UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Words and Worlds: The Construction of Context for Definite Reference

Permalink

Journal

Authors

Chambers, Craig G.
Tanenhaus, Michael K.
Eberhard, Kathleen M.
et al.

Publication Date

1998

Peer reviewed

Words and Worlds: The Construction of Context for Definite Reference

Craig G. Chambers (chambers@bcs.rochester.edu)
Michael K. Tanenhaus (mtan@bcs.rochester.edu)
Department of Brain & Cognitive Sciences: University of Rochester
Rochester, NY 14627 USA

Kathleen M. Eberhard (Kathleen.M.Eberhard.1@nd.edu)
Department of Psychology: University of Notre Dame
Notre Dame, IN 46556 USA

Greg N. Carlson (carlson@ling.rochester.edu)
Hana Filip (filip@ling.rochester.edu)
Department of Linguistics: University of Rochester
Rochester, NY 14627 USA

Abstract

Two eyetracking experiments were conducted to investigate how the domain of interpretation for referential expressions is constructed and coordinated during utterance processing. Of particular interest was how the uniqueness requirement for a definite noun phrase (e.g. 'the book') could be satisfied given a particular array of candidate referents. Experiment 1 demonstrated that the conceptual relation expressed by a preposition immediately limits attention to compatible referents and in turn facilitates definite reference to these objects. Experiment 2 showed that domains are further constrained by the judgment of which referents are compatible with an intended action, and that uniqueness can be established by such factors even when several objects compatible with the noun phrase are present in the perceptual field. The results demonstrate that domains of interpretation for referential expressions are constructed and updated dynamically as an utterance unfolds in time, taking both linguistic and nonlinguistic factors into account.

Introduction

Semantic investigations of reference have suggested that definite noun phrases carry uniqueness implications (Kadmon, 1990; Kamp & Reyle, 1993). For example, in responding to the instruction "pass me the book", a hearer must determine that only a single book is relevant, even though several books may be available for reference within the immediate situation or current discourse. The apparent lack of ambiguity in the use of definite reference reflects that conversational agents interpret these expressions with respect to some subset of the world, or domain of interpretation.

Although the process of subdividing the world of discourse is essential for effective reference, relatively little is known in detail about how interpretive domains are constructed or used in language comprehension, particularly in terms of on-line processing. The current study explores the question of how domains of interpretation are coordinated and used as an utterance unfolds. In particular, we are interested in evaluating how linguistic and nonlinguistic sources of information are coordinated in the construction of domains, and in evaluating the time course of this process. Answers to these questions are relevant for understanding the cognitive representations that mediate the link between language and the world and for evaluating linguistic theories of meaning and reference.

To explore these issues, we used an eyetracking methodology which makes it possible both to systematically manipulate properties of the visual context and to use spoken language stimuli. In this paradigm, participants wear a head-mounted eyetracker which allows unconstricted head and body movement. Eye movements are recorded as participants hear instructions to move objects in a visual display immediately in front of them. The record of eye movements is used to infer which entities in the display are considered as candidate referents over the course of interpreting the instruction. Previous studies using this paradigm have shown that the process of referential interpretation is highly incremental. In the case of definite noun phrases, an eye movement is made to a target referent as soon as it becomes unique within the given display. For example, given the instruction "Pick up the red block", eye movements to the block occur at the point of the adjective red, if the block is the only red object in the display (Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995). The focus of the current study is to evaluate the extent of the visual domain which is taken into consideration when identifying a referent for a definite noun phrase, and what factors control the scope of this domain.

Linguistic cues in domain construction

The goal of our first experiment was to evaluate how domains of interpretation are constrained on-line in response to semantic-conceptual cues in the unfolding speech. We examined the potential for a spatial preposition (e.g., 'inside') to limit the domain of interpretation for a subsequent definite noun phrase.
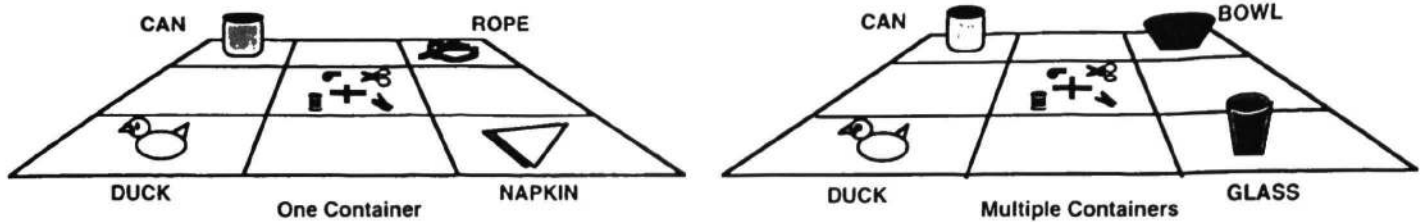
Figure 1: Example displays, experiment 1

The question of interest was whether a preposition immediately restricts the set of candidate referents for the following noun to only those objects which are compatible with the conceptual relation evoked by the preposition. In the case of the preposition 'inside', the set of compatible referents would consist of container-like objects.

The displays used are illustrated in Figure 1. In critical trials, participants were instructed to pick up an object in the display (e.g. a whistle) and put it inside another object in the display (e.g. a can). We varied the number of containers in the display such that half the displays in critical trials contained only one container, while the remaining half contained three containers. This manipulation was crossed with a second factor, namely the particular preposition used. Half the critical instructions used the preposition 'inside', and the remaining instructions used 'below'. The instructions using 'below' were control conditions, since this preposition should not restrict attention to containers. (Participants were told beforehand that 'below' corresponded to the grid square in front of the particular location).

If the relation evoked by the preposition 'inside' restricts the domain of interpretation to containers only, then facilitation in identifying the referent, as reflected in faster eye movement latencies, should be evident when displays contained only one container, relative to when multiple containers are present and to the control conditions where the preposition used was 'below'.

Results

The graphs in Figure 2 show the proportion of fixations for the various objects in the display at 33 millisecond time-slices during the critical region of the instruction. The figure contains the results for the two conditions in which the preposition used was 'inside'. Results from the condition where multiple containers were present reveals that identification of the target did not occur until after the onset of the head noun. The results for the control conditions using the preposition 'below' (not shown) mirrored this result. However, when only one container is present, eye movements to the target object began toward the offset of the preposition. The results for all conditions are represented in Figure 3 as mean eye movement latencies to the target container.

An analysis of variance performed on the latency data revealed a significant effect of preposition type when only one container was present ($F(1,11)=68.78, p<.001$). Instructions with 'inside' lead to faster latencies than those with 'below'. There was also an effect of number of containers when the preposition was 'inside' ($F(1,11)=64.58, p<.001$). Latencies to the target when multiple containers were present were longer than those when only one container was in the display. In sum, the results demonstrate that the semantic-conceptual information evoked by the preposition immediately constrains the domain of interpretation to those referents compatible with the preposition. This result is consistent with findings from previous studies which demonstrated that referential interpretation reflects a rapid
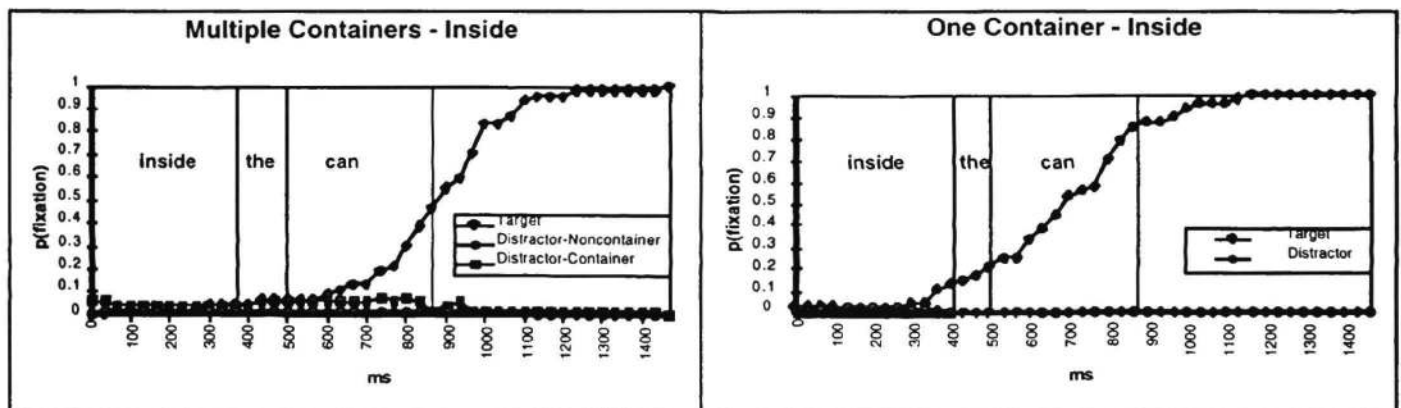

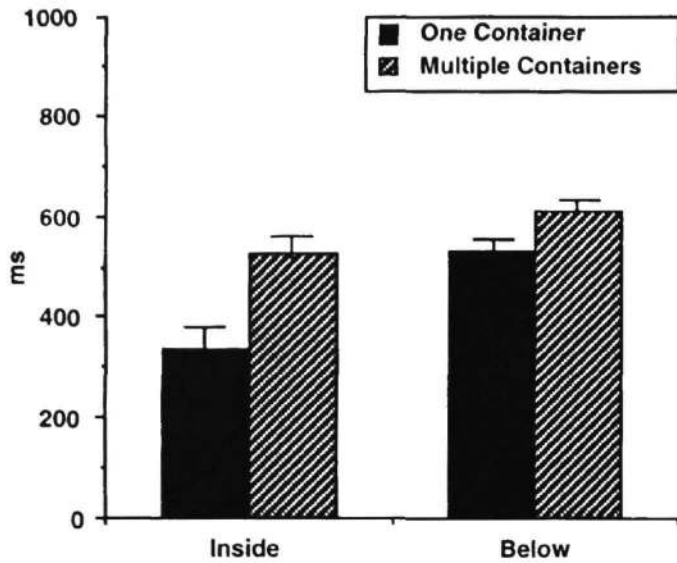**Figure 2**: Proportion of fixating objects over time, experiment 1.

Figure 3: eye movement latencies to target, Experiment 2

speeded eye movements do not reflect any process of domain construction, but rather that participants were engaging in a simple problem-solving strategy where they were trying to accommodate the instruction as quickly as possible. In other words, one might argue that the limiting information made available by the preposition is not used routinely during language comprehension, but is being consciously applied to perform the task in the present experiment. This issue is addressed in the second experiment, where we focus on the role which <u>nonlinguistic</u> information plays in limiting domains.

Pragmatic cues in domain construction

In a second experiment, we examined in greater detail the process of how uniqueness is established in definite reference. This was done by increasing the complexity of the experiment in two ways. One was to modify displays such that the uniqueness of the referent could only be established on the basis of a subset of the available referents. To achieve this, we used displays in which there were two instances of the referent of the noun in the critical noun phrase (e.g. two cans). These objects were two different sizes (see Figure 4). Instructions in this experiment were analogous to those in Experiment 1 (e.g., 'Pick up the cube. Now put it inside the can'). However, on purely perceptual grounds, the use of the definite article in these instructions violates uniqueness requirements, since there is more than one can in the display. Our second modification, which introduced the potential for uniqueness to be satisfied, was to vary the size of the moved object (e.g. cube) such that it could fit in either both cans (i.e. small cube), or only the large can (i.e. large cube). This allowed us to examine the use of a nonlinguistic constraint in the construction of domains, namely the judged compatibility of the moved object with potential locations. We hypothesized that the outcome of this compatibility assessment would exclude inappropriate locations from the domain of interpretation. The uniqueness conditions associated with definite expressions should reveal the success of this pragmatic cue in constraining the domain. For example, upon hearing "Pick up the cube. Now put it
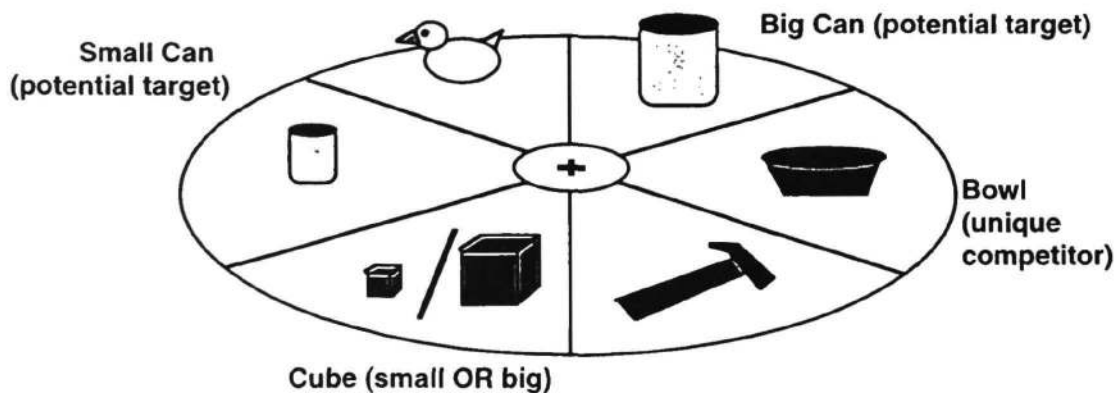
process of establishing correspondences between candidate referents and the unfolding speech.

This potential for a preposition to direct attention to certain referents is particularly interesting in light of the distinction traditionally made between referential properties of referring expressions (such as noun phrases) and predicates (such as verb phrases and prepositions). This distinction has been argued to be absolute, meaning that the potential to establish reference is a property of the class of referring expressions only (e.g. Hurford & Heasley, 1983). The present finding that prepositions may play an important function in constraining reference suggests that this distinction may be overstated.

One limitation in the current experiment is that there were no conditions present which actually <u>required</u> participants to interpret the uniqueness of the referent container with respect to only the set of containers rather than the entire array of objects in the display (including noncontainers). One possible alternative explanation of the results is therefore that

Figure 4: Example display, experiment 2

inside...", a comprehender may not simply limit attention to containers, but to only those containers which are compatible with the object in hand. If the small can is excluded since it is not compatible with the (big) cube, the domain of interpretation will contain only one can, thereby satisfying uniqueness for definite reference.

The object size manipulation was crossed with a second factor, namely the type of article used in the instructions. Half the instructions used the definite article 'the' to refer to the target container, and the other half used the indefinite article 'a'. This manipulation allowed us to evaluate whether any detected effect of domain restriction for definites is likely to be due to the use of problem-solving strategies to complete the task. If the process of restricting the domain to a compatible container results from simple problem-solving, results for definites should resemble those for indefinites. This is because the same number of compatible and incompatible objects is used in each condition. In other words, if participants are using the incompatibility of some of the containers as a means to strategically eliminate possible locations and accommodate the instruction as quickly as possible, then participants should be equally fast when a definite or indefinite is used, since the range of possible targets is the same in each case.

The displays also contained a third container (e.g. a bowl) which could contain both the large or small cube. This container was included to provide a 'genuinely' unique container referent in the display. Given previous evidence for the incremental nature of referential processing, this object may be likely to attract eye movements in the definite ('the can') conditions. For instance, upon hearing the partial instruction "put the cube inside the...", an eye movement may be made to this object since it is the only perceptually unique container in the display. A considerable number of looks to this object might be expected if semantic and pragmatic cues do not have the potential to constrain domains since reference to one of the actual targets (i.e. one of the cans) would be unlikely with a definite noun phrase.

Finally, to evaluate how uniqueness as satisfied within a pragmatically-circumscribed domain is comparable to when uniqueness is satisfied on purely perceptual grounds, the design contained control trials in which only one instance of the referent of the location noun phrase was present (e.g. 'put the pen inside the glass', where there was only one glass in the display).

Results

In collecting the eye movement data from the second experiment, the target was coded as the referent container which the participant selected to put the moved object in. As Figure 5 shows, the shortest eye movement latencies occurred in the definite article condition when the display contained only one compatible referent of the noun phrase. These latencies were faster than the analogous condition in which an indefinite was used ($F(1,4) = 8.11$, $p < .05$), and the definite condition where both referents were compatible ($F(1,4) = 13.24$, $p < .05$). These results suggest that the pragmatic compatibility assessment constrained the domain of

interpretation to only those containers which could accommodate the moved object, thereby allowing the uniqueness condition for the definite to be satisfied. The results also suggest that using an indefinite expression to identify a referent which is unique within a given domain violates expectations, as shown by increased latencies. This result shows that the domain-restriction effect found for definites is not due to problem-solving. If this were the case, participants should strategically eliminate the incompatible container as a candidate and should make comparable rapid eye movements to the target in both the definite and indefinite conditions.
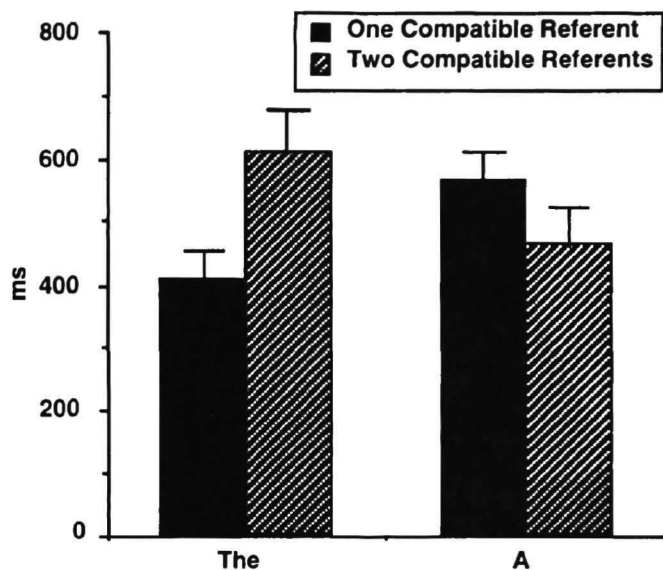

Figure 5: Eye movement latencies to target, Experiment 2

To assess how the satisfaction of uniqueness within pragmatically-circumscribed domains compares to the case where uniqueness can be satisfied on the basis of the perceptual domain alone, the definite conditions in Figure 5 were compared with the control condition in which a definite article was used with displays where only one instance of the referent container was present. As Figure 6 shows, latencies in the control condition are extremely similar to those in the condition in which two referents were present, but only one was compatible. In fact, no reliable difference was detected between these conditions ($F<1$). This similarity provides a clear illustration of how conceptual and pragmatic cues eliminate the ambiguity which would otherwise be expected when using a definite expression to refer to an entity which is not unique in the immediately relevant perceptual context.
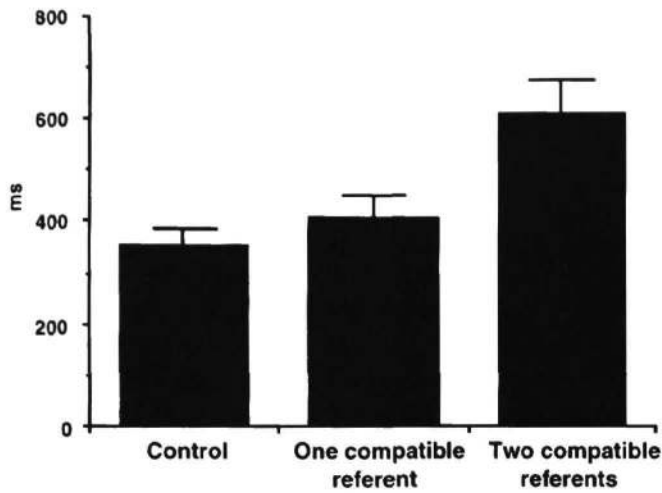
Figure 6: Eye movement latencies, definite conditions and control conditions, experiment 2

Figure 7 shows the proportion of trials containing a look to the other instance of the noun phrase referent. The results for the definite article conditions show that incidence of looks to the other instance of the named container was elevated in the condition where both referent containers were compatible with the moved object ($F(1,4) = 8.13$, $p<.05$), indicating indecision or confusion. In addition, when both referent containers were compatible, there were fewer looks to the other instance of the target referent in the indefinite condition than in the definite condition ($F(1,4)=7.68$, $p<.05$), suggesting again that the use of definite reference was incompatible with displays which did not allow the domain of interpretation to be limited to a unique referent.
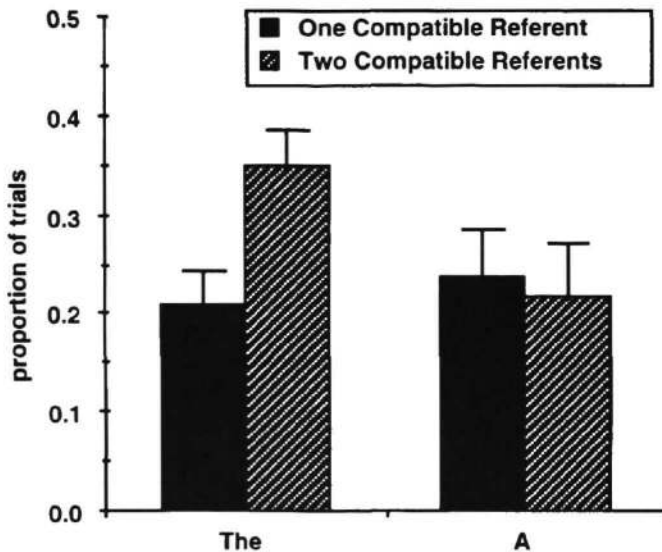
Figure 7: Proportion of trials containing a look to the alternative referent

Next we considered the proportion of trials containing a look to the competitor object in the displays. The competitor was the container in the display whose head noun

always picked out a unique referent (i.e. the bowl in Figure 4). We hypothesized that the incremental interpretation of referential expressions may cause looks to this item to increase in the definite condition relative to the indefinite condition. As Figure 8 shows, the incidence of looks to this object was greater in the definite conditions than in the indefinite conditions ($F(1,4)=6.99$, $p<.05$). However, the average incidence of looks to this object was fairly low, suggesting that the potential for conceptual and pragmatic factors to restrict reference does not permit the perceptually unique container to become particularly salient.
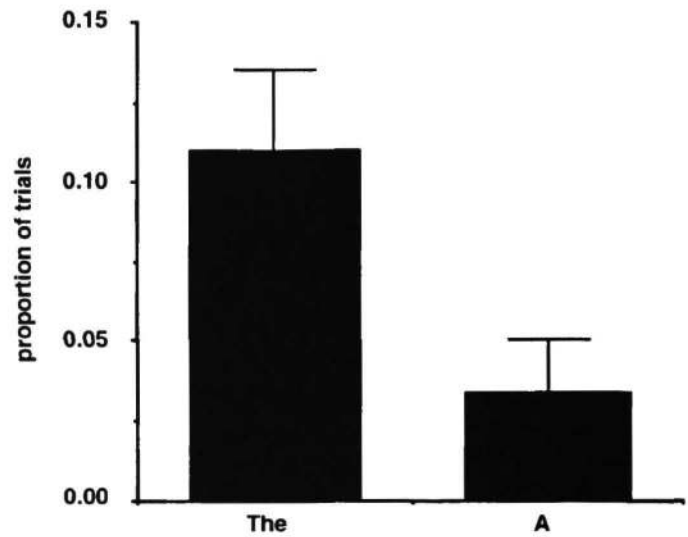
Figure 8: Eye movement latencies for definite conditions and control condition, Experiment 2

Collectively, the results of the second experiment highlight the way in which semantic and pragmatic cues operate in tandem to immediately restrict the domain of interpretation for a referring expression.

Discussion

The experimental findings suggest some interesting extensions to the issue of context and reference. The first experiment demonstrated that the process of establishing reference is facilitated by the presence of semantic-conceptual cues (in this case, prepositions) which constrain the set of candidate referents. The finding that these cues function to facilitate reference even when they are not logically or pragmatically necessary for interpreting a definite expression provides some insight into the question of how reference is routinely established with minimal effort and with no apparent delay.

The second experiment demonstrated that the semantic-conceptual cues examined in the first experiment are further constrained by pragmatic factors such as the evaluation of particular actions relevant to completing a task. The conclusion that task intentions are relevant in definite reference is consistent with some theoretical analyses (e.g.

Birner & Ward, 1994). However, the finding that this factor has an <u>immediate</u> effect in constraining the interpretation of a subsequent noun phrase is incompatible with a prevalent claim in psycholinguistics that the complexity of pragmatic inference leads to delayed processing. Rather, the results are compatible with the view that information from a wide variety of sources is used to dynamically generate and reevaluate the domain of interpretation of referential expressions.

The results of both experiments have implications for how the notions of 'context' and 'utterance' are applied in theoretical and experimental investigations of reference. Traditionally, context and utterance have been treated as largely distinct and separable constructs, where context is characterized in terms of factors <u>external</u> to the utterance. Such factors include perceptual salience (Clark, Schreuder & Buttrick, 1983), mutual gaze (Argyle & Cook, 1976), shared knowledge (Clark & Marshall, 1981), and prior discourse (Prince, 1992). The current results are inconsistent with a strict separation of context and utterance since they demonstrate conceptual information evoked by some linguistic unit (e.g. preposition) may be crucial in creating the contextual conditions which license a unique referent for a definite noun phrase in the same utterance.

While we have provided some specific evidence regarding the coordination of information and the time course of establishing domains, the current results only scratch the surface of problem. Future considerations include examining the resolution of reference in richer contextual environments, and how this process interacts with highly domain-sensitive linguistic elements such as prosodic focus and quantification operators.

Acknowledgments

References

Argyle, M., & Cook, M. (1976). Gaze and mutual gaze. Cambridge: Cambridge University Press.

Birner, B.J. & Ward, G. (1994). Uniqueness, familiarity and the definite article in English. *Proceedings of the Twentieth Annual Meeting of the Berkeley Linguistics Society* (pp. 93-102).

Clark, H.H., & Marshall, C.R. (1981). Definite reference and mutual knowledge. In A. Joshi, B. Webber, & I. Sag (Eds.), *Elements of discourse understanding*. Cambridge, UK: Cambridge University 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.

Hurford, J.R. & Heasley, B. (1983). Semantics: a coursebook. Cambridge, UK: Cambridge University Press.

Kadmon, N. (1990). Uniqueness. *Linguistics & Philosophy*, 13, 273-324.

Kamp, H., & Reyle, U. (1993). From discourse to logic. Dordrecht: Kluwer.

Prince, E.F. (1992). The ZPG letter: subjects, definiteness, and information-status. In S. Thompson & W. Mann (Eds.), *Discourse description: Diverse analyses of a fundraising text*. Amsterdam: John Benjamins.

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