Journal of Memory and Language **47,** 30–49 (2002)
doi:10.1006/jmla.2001.2832

Circumscribing Referential Domains during Real-Time Language Comprehension

Craig G. Chambers

University of Calgary

Michael K. Tanenhaus

University of Rochester

Kathleen M. Eberhard

University of Notre Dame

Hana Filip

Northwestern University

and

Greg N. Carlson

University of Rochester

A head-mounted eye-tracking methodology was used to investigate how linguistic and nonlinguistic information sources are combined to constrain referential interpretation. In two experiments, participants responded to instructions to manipulate physical objects in a visual workspace. Instructions on critical trials contained definite noun phrases preceded by spatial prepositions (e.g., "Put the cube *inside the can*"). Experiment 1 established that the lexical–semantic constraints of the preposition *inside* immediately limited attention to objects compatible with those constraints (i.e., containers), suggesting that the referential context is dynamically restructured as sentence comprehension proceeds. Experiment 2 evaluated the additional influence of nonlinguistic constraints by varying the number of container objects that were large enough to hold the object being moved. The results indicated that attention was initially restricted to only those containers large enough to accommodate the object. This outcome suggests that referential candidates are continuously evaluated in terms of their relevance for the action denoted by the unfolding utterance. Overall, the findings are consistent with an expectation-driven interpretive system that rapidly integrates linguistic information with situation-specific constraints and knowledge of possible actions.

Key Words: incremental processing; referential interpretation; context.

Many linguistic expressions can be understood only with respect to a circumscribed context or *domain of interpretation*. Whereas contextual domains were traditionally discussed in relation to deictic or indexical expressions such as *you*, *here*, and *now* (e.g., Bar-Hillel, 1954), current semantic theory acknowledges their role in the interpretation of a broader range of expressions including quantified and unquantified noun phrases, focus and tense operators, modals, adjectives, and adverbials (Barwise & Perry, 1983;

This research was supported by National Institutes of Health Grant HD-27206. We are grateful to Ellen Hogan, Jared Novick, Sarah Scheidel, Dana Subik, and Eman Moustafa for their assistance in running the experiments. We also thank three anonymous reviewers for helpful comments on an earlier version of this article.

Address correspondence and reprint requests to Craig Chambers, Departments of Psychology and Linguistics, University of Calgary, Calgary, Alberta, Canada T2N1N4. Email: craig@ucalgary.ca.

0749-596X/02 $35.00

Partee, 1989; Roberts, 1995; Rooth, 1992). Thus, understanding how contextual domains are constructed and updated is of central importance for theories of language comprehension.

Domains of interpretation are defined by various kinds of information including the physical environment, prior linguistic discourse, and knowledge shared among conversational agents (Ariel, 1998; Clark, 1996). These domains are not static but rather are modified as new information is encountered. To date, this updating process has received little, if any, explicit consideration in studies of real-time language understanding. The research described here begins to investigate this process as it relates to *referential interpretation*. In particular, we investigate how different classes of information are used on-line to define the domain of interpretation for *definite noun phrases*.

Definite noun phrases provide a natural starting point for investigating domain construction because their referents must typically be *uniquely identifiable*, requiring a comprehender to differentiate a single entity from a particular subset of alternatives in the referential environment (e.g., Barwise & Perry, 1983; Clark, Schreuder, & Buttrick, 1983; McCawley, 1979). Research examining the real-time comprehension of definite noun phrases has revealed that reference resolution is highly incremental. For example, when presented with spoken instructions to manipulate referents of modified noun phrases (e.g., "Touch *the starred yellow square*"), listeners rapidly use each succeeding modifier or noun to winnow the set of alternatives to the intended referent (Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus, 1995). In addition, the ability to establish a unique referent for a definite noun phrase is believed to have implications for other aspects of real-time language processing such as the resolution of syntactic attachment ambiguities (e.g., Altmann & Steedman, 1988; Crain & Steedman, 1985; Spivey, Tanenhaus, Eberhard, & Sedivy, in press; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; Trueswell, Sekerina, Hill, & Logrip, 1999).

Because contextual domains provide essential restrictions on the range of entities available for definite reference, we can begin to address how domains are structured in time by considering the time course with which referential candidates are introduced into or eliminated from the domain. Below, we outline three conceptions of when and how this process occurs.

One possibility, reflected in a number of psycholinguistic, computational, and theoretical models of discourse interpretation, is that contextual domains are modified at sentence or utterance boundaries. For example, "mental models" approaches to text comprehension have suggested that updating occurs when the propositional representation of a new sentence is added to the mental model (e.g., Johnson-Laird, 1983; Morrow, Bower, & Greenspan, 1989). Similarly, in centering accounts of discourse coherence and interpretation (e.g., Grosz, Joshi, & Weinstein, 1995), the relative accessibility of discourse entities is evaluated when an utterance boundary is reached. A sentence-based approach to updating is also apparent in dynamic theories of semantic interpretation (e.g., Groenendijk & Stokhof, 1991; Stalnaker, 1978) in which the meaning of a sentence is expressed as a function from one context (or "information state"), *x*, to another context, *y*.

If a contextual domain is updated only at a sentence boundary, then the domain should not be affected by processing occurring at any point before this boundary is reached. However, as mentioned above, there is evidence that contextual information can be integrated with linguistic information on a word-by-word basis (Eberhard et al., 1995). This outcome is consistent with the idea that contextual domains may be modified more continuously. If so, then it becomes important to identify the information used in this process.

One source of information may be the semantic–conceptual constraints associated with individual lexical items (cf. Ballmer, 1981; Poesio, 1995). For example, given the utterance "Put the book inside the box," the preposition *inside* may be used to immediately restrict the domain for the subsequent noun phrase to only those referential candidates with container-like properties, consistent with the "basic meaning" (Herskovits, 1986) of this preposition. A more complex alternative is one in which contextual domains continuously integrate both lexical–semantic and nonlinguistic information sources. On this account, the domain of interpretation would also reflect an evaluation of which referential entities are relevant or possible candidates for the event(s) evoked by the utterance. For instance, on hearing "Put the book inside . . .", the domain will be limited to only those containers in the immediate environment that are large enough to hold the book. The increased complexity of this alternative stems from the need to integrate general world knowledge of actions and events with the event-relevant properties, or "affordances" (see Gibson, 1977), of situation-specific objects. However, given these additional computational requirements, it is quite possible that these pragmatic considerations cannot be used to constrain domains during the early moments of processing.

To evaluate the possibilities outlined above, we used an experimental paradigm in which eye movements are monitored as participants follow spoken instructions to manipulate real-world objects in a workspace (Tanenhaus et al., 1995). This technique allowed us to directly manipulate both the perceptual and linguistic context and to obtain a continuous on-line measure of the listeners' evaluation of the referential candidates as the instruction unfolded in time. Experiment 1 investigates whether contextual domains are constrained by the semantics of lexical items as these items are encountered. Experiment 2 addresses whether lexical–semantic effects are further constrained by pragmatic factors such as the consideration of possible actions.

EXPERIMENT 1

To evaluate the effects of individual linguistic expressions on contextual domains, we presented participants with instructions containing definite noun phrases preceded by spatial prepositions (e.g., "Put the whistle *inside the can*"). Spatial prepositions were used because they restrict the type of objects or entities that may occur as their internal argument(s). For instance, *inside* typically requires its noun phrase argument to possess container-like properties. As suggested above, this information may be used to limit the referential domain to only those candidates with compatible properties. When the preposition information narrowly restricts the set of candidates in the referential domain, we would expect a relative decrease in the time required to identify a referent for a noun phrase. This is because fewer candidates will need to be evaluated against the information provided by the noun phrase. For example, given the instruction mentioned above, the referent of the noun phrase "the can" may be identified faster when it is the only container in the display (see top panel of Fig. 1) than when several containers are present (see bottom panel of Fig. 1). In contrast, no difference should be observed with an instruction such as "Put the whistle *below* the can" because the preposition *below* is compatible with all referential candidates in the display.

Method

Participants. Participants were 12 members of the University of Rochester community recruited from posted notices. All were native speakers of English and received payment for their participation.

Materials. The visual materials consisted of eight objects placed on a tabletop marked with a 5×5 grid. The center square of the grid was marked with a cross. Four small objects (a clothespin, a small pair of scissors, a whistle, and a battery) were present in the center square at the beginning of each trial. These objects functioned as the "theme" objects for the instructions (i.e., the objects to be moved). Each trial also began with four additional objects on the display table that functioned as possible "goal" objects for the instructions. Each goal object occupied a corner square of the inner 3×3 grid, as illustrated in Fig. 1. Whereas the theme objects were the same on each trial, the four goal objects varied. On critical trials, the goal objects consisted of either three containers (e.g., a box, a bowl, and a glass) and one noncontainer or three noncontainers (e.g., a board, a napkin, and a rope) and one container. The linguistic materials consisted of 16 pairs of critical instructions. The form of the two instructions in each pair was "Pick up the X and hold it over the cross. Now put it below the Y." The preposition

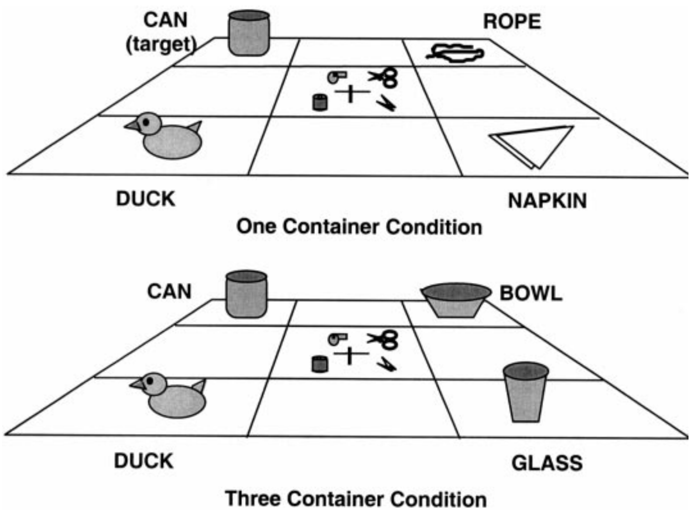

FIG. 1. Examples of experimental displays (Experiment 1).

in the second instruction was varied such that half of the trials contained *inside* and the other half contained *below*. The preposition manipulation was crossed with the display manipulation (three containers vs one container) to yield four experimental conditions. In all critical pairs of instructions, the target object referred to in the second instruction was a container. The target object appeared in four experimental trials—once in each experimental condition—and four target objects were used in total. The relative positions of target and nontarget goal objects were counterbalanced across trials.

In addition to the critical instructions, the materials contained 48 pairs of filler instructions. The filler pairs had the same form as the critical pairs except that they contained the prepositions *above* and *on* in addition to *below* and *inside*. Across all 64 pairs of instructions (16 critical plus 48 filler), each of the four prepositions occurred 16 times. In addition, 32 pairs referred to goal objects that were containers and 32 referred to goal objects that were noncontainers. All 64 pairs of instructions were presented once during an experimental session, with 2 pairs presented on each trial. On half of the 32 trials, the first pair of instructions were critical and the second pair were fillers; on the other half, both pairs of instructions were fillers.

Procedure. Participants were tested individually. They were seated in front of the display table, which was adjusted to accommodate their height and reach. They were told that they would receive instructions to move the objects on the tabletop and that they should follow the instructions in a natural manner including asking for clarification when necessary. They were then given several example instructions. After the examples, participants were fitted with a head-mounted eye tracking device (E4000, Applied Scientific Laboratories). The device consists of a lightweight eye camera and video scene camera attached to an adjustable headband. The eye camera provides an infrared image of the participant's left eye sampled at 60 Hz. Relative eye in-head position is calculated from the image by tracking the center of both the pupil and the first Purkinje corneal reflection. The video scene camera provides an image of the environment from the perspective of the participant. The scene image is displayed on a television monitor with superimposed crosshairs indicating the participant's point of fixation. A brief calibration procedure is conducted at the beginning of the experiment to map eye position coordinates onto corresponding scene image coordinates. The accuracy of the resulting eye movement record is within 1 degree of visual angle across a range of ± 20 degrees. An Hi8 videocassette recorder (VCR) is used to record the image on the television monitor along with the instructions, which are spoken by the experimenter into a microphone connected to the VCR. Software running on a personal computer allows point of gaze to be represented in an Hi-8 videotape record as a set of crosshairs superimposed on the visual scene captured by the scene camera.

A practice trial preceded the 32 experimental trials to ensure that the participants understood the procedure. The experimenter stood next to the participants and read aloud the pairs of instructions for each trial from a script. Because the first instruction in each pair directed the participants to pick up an object located in the center grid square and hold it over that square, the object being fixated at the beginning of the second instruction was equidistant from the four possible goal objects referred to in the second instruction. After both pairs of instructions were given, the experimenter and an assistant set up the display for the next trial. The accuracy of the eye movement record was monitored throughout the experiment by a second assistant, and minor adjustments were made between trials when necessary. The entire session lasted approximately 40 min.

Results and Discussion

Data were analyzed using frame-by-frame playback of the videotapes with the video and audio channels synchronized. The playback was used to locate the onsets and offsets of the spoken words in the prepositional phrases of the critical instructions. In addition, the timing and location of eye movements were scored beginning with the first fixation made 200 ms following the onset of the preposition and ending with the fixation on the goal object that preceded the reach toward it. This criterion ensured that the analysis contained only those eye movements that could plausibly have been programmed on the basis of the information in the preposition or the following speech. The locations of the eye movements were scored according to which squares in the display grid the intersection of the crosshairs appeared.

Figures 2 and 3 show the mean cumulative proportions of fixating the various objects in the display in the four experimental conditions. The vertical lines indicate the onsets of the three words in the critical region of the instruction and the offset of the final word. The zero point on the *x* axis is aligned with the onset of the article *the*; the other speech landmarks represent the average onset or offset. Fixations to nontarget objects were separated into container (distractor) and noncontainer (unrelated) objects in the three-container condition (nontarget objects in the one-container condition all were noncontainers). Figure 2 shows fixations in the two control conditions in which the preposition used was *below*. In both the one-container (top panel) and three-container (bottom panel) conditions, fixations to the target referent begin to diverge from nontargets at about 350 to 400 ms after the onset of the noun identifying the target referent. Nontarget objects were fixated before the target on only a few trials, demonstrating that participants generally waited until sufficient information was available to uniquely identify the referent before making eye movements.

Figure 3 shows the results for the conditions in which the preposition was *inside*. The results for the three-container condition (bottom panel) were similar to the pattern of fixations presented for the *below* conditions in Fig. 2. Specifically, the likelihood of fixating a target object began to diverge from the likelihood to fixate a nontarget object around 350 to 400 ms after the onset of the head noun. In contrast, in the one-container condition (top panel), fixations to the target object began to diverge from fixations to nontarget objects during the offset of the preposition. This result suggests that listeners were able to use the preposition to restrict the referential domain to the single object that was a plausible container.

To provide a statistical analysis of the data, we analyzed the cumulative proportion of fixations across 100-ms temporal windows measured relative to the onset of the article preceding the final noun. Within-subjects analyses of variance (ANOVAs) were conducted separately for each condition to determine the point at which fixations to the target object were reliably greater than fixations to other display objects.

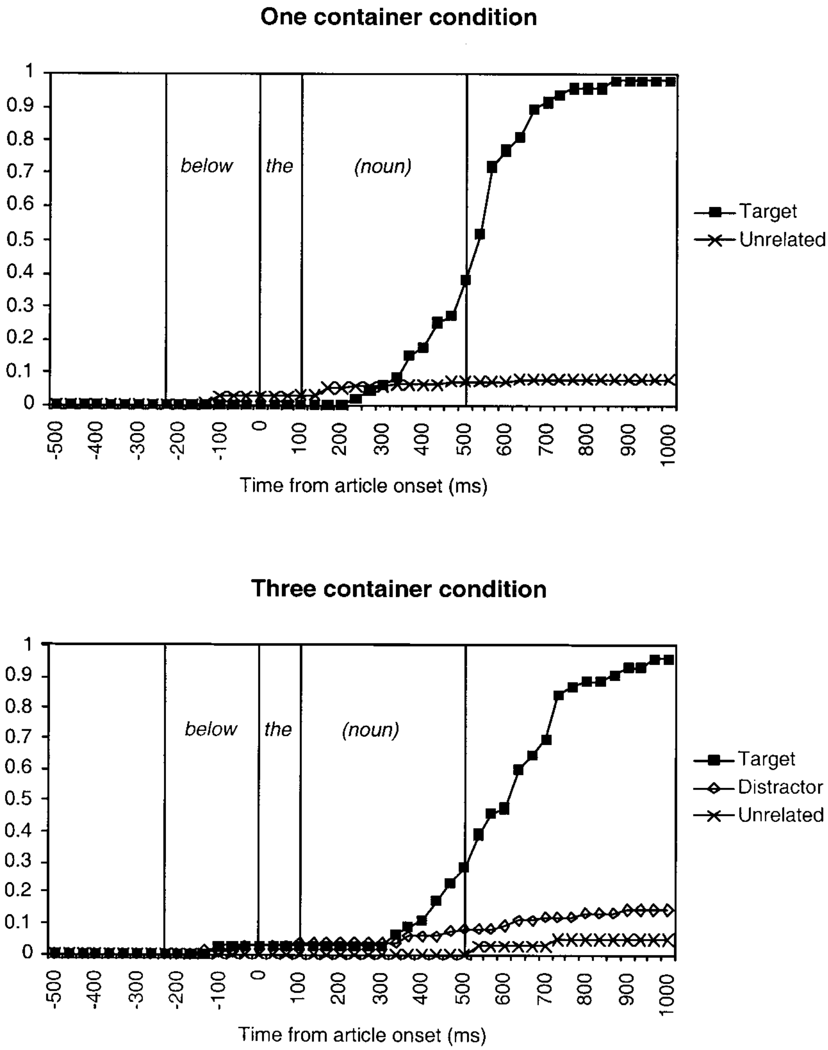

FIG. 2. Cumulative proportions of fixations to display objects, *below* conditions (Experiment 1).

The proportion data were submitted to an arcsine transformation before conducting the analysis. Because a counterbalanced design was used, only by-subjects analyses are reported (cf. Raaijmakers, Schrijnemakers, & Gremmen, 1999). We begin with the results for the *below* conditions illustrated in Fig. 2. No reliable differences were detected in any time interval prior to the 400- to 500-ms interval, at which point the proportion of fixations to the target was greater than that to noncontainer objects in both the one-container and three-container conditions, $F(1, 11) = 11.03$, $p < .01$, $MSE = .03$, and $F(1, 11) = 14.02$, $p < .01$, $MSE = .03$, respectively. The difference in the proportion of fixations to targets versus container distractors in the three-container condition was not fully reliable until the 500- to 600-ms interval, $F(1, 11) = 17.35$, $p < .01$, $MSE = .06$, although it was marginally reliable in the 400- to 500-ms interval, $F(1, 11) = 4.32$, $p = .06$, $MSE = .04$.

It has been suggested that the minimum latency to plan and launch a saccade is between 150 and 180 ms in simple tasks (e.g., Fischer, 1992; Saslow, 1967) and that intersaccadic intervals in visual search tasks fall in the range of

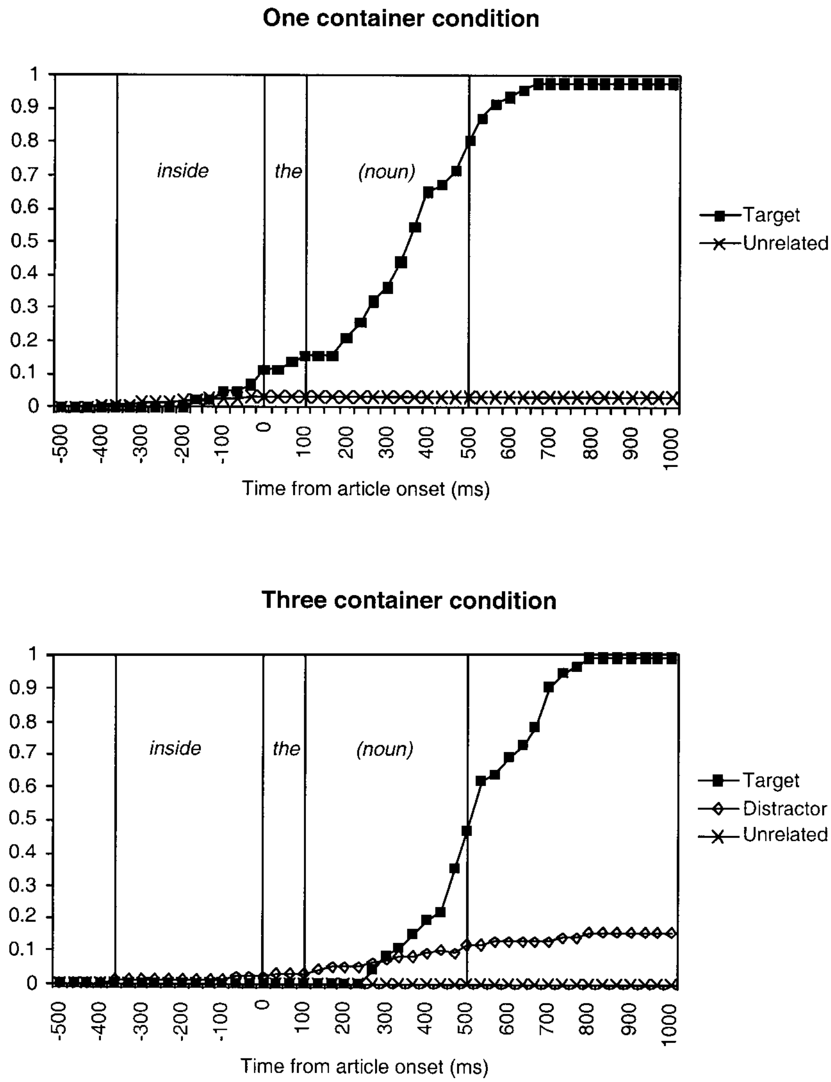

FIG. 3. Cumulative proportions of fixations to display objects, *inside* conditions (Experiment 1).

200 to 300 ms (e.g., Viviani, 1990). Thus, in the current experiment, fixations that are driven by the speech information are likely to begin about 200 ms after the relevant speech information is encountered. This estimate has been supported by the results of a number of recent studies (e.g., Allopenna, Magnuson, & Tanenhaus, 1998; Dahan, Magnuson, & Tanenhaus, 2001). Given that the average duration of the article in the critical noun phrase was only approximately 100 ms, the results suggest that the earliest fixations to the intended target in the *below* conditions were driven by the speech information encountered in the initial portion of the final noun and not by information in the preposition or the article.

We now turn to the results from the *inside* conditions illustrated in Fig. 3. A significantly greater proportion of fixations to the target than that to noncontainer objects was detected in the one-container condition in the 0- to 100-ms interval, $F(1, 11) = 5.31$, $p < .05$, $MSE = .02$. In the three-container condition, however, this difference was not reliable until the 300- to 400-ms interval, $F(1, 11) = 7.03$, $p < .05$, $MSE = .02$, although it was marginally reliable in the 200- to 300-ms interval, $F(1, 11) = 4.86$, $p = .05$, $MSE = .002$. However, the difference between the proportion of fixations to the target and that to the container distractors in the three-container condition was not reliable until the 400- to 500-ms interval, $F(1, 11) = 23.82$, $p < .01$, $MSE = .03$.

In contrast to the results from the *below* condition, the data pattern from the *inside* conditions suggests that, in the one-container condition, fixations to the target container were driven by information provided by the preposition. When three containers were present, fixations to the target were delayed until the final noun was encountered. In addition, fixations to the target container diverged from the noncontainer object earlier than from the container distractors in the three-container condition, providing additional evidence that preposition information is rapidly used to constrain reference.

In sum, the data suggest that lexical–semantic properties of prepositions dynamically restrict the domain of interpretation for a following noun phrase. The results are inconsistent with the claim that contextual domains are updated only when an entire sentence has been comprehended and suggest instead that updating is driven by the information contained in individual lexical items. This conclusion is compatible with the results of recent eye-tracking studies reported in Altmann and Kamide (1999) and Kako and Trueswell (2000) that investigated the use of *verb* information during referential processing. Altmann and Kamide (1999), for example, used a task in which participants visually inspected a semi-realistic scene while hearing sentences such as "The boy will eat the cake." They found that participants began to fixate the single edible object pictured in the display (i.e., the cake) on hearing the main verb *eat*. In contrast, eye movements to this object were delayed when the main verb could be used in conjunction with a number of display objects (e.g., *touch*). Thus, verbs and prepositions appear to be highly similar in their ability to rapidly limit attention to semantically compatible referential candidates. Because verbs and spatial prepositions both assign thematic roles and are consequently analyzed as *predicates*, it is reasonable to conclude that this ability is a property of this particular semantic class.

One question that cannot be answered from the results of Experiment 1 or the studies of verb effects mentioned above is whether nonlinguistic information is also used on-line to structure the domain of interpretation. This issue is important not only for clarifying the nature of the underlying mechanisms but also because it has been argued that the precise meaning of relational terms such as spatial prepositions cannot be specified without additional nonlinguistic information (Coventry, Carmichael, & Garrod, 1994; Herskovits, 1986). For example, in certain circumstances, it would be possible to refer to the "insides" of the noncontainer objects used in Experiment 1 such as a rope and a rubber duck. Experiment 2 addresses whether inarguably nonlinguistic constraints are also used to dynamically update domains.

EXPERIMENT 2

In this experiment, we investigate the influence of a specific pragmatic constraint on domain construction, namely the compatibility of candidate referents with the event denoted by the utterance. The question is whether this compatibility is assessed as the comprehender processes the utterance, thereby restricting the domain of interpretation to only those candidates that are compatible with the event. To answer this question, we modified the visual displays used in Experiment 1 in two ways. First, the displays were changed to contain *two* exemplars of the goal container, which differed in size. For example, if the instruction was "Put the cube inside the can," then the corresponding display contained, among other objects, one large can and one small can. Second, the size of the theme object (i.e., the cube) was varied such that in one condition the object could fit inside both goal exemplars, whereas in a second condition it could fit inside only the larger one. An example display showing both the large and small versions of the theme object is provided in Fig. 4. We reasoned that an assessment of pragmatic compatibility ought to limit attention

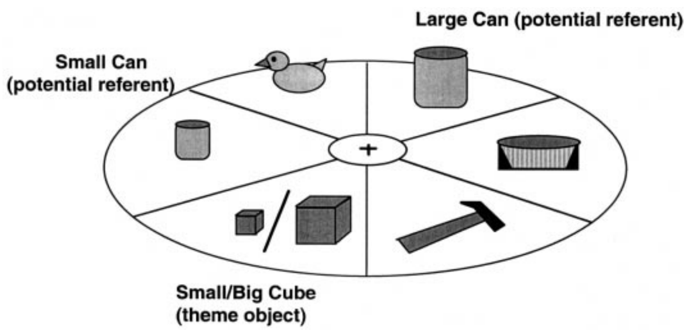

FIG. 4. Example of experimental display (Experiment 2).

to only compatible containers when the preposition *inside* was reached. If so, then the interpretation of the following definite noun phrase should be facilitated in the case where only a single goal exemplar can accommodate the theme object. This is because the smaller can will be excluded from consideration, thereby allowing the uniqueness requirement of the definite noun phrase to be met. If, on the other hand, these pragmatic considerations are not immediately available to constrain the referential domain, then the size manipulation should not produce any effect, at least during the early moments of comprehension.

Experiment 2 also addresses an important consideration regarding our previous interpretation of the eye movement data in Experiment 1. We assumed that the facilitation effect observed in the one-container condition with *inside* instructions reflects the use of preposition information to redefine the referential domain. However, an alternative explanation is that early eye movements to the target in this condition reflect a problem-solving strategy specific to the experimental task. On this account, participants are attempting to find a possible solution for the "Put the X inside . . ." command as quickly as possible, and eye movements reflect the shift of attention toward possible candidates. This interpretation still maintains that the data reflect a rapid integration of linguistic and nonlinguistic information (i.e., that preposition information can be used to direct attention to visually available containers). However, it does *not* require making the stronger assumption that, during normal processing, the preposition information establishes the semantically relevant domain in which noun phrase referents are evaluated.

The problem-solving explanation is addressed in Experiment 2 by manipulating the definiteness of the noun phrase following the preposition. In particular, the final noun phrase in the instructions was varied such that it was definite on half of the trials (e.g., "the can") and indefinite on the remaining half (e.g., "a can"). Importantly, definite and indefinite noun phrases differ in the uniqueness requirements that they place on their referents. As stated earlier, it is not felicitous to use a definite noun phrase when multiple alternatives meeting the description of the expression are present. Indefinite noun phrases, however, are routinely used for this purpose.[1]

We further clarify this manipulation in relation to the example display in Fig. 4. First, consider the interpretation of the instruction "Put the cube inside the can" when the theme object is the large version of the cube. If linguistic and pragmatic constraints are used in tandem to restrict referential domains, then the small can could be excluded from the domain of interpretation on hearing *inside*. This means that, when the following definite noun phrase is encountered, a unique referent (i.e., the large can)

[1]Throughout this study, we are concerned with the "specific" or "referential" interpretation of indefinite noun phrases in which "a Z" can be paraphrased as "one of the Z's." This is the most natural interpretation when referential alternatives have been contextually established, as in the case of our visual displays (Hawkins, 1991).

should be easily identified and the small can should receive minimal consideration. In contrast, when the theme object is the small version of the cube (i.e., the one that fits in both cans), both can exemplars will be included in the relevant contextual domain. In this case, the definite noun phrase “the can” will not have its uniqueness requirement satisfied, and the listener will have difficulty in determining which can was intended. The opposite pattern of results would be expected when indefinite versions of the instructions are used (e.g., “Put the cube inside *a can*”). Listeners should have no difficulty in interpreting the final noun phrase when the large version of the cube is used because the referential domain will be narrowed to only one can. However, when the cube can be put inside both cans, the indefinite noun phrase should be felicitous.

Thus, the linguistic domain hypothesis predicts an interaction between the number of compatible referents and the definiteness of the noun phrase. The predicted interaction occurs because the noun phrase is being initially interpreted within the circumscribed referential domain. In contrast, the problem-solving explanation predicts fast latencies whenever there is only one compatible exemplar. This prediction arises because there is only one possible action regardless of the definiteness of noun phrase.

Method

Participants. Participants were 16 native speakers of English drawn from the same population as in the previous experiment. None had participated in Experiment 1.

Materials. The table used in this experiment was similar to the one used in Experiment 1 except that the design on the surface consisted of a large circle (radius = approximately 17 cm) divided into six equal segments. A smaller circle in the center contained the fixation cross (radius = approximately 5 cm) (see Fig. 4). The circular display design was used to reduce the possibility that participants would expect the goal referent to be disambiguated by a postnominal phrase (e.g., “. . . the can *above/below/to the right of the bowl*”). A total of 12 critical displays were constructed. Each display contained six objects, one in each of the six partitions. Three of these objects were open containers, two of which were the potential goal referents. These two containers were identical except for their size (e.g., a large can vs a small can). The third container, the “unique competitor,” was a distinctly different type of container (e.g., a bowl) that was large enough to accommodate either version of the theme object. The competitor was included to evaluate the possibility that the definite article may be used to limit attention to a container that was unique in its respective category, irrespective of pragmatic plausibility. For example, on hearing *the*, reference to one of the bowls may be dispreferred because two exemplars of the category *bowl* are present. This hypothesis would predict that a significant proportion of early fixations to the competitor would be made in the definite noun phrase conditions. The presence of the competitor also reduced the likelihood that participants would expect the instruction to require them to make a decision between the large and small pair of containers.

The relative positions of the two potential referents and the competitor were counterbalanced across the 12 displays. In addition, the two potential referents were always separated by one partition in the display. The remaining three objects in the display were noncontainers. Two of these objects were not related to the instruction in any way (e.g., a duck and a hammer). The third noncontainer was the theme object (e.g., a cube) of the critical instructions described below. Two versions of each of the 12 critical displays were constructed, with each version differing in whether the theme object was small enough to fit inside both of the potential goal referents or only the large goal referent. The large and small versions of the theme object could always fit inside the competitor object.

The linguistic stimuli corresponding to the experimental displays consisted of a pair of instructions of the form “Pick up the X. Now put it inside the/a Y,” where X named a theme object and Y named the target object. On half of the experimental trials, the target noun was definite (e.g., “the can”), and on the remaining trials, it was indefinite (e.g., “a can”). The definiteness manipulation was crossed with the theme size manipulation, yielding four conditions. Four lists of trials were constructed, with each containing 12 critical trials. Three critical trials represented each of the four conditions in each list, and across all four lists, each version of the critical instructions together with each version of the critical displays occurred just once.

In addition to the experimental instructions, 48 pairs of filler instructions were constructed and added to each of the four lists. A total of 12 pairs of filler instructions followed each of the critical instructions and referred to objects in the corresponding experimental display. The remaining 36 pairs of filler instructions were associated with 18 distinct filler displays, and 2 instruction pairs were used with each display. These filler trials were randomly interposed with the experimental trials. The prepositions used in the filler instructions were varied (*beside* or *inside*) so that, within a list, each preposition occurred equally often. In addition, the types of the final noun phrases used in the fillers were varied so that the number of instructions containing definite and indefinite noun phrases in this position was the same. The fillers also equated the number of instructions in a list referring to container goal objects versus noncontainer goal objects. Finally, displays on filler trials were similar to critical displays, consisting of a mix of containers and noncontainers. However, the relative numbers of containers versus noncontainers were varied; some displays had a single container, whereas others had three exemplars of a particular container type.

Procedure. The procedure for this experiment was identical to that for Experiment 1 with the exception that the entire array of objects was changed between trials.

Results

Figure 5 shows the cumulative proportions of fixations to display objects for the conditions with definite noun phrases, and Fig. 6 shows the results for the indefinite noun phrase conditions. In each figure, the upper panel shows the condition in which only one potential goal referent could contain the theme object, and the lower panel illustrates the condition in which both potential goal referents could contain it. As before, the vertical lines indicate speech landmarks in the critical region of the instruction. The zero point on the x axis corresponds to the onset of the target noun.

As in Experiment 1, we plotted the cumulative proportions of fixations to display objects within each condition. Mean proportions were calculated for 100-ms time intervals, measured relative to the onset of the noun. The critical comparison for the current hypothesis is the point at which the proportion of fixations made to the target referent diverges from fixations made to the alternative referent (i.e., the container of the same name that was not selected as the location for the theme object). By this measure, faster reference resolution will be reflected in a relatively earlier point of divergence. Unlike Experiment 1, the pairing of displays with the experimental conditions varied across the lists to which participants were assigned. For this reason, a list factor was included in the ANOVAs (Pollatsek & Well, 1995; Raaijmakers et al., 1999). The list factor did not enter into any reliable effects or interactions. As before, the proportion data were submitted to an arcsine transformation before analysis.

We begin with the results for the conditions with definite noun phrase instructions illustrated in Fig. 5. No reliable differences were detected in the 0- to 100-ms or 100- to 200-ms interval following the onset of the final noun. However, in the 200- to 300-ms interval, the analysis revealed that, in the one compatible referent condition, the proportion of fixations to the target was marginally greater than that to the alternative referent, $F(1, 12) = 4.43$, $p = .06$, $MSE = .05$. This contrast was fully reliable in the 300- to 400-ms interval, $F(1, 12) = 7.64$, $p < .05$, $MSE = .13$. In contrast, fixations to the target were not reliably greater than those to the alternative in the two compatible referent condition until the 400- to 500-ms interval was reached, $F(1, 12) = 8.51$, $p < .05$, $MSE = .24$.

As with the definite conditions, the analyses did not reveal any significant differences in the 0- to 100-ms or 100- to 200-ms intervals in conditions with indefinite noun phrase instructions (shown in Fig. 6). However, in the 200- to 300-ms interval, fixations to the target were greater

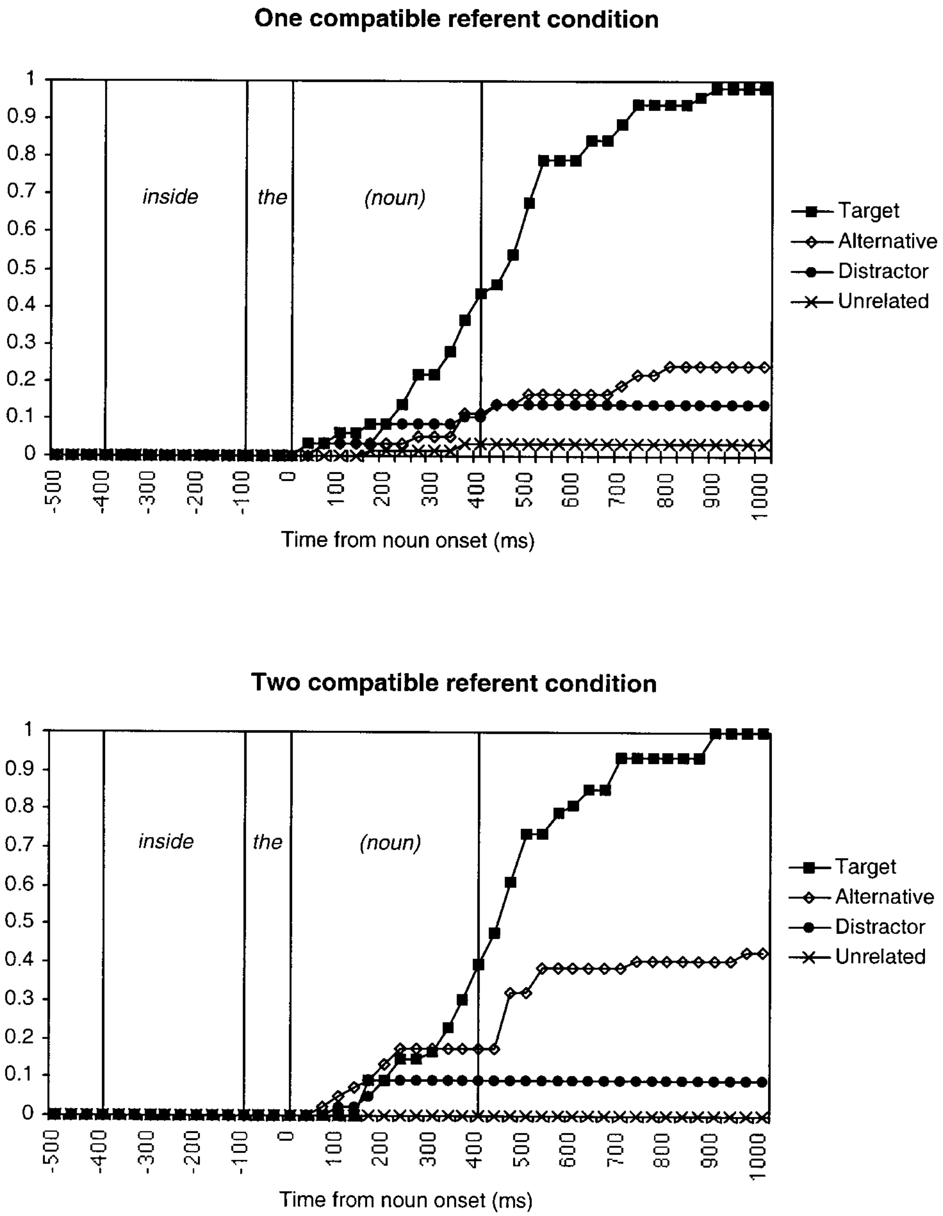

FIG. 5. Cumulative proportions of fixations to display objects, definite noun phrase conditions (Experiment 2).

than those to the alternative in the two compatible referent condition, $F(1, 12) = 5.73$, $p < .05$, $MSE = .04$. This difference did not reach significance in the one compatible referent condition until the 500- to 600-ms interval after the onset of the article, $F(1, 12) = 7.18$. $p < .05$, $MSE = .30$.

Inspection of Fig. 5 reveals that the unique competitor object (e.g., the bowl in Fig. 4) did not attract substantial fixations in advance of fixations to the target or alternative referent. As mentioned above, early looks to this object would have suggested a bias to link the unfolding definite noun phrase with object that was unique in its conceptual category. In fact,

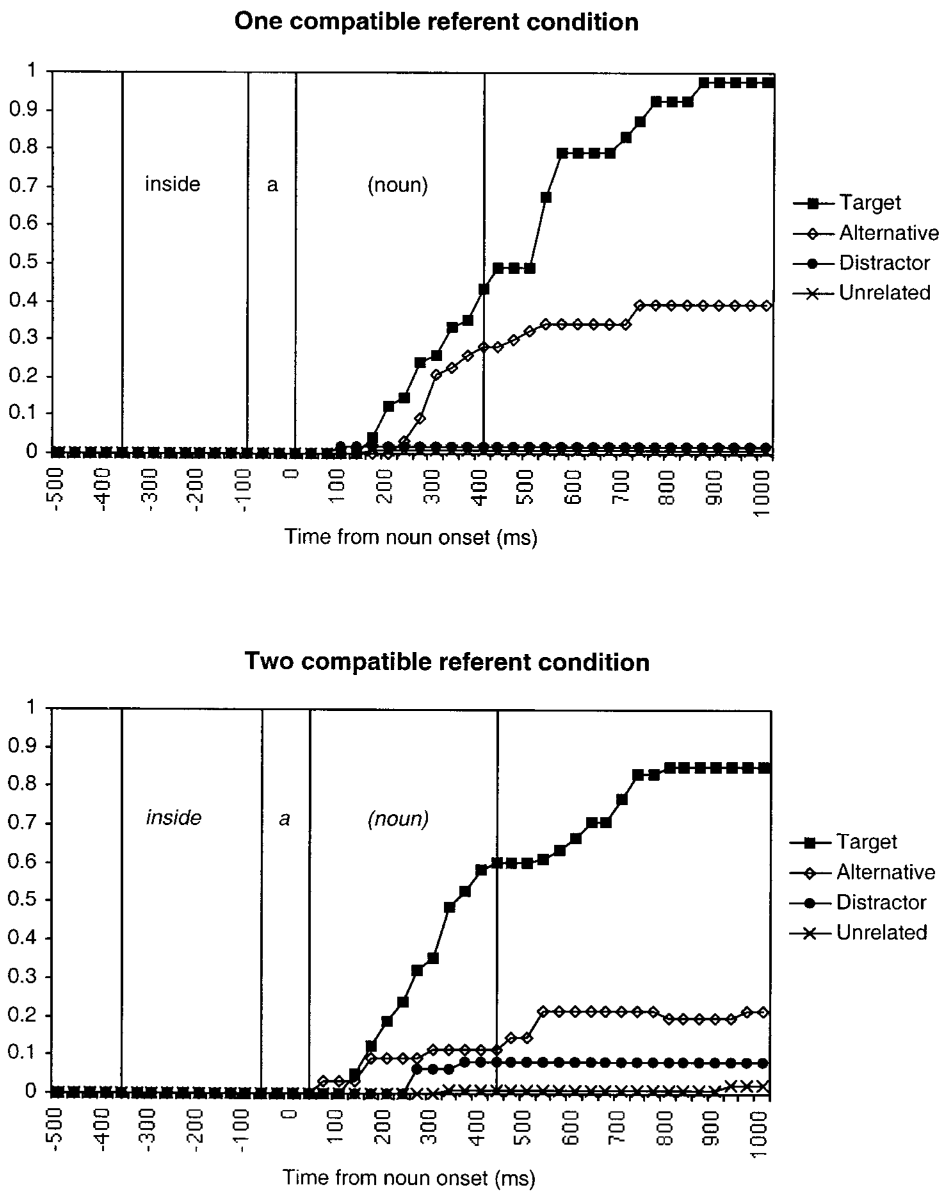

FIG. 6. Cumulative proportions of fixations to display objects, indefinite noun phrase conditions (Experiment 2).

the lack of an effect in this regard is not surprising given that modifiers such as adjectives (e.g., "the *big/small/red/other* can") are frequently used to distinguish among members of the same category. Consequently, a mechanism that directs attention to category-unique referents simply on the basis of the word *the* would likely be counterproductive in many instances. It is conceivable, however, that this kind of effect would be difficult to detect with the current experimental design because the head noun information was heard very soon after the onset of the article (approximately 100 ms).

Discussion

The results for definite instructions demonstrate that considerations of possible actions are integrated with semantic–conceptual constraints on-line to circumscribe the domain of interpretation relevant to referential interpretation. When only one potential goal was compatible with the theme object, a referent for the expression was identified earlier than when both potential goals could accommodate the theme object. In addition, when only one potential goal was compatible, reference resolution occurred sooner when the noun phrase was definite rather than indefinite. However, indefinites led to relatively fast reference resolution when the display contained two compatible goal referents. This outcome is consistent with the general proposal that definite noun phrases require their referent to be uniquely identifiable, whereas referential indefinites are used when multiple alternatives are available.

It is important to note that the pattern of results obtained in the indefinite noun phrase conditions provides evidence against a problem-solving interpretation of the data from the definite noun phrase conditions in this and the previous experiment. If eye movements reflected a strategy whereby participants were simply attempting to identify plausible goal objects independently of the content and particular constraints of the noun phrases, then the data pattern for definite and indefinite noun phrases should be similar, with earlier fixations to the target whenever only one container was a possible goal for the action. However, the results demonstrated that indefinite noun phrases had the opposite pattern of definite noun phrases. Identification of a referent occurred sooner when both containers were possible candidates, consistent with the claim that a referential indefinite noun phrase is understood to refer to one of several contextually evoked alternatives.

In sum, the results demonstrate that both linguistic and nonlinguistic constraints are rapidly used to circumscribe referential domains. However, there are two possible accounts of how or when the two types of constraints are used in this process. According to the account described above, referential domains are updated continuously, with relevant constraints being rapidly used as soon as they are encountered. Thus, on hearing “Put the cube inside . . .”, pragmatic considerations, along with the lexical–semantic constraints of the preposition, have narrowed the domain to the set of containers that may accommodate the cube. When the command continues with the definite noun phrase “the can,” and when only one can in the display can accommodate the cube, reference is quickly and unambiguously resolved. On an alternative account, action-based inferences come into play only when a unique referent for a definite noun phrase cannot be established within the domain defined by the lexical–semantic information. For example, on hearing “Put the cube inside . . .”, the lexical–semantic constraints will have restricted the referential domain to container objects in general and not only those that will contain the cube. If the command continues with the definite noun phrase “the can,” then the failure to satisfy the uniqueness constraint signaled by the definite article will trigger an “accommodation” process (e.g., Lewis, 1979) in which additional information sources, such as the compatibility of objects, are used to select a domain in which a unique referent for the definite noun phrase can be identified. This type of two-stage filtering model is similar in spirit to two-stage models that have been proposed for syntactic ambiguity resolution (e.g., Frazier & Rayner, 1982), for anaphora resolution (Gordon & Scearce, 1995), and most recently for the use of common ground in comprehension (Keysar, Barr, & Horton, 1998).

If pragmatic and linguistic constraints are rapidly integrated to restrict the initial referential domain, then the time course of definite reference resolution in the one compatible referent condition used in this experiment should be comparable to a case in which the display contains only a single candidate meeting the description of the noun phrase. If, however, pragmatic constraints are applied only during a late-occurring accommodation phase, then reference resolution should be faster when only a single candidate referent is visually available. We did not include a one-referent condition as part of the factorial design. However, we did include some filler trials in which the display contained only one exemplar of the object denoted by the final noun phrase. The full set of objects on these trials included a single target container, a second container of a different type (i.e., the unique competitor), the theme object, and three noncontainers. The theme object could be accommodated in both the target container and the unique competitor. We conducted a post hoc evaluation of the accommodation hypothesis by comparing fixation data in a baseline condition taken from these filler trials to data taken from the definite noun phrase conditions reported above.

Unlike in the previous analyses, we cannot calculate the point of divergence between the target referent and the alternative candidate because no alternative candidate exists in the one-referent baseline condition. However, a useful index of the relative time required to establish reference in each condition is provided by measuring the mean eye movement latency to the target object. These values were calculated by measuring the time, in 33-ms increments (i.e., frame units in VCR playback), between the onset of a critical point in the instruction and the final eye movement to the target referent that preceded the reach toward it. Eye movement launches were operationalized as the point at which the crosshairs left the center square.

The mean eye movement latencies for the three conditions of interest are presented in Table 1. A one-way within-subjects ANOVA performed on these data revealed a significant omnibus effect of the number of (compatible) referents, $F(2, 24) = 7.37$, $p < .01$. Pairwise contrasts confirmed the faster resolution time for the one compatible referent condition compared to the two compatible referent condition, as revealed earlier in the analysis of the proportion data, $F(1, 12) = 6.62$, $p < .05$. In addition, eye movements to the target referent were faster in the baseline (one referent) condition compared to the two compatible referent condition, $F(1, 12) = 14.08$, $p < .05$. The means also reflected a 41-ms advantage for the baseline condition over the one-container condition. However, this difference was not significant in the analysis, $F(1, 12) = 1.39$, $p > .20$. Thus, the results are most consistent with the initial domain restriction hypothesis, although we cannot completely rule out the possibility of rapid accommodation after the definite article was encountered. In any event, the finding that the times to identify a perceptually defined unique referent versus a pragmatically defined unique referent are similar underscores the speed with which nonlinguistic factors can be integrated to constrain the domain of interpretation. This outcome is particularly striking given that a fairly fine-grained evaluation process was necessary to assess the physical compatibility of theme and goal objects in the experimental conditions where the display contained more than one goal candidate. For example, although the large exemplar of a can might be the only referent able to *fully* contain the large version of the cube, the small can could nonetheless *partially* contain it such as when a corner of the cube is angled into the mouth of the can. In contrast, the physical compatibility of the theme and goal objects in the baseline condition could be established almost trivially. For instance, the required action might consist of putting a pen inside an open can when no additional container objects other than the competitor were present in the display.

TABLE 1

Mean Latencies of Eye Movement to Target Object, Measured from Onset of Noun: Experiment 2

	Mean latency (milliseconds)
One referent: Baseline	412 (25)
Two referents: One compatible	453 (19)
Two referents: Both compatible	542 (22)

Note. Standard errors are in parentheses.

In summary, the results of this experiment refine the conclusions reached in Experiment 1 by clarifying how referential domains are constrained on-line by pragmatic considerations. Objects that are not compatible with the action evoked by the unfolding instruction are significantly less likely to be considered as candidates for subsequent reference, even when they are perceptually salient and are compatible with the lexical–semantic constraint provided by the spatial preposition. Moreover, reference resolution for a definite noun phrase is not appreciably more difficult when its uniqueness is evaluated within a pragmatically defined domain rather than a more simple domain defined by perceptual information and context-independent lexical–semantic constraints.

GENERAL DISCUSSION

We began this research by considering how the domains of interpretation for linguistic reference are constructed or updated during comprehension. We identified three possibilities: (a) that domains are only updated at the closure of a linguistic unit such as a sentence or proposition, (b) that domains are updated continuously using only linguistically encoded information, and (c) that domains are updated continuously using both linguistic and linguistically relevant pragmatic constraints. We evaluated these alternatives by examining the time course with which listeners resolved definite noun phrases following spatial prepositions.

Experiment 1 demonstrated that the lexical–semantic constraints of the preposition *inside* were immediately used to restrict the referential domain to objects with container-like properties, ruling out the first hypothesis. Experiment 2 revealed that pragmatic information plays an additional role in this process, thereby eliminating the second hypothesis. In particular, we found that a referent was difficult to establish when two containers meeting the description of a definite noun phrase were present in the display, consistent with the claim that referents for definite noun phrases must be uniquely identifiable. However, when one of these containers was too small to contain the object being picked up and moved by the listener, this container was not considered during the early moments of referential processing and a unique referent was more easily established. This outcome suggests that candidate referents are evaluated in terms of their relevance to the immediate task and that this information is used in tandem with linguistic information to incrementally define referential domains.

It is important to acknowledge that the results we have presented represent only a preliminary foray into the complex question of how listeners establish and update the contextual domains used to interpret language. Nonetheless, the results have potentially far-reaching implications for models of real-time comprehension. First, the finding that semantic–conceptual information is used to rapidly narrow the domain of interpretation adds to a growing body of evidence suggesting that semantic interpretation proceeds continuously and is not directly mediated by syntactic constituency (Altmann & Kamide, 1999; Kako & Trueswell, 2000; Sedivy, Tanenhaus, Chambers, & Carlson, 1999). For example, the evidence would appear to be incompatible with a conception of reference resolution as a type of specialized subroutine that is triggered when a noun phrase is encountered (e.g., Matthews & Chodorow, 1988). More generally, the rapid uptake and availability of this information is likely to have significant implications for comprehension processes other than referential interpretation including the resolution of ambiguous words or the identification of grammatical relations (for some discussion of these issues, see Altmann & Kamide, 1999; Androutsopoulos & Dale, 2000).

In addition to questions of time course, the findings also inform current perspectives on the types of information that are coordinated during comprehension. As stated earlier, our results indicate that comprehenders evaluate referential candidates on the basis of their compatibility with the event denoted by the unfolding utterance and that this information is available during the early moments of processing. It is useful to consider this outcome with regard to current constraint-based approaches to real-time language interpretation. These approaches have emphasized the possibility that semantic and pragmatic constraints are reflected in the distributions of particular linguistic items and constructions in natural language data and that this distributional information is reflected in the mental representation of these expressions (Burgess & Lund, 1997; Landauer & Dumais, 1997). If so, then it is in turn possible that comprehenders exploit distributional regularities to effectively bypass the need to construct computationally expensive context-specific representations during processing (MacDonald, 1999). However, our findings cannot be straightforwardly captured in a processing model that does not incorporate situation-specific information from the immediate contextual environment or even one that uses this information only at a relatively late point in processing. Rather, the evidence is most compatible with a mechanism that has consistent and ready access to detailed information in the immediate contextual environment and that rapidly exploits this information to formulate hypotheses about the mapping among linguistic expressions, possible actions, and referential entities.

The claim that knowledge of actions plays a central role in language interpretation is not unique to the current study and, in fact, is a central theme in recent work by Glenberg and associates (e.g., Glenberg, 1997; Glenberg & Robertson, 1999, 2000; Kaschak & Glenberg, 2000). Most broadly, this work seeks to establish that a theory of (linguistic) meaning grounded in action is superior to one based on abstracted mental representations such as the representations derived from co-occurrences of words over the course of experience (i.e., as suggested by Burgess & Lund, 1997, and Landauer & Dumais, 1997). In relating this perspective to theories of language processing, Glenberg and Robertson (1999, 2000) proposed that successful comprehension involves three steps: (1) identifying the referents for individual expressions, (2) establishing the action-relevant properties (i.e., affordances) of these referents, and (3) coordinating the affordances of the referents to arrive at a coherent set of actions in accordance with the syntax of the sentence. We would clearly agree that these are important considerations in a theory of language processing and that they provide useful starting points for investigating language within a situated and embodied approach to cognition (cf. Clark, 1999). Our own work, however, is distinguished by the long-term goal of developing explicit models of the mechanisms that support the interpretation of language as it unfolds in time. Although it is not entirely clear how Glenberg and Robertson would expect their theory to be instantiated in models of real-time comprehension, one relevant observation based on our current findings is that the "steps" they identified should *not* be understood as fully dissociable stages of processing. In particular, the process of coordinating affordances with event information (i.e., Step 3) appears to constrain the process of referent identification (i.e., Step 1). In addition, although we would agree that information drawn from the situation-specific environment is a critical factor that is often overlooked or undervalued in contemporary models of linguistic competence and performance, we believe that it is equally important to acknowledge the role played by more static kinds of contextually derived knowledge. It would seem that even the process of associating actions with particular context-specific objects relies, to some extent, on relatively stable mental representations. For example, an individual's aggregate experience with particular objects occurring in particular events is likely to underlie his or her ability to discern "possible actions" or "affordances." Although these kinds of mental constructs are unlikely to be drawn from the distributional regularities of linguistic forms, they illustrate the need for mechanisms that can extract certain kinds of generalizations from experience and the role played by this type of contextual information in language interpretation. Further research is clearly needed to precisely determine how this information combines with the more ephemeral yet extremely salient information drawn from the immediate situation as comprehension proceeds.

A final question that we believe will be crucial to address in future studies is the extent to which the *communicative intention* underlying an utterance influences the way in which knowledge of actions is used for language comprehension. This concern stems from a consideration of the presuppositions that accompany different kinds of speech acts. In the experiments reported above, the linguistic materials consisted of *instructions* to manipulate physical objects in the immediate environment. Typically, a listener presented with these instructions would presuppose that the evoked action is capable of being performed and that the objects required to execute the action are present (Austin, 1962; Searle, 1969). Given these assumptions, the planning of the physical action can begin early on and attention can be rapidly directed to objects possessing the physical characteristics appropriate for this action. However, it is not always appropriate to assume that the denoted action is possible or that the available entities possess properties that will allow the action to be completed. Consider, for example, an interrogative version of our example experimental sentence such as "Is it possible to put the cube inside the can?" When presented with this utterance, the listener will understand his or her task to be that of assessing the possibility of performing the evoked action and then producing an appropriate response. In some cases, it may be that the can in question is, in fact, too small to accommodate the cube, requiring the listener to produce a negative response. Given that this possibility exists, there is little reason for the listener to presuppose that the action of placing the cube in the can may be performed. In fact, if this were already known, then the speaker would have no basis for asking the question. Because the listener's task is to assess the possibility for the action to be performed, it is less likely that the domain of interpretation will be initially restricted to only those containers large enough to accommodate the cube. Consequently, the final definite noun phrase in the interrogative form of the utterance is more likely to be perceived as infelicitous, even when there is only one can compatible with the cube.[2] However, this prediction rests on a key assumption, namely that the underlying communicative intention is recognized incrementally as the utterance unfolds in time. We must leave it to future research to specify the precise nature of this process and its influence on real-time referential interpretation.

[2]We conducted a preliminary test of this hypothesis with interrogative versions of the stimuli used in Experiment 2 (e.g., "Can you put the cube inside the/a can?"). The pattern of verbal response latencies did in fact suggest that, with interrogatives, information about the size of the cube was not initially used to limit attention to the compatible can. However, a direct comparison of this experiment with the results of Experiment 2 was thwarted by the fact that participants were often able to respond to the question before making any eye movements. We are currently designing studies using manipulations that will allow us to make the appropriate comparisons.

REFERENCES

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition: Evidence for continuous mapping models. *Journal of Memory and Language*, **38**, 419–439.

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*. **73**, 247–264.

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

Androutsopoulos, I., & Dale, R. (2000). Selectional restrictions in HPSG. In *Proceedings of the Eighteenth International Conference on Computational Linguistics* (pp. 698–703).

Ariel, M. (1998). The linguistic status of the "here and now." *Cognitive Linguistics*, **9**, 189–237.

Austin, J. L. (1962). *How to do things with words*. Oxford, UK: Oxford Univ. Press.

Ballmer, T. T. (1981). Context change and its consequences for a theory of natural language. In H. Parret, M. Sbisà, & J. Verschueren (Eds.), *Studies in language companion series*, Vol. 7: *Possibilities and limitations of pragmatics* (pp. 17–55). Amsterdam: John Benjamins.

Bar-Hillel, Y. (1954). Indexical expressions. *Mind*, **63**, 359–379.

Barwise, J., & Perry, J. (1983). *Situations and attitudes*. Cambridge, MA: MIT Press.

Burgess, C., & Lund, K. (1997). Modelling parsing constraints with high-dimensional space. *Language and Cognitive Processes*, **12**, 177–210.

Clark, A. (1999). An embodied cognitive science? *Trends in Cognitive Sciences*, **3**, 345–351.

Clark, H. H. (1996). *Using language*. Cambridge, UK: Cambridge Univ. Press.

Clark, H. H., Schreuder, R., & Buttrick, S. (1983). Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior*, **22**, 1–39.

Coventry, K. R., Carmichael, R., & Garrod, S. C. (1994). Spatial prepositions, object-specific function, and task requirements. *Journal of Semantics*, **11**, 289–309.

Crain, S., & Steedman, M. J. (1985). On not being led up the garden path: The use of context by the psychological parser. In D. Dowty, L. Karttunen, & A. Zwicky (Eds.), *Natural language parsing: Psychological, computational, and theoretical perspectives* (pp. 320–358). Cambridge, UK: Cambridge Univ. Press.

Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001). The time course of frequency effects in spoken word recognition: Evidence from eye movements. *Cognitive Psychology*, **42**, 317–367.

Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye movements as a window into real-time spoken language processing in natural contexts. *Journal of Psycholinguistic Research*, **24**, 409–436.

Fischer, B. (1992). Saccadic reaction time: Implications for reading, dyslexia, and visual cognition. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 31–45). New York: Springer-Verlag.

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.

Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing* (pp. 67–82). Hillsdale, NJ: Erlbaum.

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, **20**, 1–19.

Glenberg, A. M., & Robertson, D. A. (1999). Indexical understanding of instructions. *Discourse Processes*, **28**, 1–26.

Glenberg, A. M., & Robertson, D. A. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory and Language*, **43**, 379–401.

Gordon, P. C., & Scearce, K. A. (1995). Pronominalization and discourse coherence, discourse structure, and pronoun interpretation. *Memory & Cognition*, **23**, 313–323.

Groenendijk, J., & Stokhof, M. (1991). Dynamic predicate logic. *Linguistics & Philosophy*, **14**, 39–100.

Grosz, B. J., Joshi, A. K., & Weinstein, S. (1995). Centering: a framework for modeling the local coherence of discourse. *Computational Linguistics*, **21**, 203–225.

Hawkins, J. A. (1991). On (in)definite articles. *Journal of Linguistics*, **27**, 405–442.

Herskovits, A. (1986). *Language and spatial cognition: An interdisciplinary study of the prepositions in English*. Cambridge, UK: Cambridge Univ. Press.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, MA: Harvard Univ. Press.

Kako, E., & Trueswell, J. C. (2000). Verb meanings, object affordances, and the incremental restriction of reference. In *Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society*. [Online]. Available: ircs.upenn.edu/cogsci2000/prcdngs/sprcdngs/papers/kak-tru.pdf

Kaschak, M. P., & Glenberg, A. M. (2000). Constructing meaning: The role of affordances and grammatical constructions in sentence comprehension. *Journal of Memory and Language*, **43**, 508–529.

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.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, **104**, 211–240.

Lewis, D. (1979). Score-keeping in a language game. In R. Bauerle, U. Egli, & A. von Stechow (Eds.), *Semantics from a different point of view* (pp. 172–187). Berlin: Springer-Verlag.

MacDonald, M. C. (1999). Distributional information in language comprehension, production, and acquisition: Three puzzles and a moral. In B. MacWhinney et al. (Eds.), *The emergence of language* (pp. 177–196). Mahwah, NJ: Erlbaum.

Matthews, A., & Chodorow, M. (1988). Pronoun resolution in two-clause sentences: Effects of ambiguity, antecedent location, and depth of embedding. *Journal of Memory and Language*, **27**, 245–260.

McCawley, J. D. (1979). Presupposition and discourse structure. In C.-K. Oh & D. Dinneen (Eds.), *Syntax and semantics 11: Presupposition* (pp. 371–388). New York: Academic Press.

Morrow, D. G., Bower, G. H., & Greenspan, S. L. (1989). Updating situation models during narrative comprehension. *Journal of Memory and Language*, **28**, 292–312.

Partee, B. H. (1989). Binding implicit variables in quantified contexts. *Papers of the Chicago Linguistics Society*, **25**, 342–365.

Poesio, M. (1995). A model of conversation processing based on micro conversational events. In J. D. Moore & J. F. Lehman (Eds.), *Proceedings of the Seventeenth Annual Meeting of the Cognitive Science Society* (pp. 698–703). Mahwah, NJ: Erlbaum.

Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**, 785–794.

Raaijmakers, J. G. W., Schrijnemakers, J. M. C., & Gremmen, F. (1999). How to deal with "the language-as-fixed-effect fallacy": Common misconceptions and alternative solutions. *Journal of Memory and Language*, **41**, 416–426.

Roberts, C. (1995). Domain restriction in dynamic semantics. In E. Bach, E. Jelinek, A. Kratzer, & B. H. Partee (Eds.), *Quantification in natural languages* (pp. 661–700). Dordrecht, Netherlands: Kluwer.

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

Saslow, M. G. (1967). Latency for saccadic eye movement. *Journal of the Optical Society of America*, **57**, 1030–1033.

Searle, J. R. (1969). *Speech acts*. Cambridge, UK: Cambridge Univ. Press.

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

Spivey, M. J., Tanenhaus, M. K., Eberhard, K. M., & Sedivy, J. C. (in press). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology*.

Stalnaker, R. C. (1978). Assertion. In P. Cole (Ed.), *Syntax and semantics* (Vol. 9, pp. 315–332). New York: Academic Press.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information during spoken language comprehension. *Science*, **268**, 1632–1634.

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

Viviani, P. (1990). Eye movements in visual search: Cognitive, perceptual, and motor control aspects. In E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes* (pp. 353–383). New York: Elsevier.

(Received October 12, 1999)
(Revision received June 6, 2001)