

# Levels of representation in the interpretation of anaphoric reference and instrument inference

MARGERIE M. LUCAS

*Wellesley College, Wellesley, Massachusetts*

and

MICHAEL K. TANENHAUS and GREG N. CARLSON

*University of Rochester, Rochester, New York*

The level of representation accessed when inferences are made during sentence comprehension was examined. The inferences investigated included antecedent assignment for both definite noun phrase anaphors and pronouns and also instrument inferences. In making these inferences, a listener must access the inferred element, whether an antecedent or an instrument, in either a linguistic form representation or a discourse model. The level of representation involved in these inferences was determined by exploiting differences in the lexical decision and naming tasks, which were argued to exhibit differential sensitivity to representational levels. In three experiments, the priming of antecedent and instrument targets in the lexical decision task was compared with priming of the same targets in the naming task. Differences in the patterns of activation across the two tasks indicated that all three types of inferences required accessing elements in a discourse model. Three control experiments ruled out simple context or congruity checking as an explanation for our results. The following conclusions were also supported by these studies: (1) Antecedent assignment occurs immediately after processing an anaphor; (2) antecedent assignment involves inhibition for the inappropriate antecedent rather than facilitation for the appropriate antecedent; (3) although subjects do not make instrument inferences when they hear isolated sentences containing verbs that strongly imply certain instruments, the inferences are made when sentences are preceded by a context that mentions the instrument.

Understanding sentences in discourse involves integrating information in the sentence currently processed with information in a memory representation of the previous discourse. Sometimes the need for integration is explicitly marked, as when a sentence contains an anaphoric expression. Often it is less explicit (and necessary for maintaining coherence), as when actions mentioned in the sentence imply agents or instruments that were mentioned earlier in the discourse. A great deal of psycholinguistic research has been devoted to investigating how often, at what point in the sentence, and under what circumstances integrative processes occur. A basic question that has received relatively little attention, though, is that of the form of the representation involved in integrative processes. Although a common assumption is that integrating an element in the currently processed sentence requires changing the activation of an element in a memory representa-

tion of prior discourse, little experimental work has been done to investigate the form of the representation that is accessed (but cf. Cloitre & Bever, 1989; Murphy, 1985). In this article, we examine the types of representations used in accessing antecedents of definite noun phrases and pronouns as well as implied instruments. We also explore the conditions under which these inferences are made, and their time course.

The two most natural candidates for levels of representation in which prior information might be found are a representation of the surface form of a sentence and a more abstract discourse representation. One of the most robust results in the psycholinguistic literature is that memory for linguistic form is relatively transient, in contrast with memory for content or meaning (Bartlett, 1932; Johnson-Laird & Stevenson, 1970; Sachs, 1967; Wanner, 1974). Accordingly, most models of discourse processing distinguish between recovering the linguistic form of a sentence and developing a discourse model. The linguistic form representation includes the results of processing the words and the syntactic structures relating them, and it allows information about lexical form and syntax to be recalled for a short time. The discourse model includes the set of entities and relationships among them that are evoked by the discourse.<sup>1</sup> Discourse models preserve little or no information about linguistic form.

The linguistic literature on anaphora suggests that both of these types of representations might be used in dis-

---

The research presented here was supported by National Science Foundation Grants BNS-8217378 and BNS-8617738 to Michael Tanenhaus and Greg Carlson and by a Wellesley College Faculty Award to Margery Lucas. We are grateful to Robert Peterson and several anonymous reviewers for their many helpful comments on an earlier version of this paper. We also thank Cindy Nichols and Marlon Housman for their help in conducting the research and Sharon Utakis for her help in analyzing the data. Special thanks to Robyn Vallejos for her yeoman efforts in running the control experiments. Send reprint requests to Margery M. Lucas, Wellesley College, Department of Psychology, Wellesley, MA 02181.

course processing. Hankamer and Sag (1976) present evidence that anaphors can be divided into two classes, *deep* and *surface* anaphors. Surface anaphors, which include many types of verb phrase anaphora, are sensitive to the linguistic form used to introduce the referent of the anaphor. Thus there are syntactic constraints on the form of the antecedent for a surface anaphor and the antecedent must be introduced linguistically (for details see Garnham, 1986; Hankamer & Sag, 1976; Sag & Hankamer, 1984; and Tanenhaus & Carlson, in press). In contrast, deep anaphors, which include definite pronouns and definite noun phrase anaphora, are not similarly sensitive to the linguistic form of their antecedents. Deep anaphors can be used deictically to refer to salient objects or events in the environment. Deep anaphors also do not place constraints on the form of the antecedent when the antecedent is introduced linguistically. These differences between surface and deep anaphors can be explained by assuming that interpreting a surface anaphor requires the listener or reader to consult the linguistic form of the antecedent of the anaphor, whereas interpreting a deep anaphor requires consulting only discourse representations (Sag & Hankamer, 1984).

The view that definite pronouns require access to a discourse model but not to a linguistic level of representation appears open to immediate empirical challenges, for there appear to be syntactic effects on antecedent assignment. One type of constraint is most definitely linguistic in nature, and only affects pronoun-antecedent relations in the same sentence. For instance, in a sentence like "She thinks that Mary should be invited," the pronoun "*she*" cannot be interpreted as referring to Mary, since the syntactic relationship between the noun phrases "*she*" and "*Mary*" precludes making this assignment. However, the most common linguistic view is that although the syntactic relationship serves to rule out certain potential antecedents, the actual selection of the antecedent must be accomplished at a nonlinguistic level (Lasnik, 1976; but also see Reinhart, 1983, and Kuno, 1987, who argue that these constraints may be pragmatic rather than syntactic).

Another type of constraint that might appear to require the accessing of the form of the antecedent is that pronouns agree with their antecedents in gender and number. However, syntactic number typically corresponds to conceptual or notional number. In cases in which there is a mismatch between syntactic and notional number, agreement is typically with notional number. Gernsbacher (1986) contrasted three discourse situations in which the true antecedent was a multiple, collective, or generic item and found that in those situations subjects rated sentences with a mismatched plural pronoun more natural than those with a matched, singular pronoun. An example is the short discourse, "I think I'll order a margarita. I just love them," where the true antecedent is not the literal antecedent, which is singular, but a multiple item—margaritas in general.

A number of studies also suggest that memory for syntactic form is not necessary for interpreting pronouns and definite noun phrase anaphors. Memory for syntactic form

decays rapidly (Sachs, 1967; Wanner, 1974). If syntactic form were necessary for interpreting an anaphor, then comprehension should become increasingly difficult as the distance between the antecedent and the anaphor increases. However, discourse factors such as focus and referential continuity affect the ease of interpreting these anaphors more than distance (Anderson, Garrod, & Sanford, 1983; Ehrlich & Johnson-Laird, 1982; Lesgold, Roth, & Curtis, 1979). In fact, distance effects are eliminated when focus is controlled for both definite noun phrase anaphors (Lesgold et al., 1979) and definite pronouns (Anderson et al., 1983).

Although this evidence suggests that deep anaphors access a discourse model, in most studies processing-load measures such as reading time have been used, which are global measures of ease of comprehension and are not sensitive to the level of representation. An on-line measure of antecedent assignment would provide more precise and accurate information, especially if it could also distinguish among the levels of representation.

Two on-line processing measures reported in the word recognition literature as having been used successfully are lexical decision and naming. At one time, it was generally assumed that these two tasks provided similar information about word recognition (Meyer & Schvaneveldt, 1976). For example, in both tasks there are faster responses to words preceded by semantically or associatively related words than to words preceded by unrelated words, indicating that both tasks are sensitive to the activation of items in semantic memory. But recently, a number of interesting differences between the tasks have been discovered. Lexical decisions, but not naming responses, are facilitated by syntactic relationships between words, backward associations (i.e., target-to-prime associations), and changes in proportion of related to unrelated items (these differences are described in more detail in Seidenberg, Waters, Sanders, & Langer, 1984). Lexical decisions are also affected by a number of factors related to the meaning of the word, such as the likelihood that the target will be given as a response to a category name and the number of dictionary meanings it has; in contrast, naming seems to be relatively insensitive to the same factors (Balota & Chumbley, 1984; Chumbley & Balota, 1984). Also, inhibition for contextually incongruent targets has been observed primarily with the lexical decision task, and not with naming (West & Stanovich, 1982; but cf. West & Stanovich, 1986, on inhibition in naming in syntactically inappropriate contexts).

A reasonable explanation for these differences appears to be that naming is primarily sensitive to lexical processes and lexical decision to both lexical and postlexical processes. This may be because lexical decisions require one to make a conscious decision as quickly and as accurately as possible. Forster (1976, 1979) suggested that under such circumstances, a simple strategy of matching the orthographic characteristics of the visually presented word with an internal representation may be time-consuming and inaccurate, causing the lexical decision targets to undergo postaccess checks for error. Forster

argues that an effective error-checking mechanism for a discriminatory task would involve a different format from the one involved in the initial match. In a typical lexical decision task, this means using message-level processes to integrate the target with its context. Successful integrations support *yes* decisions, failures support *no* decisions. Naming, in contrast to lexical decision, does not require a conscious *yes/no* decision. Instead, it requires the accessing of procedural knowledge directly tied to the lexical representation of the word. Knowing whether or not a word makes sense in context is unlikely to facilitate this process, which may explain why naming has been shown to be less sensitive to contextual integration effects. Although there is little disagreement that lexical decision is more sensitive to such effects than naming, the differences in the two tasks are not absolute. Keefe and Neely (in press) and Norris (1987), for example, have found proportionality effects with naming under certain circumstances. Nonetheless, the generalization that lexical decision is more sensitive than naming to postlexical processes does seem to hold across most studies that have been focused on differences.

Lexical decisions, then, seem to be the result of cognitive processes that cut across levels of representation, whereas naming seems primarily sensitive to representations of lexical form. This analysis suggests a possible way of distinguishing between the levels of representation associated with anaphoric processing, through the use of differences in patterns of responses in the lexical decision and naming tasks as a diagnostic test for determining representational level. The logic behind the use of the two tasks follows from two possibilities for the mechanism involved in antecedent assignment or any other integrative process: Integration requires changing the activation level of either the lexical form associated with the relevant discourse element in a linguistic level of representation or the activation level of the conceptual form of the relevant element in a discourse model. If the conceptual form is accessed, this will be reflected in lexical decisions; if the lexical form is accessed, this will be reflected in both naming time and lexical decision (given that pure associative or intralexical priming is eliminated as a possibility). Direct experimental evidence for the representational level or levels accessed in integrative processes would provide an important constraint on models of integration.

In the experiments presented here, the lexical decision and naming tasks were used to examine the processing of definite noun phrase anaphors, pronouns, and verbs with implicit instruments. In Experiments 1A-1C, definite noun phrase anaphora were examined; in Experiments 2A-3D, pronouns and instruments were examined.

### EXPERIMENT 1 Definite Noun Phrase Anaphors

In this set of experiments, we examined two hypotheses. The first involved the time course of antecedent assign-

ment. The assumption behind the use of priming tasks to study antecedent assignment is that once antecedents are selected, they will be more activated than other discourse entities, including entities that are potential antecedents. This may be because the actual antecedent will be more highly activated or because the potential antecedents will be somewhat inhibited. Previous research has demonstrated the validity of using priming tasks in the study of anaphoric processing, but these have mostly been probe recognition studies (Cloitre & Bever, 1989; Corbett & Chang, 1983; Dell, McKoon, & Ratcliff, 1983; Gernsbacher, 1989; McKoon & Ratcliff, 1980; Walker & Yekovich, 1987; but cf. Nicol & Swinney, 1989, for a review of a few studies using the cross-modal priming task). These have shown that the actual antecedent is activated more than other potential antecedents at some point after the anaphor (usually a pronoun) has been processed. However, there is no difference in activation among potential antecedents either before or immediately after the pronoun has been read (Gernsbacher, 1989). In the first two experiments here, two different priming tasks were used to measure the activation of antecedents of definite noun phrase anaphors. Subjects listened to sentence pairs like the following:

1. Sarah could not decide whether to buy steak or hamburger.  
She finally chose the more expensive meat.

Subjects then made either lexical decisions or naming responses to visual targets that were appropriate or inappropriate antecedents. On the basis of the probe recognition studies, we expected to find that both antecedents would be activated immediately after processing the anaphor, but that only the actual antecedent would be activated sometime later.

The second hypothesis involved representational issues. We expected, on linguistic grounds, that the definite noun phrase in the second sentence—"the more expensive meat"—would be linked to an entity in a discourse model rather than a lexical representation for "steak." As discussed earlier, deep anaphors, including definite noun phrases, are not sensitive to the linguistic form of the antecedent. Also, there need not be a linguistic expression to act as antecedent in order for a definite noun phrase to be understood. For example, if a speaker who is looking at two pieces of meat says, "I'd like the bigger piece," it will be apparent from following the direction of the speaker's gaze (and the size of the meat) which piece the speaker is referring to.

We would expect, then, to find an appropriateness effect with the lexical decision task, because lexical decision is sensitive to both discourse level and lexical form. But we would expect to see an effect in naming only if lexical form is involved. Thus, by comparing the two tasks, we should be able to establish the level or levels of representation involved in anaphoric processing.

