predicted noun head in order to allow the words “the” and “LARGE” to combine with one another and with the rest of the utterance (almost any phrase structure parser would require this). The prediction of the noun may be supported by the formulaic nature of the instructions given in Experiment 3, in which all instructions were of the form “Pick up the *size-adjective color-adjective shape-noun*”. Alternatively, the free variable may be contributed as part of the interpretation of the scalar adjective *large*, as suggested by Bierwisch (1987) and Pollard & Sag (1994). Sedivy suggests in Chapter 5 that this may be what is responsible for the bias to give a contrastive interpretation to non-focused scalar adjectives, and therefore, it might be adopted to explain the results involving contrastive focus.

An advantage of building the free variable into the semantics of scalar adjectives is that it could explain why Sedivy found little or no effect of contrastive focus in Experiment 5, in which the critical adjectives were contrastively focused *color* terms. (Experiment 5 was also different from the earlier experiments in that it used household objects rather than cardboard shapes.) If color adjectives lack the free variable found in scalar adjectives, the difference between the experimental results could be explained (Sedivy offers additional possible explanations for the differences).

A disadvantage of using the free variable to account for the focus effect is that it would sacrifice Sedivy's use of the free variable to explain the bias to interpret scalar adjectives contrastively, even in the absence of contrastive focus. Something is needed to explain why contrastive interpretations are favored in *all* contexts, but are accessed *even faster* in the presence of contrastive focus. A further disadvantage of this approach (Julie Sedivy, p.c.) is that the free variable introduced by a scalar adjective does not have to be fixed by the head noun; according to context and syntactic position, the free variable may be fixed to a larger or smaller comparison class. Thus, the indeterminacy problem remains.

Note that whether the crucial free variable results from syntactic structure-building or whether it arises from the semantics of scalar adjectives, the rapid interpretation of contrastive focus is not a consequence of complex inferential processes. Rather, under either of these interpretations, rapid interpretation of contrastive focus is a consequence of the organization of syntactic and semantic knowledge, which happens to be conducive to rapid incremental interpretation in this domain. This is consistent with Sedivy's claim that computing the interpretation does not involve a complex inferential process (see above). However, it is precisely because formal linguistic representations are first computed that these effects are found, suggesting that focus interpretation relies on an initial computation of formal representations. If the rapid computation of focus is due to the convenient form of syntactic and semantic representations, then Sedivy's findings do not undermine Fodor & Katz' sceptical position on the use of world knowledge in interpretation; rather, they would just lend further support to the notion that focus is a highly grammaticized domain.

Alternatively, we may try to invoke a pragmatic account of how contrastive focus is interpreted 'prematurely'. Under this account, the semantics contributes no more than is shown in (2). The additional restriction required to identify the unique object that is contrastively large is instead supplied by pragmatics. Note that although Rooth (1992) explicitly acknowledges the need for the pragmatic restriction of the focus semantic value to a smaller set of relevant substitutes for the focused expression (see also Sedivy p.49), this is probably not enough to guarantee the effect that we need in this case. What we need is not pragmatic restriction on the *contrasting* properties (the variable P in (3)), but rather a restriction on the *non-contrasting* properties (the variable Q₁ in (3)), something which comes for free when dealing with the interpretation of complete sentences.

One possibility is that the interpretation shown in (2) is first computed, yielding an instruction to pick up the unique object which is large, in contrast with objects that are not large. This interpretation might fail, because two objects in the display are compatible with this description (the large square and the large circle), contradicting the uniqueness requirement imposed by the definite determiner (there were no plural NPs in this experiment, so the definite determiner can be treated as a singular). As a consequence, pragmatic processes (to be specified) narrow down the set of objects to the red squares, as a result of which the target is correctly identified. Since this account relies on the uniqueness presupposition of the definite determiner, it predicts that changing the determiner to an indefinite would eliminate the effect of contrastive focus (e.g. "pick up a LARGE..."). I do not know the status of this prediction. Sedivy does test examples containing indefinite NPs in Experiment 8, using yes/no questions (e.g. "Is there a tall glass?"), and finds that the bias to interpret non-focused scalar adjectives as contrastive

remains with the change in definiteness. However, this does not necessarily entail that the *additional* advantage observed when adjectives are marked with contrastive focus would also be found in indefinite NPs.

One concern with this account, however, is that it is not even clear that the interpretation of (2) should fail if two large objects are identified. Under standard assumptions, the definite determiner requires that the object picked out by the *entire* NP be unique, not that the object picked out by the contrastively focused adjective alone be unique. Therefore, the pragmatic account suggested in the previous paragraph might not be viable. (There may be a way around this, involving an interpretive mechanism which seeks to assign a unique referent to a definite determiner as soon as a contrastively focused adjective is encountered – but that would be a different mechanism from the one shown here, and would not follow from the standard semantics for definites.)

The Contrast Bias

In discussing contrastively focused scalar adjectives, the concern was to explain how an appropriately articulated semantics/pragmatics can predict how quickly the grammatically required interpretation is accessed. But we must also consider Sedivy's other main finding, that there is a bias to interpret scalar adjectives contrastively, even in the absence of contrastive focus. This bias goes beyond what the linguistic forms demand, and is reminiscent of many other interpretive biases uncovered by the literature on ambiguity resolution, which make reference to a variety of structural, semantic and probabilistic factors.

The contrast bias may be grounded in semantic or pragmatic terms, and Sedivy discusses a couple of interesting possible explanations for it. It is not my goal here to evaluate the relative merits of these different explanations, but it is important to consider the connection between the contrast bias shown in Chapter 5 and the rapid focus interpretation effects shown in Chapter 4. Collectively, the results show that contrastive interpretations are accessed quickly, and that this process is facilitated yet further by the presence of contrastive focus. Why is this additional facilitation observed? In the previous section I assumed that speakers are rapidly computing the interpretation required by contrastive focus, and I discussed how the semantic interpretation procedures could allow this to happen so quickly. However, an alternative possibility is based on the fact that the instructions with contrastive focus happen to grammatically require an interpretation which speakers are independently biased to select. Under this scenario, the 150-200ms facilitative effect of contrastive focus is really just an enhancement of the independently motivated contrast bias. This said, I cannot claim to have an explanation to offer at this point of *how* contrastive stress could induce this facilitation.

Focus Sensitive Operators

In an earlier chapter of the dissertation (Chapter 3), Sedivy uses standard reading-time and sentence completion measures to show that the processing of classic main verb/reduced relative ambiguities (descendants of the venerable *The horse raced past the barn fell* (Bever 1970)) is affected by the presence or absence of the focus sensitive operator *only*. These experiments are designed to extend and confirm a proposal by Ni et

al. (1996). Experiment 1 shows that the presence of *only* in a subject NP facilitates the reduced relative interpretation of the ambiguous verb in sentences like (4). These sentences were presented without prior discourse context. Experiment 2 used a sentence completion task in which the test sentences were preceded by a brief discourse context. The results showed that subjects were more likely to treat the ambiguous verb as part of a relative clause if the head of the relative clause picked out all existing discourse entities (5a), than if the head of the relative clause picked out a subset of the existing discourse entities (5b) (see Sedivy's summary for further discussion of these examples).

- (4) Only the businessmen loaned money at low interest were told to record their expenses.
- (5) a. All the secretaries in the company were made to take a tough computing course.
 b. All the secretaries and accountants were made to take a tough computing course.
 c. Only the secretaries prepared for the exam _____.

These findings again show that readers are quick to act upon the consequences of focus. But as above, it is worthwhile to consider how focus could be acted upon so efficiently. If we adopt Rooth's semantics for *only*, we have a straightforward account of the interpretation of sentences like (6). In addition to the ordinary semantic value (OSV), which conveys the information that dogs bark, the focus semantic value asserts that if anything satisfies the expression $bark(x)$, then it must be a dog (I set aside for present purposes the question of whether the OSV is presupposed or asserted, cf. Horn 1972, Krifka 1991; also, the OSV is represented informally, implying no specific implementation of the semantics of generic statements).

- (6) Only dogs bark.

Ordinary semantic value: bark(dogs)
 Focus semantic value: $\forall x (bark(x) \rightarrow x = dog)$

Applied to a sentence fragment like (7), we may assume the following ordinary and focus semantic values, where P is a variable ranging over properties. We may assume that this variable is subject to existential closure until it is further specified.

- (7) Only secretaries ...

Ordinary semantic value: P(secretaries)
 Focus semantic value: $\forall x (P(x) \rightarrow x = secretary)$

In order to account for Sedivy's findings, we want the semantics to yield a bizarre or uninformative interpretation in situations where the secretaries are the only discourse entity, but not in situations where the secretaries are not the only discourse entity. This would ensure that the head noun would be less likely to be treated as focus-marked in scenarios like (5a), and more likely in scenarios like (5b). Does (7) achieve this? The

ordinary semantic value may be paraphrased as “there is some property such that secretaries have that property”; the focus semantic value may be paraphrased as “there is some property such that everything that has that property is a secretary”. In order to obtain the desired result, the focus semantic value must yield an uninformative assertion in some discourse models, but not in others. In a discourse model containing nothing other than secretaries, it is trivially true that all properties apply to only secretaries, therefore the assertion is uninformative. In a discourse model containing both secretaries and accountants, it is informative to assert that there is some property which only secretaries have. Thus, if the semantics of *only N* yields an uninformative interpretation, then this may be what is needed to trigger a search for a restrictive modifier to associate the focus marking with.

Note that in this case, the work of explaining the context sensitivity effect is done exclusively by the focus semantic value. In the discussion of contrastive focus above, we saw that it was critical that (i) the free variable co-vary between the ordinary semantic value and focus semantic value, and that (ii) the free variable be forced to range over a suitably restricted set of properties. These considerations do not seem to be important in the case of expressions with *only*.

Concluding Remarks

I should reiterate my enthusiasm for this dissertation. It is an example of very careful experimental research which establishes some elegant findings of interest to semanticists. I am especially enthusiastic because it deals with an area of semantic interpretation for which we do have well worked out formal models.

As I have emphasized here, though, in explaining rapid *incremental* interpretation, our existing formal models of the interpretation of complete sentences need to be elaborated and adapted. I found the experiments in the thesis convincing, and the semantic assumptions well-motivated. The main thing that I missed was an explanation of precisely how the experimental results follow from the semantic models adopted. This does not detract from the impressiveness of the results, but this additional step is crucial if formal semantics and models of incremental interpretation are to be properly integrated.

Some of the more programmatic claims in the dissertation present the main conclusion of the experiments as a strike against strong versions of the independence of linguistic form and meaning in language processing. Sedivy also argues that the work argues against the *primacy* of form over meaning (p.140, see quote above). While it is clear that formal representations and world knowledge must interact very rapidly in order to account for Sedivy’s findings, I feel that it is hard at present to draw more specific conclusions about the precise sequence of computations that account for her findings. Under at least some scenarios, as discussed above, rapid success in interpretation may only be possible because exactly the right syntactic/semantic representations are computed first.

One additional caveat is in order. Sedivy’s results show that speakers can use knowledge of the world/discourse model very quickly in the interpretation of focus. These results go beyond what has been shown before, and challenges the scepticism that others have voiced about the usefulness of world knowledge in language processing.

However, it should be remembered that most of Sedivy's results were obtained in a paradigm which provided subjects with an *extremely* vivid and constrained discourse model (e.g. the displays shown above), which placed minimal demands on the subjects' memory or imagination. Many questions remain about how the interpretive capacities which Sedivy demonstrates can be deployed in the frequently occurring situations in which the discourse model is far more complex (i.e. real life) or less vivid (e.g. reading).

In sum, Sedivy's study has opened up an exciting new line of inquiry, which we can hope will lead to closer connections between research in formal semantics and language processing, and has set high standards of experimental care for subsequent work to live up to. I eagerly await what is to come.

Acknowledgment

I am grateful to Satoshi Tomioka and Julie Sedivy for valuable discussion of the issues raised here.

References

- Bever, T. 1970. The cognitive basis for linguistic structures. In: J.R. Hayes (ed.), *Cognition and the Development of Language*. New York: Wiley, 279-352.
- Bierwisch, M. 1988. The semantics of gradation. In M. Bierwisch & E. Lang (eds.), *Dimensional Adjectives*. Berlin: Springer Verlag.
- Clifton, C. & F. Ferreira. 1989. Ambiguity in context. *Language and Cognitive Processes*, 4, 77-103.
- Horn, L. 1972. *On the semantic Properties of Logical Operators in English*. Ph.D. dissertation, UCLA.
- Katz, J. & J.A. Fodor. 1963. The structure of a semantic theory. *Language*, 39, 170-210.
- Krifka, M. 1991. A compositional semantics for multiple focus constructions. *Proceedings of SALT I*. Ithaca, NY: Cornell University Working Papers in Linguistics.
- Ni, W., S. Crain & D. Shankweiler. 1996. Sidestepping garden paths: Assessing the contributions of syntax, semantics and plausibility in resolving ambiguities. *Language and Cognitive Processes*, 11, 283-334.
- Pollard, C. & I. Sag. 1994. *Head-driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
- Rooth, M. 1992. A theory of focus interpretation. *Natural Language Semantics*, 1, 75-116.
- Tanenhaus, M., M. Spivey-Knowlton, K. Eberhard & J. Sedivy. 1995. Integration of visual and linguistic information during spoken language processing. *Science*, 268, 1632-1634.

colin@glue.umd.edu