



ACADEMIC  
PRESS

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Cognitive Psychology 45 (2002) 447–481

---

---

Cognitive  
Psychology

---

---

[www.elsevier.com/locate/cogpsych](http://www.elsevier.com/locate/cogpsych)

# Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution<sup>☆</sup>

Michael J. Spivey,<sup>a,\*</sup> Michael K. Tanenhaus,<sup>b</sup>  
Kathleen M. Eberhard,<sup>c</sup> and Julie C. Sedivy<sup>d</sup>

<sup>a</sup> *Department of Psychology, Cornell University, 238 Uris Hall, Ithaca, NY 14853, USA*

<sup>b</sup> *Department of Brain and Cognitive Sciences, University of Rochester,  
Rochester, NY 14627, USA*

<sup>c</sup> *Department of Cognitive and Linguistic Sciences, University of Notre Dame,  
Notre Dame, IN 46556, USA*

<sup>d</sup> *Department of Psychology, Brown University, Providence, RI 02912, USA*

Accepted 17 July 2000

---

## Abstract

When participants follow spoken instructions to pick up and move objects in a visual workspace, their eye movements to the objects are closely time-locked to referential expressions in the instructions. Two experiments used this methodology to investigate the processing of the temporary ambiguities that arise because spoken language unfolds over time. Experiment 1 examined the processing of sentences with a temporarily ambiguous prepositional phrase (e.g., “Put the apple on the towel in

---

<sup>☆</sup> We are grateful to Keith Rayner, Gerry Altmann, and two anonymous reviewers for very helpful comments on the article. This research was supported by an NSF predoctoral fellowship to M.J.S. while he was at the University of Rochester, a Sloan Foundation Fellowship in Neuroscience at Cornell University, and by NIH Grant HD-27206 to M.K.T. A subset of the data for Experiment 1 was published as part of a short report (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). This article contains the first complete report of the design, methods, and results for Experiment 1; the first published report of the three-and-one-referent context from that experiment; and the first report of Experiment 2.

\*Corresponding author. Fax: 1-607-255-8433.

E-mail address: [mjs41@cornell.edu](mailto:mjs41@cornell.edu) (M.J. Spivey).

the box”) using visual contexts that supported either the normally preferred initial interpretation (the apple should be put on the towel) or the less-preferred interpretation (the apple is already on the towel and should be put in the box). Eye movement patterns clearly established that the initial interpretation of the ambiguous phrase was the one consistent with the context. Experiment 2 replicated these results using prerecorded digitized speech to eliminate any possibility of prosodic differences across conditions or experimenter demand. Overall, the findings are consistent with a broad theoretical framework in which real-time language comprehension immediately takes into account a rich array of relevant nonlinguistic context.

© 2002 Elsevier Science (USA). All rights reserved.

*Keywords:* Spoken language comprehension; Word recognition; Sentence processing; Syntactic ambiguity resolution; Modularity; Context effects; Information integration

---

## **1. Introduction**

As natural language is comprehended in real time, listeners and readers are faced with the problem of resolving ambiguities at multiple levels of linguistic representation. The problem is pervasive: From the perspective of a real-time processing system, even unambiguous words and phrases are temporarily ambiguous. For example, the beginning of the spoken word “candy” is consistent with several lexical alternatives, including the word “candle.” Similarly, the syntactically unambiguous sentence “Put the apple on the towel” contains a prepositional phrase “on the towel” that modifies the verb phrase “put,” specifying the destination or goal where the apple is to be put. However, the prepositional phrase is also temporarily consistent with an interpretation in which it modifies the noun phrase “the apple,” as it does in Example (1).

(1) Put the apple on the towel into the box.

Beginning with Bever’s classic work in the early 1970s, sentences with local syntactic ambiguities have served as the primary empirical base for developing and testing models of syntactic processing (Bever, 1970; Frazier, 1978; Frazier & Clifton, 1996; Gorrell, 1988; Kimball, 1973; Pritchett, 1992). When a sentence containing a temporary ambiguity is resolved in favor of the less preferred alternative, as in Example (1), comprehenders often experience a feeling of having been led down the “garden path.” Moreover, these preferences are systematic; there is a strong tendency for sentences with similar structures to exhibit similar preferences.

Although the presence of systematic preferences for temporarily ambiguous sentences is well documented, models of sentence processing differ in how they account for these preferences. Models of ambiguity resolution can be divided into classes along two interrelated dimensions. First, models differ in whether they assume that a single syntactic alternative is initially

considered (serial models) or whether multiple alternatives are evaluated in parallel. Second, models differ in what information is used when—in the case of serial models to determine the initial analysis and in the case of parallel models to determine the relative viability of the alternatives.

At one end of the continuum are models in which a restricted domain of information, typically syntactic constraints or a subset of syntactic constraints, plays a privileged role in initially structuring the input or ranking the alternatives. For example, in the influential Garden-path model (Frazier & Rayner, 1982), an encapsulated syntactic processor initially structures the linguistic input, making a provisional commitment to a single structure using decision principles based primarily on structural complexity. Other encapsulated subsystems or modules are assumed to be responsible for other aspects of sentence processing, including lexical access, reference resolution, and assignment of thematic roles. Information from these modules does not inform initial syntactic decisions, but is used to evaluate and, if necessary, revise initial syntactic commitments (e.g., Clifton, Speer, & Abney, 1991; Ferreira & Clifton, 1986; Frazier & Clifton, 1996; Rayner, Carlson, & Frazier, 1983; Frazier, 1987; Mitchell, Corley, & Garnham, 1992; Pritchett, 1992).

At the other end of the continuum are constraint-based models in which rich lexical representations make available multiple syntactic alternatives, which are weighted by the frequency of lexical forms and their argument structures in specific syntactic environments. The alternatives are continuously evaluated using relevant linguistic and nonlinguistic constraints such as the semantic/thematic fit between a phrase and a potential argument position and the effects of information from the discourse context (e.g., MacDonald, Pearlmutter, & Seidenberg, 1994; McRae, Spivey-Knowlton, & Tanenhaus, 1998; Spivey & Tanenhaus, 1998; Trueswell, 1996; Tanenhaus & Trueswell, 1995; see also Bates & MacWhinney, 1989; Taraban & McClelland, 1988). A central claim of these models is that the complex patterns of structural preferences and interactions with discourse and local semantic context arise from simple, domain-independent integration mechanisms, without appeal to syntactic complexity as an explanatory primitive. Other models fall somewhere in between these two classes in the degree to which they rely on structural complexity, parallel analysis, and use of multiple constraints (cf. Boland, 1997; Gibson, 1998; Gorrell, 1988; Jurafsky, 1996; Stevenson, 1995).

Although models as different as restricted-domain serial models and multiple-constraint parallel models might seem to make dramatically different predictions that would be easily testable, the models differ primarily in their claims about when in processing nonsyntactic context has its effects. These claims have often been couched within a broader debate about the extent to which processing systems are modular, i.e., informationally encapsulated in the sense proposed by Fodor (1983). For example, Ferreira and Clifton

(1986, p. 348) argued that “If the syntactic processor (or parser) is modular, it should initially construct a syntactic representation without consulting nonsyntactic information sources. . . Notice, however, that the modular view does not imply that this higher-level information is never consulted by the language processor. It is important to distinguish between *initial* and *eventual* [original emphasis] use of nonsyntactic information.”

One important source of nonsyntactic constraints is the discourse context in which the syntactic ambiguity occurs. Crain and Steedman (1985) called attention to the fact that many of the classic structural ambiguities involved a choice between a syntactic structure in which the ambiguous phrase modifies a definite noun phrase and one in which it is a syntactic complement or argument of a verb phrase. Under these conditions, the complement analysis is typically preferred. Crain and Steedman noted that one use of modification is to differentiate an intended referent from other alternatives. Thus, the sentence in Example (1) might be uttered in a context in which there was more than one apple. In such a context, the modifying phrase “on the towel” provides information about which of the apples is intended. Crain and Steedman proposed that listeners might initially prefer the modification analysis to the complement analysis in situations that provided the appropriate referential context. Moreover, they suggested that referential fit to the context, rather than syntactic complexity, was the factor controlling syntactic preferences.

Numerous empirical studies have now been conducted to evaluate the extent to which initial parsing decisions are influenced by referential context, beginning with studies by Altmann and Steedman (1988) and Ferreira and Clifton (1986). (For recent reviews, see Altmann, 1996; Gibson & Pearlmutter, 1998; Spivey & Tanenhaus, 1998; Tanenhaus & Trueswell, 1995.) Nearly all of these studies have used printed text in which a discourse context is created by setting up a scenario and reading time is measured for critical regions of a sentence with a local ambiguity. This work has used text not because the psycholinguistic community was primarily interested in reading per se, but rather because the theoretical questions required response measures that can provide fine-grained temporal information about ambiguity resolution. Self-paced reading and especially monitoring eye fixations during reading provide this kind of information because processing difficulty can be measured for each word in a sentence (Rayner, 1998).

Although studies of syntactic ambiguity resolution using reading paradigms have provided, and continue to provide, invaluable information about the role of context in sentence processing, they also have some intrinsic limitations. One limitation is that reading time measures are primarily restricted to providing information about processing difficulty. That is, they do not provide information about *what* is being processed or *how* it is being processed, but merely indicate whether the processing required additional time compared to some baseline. A second limitation arises because, in reading, the linguistic expressions in the text create or evoke the referential

context for a sentence. However, it is important not to confuse the referential context for a sentence or utterance with the preceding linguistic context provided by the text. It is widely known that the relevant notion of “context” for a sentence cannot be equated with the preceding linguistic context, but also includes the accessible entities and properties in the interlocutors’ environment, as well as the set of presuppositions shared by discourse participants (cf. Clark, 1992): Semantic interpretation does not appear to distinguish the two. So, for instance, a quantifier such as “most” in a sentence like “Most are made of glass” can be uttered equally well within view of, say, a collection of vases the interlocutors are examining, as in the context of a sentence such as “Royal Doulton vases are hand-crafted. Most are. . .” when no such vases are anywhere in sight. More generally, the relevant notion of referential context that applies to all aspects of reference, including deictic devices, such as demonstratives, pronouns, tense, and deictic words such as “come,” “go,” “behind” and so forth, does not distinguish between information introduced linguistically, salient information in the environment, and even between presuppositions shared between conversational participants and created by the sentence being uttered.

This broader notion of context has important theoretical and methodological consequences. From a theoretical perspective, thinking of context in terms of linguistic expressions alone is likely to be misleading. From a methodological perspective, it is difficult to distinguish limitations on context effects that are intrinsic to reading from those that are due to the architecture of the language processing system. For example, effects of context in reading might be relatively weak because reading requires shifting focal attention throughout the text while maintaining prior information in memory. In addition, we do not know what information is salient to the participant, what his or her behavioral goals are, and what information in the context is deemed relevant at the point of ambiguity.

More generally, Clark and his colleagues have challenged the notion of context used in most psycholinguistic studies as poorly defined (cf. Clark & Carlson, 1982) and questioned whether the results of experiments conducted with traditional paradigms using relatively “decontextualized” materials will generalize to more normal language use. As a result, many psycholinguists interested in situated language have increasingly turned to paradigms in which conversational participants cooperate with one another in relatively well-defined tasks, typically with real-world referents and circumscribed behavioral goals (Clark, 1992, 1996; see also Barsalou, 1999; Glenberg & Robertson, 1999; Zwaan, 1999). In these situations, the context for comprehension is well defined. However, traditional on-line measures of processing are not well suited to studying language processing in natural tasks with real-world referents.

As a consequence, research in language processing has been largely divided into two broad traditions along methodological and theoretical lines.

One tradition, dubbed by Clark (1992) as the “language-as-action” tradition, focuses on spoken language processing in interactive settings, with real-world referents and clearly defined behavioral goals, using largely off-line methods. The other tradition, which Clark (1992) dubs the “language-as-product” tradition, uses response measures that are closely time-locked to the linguistic input in order to develop and evaluate detailed mechanistic processing models, using largely decontextualized language in which participants are told to “comprehend” the linguistic input, but not actually use it for a goal-directed behavior.

The experiments reported here examined the effect of referential context on syntactic ambiguity resolution using a paradigm that preserves important aspects of the typical language-as-action situation while monitoring comprehension with the temporal precision of the finest grained on-line response measures used in the language-as-product tradition. Our participants followed spoken instructions such as “Put the apple on the towel in the box” to manipulate objects in a visual workspace. While this task did not involve fully interactive conversation, it preserved three important assumptions of the typical language-as-action paradigm: (1) spoken language is the medium of communication, (2) the language takes place within a well-defined context, and (3) the participants have clear behavioral goals. Referential context was manipulated by varying the objects in the workspace, e.g., by having one or more apples in the display.

We monitored comprehension by recording participant’s eye movements using a lightweight eye-tracker mounted on a headband (Tanenhaus et al., 1995, 1996). In a pioneering experiment, Cooper (1974) demonstrated that eye movements to pictures are closely time-locked to relevant information in a spoken story. Subsequent research incorporating instructions and actions, initiated by Tanenhaus et al. (1995), showed that eye movements provide useful insights into the time course of reference resolution (Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus, 1995; Sedivy, Tanenhaus, Chambers, & Carlson, 1999; Trueswell, Sekerina, Hill, & Logrip, 1999; see also Altmann & Kamide, 1999; Arnold, Eisenband, Brown-Schmidt, & Trueswell, 2000, for results in a passive listening context), while providing sufficient temporal resolution to measure lexical access in continuous speech (Allopenna, Magnuson, & Tanenhaus, 1998; Spivey & Marian, 1999; Spivey-Knowlton, 1996).

## **2. Experiment 1**

The goal of this experiment was to determine whether a behaviorally relevant nonlinguistic context would affect syntactic ambiguity resolution when the context supported the normally less preferred syntactic structure. We focused on prepositional phrase ambiguities such as those illustrated in

Example (2), in which the prepositional phrase (“on the towel”) could either modify the noun phrase, (“the apple”), in which case it is an optional “adjunct” phrase or it could introduce a goal argument as in “Put the apple on the towel” [In Example (1), the attachment is syntactically disambiguated by the second prepositional phrase, whereas in Example (2), the sentence remains syntactically ambiguous]:

(2) Put the apple on the towel in the box.

Readers and listeners have a strong preference to initially interpret prepositional phrases such as these as goal arguments (e.g., Rayner et al., 1983; Ferreira & Clifton, 1986). Constraint-based models explain this preference as arising from an intersection of local constraints, e.g., a prepositional phrase introduced by “on” following the verb “put” nearly always introduces a goal argument (MacDonald et al., 1994; Trueswell & Tanenhaus, 1994). All structurally based theories predict an argument preference, although for different reasons. According to some models, the goal argument is preferred because it is syntactically simpler, involving construction of fewer syntactic categories (e.g., Frazier, 1987). Other models assume that there is a general preference for arguments over adjuncts (e.g., Abney, 1989; Pritchett, 1992).

In contrast to both constraint-based theories and structural theories, Crain and Steedman (1985) proposed that the preference for an argument arises due to nonlocal discourse factors interacting with linguistic presuppositions. In particular, given that a definite noun phrase presupposes the existence of a unique entity, modifying a definite noun phrase is most felicitous when the referent would otherwise not be unique in some domain, i.e., when there is more than one apple in the context. Crain and Steedman argued that listeners initially compute parallel structures and then make the simplest assumption necessary to integrate the linguistic input into a continuously updated discourse model. Simplicity was defined in terms of new presuppositions. In the absence of a referential context that requires modification to establish uniqueness, comprehenders would assume that there was a single entity, making modification redundant, thus favoring the argument analysis. A strong prediction stemming from this approach, then, is that when the context introduces multiple referents, listeners will prefer the modification analysis because no additional presuppositions are necessary and modification is required to establish a unique referent (Altmann & Steedman, 1988).

The Crain and Steedman proposal has motivated numerous experiments during the past decade, most of which have compared predictions made by the discourse-based referential approach with predictions made by serial parsing models in which referential context is used only to evaluate and, if necessary, revise an initial structure assigned according to simplicity-based structural principles. The typical study examines temporarily ambiguous sentences such as the sentence illustrated in Example (2) in discourse con-

texts that introduce either one or two potential discourse referents for the definite noun phrase (cf. Spivey & Tanenhaus, 1998). The primary question has been whether the multiple-referent contexts eliminate processing difficulty for the otherwise less-preferred modification analysis. The literature on this topic is extensive (for a review, see Spivey-Knowlton & Tanenhaus, 1994; Altmann, 1996), so here we restrict our focus to studies examining the effects of referential context on the resolution of prepositional phrase ambiguities.

In a typical study, a discourse context consisting of one or more sentences introduces either a single referent or multiple potential referents for a definite noun phrase in a subsequent target sentence. The target sentence contains a prepositional phrase that is temporarily ambiguous between being a noun phrase modifier and a goal argument for the verb. Disambiguating information is provided by the semantic content of the noun in the prepositional phrase, e.g., “The man fixed the rusty door with the lock/screwdriver” (from Altmann & Steedman, 1988), the presence of a second prepositional phrase, e.g., “George placed the record on the shelf onto the turntable” (from Ferreira & Clifton, 1986) or by a combination of the two, e.g., “John put the book on the Civil War on the table” (from Britt, 1994). Processing difficulty during reading is measured by monitoring eye movements or by using self-paced reading.

The generalization emerging from this literature is that referential contexts can influence attachment preferences for prepositional phrases (Altmann & Steedman, 1988); however, the strength of the context effect interacts with lexical information (Britt, 1994; Spivey-Knowlton & Sedivy, 1995). For example, Britt (1994) presented sentences such as “John put the book on the Civil War on the table” in discourse contexts that either introduced one potential referent (e.g., a book about the Civil War) or two-referents (e.g., a book about the Civil War and a book about gardening). In addition, Britt manipulated the argument requirements of the verb in the target sentence. Half of the sentences used verbs such as “put” that require a goal argument and half used verbs such as “drop” for which the goal argument was optional. The two-referent context eliminated processing difficulty for the prepositions modifying noun phrases for verbs with optional, but not obligatory, goal arguments (see also Liversedge, Pickering, Branigan, & van Gompel, 1998). These results are compatible with constraint-based models in which effects of contexts are immediate, but they interact with other constraints (Spivey-Knowlton & Sedivy, 1995). However, they are also compatible with structure-based serial processing theories in which obligatory argument assignment precedes use of context. Moreover, because disambiguation comes late in the prepositional phrase, it is possible to argue that the preposition was initially treated as introducing a goal argument, then rapidly revised based on lexical information (Frazier & Clifton, 1996). Ruling out a rapid revision hypothesis is further complicated because



the ambiguous sentences used by Britt (1994) were not compared to unambiguous baseline control sentences with similar structure and content. Finally, in materials such as those used by Britt and others, the sense of the preposition is often confounded with its syntactic/thematic role. For example, in a fragment such as “John put the book on the Civil War on the table,” “on” specifies a location when it introduces the goal argument “on the table” but not when it modifies the noun phrase “the book.”

It is important to carefully consider the nature of the referential mechanisms that are suggested by these data. Crain and Steedman (1985) originally motivated their referential account as a potential explanation for commonly observed parsing preferences, typically involving a preference for a simple unmodified noun phrase over a modified noun phrase. It was argued that a modified definite noun phrase presupposes a discourse context which contains multiple discourse entities that can be referred to by the noun alone, with the modifier providing distinguishing information. Such a discourse context is more complex than the discourse context for simple definite noun phrases, presumably accounting for the garden path effect. The demonstration that garden path effects did not occur with appropriately supporting contexts for modification has been argued to support this account. However, while the standard referential manipulations clearly do show that referential contexts exert powerful influences on parsing, they do not necessarily show that the effect is driven by presuppositions associated with modifiers. Rather, the results could have arisen from the incremental nature of establishing reference on-line, independently of any linguistically specific presuppositions (see also the Principle of Referential Failure from Altmann, 1987).

To illustrate, let us consider what happens in a situation of referential indeterminacy that does not involve syntactic ambiguity, as was the case in studies reported by Eberhard et al. (1995). When presented with a visual display, subjects showed clear evidence of attempting to establish reference on-line, in immediate response to incoming speech input. For instance, upon hearing the instruction “Touch the starred yellow square” in a display containing only one starred object, subjects launched an eye movement to the target shortly after the word “starred.” However, when presented the same instruction with a display that contained multiple starred objects, subjects either delayed eye movements to the target until disambiguating information was provided or scanned the class of items consistent with the input at that point. Thus, there is evidence that in the face of referential ambiguity, subjects anticipate further disambiguating information. These simple contexts are essentially analogous to the referential manipulations. In the typical two-referent contexts, referential indeterminacy is created, with two possible referents consistent with the linguistic input at a certain point. Thus, the preference for a modification reading of the ambiguous string is presumably driven by the anticipation of further disambiguating

information. The resulting effect is powerful and clearly has implications for distinguishing among theories of sentence processing. However, it does not directly support a view in which referential factors are either partly or wholly responsible for garden path effects in the absence of a context that creates referential indeterminacy. This is a subtle, but crucial point, especially as it becomes increasingly important for the field of sentence processing to precisely specify the interaction of mechanisms involved in on-line processing.

Evidence for the presuppositional source for the typical garden path effect can be seen by manipulating the definiteness of the noun phrase, as the referential effects are argued to result from presuppositions of uniqueness associated with definite noun phrases. Spivey-Knowlton and Sedivy (1995) showed that parsing preferences for ambiguously attached PPs were affected by the definiteness of the noun phrase (interacting with verb-based constraints), lending support to the claim that linguistic presuppositions are implicated in garden path effects. One might argue that such intrasentential effects could have derived from local statistical generalizations, such as the probability of encountering a modification phrase contingent on the definiteness of the noun phrase. A local statistical account is more difficult to provide when the referential constraints are nonlocal (i.e., originating from outside of the sentence, in the general discourse context). One goal of Experiment 1 was to establish whether nonlocal effects of referential context could be seen on-line even in the absence of referential indeterminacy (e.g., the three-and-one-referent context described below).

In this experiment, instructions such as those in Examples (3) and (4) were presented in visual contexts corresponding to the referential manipulations used in studies with text (e.g., Altmann & Steedman, 1988; Spivey & Tanenhaus, 1998; Spivey-Knowlton, Trueswell, & Tanenhaus, 1993). Note that the preposition “on” specifies a location for both the modification and argument analysis, so that sense is not confounded with syntactic and/or thematic role.

(3) Put the apple on the towel in the box. (syntactically ambiguous)

(4) Put the apple that’s on the towel in the box. (syntactically unambiguous)

The one-referent context contained an apple on a towel, a pencil, another towel, and a box, whereas the two-referent context contained an apple on a towel, another apple on a napkin, another towel, and a box; see Figs. 1A and B.

In the one-referent context (Fig. 1A), upon hearing “Put the apple,” in the instruction in Example (3), the participant can immediately identify the object to be moved because there is only one apple. Thus, shortly after hearing “apple,” he/she is likely to make an eye movement to fixate on the apple. The participant is then likely to assume that “on the towel” is specifying the goal of the putting event, thus making the empty towel relevant

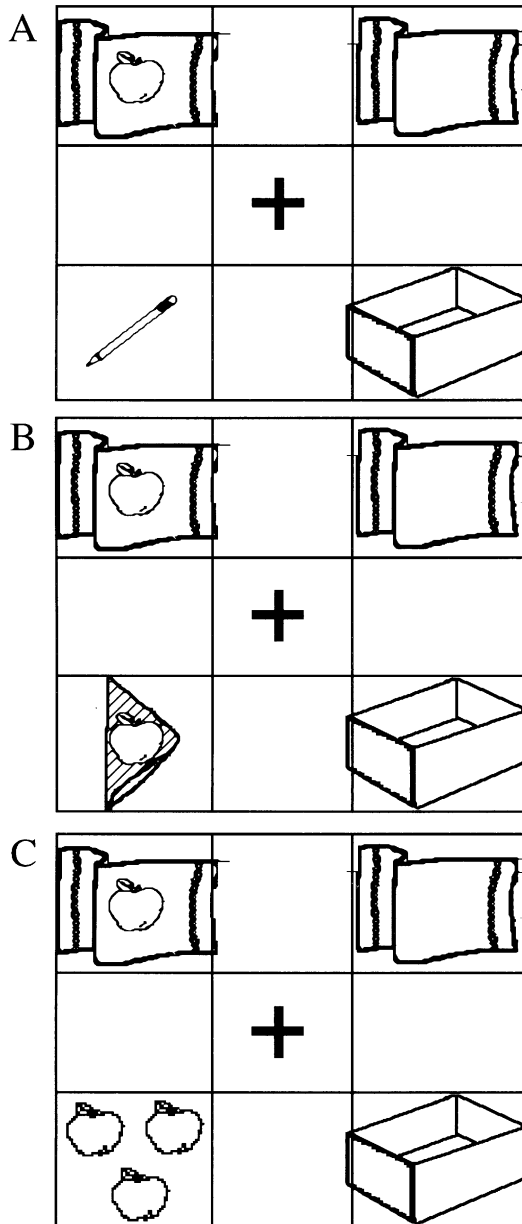


Fig. 1. Schematic example of the display conditions for Experiments 1 and 2: one-referent context (A), two-referent context (B), and the three-and-one-referent context (C).

for the action. As a result, attention should shift to the empty towel on at least a portion of the trials, resulting in an eye movement to it. In contrast, attention should be less likely to shift to the empty towel in the unambiguous instruction [Example (4)] because it never becomes relevant. Thus, a “garden path” due to initially misinterpreting the prepositional phrase as the goal argument should be reflected in more looks to the empty towel in the ambiguous instruction compared to the unambiguous instruction.

In the two-referent context (Fig. 1B), the referent of “the apple” will be temporarily ambiguous between the two apples. Under these conditions, the participant is likely to look at one or sometimes both of the apples, as in other situations involving referential indeterminacy (Eberhard et al., 1995). In the unambiguous instruction, the participant should immediately interpret the prepositional phrase as modifying the noun phrase, specifying which apple is intended. If the wrong apple was initially fixated, this should result in an eye movement to the correct apple shortly after hearing “towel.” In the ambiguous instruction, if the prepositional phrase is initially interpreted as introducing a goal argument then the participant should be more likely to look at the empty towel compared to the control condition—similar to the expected results for the one-referent condition. However, if the prepositional phrase is immediately interpreted as modifying the noun phrase, as predicted by a referential account (e.g., Altmann & Steedman, 1988; Crain & Steedman, 1985), then the patterns and timing of fixations during the ambiguous instruction should be similar to those during the unambiguous instruction.

The one-referent and two-referent contexts are visual analogs of the referential contexts that have been used in the reading literature. Crucially, we also included a “three-and-one-referent” context, in which the pencil in Fig. 1A was replaced with a group of three apples (Fig. 1C). This context contains multiple referents, but one of them, the single apple, is uniquely identifiable because it is by itself instead of in a group. In this context, upon hearing “put the apple. . .,” participants should be able to identify the single apple as the likely referent because if the other apples were being referred to, an indefinite reference would typically have been made (e.g., “an apple” or “one of the apples”). However, saying “the apple,” without modification, to refer to the single apple sounds distinctly odd in this context. Presumably, this derives from the meaning of definite noun phrases, which presupposes the existence of a unique object that corresponds to the descriptive content of the noun phrase. Thus, felicitous use of “the apple” would be limited to a context where there is only one apple, not merely one identifiable apple (see Roberts, 2000, for arguments that uniqueness presuppositions for definites pertain to the existence of unique objects, not unique identifiability). Thus, using the unmodified definite noun phrase “the apple” seems to be incorrectly committing the speaker to the existence of a unique apple in a display that also contains three other apples. If specific presuppositions associated

with linguistic form are used on-line in resolving syntactic ambiguity, such three-and-one-referent contexts should result in subjects interpreting the ambiguous phrase “on the towel” as a noun modifier, even though they are able to identify a unique referent on the basis of hearing “the apple” alone.

Equally importantly, the three-and-one-referent condition also provides a critical control for a possible alternative explanation should the two-referent context eliminate looks to the incorrect goal. In particular, it is possible to argue that in the two-referent condition, subjects were in fact temporarily garden-pathed, treating the prepositional phrase as the goal. However, the presence of two potential referents might have masked the effect of this garden path because activation of a potential fixation to the alternative referent was greater than activation of a potential fixation to the goal.

The argument goes as follows. Execution of eye movements is necessarily sequential; however, there is evidence that some planning of saccades may take place in parallel (for reviews, see Findlay & Walker, 1999; Desimone & Duncan, 1995; see also Becker & Jürgens, 1979; Rayner, 1998). When there are two potential referents for a word or phrase in a visual context, there is clear evidence that (1) the probability of fixating one of the candidate referents at a particular point in time is related to the linguistic evidence for that referent and (2) the presence of a potential competitor delays fixations to the correct referent, suggesting that the alternatives are competing with one another. For example, Spivey-Knowlton (1996) and Allopenna et al. (1998) showed that with instructions such as “Pick up the candy,” fixations early in the speech stream are equally likely to a piece of candy and a candle, when both are present in the display. Moreover, the presence of a competitor affects even the initial fixations to the correct referent, shifting their latency distribution downstream. If we assume that the planning of each saccade is a parallel process, such that internal representations of behaviorally relevant objects or spatial locations simultaneously compete for control of the eye movement system, then we might expect that a participant who looks at one of the potential referents has also experienced partial activation on the eye movement salience map for the alternative referent as well. That is, if at a certain point in the speech stream two objects are equally viable as saccade targets, and an eye movement is launched to one of them, it is likely that some degree of preparation for saccades to both objects had actually taken place.

In a directly relevant study, Eberhard et al. (1995) examined reference resolution for definite noun phrases, which were disambiguated by a postnominal relative clause. For example, an instruction such as “Put the five of hearts that is below the eight of clubs above the three of diamonds” was presented in visual contexts in which there were two fives of hearts, but only one was below an eight of clubs. During the noun phrase, subjects often made a saccade to each of the potential referents. Consider now the situation in

which a syntactically ambiguous instruction (e.g., “Put the apple on the towel in the box”) is presented with a display containing two apples (Fig. 1B). The subject is likely to fixate on one of the apples, shortly after hearing “the apple.” The other apple remains a good candidate for a fixation. Thus, at the time that the prepositional phrase occurs, a potential saccade to the incorrect goal (e.g., the empty towel) would be competing with a potential fixation to the other apple. In order to rule out this alternative interpretation, it is necessary to create a situation where there is a referential context that supports modification yet the likelihood of fixating an alternative referent is minimal at the place in the speech stream where the prepositional phrase occurs. The three-and-one-referent context should create just this situation. The referential context supports modification, on the more general pragmatic view, yet the use of a definite article makes it unlikely that the alternative referent group (e.g., the three apples) will attract many fixations.

### 2.1. Method

*Participants.* Six members of the University of Rochester community were paid for participating in the experiment.

*Stimuli and procedure.* All instructions were read out loud from a script. The experimental instructions used the prepositional phrase attachment ambiguity as described above, with either the form “Put the [x] on the [y] in the [z]” or “Put the [x] in the [y] on the [z].” In the experimental instructions, the target object and its distractor (e.g., the apple and the pencil) were always on the left half of the display (in the upper or lower square), and the possible goal objects (e.g., the box and the empty towel) were always on the right half of the display (in the upper or lower square). However, all other, nonexperimental, instructions were equally likely to involve moving an object in one of the four squares leftward or rightward.

At the beginning of a trial [containing a triplet of movement instructions, as in Example (5)], the objects were placed on the four corners of the workspace, as illustrated in Fig. 1. Participants were allowed to view this placement of objects, thus allowing them a few seconds worth of preview. The first instruction was always “Look at the cross.” The critical trials involved initial instructions with either ambiguous prepositional phrases such as that in Example (3) or unambiguous prepositional phrases such as that in Example (4) and two additional filler instructions that followed the experimental instruction. A total of 36 trials (or instruction triplets) were used, beginning with a filler trial and then alternating between experimental and filler trials. Eighteen of the trials began with experimental instructions, and the other 18 had entirely filler instructions. Thus, 90 of the 108 instructions (excluding “Look at the cross”) were filler instructions.

Each type of context was presented six times, three times each with an ambiguous and unambiguous instruction. A typical instruction set is pre-

sented in Example (5) (corresponding to the one-referent context in Fig. 1A). Note that the critical instruction was always the first instruction in the set. Any one visual context was only presented once to each participant. Each of the 18 instructions was rotated across six presentation lists, where a presentation list was generated by rotating instruction type (ambiguous and unambiguous) through the three types of visual contexts (one-referent, two-referent, and three-and-one-referent). Moreover, the objects themselves were reused for experimental and filler displays, so that all of the objects from experimental trials were also seen in filler trials.

(5) Look at the cross.

Put the apple on the towel in the box.

Now put the pencil on the other towel.

Now put it in the box.

In constructing the stimulus lists, great care was taken to avoid predictable contingencies in the instructions. For example, so that the first of a pair of prepositions did not always modify the noun phrase, 12% of the filler instructions involved initial prepositional phrases that modified the verb phrase (e.g., “Put the spoon in the cup on the saucer,” in a display containing one spoon resting on the table by itself and two cups, one of which was on a saucer). Additionally, 12% of the fillers used the preposition “on” to denote something other than an “on top of” relationship (e.g., “Now put the bowl on the saucer on the right,” where there were two saucers, one on the left and one on the right). The remaining 76% of the filler instructions contained only a single prepositional phrase, introducing a goal argument (e.g., “Put the pencil on the napkin”). Three of the six two-referent-context trials had filler instructions that referred to the alternate referent, and four of the six three-and-one-referent-context trials had filler instructions that referred to one or all of the alternate referents. The displays themselves were also controlled to avoid predictable circumstances. Of the 18 filler displays, 5 had pairs of noncontainer objects (e.g., pencils), 5 had pairs of container objects (e.g., bowls), 3 had triplets of noncontainer objects, and 5 had only singleton objects. In none of these filler displays did the first movement instruction involve a member of the pair or triplet.

## 2.2. Results

The data were analyzed using a Sony Hi-8 VCR with 30-Hz frame-by-frame playback and synchronized audio and video. The mean onset and offset for each of the words in the experimental instructions were determined and converted to milliseconds from the onset of the instruction. The mean length of the ambiguous instructions was 1875 ms compared to 2200 ms for the unambiguous instructions. This difference was due to the word “that’s” in the unambiguous instruction.

The initiation time for each eye movement during the instruction was determined. The onset of a saccadic eye movement was operationalized as the video frame in which the eye position left the square containing the object or cross from the previous fixation en route to the next fixation. Fixations were coded until the participant grasped the target referent. Analyses of variance were conducted across subjects ( $F1$ ) and items ( $F2$ ). However, due to track losses, three of the experimental instructions had missing data for at least one cell in the factorial design. Therefore, only the remaining 15 items were used in the by-items analyses.

Fig. 2A presents the proportion of trials in which participants looked at the distractor object(s), e.g., the pencil in Fig. 1A, the apple on the napkin in Fig. 1B or the cluster of three apples in Fig. 1C. Each proportion reported in

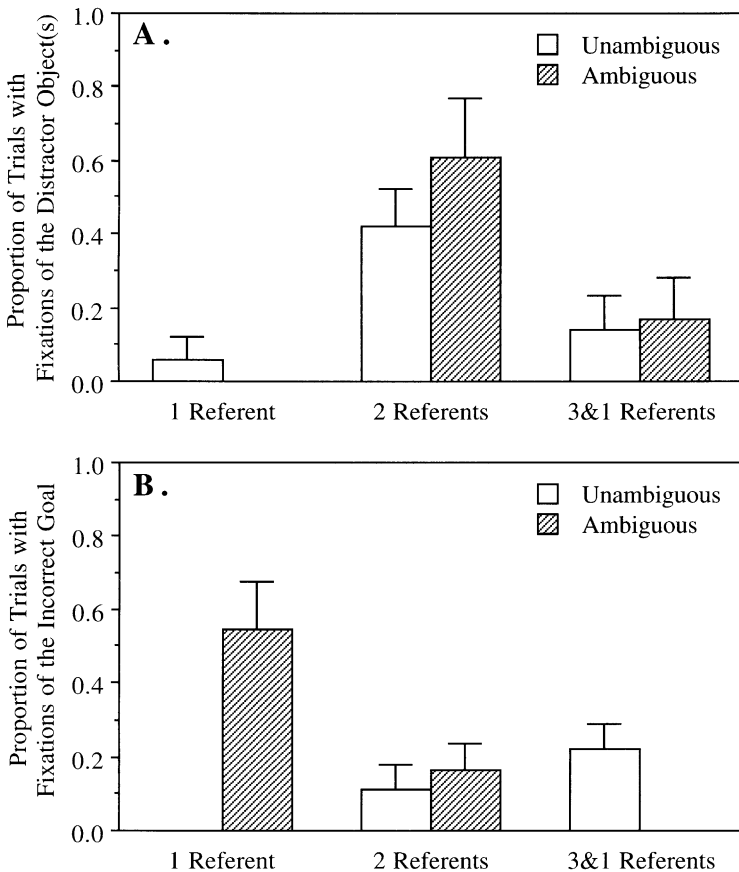


Fig. 2. Experiment 1: Proportion of eye movements to the distractor object (A) and to the incorrect goal (B) for the one-, two-, and three-and-one-referent contexts.



Fig. 2 is based on 16–18 data points. Due to the referential ambiguity, participants looked at the distractor object most frequently in the two-referent context, as indicated by a significant main effect of context [ $F(2, 10) = 25.15$ ,  $MSE = .0304$ ,  $p < .01$ ;  $F(2, 28) = 11.56$ ,  $MSE = .1500$ ,  $p < .01$ ]. Pairwise Tukey tests indicated that there were more looks to the distractor object in the two-referent context than in either the one-referent context ( $p < .01$ ) or the three-and-one-referent context ( $p < .05$ ), but the one- and three-and-one-referent contexts were not significantly different from one another ( $p > .1$ ). No other effects were obtained in the analysis of variance for fixations of the distractor object.

Fig. 2B presents the proportion of trials in which participants looked at the incorrect goal, e.g., the empty towel in Figs. 1A–C. In the one-referent context, the prepositional phrase was frequently misinterpreted as introducing a goal argument, as indicated by participants frequently fixating the incorrect goal object when the prepositional phrase was ambiguous, but never fixating it when the prepositional phrase was unambiguous. In contrast, as predicted by referential theory, the referential ambiguity in the two-referent context resolved the syntactic ambiguity toward a noun phrase modification, and participants rarely looked at the incorrect goal object—and they did so equally in the ambiguous and unambiguous conditions. Importantly, the three-and-one-referent context, despite not exhibiting an effect of referential ambiguity (see Fig. 2A), managed nonetheless to steer participants toward noun phrase modification, as indicated by very few fixations of the incorrect goal in that context (Fig. 2B). In the analysis of variance, there were only marginal main effects of context [ $F(2, 10) = 3.47$ ,  $MSE = .0257$ ,  $p < .1$ ;  $F(2, 28) = 4.93$ ,  $MSE = .0968$ ,  $p < .05$ ] and of ambiguity [ $F(1, 5) = 4.80$ ,  $MSE = .0298$ ,  $p < .1$ ;  $F(1, 14) = 4.70$ ,  $MSE = .1159$ ,  $p < .05$ ]; however, the interaction between the two was robust [ $F(2, 10) = 13.08$ ,  $MSE = .0346$ ,  $p < .01$ ;  $F(2, 28) = 13.33$ ,  $MSE = .1159$ ,  $p < .01$ ]. This interaction between context and ambiguity, indicating that the visual context was influencing syntactic ambiguity resolution, was also reliable when only one- and two-referent contexts were compared [ $F(1, 5) = 8.24$ ,  $MSE = .0435$ ,  $p < .05$ ;  $F(1, 14) = 28.00$ ,  $MSE = .0595$ ,  $p < .01$ ], as well as when only one- and three-and-one-referent contexts were compared [ $F(1, 5) = 26.45$ ,  $MSE = .0333$ ,  $p < .01$ ;  $F(1, 14) = 20.40$ ,  $MSE = .1381$ ,  $p < .01$ ]. When the two- and three-and-one-referent contexts were compared, the interaction was not significant [ $F(1, 5) = 3.31$ ,  $p > .1$ ;  $F(2 < 1)$ ].

An in-depth analysis of the timing of eye movements with respect to the speech stream is presented below, in a combined analysis of Experiments 1 and 2. At this point, an approximate timing of the fixations is sufficient to understand the typical patterns of eye movements. The vast majority of fixation patterns in the one-referent context began with a saccade to the target referent (e.g., the apple on the towel) about 500 ms after the end of the target

referent word, e.g., “apple.” With the unambiguous instruction, participants continued fixating inside the square containing the target referent until about 700 ms after the end of the sentence, at which point they fixated the correct goal location (e.g., the box). In contrast, with the ambiguous instructions, participants frequently (55% of the time) made a saccade out of the target referent region and into the incorrect goal region about 500 ms after the end of the head noun of the first prepositional phrase (e.g., “towel”). About 500 ms later, they would refixate the target referent, and about 800 ms after that they would fixate the correct goal (about 1300 ms after the end of the sentence).

In contrast to the one-referent context, the two-referent context showed no difference in fixation patterns or timing of the fixations, for the ambiguous and unambiguous instructions. In both contexts, participants looked at one of the potential referents shortly after hearing the first noun phrase (e.g., “the apple”). If the initial fixation was to the distractor referent (e.g., the apple on the napkin), which occurred on approximately 50% of the trials, then the eyes shifted to the target referent after the prepositional phrase, making the mean saccade latency to the target referent about 1100 ms after the end of the target referent word (e.g., “apple”). Participants rarely looked at the incorrect goal (14% of the trials), and there was no difference between the ambiguous and unambiguous instructions. Finally, the timing of the saccade to the correct goal (e.g., the box) was about 900 ms after the end of the sentence for both ambiguous and unambiguous instructions.

Similar to the two-referent context, the three-and-one-referent context showed no difference in the pattern or timing of eye movements between the ambiguous and unambiguous conditions. The typical eye-movement pattern for the three-and-one-referent context involved an eye movement directly to the lone target referent about 600 ms after the target referent word. Thus, saccade latencies to the target referent (the lone apple) resembled those of the one-referent context, indicating that the definite reference “the apple” was sufficient to uniquely identify the referent. Participants rarely (15% of the time) looked at the three apples. However, despite the fact that this context made the referent uniquely identifiable early on in the instruction, participants rarely looked at the incorrect goal (the empty towel). This is seen in the latter half of the eye-movement pattern, which resembled the two-referent context: participants looked at the incorrect goal 0% of the time in the ambiguous instruction and 22% of the time in the unambiguous instruction (this difference was not statistically reliable). About 700 ms after the end of the sentence, participants fixated the correct goal. Thus, both ambiguous and unambiguous instructions in the three-and-one-referent context elicited an overall eye-movement pattern most similar to that for the unambiguous instruction in the one-referent context, where participants looked first at the target referent and then upon hearing “in the box,” looked at the correct goal. This suggests that the decision to modify the noun phrase

“the apple” is not purely due to the presence of referential indeterminacy (e.g., Altmann & Steedman, 1988). Rather, it reflects on-line access to specific presuppositions associated with definiteness and modification.

This three-and-one-referent context also serves as a control for a possible alternative explanation for the eye movement pattern in the two-referent context. It could be argued that, in the two-referent context, participants were in fact briefly considering a goal interpretation of “on the towel,” but this did not result in an eye movement to the other towel because they had already programmed, or were developing activation on a “salience map” for, an eye movement to the distractor referent. However, in the three-and-one-referent context, participants typically looked immediately at the lone apple upon hearing “Put the apple” and continued to fixate on the referent until shortly after hearing “in the box.” Thus, given the sensitivity of the eye movement measure, any tendency to interpret the prepositional phrase as introducing a goal should have resulted in more fixations to the empty towel in the ambiguous compared to the unambiguous conditions. However, participants never looked at the incorrect goal (e.g., the empty towel) when the instruction was syntactically ambiguous.

### *2.3. Discussion*

Taken together these results clearly demonstrate that whether listeners initially interpreted the prepositional phrase as a goal argument or a noun phrase modifier was determined by the referential context established by the set of objects in the display. In the one-referent context, clear evidence that participants initially assumed a goal argument analysis came from frequent eye movements to the incorrect goal (i.e., the empty towel). However, in the two-referent and three-and-one-referent contexts, participants clearly assumed a modification analysis, showing the same pattern and timing of eye movements in both ambiguous and unambiguous instructions. In these contexts, participants were no more likely to look at the incorrect goal in the ambiguous versus unambiguous conditions, and they used the information in the prepositional phrase as a noun-phrase modifier instead of a goal argument. Thus, these results provide strong evidence against serial parsing models in which initial syntactic decisions are guided solely by syntactic information (e.g., Frazier, 1987; Frazier & Clifton, 1996). They also provide evidence against models in which the parsing of obligatory arguments is unaffected by input from the context (e.g., Britt, 1994; Pritchett, 1992). Rather, the results strongly support models of ambiguity resolution that allow for immediate interaction between context and syntactic processing (e.g., MacDonald et al., 1994; Spivey & Tanenhaus, 1998; Tanenhaus & Trueswell, 1995).

However, there is an important issue that needs to be addressed before it can be safely concluded that rapid integration of the visual context with the

unfolding linguistic input was responsible for the effects observed in this experiment. In Experiment 1, the instructions were read “live” from a script by the experimenter (M.J.S.). Although the experimenter could not see the display while reading the instructions, he did set up the display before each instruction sequence. Thus, it is possible that some intonational pattern, or variation in speed, in the instructions was responsible for the effects. Alternatively, the experimenter might have unconsciously transmitted some other cues. Thus, it was essential to determine whether the results could be replicated using prerecorded instructions in which exactly the same speech stream was used in different display conditions.

### 3. Experiment 2

Experiment 2 used the same stimuli and instructions as Experiment 1, but with prerecorded instructions that were played to the participant during the experimental session. Using prerecorded instructions rules out the possibility of experimenter bias. Moreover, the same acoustic signal was used across the different contexts, excluding any bias from speech intonation and timing during “live” presentation of the instructions.

#### 3.1. Method

*Participants.* Six members of the University of Rochester community were paid for participating in the experiment.

*Stimuli and procedure.* The set of 18 experimental trials (permuted across six stimulus lists) from Experiment 1 was used, with 18 filler trials intervening each experimental trial. (In contrast to Experiment 1, the filler trials in this experiment were experimental trials for an unrelated experiment; however, the overall distribution of single- and double-argument verb frames was the same as in Experiment 1.) Each trial had four instructions, beginning with “Look at the cross” followed by the critical instruction and two filler instructions. The 18 experimental trials exhibited one of three contexts (one-referent, two-referent, and three-and-one-referent) crossed with two instruction types (ambiguous and unambiguous). This allowed 3 trials per condition per participant. The 6 stimulus lists allowed each stimulus item to appear in each of its 6 conditions for different participants.

The prerecorded instructions were played back on a tape recorder, pausing between instructions. The ambiguous prepositional phrase attachment instructions (e.g., “Put the apple on the towel in the box”) were digitally converted from the unambiguous versions (e.g., “Put the apple that’s on the towel in the box”) by editing out “that’s.” (This digital editing resulted in a natural-sounding sentence. Informal tests with graduate students from

the second author's lab showed that listeners could not distinguish among spliced and unspliced versions of the instructions.) Three of the stimulus lists used exactly the same audio recording of critical instructions, with only the visual displays changing across those three lists for the prepositional phrase attachment trials. A second recording of the critical instructions, implementing the attachment ambiguity manipulation (“that’s” or no “that’s”), was constructed for the remaining three stimulus lists.

### 3.2. Results and discussion

The initiation time for each eye movement during the instruction was determined as in the previous experiment. The onset of a saccadic eye movement was operationalized as the video frame in which the eye position left the square containing the object or cross from the previous fixation en route to the next fixation. Statistical analyses were conducted across subjects ( $F1$ ) and items ( $F2$ ). However, due to track losses, seven of the experimental instructions had missing data for at least one cell in the factorial design. Therefore, only the remaining 11 items were used in the by-items analyses.

The pattern and timing of fixations closely paralleled those obtained in Experiment 1. The proportions of eye movements to the distractor objects (Fig. 3A) resembled those seen in Experiment 1 (Fig. 2A). Each proportion reported in Fig. 3 is based on 15–18 data points. There was a main effect of context [ $F1(2, 10) = 14.52$ ,  $MSE = .0517$ ,  $p < .01$ ;  $F2(2, 20) = 5.71$ ,  $MSE = .2227$ ,  $p < .02$ ], but the other effects did not approach significance. As before, pairwise Tukey tests indicated that there were more looks to the distractor object in the two-referent context than in either the one-referent context ( $p < .05$ ) or the three-and-one-referent context ( $p < .05$ ), but the one- and three-and-one-referent contexts were not significantly different from one another ( $p > .1$ ).

Similarly, the proportions of eye movements to the incorrect goal (Fig. 3B) also resembled those seen in Experiment 1 (Fig. 2B). There was a main effect of ambiguity [ $F1(1, 5) = 8.08$ ,  $MSE = .0421$ ,  $p < .05$ ;  $F2(1, 10) = 10.21$ ,  $MSE = .1485$ ,  $p < .01$ ] and an interaction between context and ambiguity [ $F1(2, 10) = 5.11$ ,  $MSE = .0875$ ,  $p < .05$ ;  $F2(2, 20) = 12.37$ ,  $MSE = .0894$ ,  $p < .01$ ], showing that syntactic ambiguity caused frequent fixations of the incorrect goal in the one-referent context but not in the other contexts. This interaction was seen again when only one- and two-referent contexts were compared [ $F1(1, 5) = 11.62$ ,  $MSE = .0623$ ,  $p < .02$ ;  $F2(1, 10) = 45.00$ ,  $MSE = .0409$ ,  $p < .01$ ] and when only one- and three-and-one-referent contexts were compared [ $F1(1, 5) = 6.60$ ,  $MSE = .0928$ ,  $p < .05$ ;  $F2(1, 10) = 13.91$ ,  $MSE = .1045$ ,  $p < .01$ ]. When the two- and three-and-one-referent contexts were compared, the interaction was not significant ( $F1 < 1$ ;  $F2 < 1$ ).

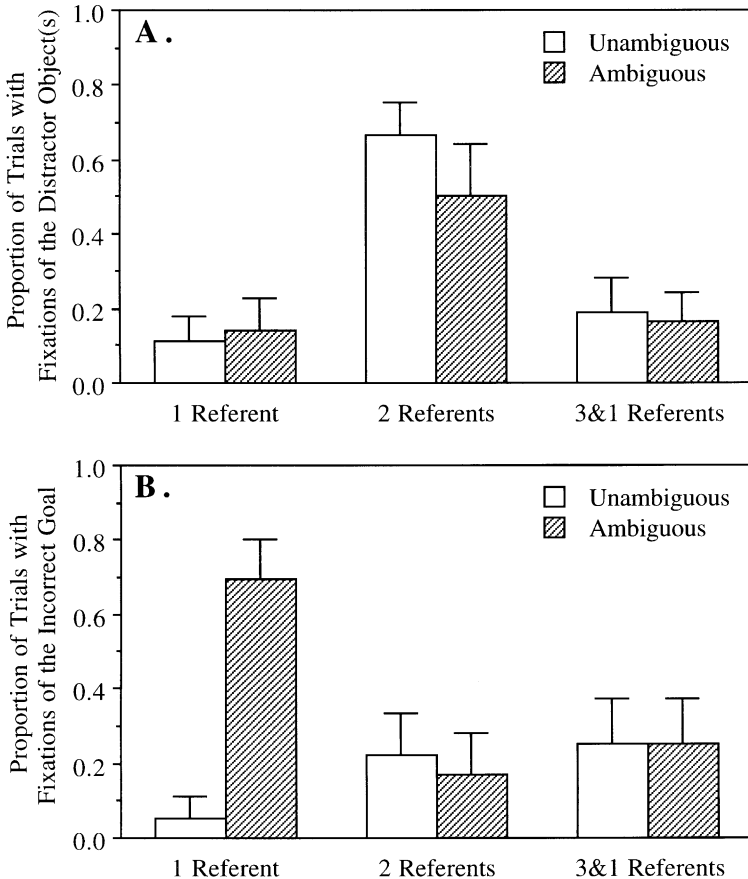


Fig. 3. Experiment 2: Proportion of eye movements to the distractor object (A) and to the incorrect goal (B) for the one-, two-, and three-and-one-referent contexts.

In the one-referent context with the ambiguous instruction, participants typically (69% of the time) looked first at the target referent (e.g., the apple) 642 ms after the end of “apple” and then looked at the incorrect goal (the upper right towel) 510 ms after hearing “towel.” In contrast, with the unambiguous instruction, participants generally did not look away from the target referent region until 495 ms after hearing the word “box,” at which point they looked directly to the correct goal (e.g., the box).

As in Experiment 1, the timing of eye movements relative to the speech stream was nearly identical for ambiguous and unambiguous instructions in the two-referent context. This suggests that participants were interpreting the prepositional phrase (“on the towel”) as a noun-phrase modifier (instead

of as a goal argument of the verb) equally quickly in both ambiguous and unambiguous instructions.

The three-and-one-referent context (e.g., where the target apple is accompanied by a set of three apples) also showed no effect of syntactic ambiguity. In both ambiguous and unambiguous instructions, participants looked at the incorrect goal on 25% of the trials. Timing of the eye movements to the target referent and then to the correct goal, relative to when those critical words occurred, was similar for the ambiguous and unambiguous instructions and comparable to that observed in Experiment 1.

### 3.3. Combined analysis of Experiments 1 and 2

The timing and pattern of fixations were similar across Experiments 1 and 2. In a combined analysis of variance of looks to the incorrect goal, with Experiment as a between-subjects factor, we observed the main effects of context [ $F(2, 20) = 3.74$ ,  $MSE = .0484$ ,  $p < .05$ ;  $F(2, 14) = 2.97$ ,  $MSE = .1086$ ,  $p = .08$ ] and ambiguity [ $F(1, 10) = 12.85$ ,  $MSE = .0359$ ,  $p < .01$ ;  $F(2, 1, 7) = 5.6$ ,  $MSE = .1190$ ,  $p < .05$ ], as well as a robust interaction of the two [ $F(2, 20) = 14.03$ ,  $MSE = .0610$ ,  $p < .001$ ;  $F(2, 14) = 18.63$ ,  $MSE = .0878$ ,  $p < .001$ ]. (Only eight items had no missing cells in both experiments.) Crucially, neither the main effect of Experiment nor its interactions with other factors approached significance (all  $ps > .1$ ). Thus, we were able to combine the data from the two experiments in order to provide more detailed information about the timing of eye movements as the instructions unfolded. Figs. 4–6 present the proportions of trials with fixations of the various objects in the display as the instructions unfolded in 33-ms time slices for each experimental condition.

Inspection of the time course graphs reinforces the conclusions from the data patterns summarized in Figs. 2 and 3. In all of the conditions, participants began to fixate the target referent early on in the instruction, then began shifting their gaze to the goal around the end of the instruction. However, the pattern and timing of the fixations differed across conditions. This is most striking for the one-referent context (Fig. 4), where the timing and pattern of fixations clearly differed for the ambiguous and unambiguous instructions. For the unambiguous condition, the proportion of fixations on the referent continued to rise until the beginning of the noun specifying the goal (e.g., the word “box”), at which point fixations began to shift from the referent to the goal (e.g., from the apple to the box). There are relatively few fixations to either the distractor (e.g., the pencil) or the incorrect goal (e.g., the empty towel). There are more fixations of the correct goal than of the target referent beginning at 900 ms from the end of the sentence. In contrast, for the ambiguous condition, about 15% of the fixations are on the incorrect goal (e.g., the empty towel), beginning with the noun in the first prepositional phrase (e.g., “towel”) and continuing for at least 2000 ms. Fixations

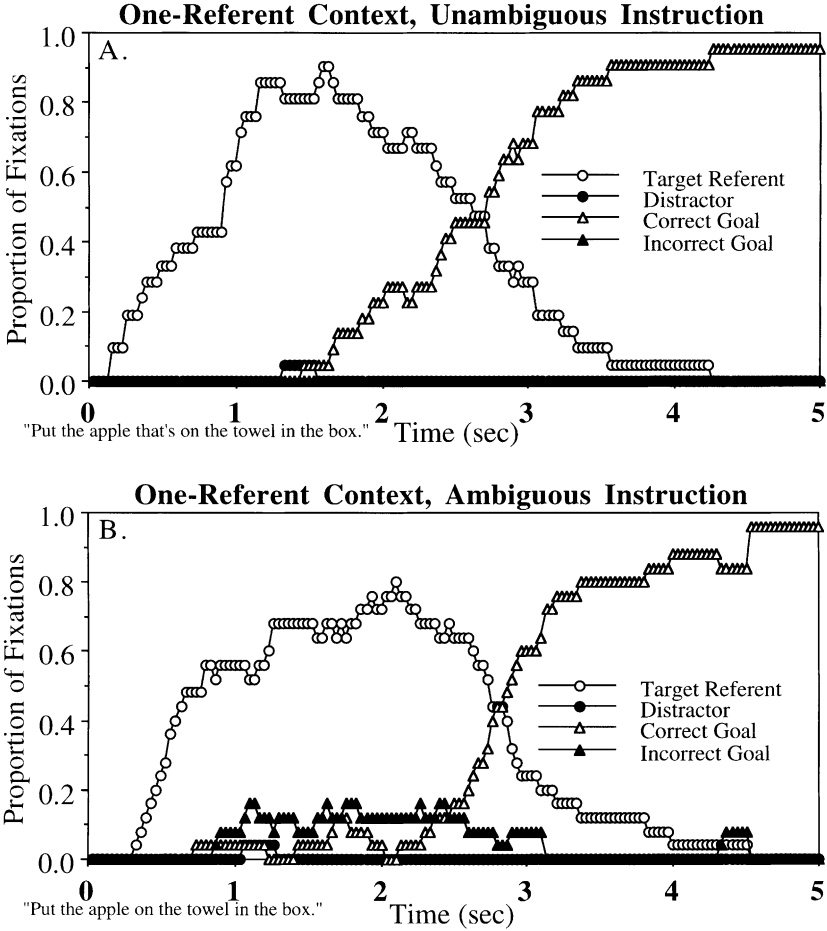


Fig. 4. One-referent context: Proportion of trials in which eye position was in each of the four regions as the instruction unfolded over time. The spoken instruction (in quotes) is aligned with its average duration along the timeline, with the measured onset of the target referent word (e.g., “apple”) at the beginning of the timeline (combined results from Experiments 1 and 2).

of the target referent do not peak until about 500 ms after the offset of the goal word (e.g., “box”), and the curves for the target referent and the correct goal fixations do not cross until about 1500 ms after the end of the sentence.

In order to provide additional statistical evidence for the presence of a garden path effect in the one-referent context, we compared the proportion of fixations to the correct goal object (e.g., the box) and the incorrect goal object (e.g., the towel), averaged over the time frame beginning with the onset of the word “towel” and continuing for 1500 ms. In this temporal win-



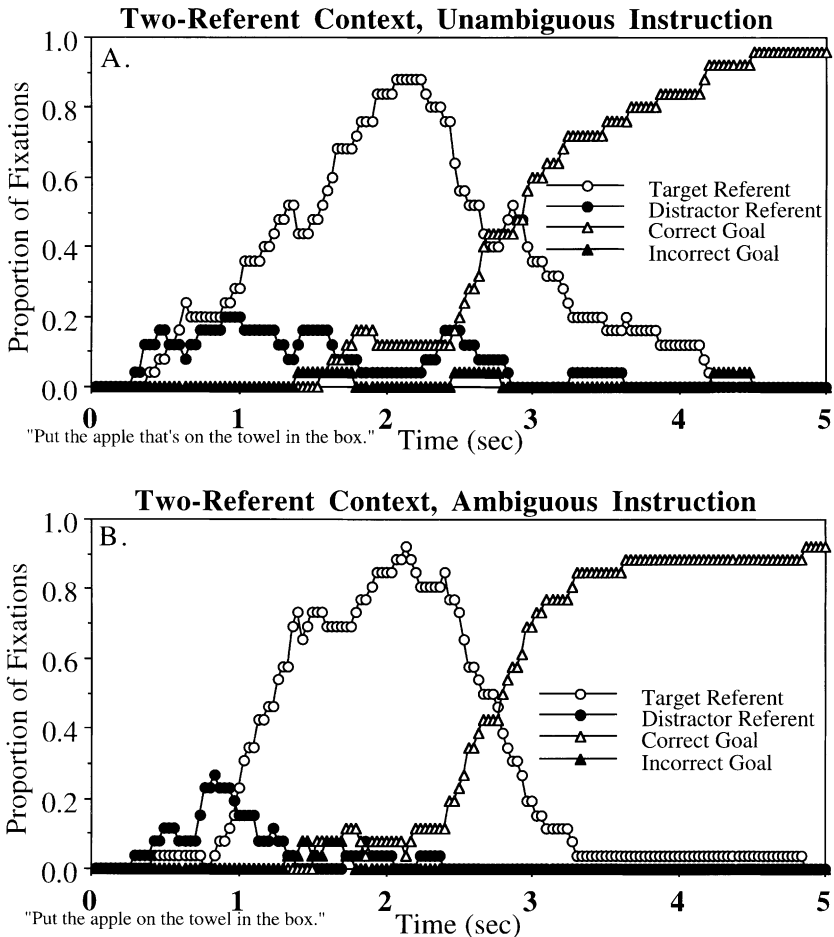


Fig. 5. Two-referent context: Proportion of trials in which eye position was in each of the four regions as the instruction unfolded over time. The spoken instruction (in quotes) is aligned with its average duration along the timeline, with the measured onset of the target referent word (e.g., “apple”) at the beginning of the timeline (combined results from Experiments 1 and 2).

dow, the proportion of fixations on the incorrect goal was greater than the proportion of fixations on the correct goal;  $t(11) = 2.24$ ,  $p < .05$ . In contrast, there was no temporal window in the unambiguous instruction in which proportion of fixations on the incorrect goal exceeded proportion of fixations on the correct goal.

For the two-referent context (Fig. 5), proportion of fixations on the target referent peaked about 500 ms after the end of the goal word (e.g., “box”) for the ambiguous condition and about 300 ms after the goal word for the unambiguous instruction. From the end of the noun introducing the referent

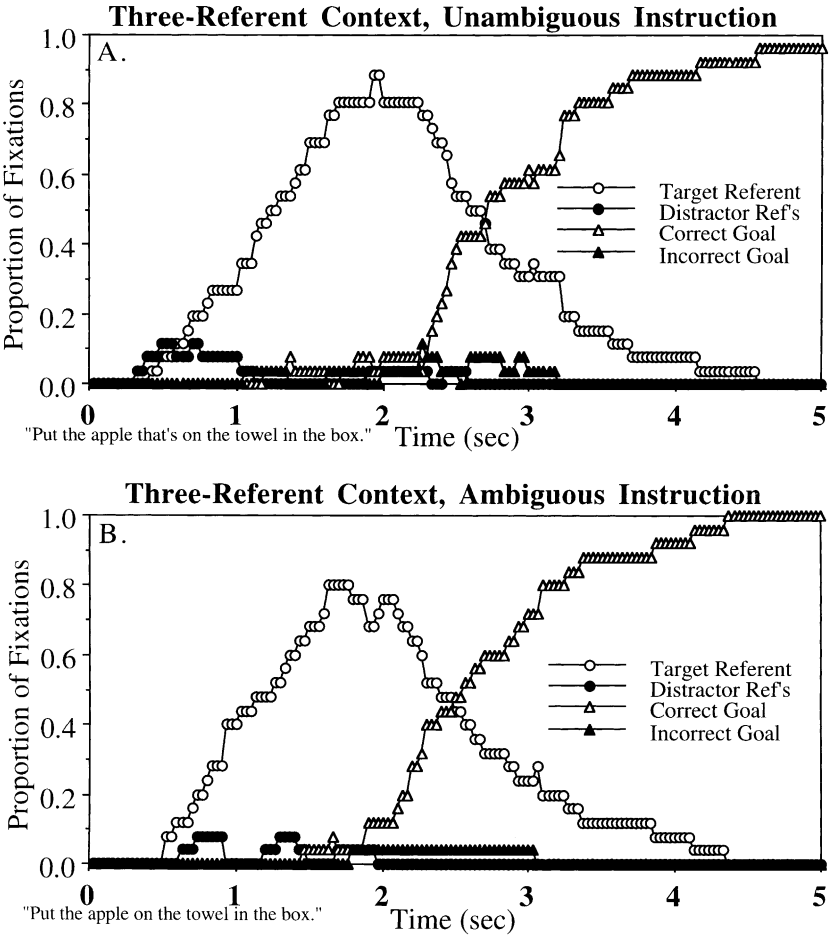


Fig. 6. Three-and-one-referent context: Proportion of trials in which eye position was in each of the four regions as the instruction unfolded over time. The spoken instruction (in quotes) is aligned with its average duration along the timeline, with the measured onset of the target referent word (e.g., “apple”) at the beginning of the timeline (combined results from Experiments 1 and 2).

(e.g., “apple”) until the onset of the goal word (“box”), fixations to the distractor referent were common. About 52% of the fixations were to the distractor referent in the unambiguous condition compared to 55% in the ambiguous condition. There were relatively few fixations of the incorrect goal in either condition. Proportion of fixations on the correct goal exceeded proportion of fixations on the target referent 1133 ms after the end of the sentence in the unambiguous condition and 1266 ms after the end of the sentence in the ambiguous condition. This may suggest a mild ambiguity effect

in the two-referent contexts; however, more importantly, this crossover latency in the ambiguous two-referent condition is delayed by 366 ms compared to the crossover latency in the unambiguous one-referent condition. This is most likely due to the fixations of the distractor referent. The timing of these fixation patterns confirms our concerns that competition from the distractor referent could have masked fixations to the incorrect goal, complicating the interpretation of the two-referent context results.

Further confirmation comes from examining the timing of fixations to the target referent in the ambiguous one- and two-referent contexts. For the one-referent context, 40% of the fixations were on the apple at the onset of “towel” and 56% at the offset of “towel” and throughout the following preposition. Assume that no other spatial location for a potential saccade is active for these particular trials at this point in the instruction. Further assume that on most of the remaining trials, a fixation to the apple has already been programmed or is in the process of being programmed, since nearly all participants fixated on the apple on each trial before making another fixation. This means that there is maximum potential for a saccade to the incorrect goal, if participants are being garden-pathed, on approximately 56% of the trials. Participants fixated the incorrect goal on 62% of the trials in the ambiguous one-referent condition, suggesting that they were nearly always misinterpreting the prepositional phrase as a goal. Now consider the two-referent condition. At the onset of “towel,” only about 20% of the fixations were on one of the potential referents, providing further evidence for competition between the multiple referents. By the end of “towel,” about 40% of the fixations were on one of the apples and only about 20% of these fixations were on the target referent (the apple on the towel), with the percentages rising to 30% by the end of the next preposition. Moreover, about 15% of the fixations were still on the distractor referent. Thus one could argue that there was relatively little opportunity for a garden path interpretation to result in a saccade to the incorrect goal. Even if the prepositional phrase was being interpreted as a goal, a potential fixation to that object would have to compete with potential fixations to the referents. Note, however, that the rapid drop in fixations of the distractor referent (apple on the napkin) beginning at “towel” suggests that the prepositional phrase was immediately interpreted as modifying “the apple” rather than as a goal.

In the three-and-one-referent context (Fig. 6), fixations of the target referent peaked during the goal word (e.g., “box”), only about 100 ms later than in the unambiguous one-referent condition. Although there was an occasional fixation to the distractor referents (e.g., the three apples), there were far fewer fixations than in the two-referent condition. Moreover, by the offset of the first prepositional phrase, there were hardly any fixations on an object other than the target referent. Thus, it is unlikely that competition from another potential referent was preventing a garden-path-induced fixation of the incorrect goal. There are only a few fixations to the incorrect

goal, and no hint of a difference between the ambiguous and unambiguous conditions. Fixations of the correct goal became more frequent than fixations of the target referent 933 ms after the end of the sentence for the unambiguous condition and 900 ms after the end of the sentence for the ambiguous condition, very similar to the crossover latency seen in the one-referent context with unambiguous instructions.

Crucially, subjects were typically fixating on the target referent as they began to hear the first locative prepositional phrase. At the onset of “towel,” about 15% of the fixations were on the target referent, rising to about 30% at the end of “towel,” and 45% by the offset of the next preposition. No other objects were being fixated at this point. Although this is less than the 56% observed for the one-referent condition, it clearly provides ample opportunity to observe a garden path effect. However, participants looked at the incorrect goal on only 12% of the trials compared to 62% for the one-referent context.

In sum, the temporal analysis of fixations as the instruction unfolded provides clear support for the claim that the prepositional phrase in the one-referent context was initially misinterpreted as a goal argument in the ambiguous instructions. The evidence comes from more looks to the incorrect goal, and delayed looks to the correct goal, compared to the unambiguous instruction. In the two-referent context, there was the suggestion of a slightly delayed interpretation in the ambiguous instruction compared to the unambiguous instruction, but strong evidence that the prepositional phrase was being interpreted as a noun phrase modifier. The pattern of fixations over time was similar for the ambiguous and the unambiguous instruction and there were few looks to the incorrect goal. However, in the ambiguous condition, the timing of fixations to the distractor referent and late fixations of the correct goal, compared to the unambiguous one-referent condition, raised the possibility that the presence of multiple referents might have inhibited the eye movements that would reveal a temporary misanalysis. Importantly, in the three-and-one-referent context, there was no suggestion of an ambiguity effect and, with few looks to the distractor referents, there was sufficient opportunity for saccades to be programmed to the incorrect goal, yet very few were observed.

## **4. General discussion**

### *4.1. Summary*

Most research on how discourse context influences syntactic ambiguity resolution has used written sentences, primarily because printed stimuli allow for use of response measures with the temporal grain necessary to distinguish among competing models. At issue has been whether the linguistic

input is initially structured by an encapsulated processing system in which syntactic information plays a privileged role in initially structuring the input. However, the focus on reading paradigms raises questions about the extent to which the pattern of context effects in the literature is due to the overall architecture of the language processing system in general, as is commonly assumed, in which case the same pattern should be observed in spoken language with copresent visual contexts or, alternatively, to limits on the strength and saliency of the discourse contexts that are typically used in reading studies.

We explored a paradigm for studying spoken language comprehension in which participants followed instructions to pick up and move real objects in a workspace, while eye movements were monitored using a lightweight headband-mounted eye tracker. Contexts were manipulated by arranging the objects in the visual workspace (Cooper, 1974; Tanenhaus et al., 1995). Under these conditions, the context was available to be interrogated by participants as the instruction unfolded, and it was clearly relevant to their behavioral goals. We assumed that listeners would shift their attention to objects that became relevant as the instruction unfolded. Attentional shifts are typically accompanied by a saccadic eye movement to the attended location in space (Hoffman, 1998), and many of the same cortical regions are involved in the two processes (Corbetta et al., 1998). Thus, eye movements are likely to be closely time-locked to comprehension processes.

We examined the comprehension of instructions such as “Put the apple on the towel in the box,” in which the prepositional phrase modifying the noun (“on the towel”) was temporarily ambiguous between a noun phrase modifier and a verb phrase complement introducing an obligatory goal argument (e.g., “Put the apple on the towel”). All parsing models which assume that structurally based principles alone guide initial ambiguity resolution predict that listeners will initially follow the argument analysis. Reading experiments with similar materials have shown that, whereas referential contexts supporting noun phrase modification can eliminate the initial verb phrase preference under some circumstances, referential contexts are ineffective for verbs with obligatory goal arguments, such as “put” (Britt, 1994). In contrast, the present results show clear evidence for strong and immediate referential context effects. When the context did not provide referential support for modification (e.g., an apple on a towel, a pencil, an empty towel, and a box), listeners initially assumed that the prepositional phrase “on the towel” was introducing the goal, as indicated by frequent fixations on the empty towel compared to an unambiguous baseline (“Put the apple that’s on the towel in the box”). However, in contexts providing support for modification (e.g., one apple on a towel accompanied by another set of one or more apples) the pattern and timing of eye movements were similar for the ambiguous and unambiguous conditions, and listeners clearly used the prepositional phrase to disambiguate which apple was intended.

This basic pattern of results was first presented in a brief report by the current authors (Tanenhaus et al., 1995) and recently replicated by Trueswell et al. (1999). The present article makes three important new contributions. First, the presence of the three-and-one-referent context allows for the conclusion that referential effects in syntactic ambiguity resolution reflect not only a highly incremental process of establishing reference, but also, more specifically, the on-line use of linguistically coded presuppositions. That is, it distinguishes between a strictly referential account, couched in terms of identification of a referent (e.g., Ni, Crain, & Shankweiler, 1996), and a more general pragmatic account in which expectations for modification affect ambiguity resolution. Second, and perhaps most important, the comparisons of the two-referent and one-referent contexts in Tanenhaus et al. (1995) and Trueswell et al. (1999) were insufficient to rule out the possibility that the effect of a temporary garden path on eye movements was masked by inhibition from potential eye movements to an alternative referent. However, the pattern of eye movements in the three-and-one-referent condition allowed us to rule out this type of explanation. Third, we ruled out the possibility that either experimenter bias or prosodic information in the speech stream, rather than the visual context, was responsible for the context effects on ambiguity resolution. Taken together, the results reported here provide the most definitive evidence to date that salient, relevant, and copresent visual context can completely eliminate even the strongest syntactic preferences.

#### *4.2. Conclusions*

Our results have important implications for research in syntactic processing and more generally for research in language comprehension. On the methodological side, the results reported here, along with other recent work by us and our colleagues (e.g., Allopenna et al., 1998; Eberhard et al., 1995; Sedivy et al., 1999; Spivey & Marian, 1999; Trueswell et al., 1999) demonstrate that it is possible to use eye movements to study spoken language comprehension using natural tasks and real-world referents, with the temporal precision of the finest grain response measures used to study reading. This has the potential of shedding new light on the processes underlying spoken language comprehension and production (see Keysar, Barr, & Horton, 1998; Hanna, Tanenhaus, Trueswell, & Novick, 2000) because many of the central issues identified within the language-as-action tradition cannot be studied using text or by monitoring eye movements to visual displays during passive listening. At the same time, fully exploiting the potential of the methodology will require us to develop more detailed linking hypotheses between underlying comprehension processes and patterns of fixations.

It is important to note that following spoken instructions and reading text are each natural comprehension situations, but they represent different

ends of a language comprehension continuum. Thus, it is risky to draw conclusions about the organization of the language comprehension system using results exclusively from either situation. Consider the specific case focused on here, namely how parsing decisions are affected by context. Numerous recent studies using reading-time paradigms have demonstrated that a variety of contextual constraints, including discourse context, interact with lexical constraints during syntactic ambiguity resolution (e.g., Boland, 1997; Boland, Tanenhaus, Garnsey, & Carlson, 1995; Britt, 1994; Garnsey, Pearlmutter, Myers, & Lotocky, 1997; MacDonald, 1993; McRae et al., 1998; Spivey & Tanenhaus, 1998; Trueswell, 1996). Whereas the interpretation of these results remains controversial, the empirical generalization is clear. In the present work, we found that appropriate referential contexts completely eliminated the bias for a verb phrase attachment for a prepositional phrase, following the verb “put,” which obligatorily takes a goal argument and thus has a strong lexical bias in favor of the verb phrase attachment. Taken alone, these data might be used to argue for a theory in which context completely determines ambiguity resolution (e.g., Ni et al., 1996). Conversely, the fact that referential context does not eliminate the verb phrase preference for verbs like “put” during reading was used to argue that argument assignment takes place without reference to context (Britt, 1994; Frazier, 1999; Frazier & Clifton, 1996; Liversedge et al., 1998).

We make the standard assumption that a single processing system supports comprehension across different modalities and different comprehension environments. When the results are combined, then, it is clear that the strength of global contextual constraints and of local lexical constraints varies across the two types of situations: local lexical constraints have stronger effects in reading, where the context is presented linguistically, whereas contextual constraints have stronger effects in spoken language when the context is presented visually and thus is copresent with the language and relevant to the behavioral goals of the comprehender.

Although this generalization is clearly consistent with the spirit of constraint-based models, such models might seem to have difficulty accounting for the fact that the strong lexical constraints and cooccurrences associated with the verb “put” were completely masked by the effect of context in the present experiments. There are two likely factors contributing to this observation, each of which we suggest is partially responsible for these results. The first factor is that the sense or role introduced by the ambiguous preposition, in this case “on,” was not confounded with the type of attachment—as it has been in previous studies. In the present studies, the temporarily ambiguous phrase introduced by “on” would have referred to a location regardless of whether it modified the Theme (i.e., the apple that was on the towel) or introduced the Goal, (i.e., where the apple was to be put). Although constraint-based models typically assume that verb preferences affect syntactic attachments, a complementary hypothesis is that they also

bias semantic/conceptual roles, such as introducing a location or instrument. Since these roles were not confounded with attachment type, the actual verb bias associated with “put [NP] on” might be weaker in our study than in studies in which the role and attachment for a phrase introduced by a preposition are confounded (e.g., Britt, 1994).

In order to understand the second factor, it is important to return to the broader notion of context, introduced earlier in this article. Natural language is inherently referential; it is about objects and events that are external to the linguistic expressions themselves. All linguistic expressions introduce, refer to or modify mental representations of entities, events, and their properties. Crucially, reference must be relativized to a domain of interpretation, which provides the context for reference resolution. This domain can include information introduced by the discourse, salient objects in the environment, and shared presuppositions between participants in a conversation. In relatively impoverished contexts similar to those typical of most reading experiments, much of the context must be created from the linguistic expressions as they are processed. For example, the first mention of the noun phrase “the apple” introduces an apple into the model, presumably a typical apple. Likewise, the verb “put” introduces a putting event, most likely a typical putting event, which will include a to-be-specified theme and goal. Under these circumstances, it makes sense that typicality information about entities, events, and cooccurring linguistic expressions will play a major role in creating the context that comprises the referential domain. However, when the context is more circumscribed, and the domain of interpretation is constrained by the visual context, the properties of the actual referents and plausible actions become part of the model. Thus, “the apple” refers to a particular apple in the display. Likewise, “put” refers to a particular type of intended action that involves moving a particular apple to one of a set of possible locations defined in the referential domain. Under these circumstances, the properties of the extralinguistic referential domain become much more important and salient than in a more impoverished context.

Clearly there is much to be learned about how listeners circumscribe referential domains and how linguistic knowledge is coordinated with extralinguistic information. However, the current results strongly suggest that approaches to ambiguity resolution that assign a central role to encapsulated linguistic subsystems are unlikely to prove fruitful. More promising are theories in which grammatical knowledge incorporates contextual parameters, thus supporting processing systems that coordinate linguistic and nonlinguistic constraints as the input is processed (cf. Sag & Wasow, 1999). Developing and evaluating such theories will require integrating the language-as-product and language-as-action research traditions. The methodological approach adopted here is an important step in this direction.



## References

- Abney, S. P. (1989). A computational model of human parsing. *Journal of Psycholinguistic Research*, 18, 129–144.
- Allopenna, P., Magnuson, J., & Tanenhaus, M. (1998). Tracking the time course of spoken word recognition using eye-movements: Evidence for continuous mapping models. *Journal of Memory and Language*, 38, 419–439.
- Altmann, G. (1987). Modularity and interaction in sentence processing. In J. Garfield (Ed.), *Modularity in knowledge representation and natural language processing* (pp. 428–444). Cambridge, MA: MIT Press.
- Altmann, G. (1996). Accounting for parsing principles: From parsing preferences to language acquisition. In T. Inui & J. McClelland (Eds.), *Attention & performance XVI*. Cambridge, MA: MIT Press.
- Altmann, G., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73, 247–264.
- Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, 30, 191–238.
- Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J. C. (2000). The rapid use of gender information: Evidence of the time course of pronoun resolution from eyetracking. *Cognition*, 76, B13–B26.
- Barsalou, L. W. (1999). Language comprehension: Archival memory or preparation for situated action? *Discourse Processes*, 28, 61–80.
- Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing*. New York: Cambridge University Press.
- Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision Research*, 19, 967–983.
- Bever, T. (1970). The cognitive basis for linguistic structures. In Hayes (Ed.), *Cognition and the development of language*. New York: Wiley.
- Boland, J. E. (1997). The relationship between syntactic and semantic processes in sentence comprehension. *Language and Cognitive Processes*, 12, 423–484.
- Boland, J. E., Tanenhaus, M. K., Garnsey, S. M., & Carlson, G. N. (1995). Verb argument structure in parsing and interpretation: Evidence from wh-questions. *Journal of Memory and Language*, 34, 774–806.
- 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.
- Clark, H. H. (1992). *Arenas of language use*. Chicago: University of Chicago Press.
- Clark, H. H. (1996). *Using language*. New York: Cambridge University Press.
- Clark, H. H., & Carlson, T. B. (1982). Hearers and speech acts. *Language*, 58, 332–373.
- Clifton, C., Speer, S., & Abney, S. P. (1991). Parsing arguments: Phrase structure and argument structure as determinants of initial parsing decisions. *Journal of Memory and Language*, 30, 251–271.
- Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language. *Cognitive Psychology*, 6, 84–107.
- Corbetta, M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Ollinger, J. M., Drury, H. A., Linenweber, M. R., Petersen, S. E., Raichle, M. E., Van Essen, D. C., & Shulman, G. L. (1998). A common network of functional areas for attention and eye movements. *Neuron*, 21, 761–773.
- Crain, S., & Steedman, M. (1985). On not being led up the garden path. In D. R. Dowty, L. Karttunen, & A. M. Zwicky (Eds.), *Natural language parsing*. Cambridge, MA: Cambridge University Press.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193–222.

- Eberhard, K., Spivey-Knowlton, M., Sedivy, J., & Tanenhaus, M. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research*, 24, 409–436.
- Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, 25, 348–368.
- Findlay, J. M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences*, 22, 661–721.
- Fodor, J. A. (1983). *Modularity of mind*. Cambridge, MA: MIT Press.
- Frazier, L., 1978. *On comprehending sentences: Syntactic parsing strategies*. Unpublished Ph.D. dissertation, University of Connecticut.
- Frazier, L. (1987). Theories of syntactic processing. In J. Garfield (Ed.), *Modularity in knowledge representation and natural language processing*. Cambridge, MA: MIT Press.
- Frazier, L. (1999). *On sentence interpretation*. The Hague: Kluwer.
- Frazier, L., & Clifton, C. (1996). *Construal*. Cambridge, MA: MIT Press.
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14, 178–210.
- Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, 37, 58–93.
- Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68, 1–76.
- Gibson, E., & Pearlmutter, N. (1998). Constraints on sentence comprehension. *Trends in Cognitive Science*, 2, 262–268.
- Glenberg, A. M., & Robertson, D. A. (1999). Indexical understanding of instructions. *Discourse Processes*, 28, 1–26.
- Gorrell, P. G. (1988). *Studies of human syntactic processing: Ranked-parallel versus serial models*. Ph.D. dissertation, University of Connecticut.
- Hanna, J. E., Tanenhaus, M. K., Trueswell, J. C., & Novick, J. (2000). *Effects of linguistic form and common ground on circumscribing referential domains*. University of Rochester manuscript.
- Hoffman, J. E. (1998). Visual attention and eye movements. In H. Pashler (Ed.), *Attention* (pp. 119–153). Hove, UK: Psychology Press/Erlbaum.
- Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, 20, 137–194.
- Keysar, B., Barr, D. J., & Horton, W. S. (1998). The egocentric basis of language use: Insights from a processing approach. *Current Directions in Psychological Science*, 7, 46–50.
- Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition*, 2, 15–47.
- Liversedge, S. P., Pickering, M. J., Branigan, H. P., & van Gompel, R. P. G. (1998). Processing arguments and adjuncts in isolation and context: The case of by-phrase ambiguities in passives. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 461–475.
- MacDonald, M. C. (1993). The interaction of lexical and syntactic ambiguity. *Journal of Memory and Language*, 32, 692–715.
- MacDonald, M. C., Pearlmutter, N., & Seidenberg, M. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101, 676–703.
- McRae, K., Spivey-Knowlton, M., & Tanenhaus, M. (1998). Modeling the effects of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, 37, 283–312.
- Mitchell, D., Corley, M., & Garnham, A. (1992). Effects of context in human sentence parsing: Evidence against a discourse-based proposal mechanism. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 69–88.

- Ni, W., Crain, S., & Shankweiler, D. (1996). Sidestepping garden paths: Assessing the contributions in syntax, semantics and plausibility in resolving ambiguities. *Language and Cognitive Processes*, 11, 283–334.
- Pritchett, B. L. (1992). *Grammatical competence and parsing performance*. Chicago: University of Chicago Press.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372–422.
- Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior*, 22, 358–374.
- Roberts, C. (2000). Uniqueness in definite noun phrases. Unpublished manuscript.
- Sag, I. A., & Wasow, T. (1999). *Syntactic theory: A formal introduction*. Stanford, CA: CSLI.
- Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, 71, 109–147.
- Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science*, 10, 281–284.
- Spivey, M. J., & Tanenhaus, M. K. (1998). Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 1521–1543.
- Spivey-Knowlton, M. J. (1996). *Integration of visual and linguistic information: Human data and model simulations*. Ph.D. dissertation, University of Rochester.
- Spivey-Knowlton, M. J., & Sedivy, J. C. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition*, 55, 227–267.
- Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1994). Referential context and syntactic ambiguity resolution. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing*. Hillsdale, NJ: Erlbaum.
- Spivey-Knowlton, M. J., Trueswell, J. C., & Tanenhaus, M. K. (1993). Context effects in syntactic ambiguity resolution. *Canadian Journal of Experimental Psychology*, 47, 276–309.
- Stevenson, S. (1995). Competition and recency in a hybrid network model of syntactic disambiguation. *Journal of Psycholinguistic Research*, 23, 295–322.
- Tanenhaus, M. K., & Trueswell, J. (1995). Sentence comprehension. In J. Miller & P. Eimas (Eds.), *Handbook of cognition and perception*. San Diego, CA: Academic Press.
- Tanenhaus, M. K., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of visual and linguistic information during spoken language comprehension. *Science*, 268, 1632–1634.
- Tanenhaus, M. K., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1996). Using eye-movements to study spoken language comprehension: Evidence for visually-mediated incremental interpretation. In T. Inui & J. McClelland (Eds.), *Attention & performance XVI: Integration in perception and communication* (pp. 457–478). Cambridge, MA: MIT Press.
- Taraban, R., & McClelland, J. (1988). Constituent attachment and thematic role expectations. *Journal of Memory and Language*, 27, 597–632.
- Trueswell, J. C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language*, 35, 566–585.
- Trueswell, J. C., & Tanenhaus, M. K. (1994). Toward a lexicalist framework of constraint-based syntactic ambiguity resolution. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing*. Hillsdale, NJ: Erlbaum.
- Trueswell, J. C., Sekerina, I., Hill, N., & Logrip, M. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, 73, 89–134.
- Zwaan, R. A. (1999). Embodied cognition, perceptual symbols, and situated models. *Discourse Processes*, 28, 81–88.