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 sen-

0749-596X/02 $35.00

mains 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

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.

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 la-

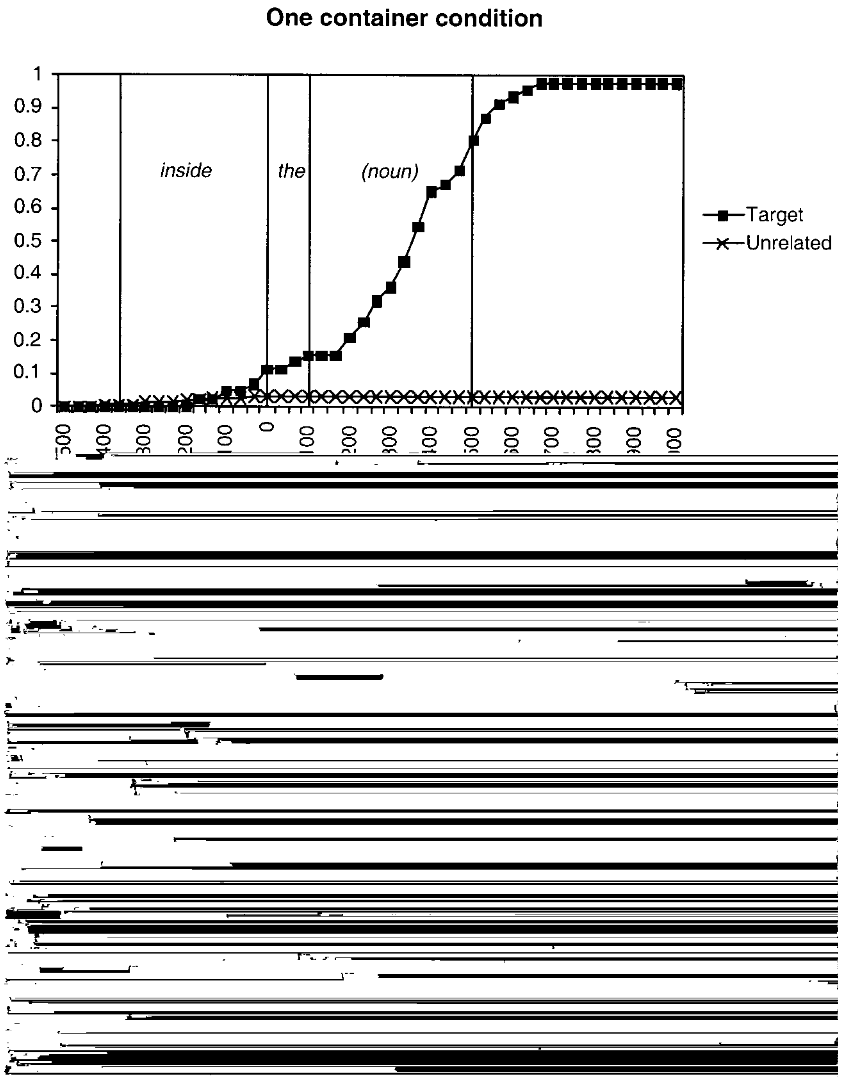

FIG. 3. Cumulative proportions of fixations to display objects, *inside* conditions (Experiment 1).

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 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 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

thof taw 11.448cle

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

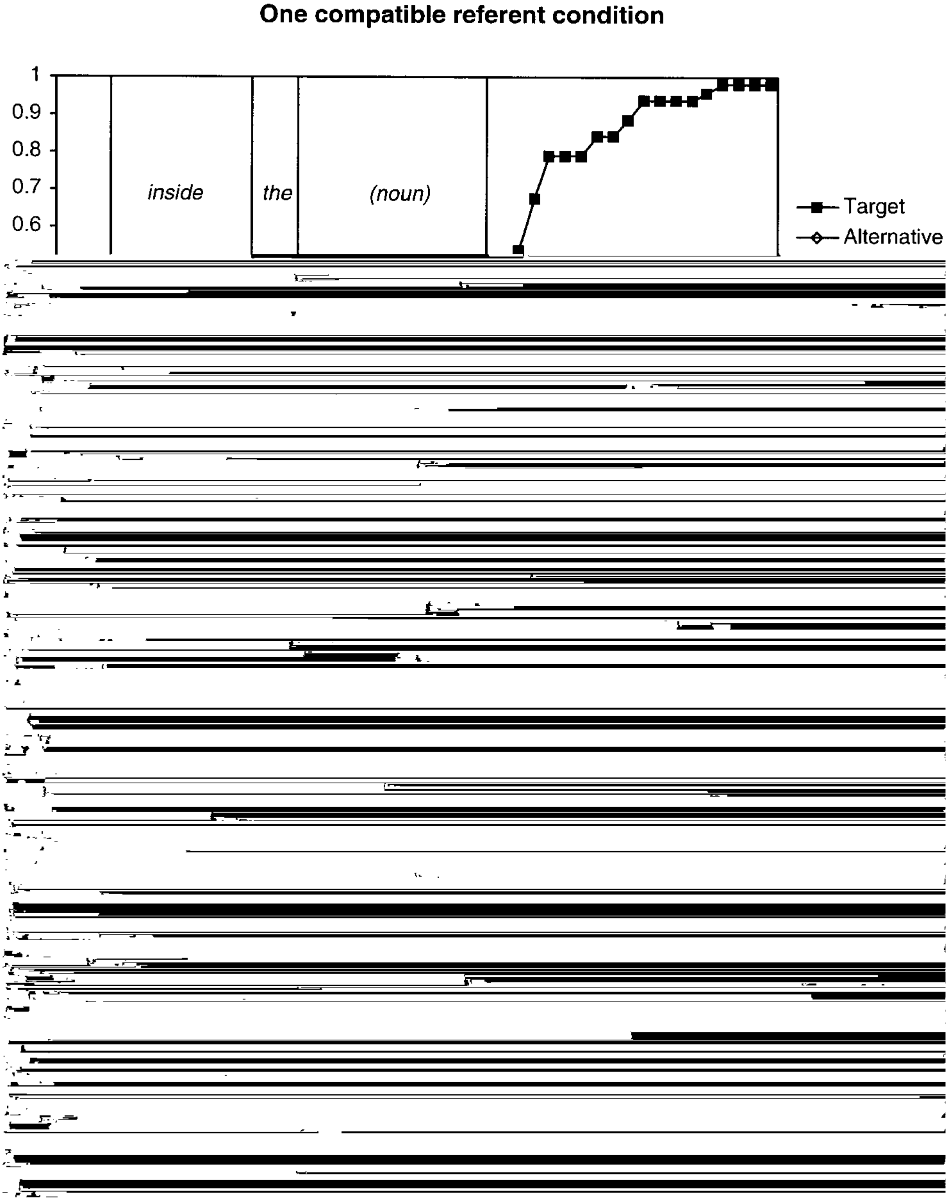

FIG. 5. Cumulative proportions of fixations to display objects, definite noun phrase conditions (Experiment 2).

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, the lack of an effect in this regard is not surprising given that modifiers such as adjectives

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.

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* effectively bypass the need to construct computationally expensive context-specific representapose 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.

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.

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-dimen-

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)