North East Linguistics Society

Volume 15
Issue 1 *NELS 15*

Article 25

1985

Processing Deep and Surface Anaphors

Michael K. Tanenhaus
University of Rochester

Greg N. Carlson
Wayne State University

Processing Deep and Surface Anaphors[1]

Michael K. Tanenhaus
University of Rochester

Greg N. Carlson
Wayne State University

In this paper we report several experiments designed to investigate the nature of the representation in which different types of anaphors find their antecedents. We are assuming a common view of language with at least three different levels of representation of immediate relevance. First, there is the syntactic/morphological form of a given linguistic expression. That form has associated with it a linguistic meaning which we will provisionally assume to be represented by some type of logical form (see e.g., Sag (1976). Finally the form of an expression by virtue of its meaning has a certain denotation in the world or some model of it. We assume forms and meanings to be parts of a specific linguistic system; the denotations, though, are extralinguistic. We make no claims about the psychological nature of linguistic meaning, but we do assume that people conceptually represent the denotations of linguistic expression in some general propositional way.

The assumption that has been guiding our work is that there is a one-to-one correspondence between levels (or types) of representation in a correct linguistic competence theory and the levels (or types) of representation that are used during language production and comprehension. It seems to us that this is the weakest claim that one can accept and still maintain that linguistic theory is theoretical cognitive psychology in the sense of Chomsky (1972).

The phenomenon of anaphora is of interest because in order to interpret an anaphor the reader or listener must associate the anaphor with its antecedent by linking or replacing it with the mental representation of its antecedent. If the properties of the 'reactivated' antecedent can be observed, we may be able to shed light on the representations being computed during comprehension. This is especially true given that there is linguistic evidence suggesting that different types of anaphors may find their antecedents in different levels of representation, with some anaphors being initially associated with linguistic representations and others with conceptual representations (Hankamer and Sag, 1976).

In the experiments described below, we will be attempting to distinguish between two alternative hypotheses about the form of the representation to which an anaphor is associated in processing. The first hypothesis is that an anaphor is associated with the linguistic representation of its antecedent (either its form or meaning); the second hypothesis is that an anaphor is associated with the conceptual representation (the denotation) of its antecedent.

We begin by exploring the processing of pronouns and definite noun phrases. A number of studies in the psycholinguistic literature have demonstrated that the antecedent of a definite pronoun or noun phrase becomes reactivated as a consequence of interpreting the anaphoric expression (Chang, 1980; Dell, McKoon and Ratcliff, 1983; Shillcock, 1982). For example, Chang (1980) had subjects read a sentence and then verify whether or not a probe word occurred in the sentence. Probe verification latencies were facilitated when the probe word was the antecedent of a pronoun which occurred later in the sentence.

While experiments such as these demonstrate that the antecedent of an anaphor is being reactivated as a result of associating an antecedent with an anaphor, they do not identify the level of representation of the antecedent. To illustrate this point consider a study by Tanenhaus and Senytka (reported in Tanenhaus, Carlson & Seidenberg, 1985). In this experiment, subjects were presented with pairs of sentences one word at a time at a rate of 400 msec per word. Following each sentence subjects made a lexical decision to a target word. The sentences were constructed so that there were two possible antecedents of a pronoun

until the last word of the sentence which served to disambiguate the interpretation of the pronoun. The target was either the appropriate or inappropriate antecedent. The sentences in (1) illustrate the materials.

(1) a. The fire raged through the forest.
b. It was destructive.
c. It was destroyed.

The context sentence (1a) introduced two possible antecedents 'fire' and 'forest' for a pronoun in the second sentence. The target word, presented for lexical decision was 'fire'. In sentence (1b) 'fire' was the appropriate antecedent for the pronoun 'it', whereas in sentence (1c) 'forest' was the appropriate antecedent and 'fire' was inappropriate.

Lexical decisions to appropriate antecedents were significantly faster than lexical decisions to inappropriate antecedents, 629 to 667 msec, respectively. These results indicate that when a pronoun is interpreted, its antecedent becomes activated in memory. However, there are at least three possible types of representations that might be becoming activated: The discourse entity introduced by the phrase 'the fire', the linguistic meaning of 'the fire', or the linguistic form of the phrase 'the fire'. The activation of all or any of these representations might facilitate lexical decisions to the word 'fire'. What is needed then is a methodological tool for distinguishing among these levels of representation.

One possibility is provided by comparing two different tasks that have been widely used in the word recognition literature: lexical decision and naming. In the lexical decision task subjects are asked to decide whether or not a string of letters comprises a familiar English word. Non-words are typically word-like pseudowords such as 'kire'. In the naming task sujects are asked to read a word aloud. At one time it was believed that these two tasks provided similar information about word recognition. For example, in both tasks a word is responded to more rapidly when it is preceded by a semantically or associatively related word (e.g., bank-money) than when it is preceded by an unrelated word (e.g., barn-money). Recently, however, a number of interesting differences between the tasks have been discovered. These differences are summarized in Seidenberg, Waters, Sanders, and Langer (1984). The amount of priming obtained with the lexical decision task but not the naming task varies with the proportion of related items. Lexical decision times but not naming times are faster when a word and its prime form a grammatical phrase than when they do not. Backwards priming, i.e., priming when the target is associated with the prime, but not vice versa (e.g., fly-fruit), obtains with lexical decision but not with naming. The most reasonable explanation for these differences appears

to be that naming is sensitive only to lexical processing. That is, any variable that increases or decreases the availability of a word will either facilitate or interfere with the time it takes to name the word. In contrast, lexical decision is sensitive to both lexical and post-lexical factors. In particular the ease with which a word can be integrated with its context influences the speed with which subjects can make a 'yes' decision to that word. The key seems to be that lexical decision requires the subject to make a conscious decision, and decisions seem to be cognitive processes that cut across levels of representation, whereas naming a word requires accessing procedural knowledge directly tied to the lexical representation of the word. This analysis suggests a possible way of distinguishing between the hypothesis that pronouns are associated with linguistic forms and the hypothesis that pronouns are associated with denotations or discourse entities. If the former hypothesis is true, both lexical decision and naming times should be facilitated when the target word introduces the antecedent of the pronoun. If, however, the discourse entity but not the word itself is being reactivated, then antecedent effects should be seen only with the lexical decision task.

We investigated this possibility in an experiment using definite noun phrase anaphora conducted in collaboration with Margery Lucas. This experiment used cross-modal presentation. Subjects listened to pairs of sentences over headphones and then made a lexical decision to or named a visually presented target word. The first sentence in each set introduced two possible antecedents for a definite noun phrase in the second sentence. Unlike the experiment with Senytka, the disambiguating information came before the anaphoric phrase. Example materials are presented in (2). The target word was 'steak'.

(2) a. Mary was trying to decide whether to buy steak or hamburger.
b. She finally chose the more expensive meat.
c. She finally chose the less expensive meat.
d. She finally went to another store.

The results for the two tasks are presented in Table 1. There was an extremely large appropriateness effect for the lexical decision task with lexical decisions to 'steak' being 137 msec slower when it was the inappropriate antecedent 'meat' than when it was the appropriate antecedent. Also, lexical decisions to 'steak' were not significantly faster when it was the appropriate antecedent compared to a control condition, sentence (2d), in which the second sentence did not contain a definite noun phrase. In contrast, the naming times to 'steak' were only 5 msec slower when it was inappropriate compared to when it was appropriate. The enormous congruity effect obtained with the lexical

decision task underscores the analysis of the task presented earlier. Apparently subjects found it difficult to respond 'yes' to the target word when it was clearly the wrong antecedent. The magnitude of the effect with these materials can possibly be attributed to subjects having to make an inference to decide which meat Mary bought. Having rejected steak as the antecedent of 'less expensive meat', they now had to say 'yes' to steak to make a correct lexical decision. The result is a type of cognitive Stroop effect.[2]

While this effect is striking, our main interest here is in the naming condition. Does interpreting a pronoun reactivate the word that introduced its antecedent? The answer would appear to be that it does not. Thus we can tentatively conclude that interpreting a pronoun involves associating the pronoun with a post-lexical representation. There are several reasons why this conclusion can only be considered tentative. First, it rests on a null result. It is possible that the experimental design was not sensitive enough to detect a very small effect. This is of particular concern because priming differences with naming are frequently fairly small and there was a small (5 msec) difference. Secondly, it seems likely that activating a discourse entity or denotation might result in some lexical activation. Clearly, discourse entities must be concepts of some form and activating a concept can activate a word; such priming is clearly necessary for language production. Thus a small amount of lexical activation might well be observed even if pronoun interpretation makes use of conceptual and not linguistic representations. This is to say that the logic of contrasting lexical decision and naming may not be quite as clear-cut as one would like. Given these caveats, then, we will tentatively conclude that during comprehension prounouns find their antecedents in conceptual representations and not linguistic representations.

We turn now to verb phrase anaphors. Hankamer and Sag (1976) provided convincing evidence that anaphors can be grouped into two classes, which they labelled deep anaphors and surface anaphors. Roughly drawn, the distinction seems to be that surface anaphors require their antecedents to be linguistic units, either surface structure units or semantic units, whereas deep anaphors place no linguistic restrictions on the form of their antecedents. Thus surface anaphors cannot be used deictically, whereas deep anaphors can. Moreover, surface anaphors require their antecedents to be constituents. In contrast, when the antecedent of a deep anaphors is introduced linguistically it need not form a constituent. Thus, both sentences (3c) and (3d), which contain a surface and deep anaphor, respectively, are acceptable when sentence (3a) introduces the antecedent for the anaphor. However, only sentence (3d) is acceptable when the antecedent is introduced in sentence (3b). The problem is that the antecedent 'take the wood to the shed' is not a linguistic unit in (3b).

(3) a. Somebody had to take the wood to the shed.
b. The wood had to be taken to the shed by someone.
c. Sally finally agreed to.
d. Sally finally agreed to do it.

Hankamer and Sag (1976) originally proposed that surface anaphors found their antecedents in surface structure and deep anaphors in deep structure. More recently, however, they have argued that the antecedents of deep anaphors are conceptual representations or discourse entities rather than linguistic forms. This better accounts for the fact that deep anaphors can be used deictically and is consistent with current grammatical theory which does not propose a common underlying representation for active and passive sentences. They have also presented arguments which suggest that the antecedents of surface anaphors are logical form representations rather than surface or S-structure representations.

We have been exploring the deep and surface anaphora distinction in a series of recent experiments. The goal of this research has been to answer two questions: (1) Do deep and surface anaphors find their antecedents in different forms of representation, with surface anaphors taking linguistic antecedents and deep anaphors taking conceptual antecedents?; and (2) If surface anaphors take linguistic antecedents, what is the nature of the antecedent?

In an experiment recently completed we presented subjects with pairs of sentences on each trial. The first sentence was a context sentence, and the second sentence was a target sentence. The subject's task was to decide as quickly as possible whether the target sentence made sense given the context sentence. On about one-third of the trials the target sentence did not make sense. An example sequence is given in:

(4) John couldn't dunk a basketball.
He was proud that he was able to.

On the test trials the context sentence introduced an antecedent for either a surface or a deep anaphor in the second sentence. The materials were drawn from twenty-four sentence sets. Each set contained two versions of a context sentence, a parallel and non-parallel version, and two versions of a target sentence: a deep anaphor and a surface anaphor version. In the parallel context sentence the antecedent was a linguistic constituent. Non-parallelism was introduced either by using a passive context sentence or by nominalizing the verb phrase that introduced the antecedent. Examples of passive and nominalization sets are given in (5a-d) and (5a-d), respectively.

(5) a. Someone took the wood out to the shed last night.
b. The wood was taken out to the shed last night.
c. Tom told us that Sally did.
d. Tom told us that Sally did it.

(6) a. Sally gets enraged when anyone even mentions Judy's name.
b. Just the mention of Judy's name is enough to enrage Sally.
c. But Tom went ahead and did anyway.
d. But Tom wen ahead and did it anyway.

The results are presented in Table 2. Consider the passive sets first. The type of antecedent, parallel or non-parallel, did not reliably influence reaction times to deep anaphors. In addition, there was no difference in the proportion of deep anaphors that were judged to be nonsensical in the two conditions. In contrast, the type of antecedent had a dramatic effect on the surface anaphors. Surface anaphors which were judged to make sense were responded to 929 msec faster in the parallel antecedent condition. Moreover, 29% of the surface anaphors with non-parallel antecedents were judged to be nonsensical compared to only 6% of the surface anaphors with parallel antecedents. The results are not so clear cut for the nominalization condition. Again, there was no reaction time difference between deep anaphors with parallel and non-parallel antecedents. However, more deep anaphors with nominalized antecedents were judged to be nonsensical than deep anaphors with parallel antecedents (8% and 17%, respectively). For the surface anaphors, there was not a significant reaction time difference, however, there was a large effect of parallelism on the proportion of surface anaphors with parallel antecedents were judged to be nonsensical on 7% of the trials, whereas surface anaphors with non-parallel antecedents were judged to be nonsensical on 32% of the trials.

The results in the passive condition clearly suport the hypothesis that surface anaphors are finding their antecedents in some level of linguistic representation where syntactic form is important, whereas deep anaphors are finding their antecedents in conceptual representations where stylistic differences would not be preserved. The nominalization condition offers an even stronger test. If deep anaphors are really finding their antecedents in conceptual representations, then as long as there is a plausible antecedent it shouldn't matter whether the antecedent was introduced in a verb phrase or in noun phrase. The results are encouraging for this position; however, a replication with more materials is clearly necessary given that there was some suggestion of a difference in the proportion of parallel and non-parallel

antecedents that were judged to be nonsensical. It will also be important to replicate this study embedding the materials in more natural texts given the results of an experiment by Murphy (1982) that failed to find significant parallelism effects.

We have conducted two additional types of experiments in order to investigate the hypothesis that deep and surfact anaphors find their antecedents in different levels of representation. In one experiment an intervening sentence is interposed between the antecedent of a dep or surface anaphor. The prediction is that intervening material should interfere with comprehension of the surface anaphor but not the deep anaphor under conditions where the intervening sentence does not introduce possible antecedents. The reasoning is based on the well-established psycholinguistic finding that memory for the linguistic form of a sentence decays rapidly. In several experiments we have found that surface anaphors are comprehended reliably faster than deep anaphors when the antecedent is introduced in the sentence immediately preceding the sentence containing the anaphor. However, when a sentence intervenes between the anaphor and it antecedent, comprehension of a surface but not a deep anaphor is slowed. Thus, comprehension of sentence (7c) but not (7d) is slowed when sentence (7b) follows sentence (7a), which introduces the antecedent for the anaphors in sentence (7c) and (7d).

(7) a. Mary quit her job yesterday.
b. The job was extremely demeaning.
c. Now Harry also has the courage to.
d. Now Harry also has the courage to do it.

However, Greg Murphy, has pointed out that some of our materials introduced implausible but syntactically possible antecedents for surface anaphors. Thus for example, we have included sentence sets similar to those in (8) in our experiments.

(8) a. Harry has been trying to run a five minute mile for three years.
b. Three years is a long time to pursue any goal.
c. Yesterday he finally did.
d. Yesterday he finally did it.

In the intervening sentence (8b), 'pursue any goal' is a syntactically possible antecedent for the verb phrase ellipsis anaphor in (8c). We are presently replicating the intervening sentence findings using materials such as those in (7) with the "makes sense" judgment task used in the parallelism study and in a reading time study in collaboration with Murphy.

PROCESSING DEEP AND SURFACE ANAPHORS

Finally, we have conducted several experiments in which we varied the syntactic form of the antecedent sentence by using verb-particle structures in which the particle either immediately followed the verb or occurred after a noun phrase. Subjects read the antecedent sentence, then read a sentence that contained either a surface or deep anaphor. After they finished reading the second sentence, they were timed as they decided whether or not a verification sentence was true or false. The verification sentence had either a shifted or an unshifted particle. Sample materials are presented in:

(9) a. Somebody had to pick up Bill's brother at the airport.
b. Somebody had to pick Bill's brother up at the airport.
c. After much effort, we finally convinced Sally to.
d. After much effort, we finally convinced Sally to do it.
e. (verification sentence) Somebody had to pick up Bill's brother.

In several experiments we found that the form of the verification sentence affected the time it took subjects to respond to the verification sentence for surface anaphors but not for deep anaphors, with verification sentences that matched the verb particle order of the antecedent being verified faster than mismatching sentences. This result seemed to provide direct evidence that readers were reactivating the linguistic form of the antecedent when interpreting surface anaphors but not deep anaphors. However, in several recent studies we have failed to replicate this effect, finding instead a mismatch effect for both deep and surface anaphors. Chuck Clifton and Lyn Frazier (personal communication) using a similar design have also found mismatch effects for both deep and surface anaphors. At this point it is probably best to assume that the results of our earlier experiments were misleading.

Where does this leave the hypothesis that deep anaphors are associated with conceptual antecedents and surface anaphors linguistic antecedents? Perhaps the safest conclusion that can be drawn at the moment is that the hypothesis remains viable but needs further empirical test before it can be accepted or rejected. The parallelism effects reported above are encouraging for the hypothesis as are the intervening sentence results. However, the failure to replicate the particle shift experiments leaves us without any direct evidence about what gets reactivated when a verb phrase anaphor is interpreted.

Table 1

Results of the definite noun phrase anaphora experiment with the lexical decision and naming tasks.

Condition	Lexical Decision Latencies	Naming Latencies
Appropriate Antecedent	720 msec	525 msec
Inappropriate Antecedent	857 msec	529 msec
Control	742 msec	551 msec

Table 2

Results of the paralellism study using the judgment task: Reaction time to judge the anaphor as sensible and percentage of anaphors judged to be non-sensible.

	Passive Stimuli	
Type of Anaphor	Parallel Antecedent	Nonparallel Antecedent
Surface	2084 msec (6%)	3013 msec (29%)
Deep	2418 msec (7%)	2360 msec (6%)
	Nominalization Stimuli	
Surface	2183 msec (7%)	2236 msec (32%)
Deep	2382 msec (8%)	2237 msec (17%)

PROCESSING DEEP AND SURFACE ANAPHORS

Footnotes

1. The research presented in this paper was supported by NSF grant BNS-8217378 to the authors.

2. The classic Stroop phenomenon obtains when color names, e.g., 'red', are printed in an antagonistic color. In these circumstances, people find it difficult to rapidly name the color the word is printed in because of interference from the word.

References

Carlson, G. & Tanenhaus, M.K. (1984). Lexical meanings, structural meanings, and concepts. In J. Drogo, V. Mishra, & D. Testen (Eds.) Parassession on Lexical Semantics. Chicago: Chicago Linguistic Society.

Chang, F.R. (1980). Active Memory processes in visual sentence comprehension: Clause effects and pronominal reference. Memory & Cognition, 8, 58-64.

Chomsky, N. (1972). Language and Mind. New York: Harcourt Brace Johanovich, Inc.

Dell, G.S., McKoon, G., & Ratcliff, R. (1983). The activation of antecedents information during the processing of anaphoric reference in reading. Journal of Verbal Learning and Verbal Behavior, 22, 121-132.

Hankamer, J. & Sag, I. (1976). Deep and surface anaphora, Linguistic Inquiry, 7, 391-426.

Murphy, G. (1982). Understanding anaphora. Stanford University unpublished doctoral dissertation.

Sag, I. (1976). Deletion and logical form. MIT doctoral dissertation distributed by the Indiana University Linguistics Club.

Shillcock, R. (1982). The on-line resolution of pronominal anaphora. Language and Speech, 25, 385-401.

Tanenhaus, M.K., Carlson, G., & Seidenberg, M.S. (1985). Do listeners compute linguistic representations? In D. Dowty, L. Kartunnen, and A. Zwicky (Eds.) Natural Language Parsing: Theoretical, Psychological, and Computational Perspectives. Cambridge: Cambridge University Press.

Seidenberg, M.S., Waters, G.A., Sanders, M., & Langer, P. (1984). Pre- and postlexical loci of contextual effects in word recognition. Memory & Cognition, 12, 315-328.