

McDonald, J. L., & MacWhinney, B. J. (1995). The time course of anaphor resolution: Effects of implicit verb causality and gender. *Journal of Memory and Language*, 34, 543-566.

Abstract

This paper examines the time course of the use of implicit verb causality and gender agreement information in determining the antecedent of a pronominal anaphor. A crossmodal probe paradigm measured the activation of candidate antecedents at four time points during comprehension. Sentences contained two participants, an implicit causality verb, and a subsequent “because”-clause (e.g., *Gary amazed Ellen time after time because he was so talented.*). In Experiment 1, a general advantage for the first mentioned participant was nullified immediately upon encountering the pronoun if the implicit causality of the verb favored the second participant. Experiment 2 showed that this effect was due to anaphor resolution processes, and not priming from the verb. Experiment 3, which deleted the causal conjunction “because” and placed the pronoun in a separate sentence, continued to find an effect of implicit causality. In addition, the first mentioned advantage was nullified at the pronoun for sentences *without* a gender contrast. The overall results are consistent with a model of anaphor resolution in which resolution processes are initiated immediately upon encountering the pronoun, but cannot be completed until information becomes available from other processes. Sentence structure may influence the timing of the availability of different types of information.

In order to understand a sentence containing a pronoun, the listener must correctly assign an antecedent to the pronoun. For example, to understand the sentence *Peter annoyed Dennis occasionally, because he was so talkative*, the listener must decide whether “he” refers to “Peter” or “Dennis.” The favored interpretation of this sentence is that Dennis was annoyed at Peter, because Peter talked too much. However, it is also possible to think of a situation where the opposite assignment would be correct. For example, what if Peter were a mischievous person, who enjoyed annoying people whenever they started to monopolize conversations? In this case, Peter would chose to annoy Dennis, precisely because Dennis talked too much and the correct antecedent of “he” would be “Dennis” rather than “Peter.”

The process by which an antecedent is assigned to an anaphor is called anaphor resolution. Recent studies of anaphor resolution have consistently indicated that the anaphor does not simply and directly activate a single possible antecedent; rather, the occurrence of an anaphor activates several competing candidate antecedents as well as co-occurring concepts and lexical associates of these candidate antecedents. For example, using a probe inserted at various points in a sentence, Dell, McKoon, and Ratcliff (1983) found that a noun anaphor activated both its antecedent and other concepts that had occurred in the same phrase soon after (250 msec) the occurrence of the anaphor. At later probe points, the antecedent remained active, while the relative availability of the other concepts that had occurred in the same phrase decreased.

Similar results have been found in studies of pronoun anaphors. The occurrence of an anaphoric pronoun appears to activate several candidate antecedents immediately at or after the pronoun (Corbett & Chang, 1983; Gernsbacher; 1989; MacDonald & MacWhinney, 1990; Smyth, 1986; Tyler & Marslen-Wilson, 1982). However, several studies have found

that at various time lags after the pronoun, the correct antecedent has become more available than the incorrect one (Chang, 1980; Gernsbacher, 1989; MacDonald & MacWhinney, 1990; Stevenson, 1986). Gernsbacher (1989) has proposed that the change in activation level of candidate antecedents over time can be explained by the enhancement of the correct antecedent and the active suppression of incorrect candidates. Using several controls, MacDonald & MacWhinney (1990) have shown both of these processes operate for pronominal anaphors.

The probe recognition methodology. Many methodologies have been used to examine anaphor resolution processes. Some of these methodologies involve off-line or global measures of processing, such as reading times or antecedent naming latencies, while others try to measure the on-line use of information through eye movements or probe recognition. Since we are particularly interested in the time course of anaphor resolution, the experiments in this paper use the on-line measure of a probe recognition task. The particular form of the probe recognition task we have selected is the crossmodal version in which the sentence is presented aurally and the probe is presented visually. This technique contrasts with the version of the probe recognition task in which both the sentence and the probe are presented visually. A direct comparison of these two methods (cf. MacDonald (1986) and MacDonald & MacWhinney (1990)) indicated that the crossmodal technique was superior to the unimodal technique in terms of providing a sensitive measure of the on-line time course of the processing of pronominal anaphors.

It is important to note that the probe recognition task is not exclusively sensitive to anaphor resolution processes; rather, it is responsive to all processes that affect the activation of a particular participant. For example, the accessibility of all participants will be affected by concurrent processing load, with faster latencies occurring under smaller processing loads. In addition, probe recognition latencies of individual participants will be

differentially affected by processes involved in stage-setting (Chafe, 1972) and perspective shift (MacWhinney, 1977) which may cause a certain word or participant to move in or out of focus. In particular, first participants appear to be a starting point or foundation for building a comprehension structure (Gernsbacher, 1990; MacWhinney, 1977), causing them to have increased activation levels over other participants in the same sentence (Corbett & Chang, 1983; Gernsbacher, 1989; Gernsbacher & Hargreaves, 1988; Gernsbacher, Hargreaves & Beeman, 1989). However, perspective can be split between participants in structures such as passives (MacWhinney, 1977), relative clauses (MacWhinney & Pléh, 1988), or verbs with stimulus-experiencer implicit causality (MacWhinney, 1985). In these other structures, the default high activation of the first participant is modified by other processes to produce final activation levels. Thus, the activation of participants due to other factors such as first mention or perspective split must be kept in mind when interpreting activation levels in probe recognition experiments on anaphor resolution.

In the current experiments, two specific factors involved in anaphor resolution: (1) the semantics of implicit verb causality and (2) gender agreement between the pronoun and candidate antecedents, are added to the background factors of processing load and first participant advantage. Information from all these factors will cooperate and compete in determining the final level of participant activation in the probe recognition task. The anaphor resolution factors of implicit verb causality and gender agreement were specifically chosen because while both these factors have been investigated previously, no prior study has looked at the combination of both these factors using an on-line methodology. With an on-line methodology we should be able to establish both when during the course of a sentence people make use of each type of information, and how the information from the sources are integrated. Below we review the literature pertinent to the effects these two factors have upon anaphor resolution.

Implicit verb causality. Certain verbs have a clear causal directionality implicitly encoded in their meaning. For some verbs this causality favors the subject of the verb as the causer, for other verbs it favors the object. Consider the verb “amaze.” In a sentence such as *Gary amazed Alan*, we usually assume that it is Gary who is amazing rather than to Alan who is amazeable. In other words, it is Gary who is the causer of the amazement. For the verb “admire,” the opposite is the case. In a sentence such as *Karen admired Becky*, the cause of the admiration is usually attributed to Becky who is admirable rather than to Karen who is admiring. For verbs of the “amaze” type, known as Stimulus-Experiencer (S-E) verbs, people tend to attribute causality to the subject; for verbs of the “admire” type, known as Experiencer-Stimulus (E-S) verbs, people tend to attribute causality to the object (Au, 1986; Brown & Fish, 1983; Garvey & Caramazza, 1974).

The effects of the implicit causality of a verb on anaphor resolution are illustrated most clearly in what we will call a “because-clause” -- a postposed subordinate clause that begins with a causal conjunction such as “because” and contains a pronominal anaphor. For example, in the sentence *Gary amazed Alan, because he was so talented*, the implicit causality of the verb “amaze” is established in the first clause and is then carried over to the because-clause where it should lead the listener to favor “Gary” as the antecedent of “he”. Similarly, in the sentence *Karen admired Becky, because she was so intelligent*, the implicit causality of the verb “admire” should cause people to favor “Becky” as the antecedent of “she”. This prediction has been supported in past research using sentences of this type. For example, when the time to read such sentences and name the pronominal antecedent was measured, people responded faster to sentences where the implicit verb causality was consistent rather than inconsistent with the assignment of the pronoun as indicated by other information in the sentence (Caramazza, Grober, Garvey & Yates, 1977; Caramazza & Gupta, 1979; Garnham & Oakhill, 1985; Vonk, 1985a; 1985b). Thus, it appears that people do make use of implicit verb causality in assigning antecedents to pronouns in because-

clauses. The use of this type of information does not seem to be under strategic control by the subject (Garnham, Oakhill & Cruttenden, 1992; McKoon, Greene, & Ratcliff, 1993). However, the time course and the details of how this information is used have not yet been established.

Gender agreement. When only one potential antecedent of a pronoun has the same gender as the pronoun, it is possible to choose the antecedent correctly on the basis of gender alone. Potentially, the information necessary for this selection is immediately available as soon as the pronoun is encountered. That is, for a sentence such as *Gary amazed Ellen, because he was so talented*, as soon as “he” is encountered, gender agreement indicates “Gary” as the only possible correct antecedent. However, previous research has indicated that, while gender agreement may eventually influence pronoun assignment, this information may not be used immediately upon encountering the pronoun. This research is reviewed below.

Initial research on this topic found faster latencies to sentences where gender agreement could resolve pronoun assignment than to sentences where it could not (Caramazza et al., 1977; Ehrlich, 1980; Stevenson & Vitkovitch, 1986). These studies asked subjects to press or release a button after they had read the sentence and determined the referent of the pronoun. Results collected by using this end-of-the-sentence technique do not indicate clearly whether gender agreement helps on-line with the initial process of referent assignment during reading, or with subsequent processes of antecedent naming or verification. Later experiments which separated reading times from times to name or verify the antecedent usually found that gender agreement effects were localized in naming or verification times, but not in initial reading times (Smyth, 1986; Stevenson & Vitkovitch, 1986; Vonk, 1985a; 1985b), although at least one study found gender agreement effects on both measures (Garnham & Oakhill, 1985).

When reading times on such sentences were broken down by the more fine-grained measure of eye movements, it was again found that initial pass reading time on sentences with gender cues did not differ from those without gender cues (Vonk, 1984). However, the pronoun itself was more likely to be fixated and more time was spent on it when there was a gender contrast. Thus, for the antecedent naming task used by Vonk (1984), an immediate effect of gender information was found at the pronoun. However, these results did not replicate in an antecedent verification task (Vonk, 1985b).

Use of gender information in anaphor resolution has also been examined with a probe recognition task. In general, when probes are inserted immediately after a pronoun, no advantage is found for a candidate antecedent that matches the pronoun in gender over those which do not (Gernsbacher, 1989; MacDonald & MacWhinney, 1990; Tyler & Marslen-Wilson, 1982). However, when the probes are inserted further downstream from the pronoun, some studies have found an advantage for the candidate that agrees in gender with the pronoun (Chang; 1980; Gernsbacher, 1989; MacDonald & MacWhinney, 1990; McKoon et al., 1993; Stevenson, 1986). However, another study (Tyler & Marslen-Wilson, 1982) found no such advantage at any point.

In an attempt to resolve these contradictory results about gender information, Garnham et al. (1992) conducted a series of experiments varying the demands for immediate pronoun resolution. They found that gender only had an effect on reading times when the task consistently and saliently demanded pronoun resolution. If the demands were not so salient, gender information did not effect reading times. Thus, Garnham et al. (1992) suggest that the use of gender information in anaphor resolution is under subject control, and only used when this strategy aids performance in the particular experimental task. This conclusion is in line with the work of others, which suggest that anaphor resolution does not necessarily

automatically begin immediately at the pronoun; rather this may only occur under certain circumstances, such as experiments where a comprehension question about the pronominal referent occurs after every sentence (Greene, McKoon & Ratcliff, 1992; McKoon et al., 1993).

Given the many different kinds of experimental methodologies used in the above research, it is not surprising that the picture of gender use in anaphor resolution is not totally clear. It may be influenced by the experimental task or materials. However, the majority of the studies indicate that gender information is not used when it is first potentially available, i.e., at the pronoun. On the other hand, there is abundant evidence to indicate that gender information influences later interpretations and later responses to overt sentence interpretation tasks.

Given that little is known about the time course of implicit verb causality information, and that there are few on-line cross-modal studies of gender agreement on anaphor resolution (MacDonald & MacWhinney, 1990; Tyler & Marslen-Wilson, 1982), the current paper investigates both these factors. Three experiments using the cross-modal probe recognition paradigm are performed. Experiment 1 examines the influence of the implicit causality of mental verbs and gender agreement in sentences where the anaphoric pronoun is contained in a subsequent because-clause. Experiment 2 provides a control condition, using sentences containing verbs of implicit causality without an anaphoric pronoun. Experiment 3 looks at the effects of implicit causality and gender agreement when the anaphoric pronoun occurs in a separate sentence. Recall that the probe recognition methodology will be affected by factors such as concurrent processing demands and the first mention advantage. We will be looking for changes in the relative activations of participants caused by implicit causality and gender agreement on top of these expected background effects.

Experiment 1

Method

Subjects. Subjects were 128 fluent speakers of English.

Stimuli. Thirty-two test sentences were constructed. Each consisted of two clauses -- an initial clause, which contained two names, a causal verb and an adverbial phrase, and a subsequent explanatory because-clause containing a pronominal anaphor. For example, in the sentence *Gary amazed Alan time after time, because he was so talented*, the two nouns are “Gary” and “Alan”; the causal verb is “amazed”; the adverbial phrase is “time after time”; the because-clause is “because he was so talented”; and the pronominal anaphor is “he.” All stimuli followed the abstract outline give in Table 1.

1 *2* *3* *4*

Name1 Verbed Name2 Adv-phrase, because Pro was so Adj

Table 1: Example stimulus sentence with location of four time delays

Half of the test sentences contained Stimulus-Experiencer (S-E) verbs such as “amazed”; half contained Experiencer-Stimulus (E-S) verbs such as “admired”. Eight verbs of each type were chosen from those given in Au (1986) and each verb occurred in two sentences. Each sentence had a same gender participants version (e.g. *Gary amazed Alan...*), and an opposite gender version (e.g., *Gary amazed Ellen...*).

During the aural presentation of the sentence, a probe word appeared on the computer screen, and subjects were to decide whether the probe word had already been mentioned in the current sentence. For test sentences, half of the time the probe was the first name, and half of the time it was the second name. Probes could occur at four different points during the aural presentation of the sentence. Delay 1 was 100 msec after the conclusion of the second name. Delay 2 occurred immediately after the pronoun. Delay 3 was 200 msec after the pronoun. Delay 4 was at the end of the sentence. Table 1 shows the approximate placement of the four time delays.

All four of the above factors were completely crossed, yielding a 2 (verb type: S-E vs. E-S) by 2 (gender of participants: same vs. different) by 2 (probe: Name 1 vs. Name 2) by 4 (time delay: 100 msec downstream from end of Name 2, immediately at the end of the pronoun, 200 msec downstream from the end of the pronoun, and end of the sentence) design, for a total of 32 conditions. Each subject heard one test sentence in each of the 32 conditions and no test sentence was heard more than once by any one subject. Across each group of 16 subjects, each sentence was systematically stepped through the various levels of gender, probe and delay.

Several steps were taken to keep test sentences as uniform as possible. The two names within a sentence were matched for number of syllables, and all names were fairly common American names. Male and female names used in the gender manipulation within a particular sentence were matched as closely as possible for phonetic properties (e.g., Alan/Ellen; Andy/Amy; Kevin/Karen). The adverbial phrase intervening between the second name and the second clause was between two to five syllables, as was the adjective at the end of the sentence. All test sentences are given in the Appendix.

In addition to the 32 test sentences, three practice and 40 filler sentences of similar form were included. These sentences also had two nouns and a verb in the main clause and a postposed because-clause with a pronominal anaphor. Although they did not all use the clear causality S-E or E-S verbs of the experimental sentences, the filler sentences were constructed so that half of the time the first name would be favored as the causer, and the other half of the time the second name would be favored as the causer. The gender of the referents and the time delays were varied as in the test sentences. However, in the filler sentences, the probe word appearing on the screen was either a third name that did not appear in the sentence (40% of filler sentences), a verb that had appeared in the sentence (40% of the fillers), or a verb that was not in the sentence (20% of the fillers). Thus, over the 72 test and filler sentences that subjects heard, 2/3 had probe words that were names and 1/3 had probe words that were verbs. Over all probe words, 2/3 had actually appeared in the sentences and 1/3 had not.

Test and filler sentences were given in a different computer-controlled randomized order for each subject. These sentences were preceded by three practice sentences of similar design to get subjects used to the instructions and apparatus.

The first and second phrases of each sentence were first recorded onto a tape cassette at a normal rate of presentation with normal intonation. These phrases were then digitized and stored on an IBM AT's hard disk using an analog to digital conversion board (Data Translation's DT 2821 board). In each first phrase (e.g. *Gary amazed Alan time after time*), a mark recognizable to the computer, but undetectable to the subjects, was placed immediately after the last sound of the second name. This was done by examining the sound spectrograph display for each phrase and marking the place of least activity between the second name and the following word. In each second phrase (e.g. *because he was so talented*), a mark was placed immediately after the pronoun using the same procedure.

These marks were then used by the program to determine the onset of the probe at the various time delays.

Sentences were formed by digitally splicing together the first and second phrase appropriate to each sentence. There were two possible first phrases for each test sentence with either same-gender referents or different-gender referents (e.g. *Gary amazed Alan time after time*, vs. *Gary amazed Ellen time after time*). The same second phrase was combined with each of these versions (e.g. *because he was so talented*).

Comprehension questions of the form "Who was so ADJECTIVE?" (e.g., *Who was so talented?*) were devised for test, filler, and practice sentences. The questions were designed to query the assignment of an antecedent to the pronoun. For S-E verbs this was the first name and for E-S verbs this was the second name.

Procedure. Subjects were seated in front of the computer screen and told to position the fingers of one hand on a three-buttoned button box. They were told that they would be hearing aurally presented sentences, and that, at some point during the presentation of the sentence, a word would appear on the computer screen. If the word had occurred in the current sentence up to that point, they were to push the button marked YES; if the word had not occurred in the sentence, they were to push the button marked NO. Responses were to be as quick and accurate as possible. Subjects were told not to wait to the end of the sentence to make their decision--if the word had not occurred in the sentence by the time it appeared on the screen, it would not occur later on. Both reaction time and judgment response were recorded for the probe word task by the computer.

After the end of the aural presentation of each sentence, and after the subject had made a decision regarding the visually-presented probe word, the comprehension question appeared

on the screen. Subjects gave their response out loud. After the experimenter recorded this response, subjects pushed the middle, unmarked button of the button box to get the next sentence.

Data trimming. Before analyzing the probe reaction time data, the mean and standard deviation for the correct responses to test sentences for each individual were computed. The data for each subject was then trimmed so that latencies that were greater than three standard deviations away from the mean were replaced by that limit. Reaction times to items that had incorrect probe judgments were replaced by the mean correct response time for that subject. Six new subjects were substituted for subjects who had more than three altered latencies, or who had an average correct reaction time greater than 2 seconds on target sentences. Trimming and replacement procedures on the final 128 subjects resulted in the alteration of 123 latencies or 3.0% of the total data.

Results

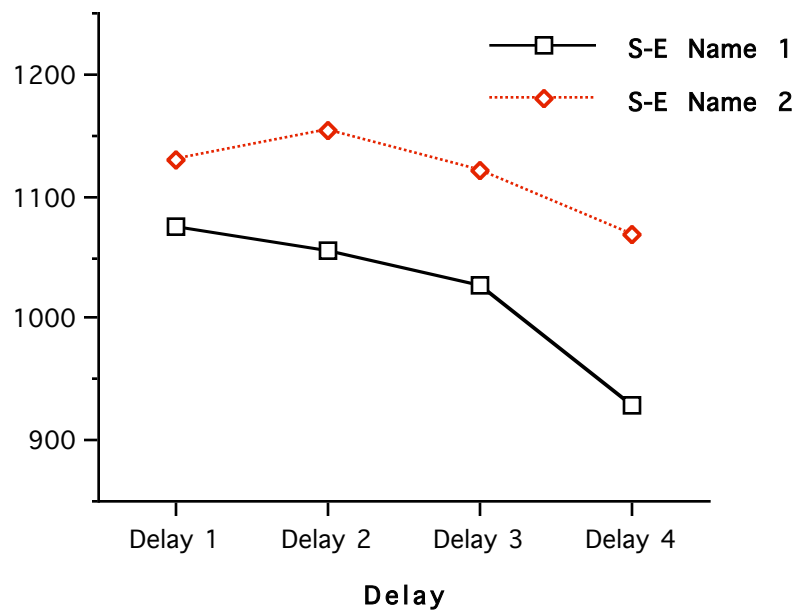
Separate analyses were conducted on the three dependent variables: (1) probe reaction time, (2) probe error rate, and (3) comprehension question error rate. Subject analyses used a 2 (verb type) x 2 (gender) x 2 (probe) x 4 (delay) within-subject design. Item analyses used the same design, with the exception that verb type was a between-item variable.

Probe reaction time. The expected main effects due to concurrent processing demands and first mention advantage occurred. Concurrent processing demands would decrease throughout the presentation of a sentence, and be particularly low at its end. The main effect of time delay ($F(3,381)=24.0, p<.001$; $F(3,90)=22.0, p<.001$) resulted from exactly this pattern (Delay 1: 1094 msec, Delay 2: 1082 msec, Delay 3: 1070 msec, Delay 4: 972 msec). The first mention advantage, well established in other literature (Corbett & Chang, 1983; Gernsbacher, 1989; Gernsbacher & Hargreaves, 1988; Gernsbacher, Hargreaves &

Beeman, 1989), also occurred here ($F(1,127)=69.5, p<.001$; $F(1,30)=27.4, p<.001$), with a nearly 80 msec advantage of the first participant (1015 msec) over the second participant (1094 msec). In addition, there was a main effect of verb type ($F(1,127)=6.7, p<.05$; $F(1,30)=5.9, p<.05$) reflecting somewhat slower latencies to sentences with S-E verbs (1070 msec) than to those with E-S verbs (1039 msec) and a marginal main effect of gender in the subjects analyses ($F(1,127)=3.7, p<.06$; $F(1,30)=1.6, n.s.$) as a result of slower response times for same-gender sentences (1067 msec) than for different-gender sentences (1044 msec). These main effects were qualified by two three-way interactions.

There was a significant three-way interaction between verb type, probe, and delay ($F(3,381)=4.0, p<.01$; $F(3,90)=3.9, p<.05$). As can be seen in Figure 1, the first participant advantage was present at all time delays for S-E verbs (Figure 1a), but disappeared at delay 2, immediately after the pronoun is encountered, and at delay 4, end of sentence, for E-S verbs (Figure 1b). The loss of the first participant advantage for these two time delays for E-S verbs is caused by the increased accessibility of the second participant--indeed, the second participant is faster for E-S verbs than S-E verbs for these two delays (delay 2: $F(1,127)=9.6, p<.005$; $F(1,31)=12.0, p<.005$; delay 4: $F(1,127)=23.9, p<.001$; $F(1,31)=13.0, p<.005$), while there is no significant difference in reaction time to second participants across verb types for delays 1 and 3. In addition, within the E-S verb type, the latency at delay 2 is faster than that at the preceding ($F(1,127)=9.4, p<.005$; $F(1,15)=7.3, p<.05$) and following test point ($F(1,127)=6.6, p<.05$; $F(1,15)=4.2, p<.06$); the latency at delay 4 is also faster than that at its preceding test point ($F(1, 127)=37.6, p<.001$; $F(1,15)=57.2, p<.001$) Thus, to summarize, in an E-S sentence such as *Karen admired Becky openly, because she was so intelligent*, there is an initial first participant advantage for "Karen" which is cancelled at the pronoun due to increased accessibility for the second participant "Becky." The first mention advantage is re-established 200 msec later at delay 3, only to disappear again by the end of the sentence. This reaction time pattern for E-S verbs

seems to indicate that while implicit causality immediately affects participant accessibility at the pronoun, this effect may fade during the rest of the sentence while other processing takes place. Implicit causality again becomes important during end of sentence wrap-up procedures.



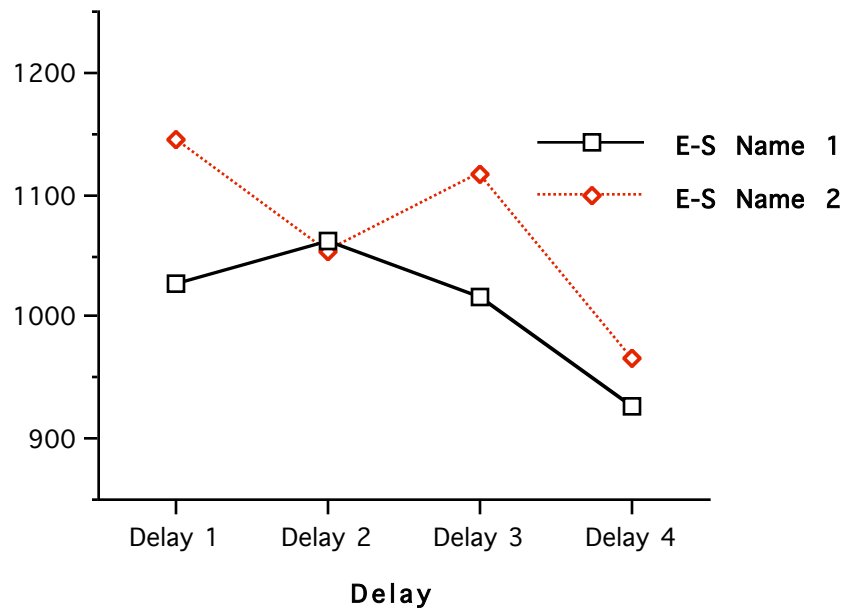
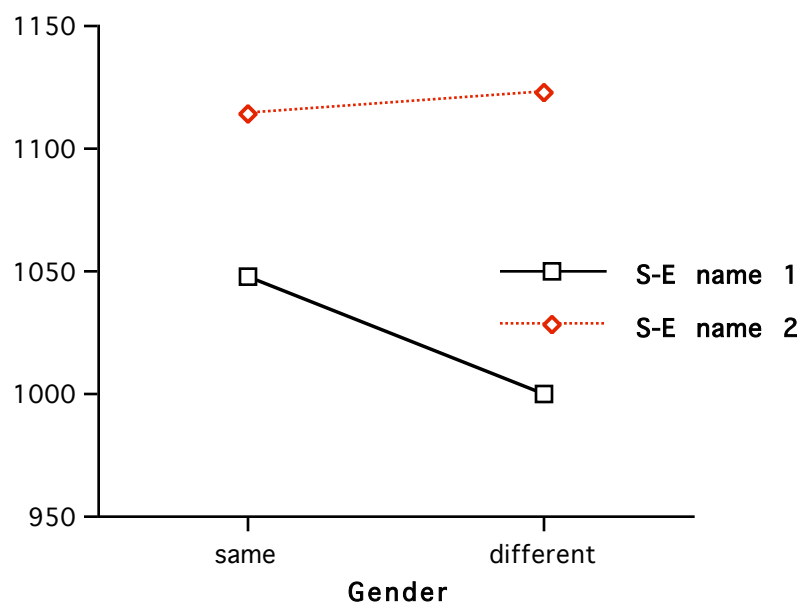


Figure 1: Experiment 1: Interaction of verb type, probe and delay. a) S-E verbs; b) E-S verbs



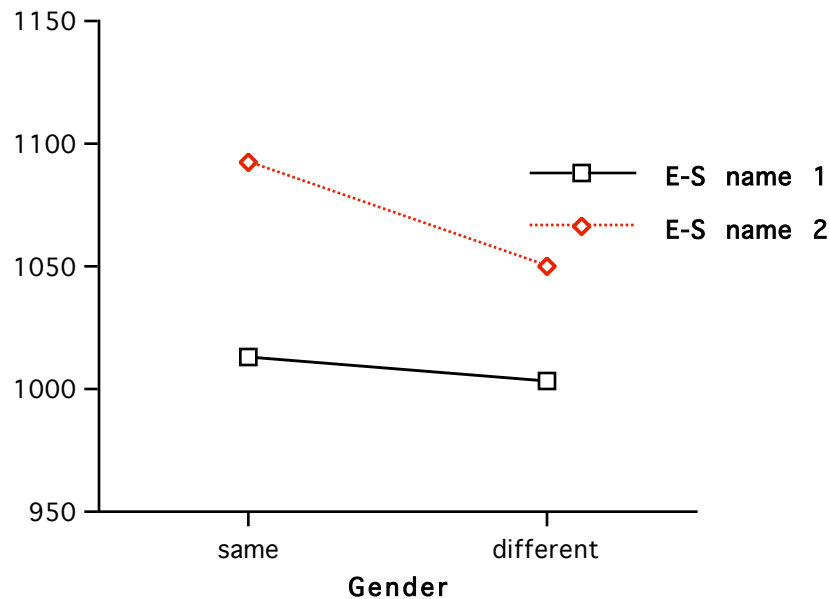


Figure 2.: Experiment 1: Interaction of gender, verb type, and probe

Verb type also interacted significantly with probe and gender in the subjects analysis ($F(1,127)=4.6, p<.05$; $F(1,30)=1.0, n.s.$). This interaction, shown in Figure 2, shows that the availability of gender information magnified the effect of verb causality in increasing the accessibility of the pronominal referent. That is, for S-E verbs the first participant was more available when a gender contrast was present than when it was not, and for E-S verbs, the second participant was more available when gender contrast was present than when it was not. As noted above, however, this tendency did not generalize across items.

There were no interactions involving both gender and delay. Thus, the statistical analysis does not support any differential use of gender information at the different time delays to resolve the anaphor. This result is consistent with previous research that has not found immediate use of gender information at the pronoun (Gernsbacher, 1989; MacDonald & MacWhinney, 1990; Tyler & Marslen-Wilson, 1982). However, an examination of the

uncollapsed cell means hints at the possibility of some gender effects at the pronoun. For example, for S-E verbs the second participant in the gender different condition is the only one to decrease in accessibility between delay 1 (before the pronoun--1120 msec) and delay 2 (immediately after the pronoun--1181 msec). That is, “Ellen” in *Gary amazed Ellen time after time, because he was so talented* is suppressed after the gender mismatched pronoun. Similarly, for E-S verbs the first participant in the gender different condition is the only one to show suppression between delays 1 (1012 msec) and 2 (1072 msec). Again, the gender mismatched pronoun is suppressed (e.g., “Kevin” in *Kevin admired Becky openly, because she was so intelligent.*)

Any hints we have of such gender effects in this study are not statistically significant; thus, these tendency shown in the means should not be taken too seriously. One reason that this and other studies have failed to find reliable use of gender information at the pronoun may have to do with whether or not subjects chose to use an interpretation strategy that focuses on gender information as useful (Garnham et al., 1992). Since the comprehension question always queried the referent of the pronoun, some subjects may have decided to use gender information as soon as it became available to resolve the anaphor they knew they would be questioned about later. However, other subjects may have chosen not to use this strategy. Thus, statistical analysis would fail to find any consistent effects of gender interacting with delay across subjects.

Probe judgment error rates . Error rates were low on the probe judgment task, averaging about 1.4%. There was one significant two-way interaction between verb type and delay ($F(3,381)=3.2, p<.05$; $F(3, 90)=2.7, p<.05$). There was a generally higher error rate to S-E verbs except at the third delay. The higher error rate to probes in sentences with S-E verbs is in accord with the main effect of verb type in the latency analysis caused by slower

reaction times to probes in S-E sentences; however the reason for a decreased error rate for this verb type at delay 3 are not clear.

Comprehension question error rates. Answers to the comprehension questions were scored as correct if they followed the indications of the verb semantics. This means that the first noun was scored as correct for S-E verbs, and the second noun was scored as correct for E-S verbs. Naming the other name in the sentence, a name not in the sentence, or failing to answer the question was scored as incorrect. Because of a recording error, all comprehension question answers for one subject were lost. In order to analyze the data in the within-items items analysis, these data were considered as all correct, rather than missing.

Error rates to the comprehension questions were rather high, around 17%. Verb type was significant ($F(1,127)=96.8, p<.001$; $F(1,30)=54.6, p<.001$), with a higher error rate to S-E verbs (23%) than to E-S verbs (10%), again showing subjects had more difficulty with the former sentences. Delay was marginally significant ($F(3,381)=2.4, p<.07$; $F(3,90)=3.4, p<.05$), with the highest rate at delay 3 (delay 1: 16%, delay 2: 15%, delay 3: 19%, delay 4: 17%). Gender was highly significant ($F(1,127)=139.4, p<.001$; $F(1,30)=89.9, p<.001$), with more errors when there was not a gender contrast (24%) than when such a contrast was present (10%). Verb type interacted with both gender ($F(1,127)=53.0, p<.001$; $F(1,30)=36.4, p<.001$) and probe ($F(1,127)=36.4, p<.001$; $F(1,30)=32.0, p<.001$). These two-way interactions were qualified by a three-way interaction between verb type, gender and probe ($F(1,127)=7.6, p<.01$; $F(1,30)=4.6, p<.05$). As shown in Table 2, people made more errors when the probe word was not the correct answer (i.e., when name 2 was the probe for S-E verbs, and name 1 for E-S verbs). This tendency was greatly attenuated when a gender contrast was present. Thus, subjects

appear to have a tendency to repeat the probe word, but this tendency is reduced when gender agreement can be used to answer the comprehension question.

VERB TYPE		GENDER	
<u>Stimulus-Experiencer verbs</u>		same	different
PROBE	Name 1	28	11
	Name 2	41	14
<u>Experiencer-Stimulus verbs</u>			
PROBE	Name 1	16	8
	Name 2	10	6

Table 2: Experiment 1: Error rates to comprehension questions with verb type by gender by probe interaction

Summary. The results of the reaction time analysis indicate that the activation of candidate referents is affected by many factors. First, latencies to probes in general decrease under decreasing processing load. This is shown by faster reaction times at the end of the sentence, and may also be the reason for generally faster latencies to sentences with E-S verbs and different gender sentences. Second, there was a pervasive advantage for the first participant over the second participant. On top of these two effects were changes in accessibility caused by implicit verb causality. For E-S verbs we saw an increase in the availability of the second candidate (the candidate favored by implicit causality) focussed at the pronoun and returning at the end of the sentence, such that the usual first participant advantage disappeared. We did not see a corresponding increase in the availability of the first noun for S-E verbs at the pronoun, probably because this candidate already had an activation advantage. The fact that implicit causality had its effect at the pronoun for E-S verbs indicates that this information is immediately available for anaphor resolution. That this increase in accessibility went away at delay 3 before reappearing at the end of the sentence may indicate that after some immediate processing at the pronoun the use of this information is temporarily suspended, until it again is used by end of sentence wrap-up procedures.

Note that at no time for E-S verbs was there an accessibility advantage for the referent (the second participant) over the non-referent (the first participant). The processes that occur at the pronoun and at the end of the sentence only cancelled out the first mention advantage. Since anaphor resolution is only one of the processes influencing participant reaction times on these sentences, we must look for relative changes in participant activation from their background state to infer the occurrence of anaphor resolution processing. These processes might not necessarily yield an absolute advantage for the referent at any particular point in time when they are engaged. This is not to say that at some point in time referents would not be more available than non-referents, just that this is not necessary to infer that anaphor resolution processes are engaged (although not necessarily completed).

Finally, while the subjects analysis indicated that gender contrast helped in general to promote the gender matched participant, there was no statistically significant evidence of gender interacting with delay in anaphor resolution. Individual cell means did show a tendency for candidate antecedents disagreeing in gender with the pronoun to be suppressed at the pronoun. The lack of statistical significance of this effect may be due to strategy differences across subjects.

Experiment 2

Experiment 1 showed an effect of implicit verb causality on the activation of the participants immediately at the pronoun. However, this effect in Experiment 1 could be interpreted as being stimulated either by (1) the process of anaphor resolution or by (2) anticipatory priming of referents on the basis of the implicit causality of the verb. We can refer to these two alternatives as the resolution account and the anticipation account. The

resolution account holds that listeners wait to hear the pronominal anaphor and then begin to activate the possible referents of the pronoun. In this account, verb type differences come into play after the pronoun is recognized and while the antecedent is being determined. By contrast, the anticipation account holds that the implicit causality of the verb operates quickly to increase the activation level of the stimulus noun that will then serve as the likely causer and antecedent of any subsequent pronoun. In this account, implicit verb causality affects the relative activation of participants as a result of the processing of the verb and its object, and not as a part of the anaphor resolution process.

In order to evaluate these two competing accounts, we used the S-E and E-S verbs from Experiment 1 in sentences that did not contain an anaphoric pronoun. As in the first experiment, we measured changes in the activation of the participants over the course of the sentence. If anticipation was leading to the activation of the stimulus noun in Experiment 1, we would expect implicit causality to affect reaction times in this experiment. However, if implicit causality only comes into play during anaphor resolution, there should be no effects of verb causality on participant activation when no pronouns are present.

Method

Subjects. Subjects were 112 fluent speakers of English, none of whom had participated in the first experiment.

Stimuli. The stimulus sentences were identical to those in Experiment 1, except the because-clause in each sentence was replaced by a prepositional phrase which did not contain a pronoun. Filler items were also changed to this format. Thus, all sentences were of the following form:

Name1 Verbed Name2 Adverb phrase Preposition Determiner Adjective Noun

e.g. Gary amazed Alan time after time at the juggling competition.

The sentences were formed by digitally splicing the prepositional phrase in after the main clause of Experiment 1 sentences. In addition to the marker already in place after the second name in the initial phrase, a second computerized marker was placed after the determiner in the prepositional phrase. Corresponding to the four time delays used in Experiment 1, time delays in Experiment 2 occurred 100 msec after second name (Delay 1), at the mark after the determiner (Delay 2), 200 msec downstream from the mark after the determiner (Delay 3), and at the end of the sentence (Delay 4).

Because the new sentences did not contain a pronoun, the question after each sentence could no longer query the interpretation of the pronoun. Thus, a new set of questions was devised. For each verb, one question asked for the subject of the sentence, (i.e., "Who VERBed?" E.g., *Who amazed?*), while the other question asked for the object of the sentence, (i.e., "Who was VERBed?" E.g., *Who was amazed?*)

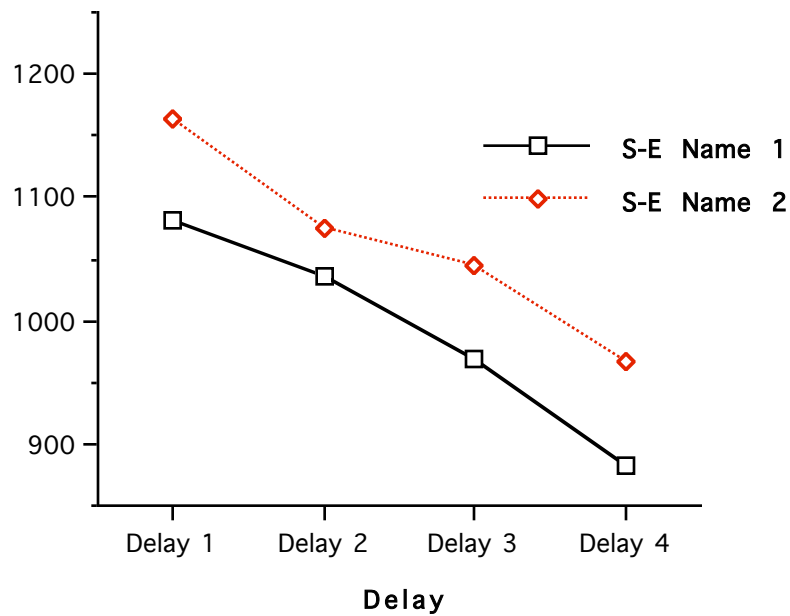
All other details of design and procedure were identical to Experiment 1. Two subjects with more than three errors or an average correct reaction time greater than 2 seconds on target sentences were replaced. Data trimming procedures resulted in replacing 3.2% of the probe reaction times in this experiment.

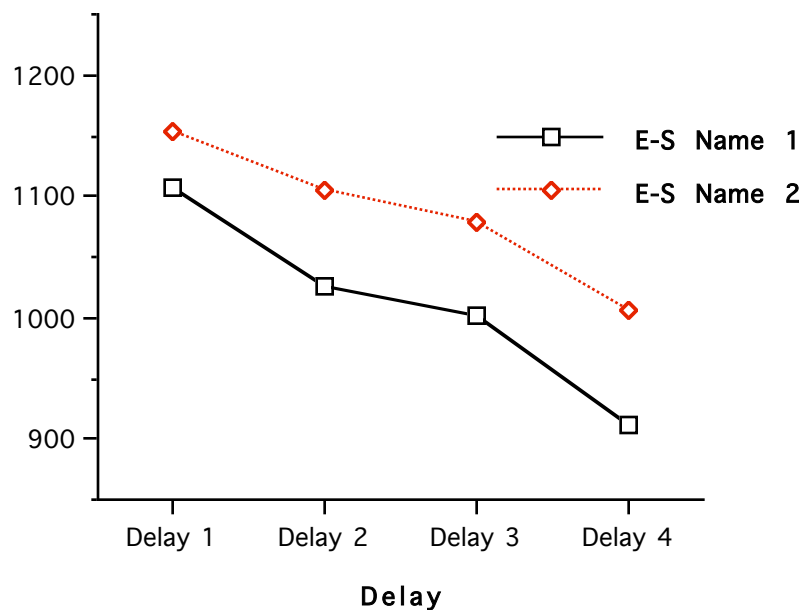
Results

Probe reaction time. As in the first experiment, the background effects of concurrent processing load and first mention advantage were significant. The significant effect of delay ($F(3,333)=27.3, p<.001$; $F(3,90)=27.9, p<.001$), showed decreasing reaction times as processing load decreased across the sentence (Delay 1: 1127 msec; Delay 2: 1060 msec; Delay 3: 1023 msec; Delay 4: 942 msec). The strong main effect of first mention ($F(1,111)=46.0, p<.001$; $F(1,30)=17.8, p<.001$), reflected the 72 msec advantage the first

name (1002 msec) had over the second (1074 msec). There was also a significant effect of verb type in the subjects analysis ($F(1,111)=4.1, p<.05$), with latencies to S-E verbs (1028 msec) slightly faster than to E-S verbs (1049 msec). However, this effect was not significant in the items analysis ($F(1,30)=2.8, p<.11$). No other effects or interactions were significant.

In order to compare latencies in this experiment to the significant three way interaction between verb type, probe and delay in Experiment 1, the corresponding cell means are graphed in Figure 3. Note that there is no evidence for the enhancement of the second participant relative to the first at any time delay for the E-S verbs.





Experiment 2: Interaction between verb type, probe and delay. a) S-E verbs; b) E-S verbs

Thus, when there is no pronominal anaphor present, a first participant advantage is seen for both verb types. There is no effect of verb type or gender on the relative activation of the two participants. The absence of an interaction with verb causality in this experiment suggests that the effect found in Experiment 1 is due to anaphor resolution processes, rather than to anticipatory effects from processing the implicit causality of verbs. Because this study included only sentences without because-clauses, and used different comprehension questions, it does not rule out the possibility that anticipatory effects were controlled in a strategic way in Experiment 1. It only provides evidence against the possibility that anticipatory effects are automatically generated whenever verbs with implicit causality are processed.

Probe judgment error rates. Probe recognition errors were low, averaging about 1.8%. A significant interaction occurred between verb type, gender, and probe ($F(1,111)=17.6$, $p<.001$; $F(1,30)=8.6$, $p<.01$). When there was not a gender contrast, more errors were

made on the participant not favored by verb semantics; when there was a gender contrast, more errors were made on the probe that matched verb semantics. The reasons for this odd interaction are not clear, and the low overall error rate suggests that we should not overinterpret these patterns. Note however, the results for the gender contrast condition go against that which would be predicted by the anticipatory account.

Comprehension question error rates. Average error rates to the comprehension questions were around 17.4%. Subjects made more errors to comprehension questions after S-E verbs (20%) than E-S verbs (15%) ($F(1,111)=15.3, p<.001$; $F(1,30)=5.1, p<.05$). Recall that since the questions in this experiment queried the subject half the time and the object half the time, this trend does not reflect a tendency for people to give the second name as the answer. Rather, it appears that people simply had more trouble understanding the sentences containing S-E than E-S verbs.

There was a three-way interaction between verb type, gender and probe ($F(1,111)=5.2, p<.05$; $F(1,30)=5.0, p<.05$). This interaction, shown in Table 3, is similar to the one found for probe judgment. This interaction is very different in character from the one found in Experiment 1 where error rates were greatly attenuated by gender information. In the current experiment gender did not help determine the correct answer to the comprehension question.

VERB TYPE

Stimulus-Experiencer verbs

		GENDER	
		same	different
PROBE	Name 1	18	22
	Name 2	20	19

Experiencer-Stimulus verbs

PROBE	Name 1	17	14
	Name 2	14	15

Table 3: Experiment 2, Error rates to comprehension questions; Verb type by gender by probe interaction

Summary. The activations of the two participants in the sentences in Experiment 2 were influenced by the same first two general factors mentioned in the summary of Experiment 1. Participants became more accessible under decreasing memory load, as at the end of the sentence, and the first participant was more accessible than the second. There were not additional effects of implicit verb causality or gender on the relative activation of the two participants.

Experiment 3

The results of the first two experiments indicate that verb information is used immediately in anaphor resolution processes and that the effect is not an anticipatory one, but one occurring during the resolution process. However, the kind of verb information tested, implicit causality, is not solely a function of the verbs themselves. Rather, the causality of these verbs emerges when they are used in combination with the conjunction “because” which joins the main and subordinate clauses. For example, Ehrlich (1980) found effects of implicit causality on anaphor resolution when the clauses were joined by “because,” weaker effects for “but,” and no effects for clauses joined by “and”. Thus, effects of implicit verb causality on participant activation during anaphor resolution may weaken or disappear when “because” does not conjoin the two clauses. By applying the more on-line methodology of the crossmodal probe recognition technique, we can extend the results of Ehrlich (1980) and further understand the time course of processing in sentences with implicit causality verbs that do not have because-clauses. To do this, we eliminated the conjunction “because” between the two clauses of Experiment 1, and presented each clause as an independent sentence (e.g., *Gary amazed Alan time after time. He was so talented*).

Elimination of the conjunction also necessarily involves breaking up one sentence into two sentences. There is abundant evidence that extra processing occurs at the ends of sentences, resulting in loss of verbatim information (Daneman & Carpenter, 1983; Jarvella, 1971; Just & Carpenter, 1980). Thus, the verbatim traces of potential antecedents may be weakened by the sentence boundary, making them less available to the anaphor resolution processes. Indeed, Daneman and Carpenter (1980) have shown that the more sentence boundaries that intervene between an antecedent and a pronoun, the more difficult correct assignment becomes, especially for people with small reading or listening spans. More intervening clause boundaries also cause anaphor resolution to take longer, as measured by sentence reading times (Clark & Sengul, 1979) and eye fixation data (Ehrlich and Rayner, 1983). In addition, sentence boundaries may cause information about the implicit causality verb in the main clause to be attenuated, which in turn may further diminish the effects of implicit causality on anaphor resolution. Indeed, Millis and Just (1994) found that verbs in first clauses were less accessible at the end of a second clause when the two clauses were in separate sentences than when they were connected by "because."

While nearly all studies of anaphor resolution with implicit verb causality have had the verb in the same sentence as the pronoun, two studies put the pronoun in a separate sentence from the implicit causality verb. One study used only S-E verbs and only same-gender participants, which made the first participant in the context sentence the favored antecedent for the pronoun in all target sentences (Hudson, Tanenhaus & Dell, 1986). This preference was borne out by end of sentence sensibility judgments. Reading times on the second half of the target sentences also showed longer latencies when the information contained in this portion of the sentence indicated that the first participant was not the referent of the pronoun. Hudson et al. (1986) interpreted this result to indicate that readers had immediately assigned the pronoun to the first participant. However, whether this is due to

the implicit causality of the verb, or a general first noun preference is not clear. In addition, the reading time measure which showed this difference was three words downstream from the pronoun, so it was not really an on-line measure of pronoun assignment.

The second study which put the pronoun in a separate sentence from the implicit causality verb used both S-E and E-S verbs, but always had a gender contrast between the participants (McKoon et al., 1993, Experiments 5 and 6). In this study, effects of implicit causality were only found for E-S verbs; no gender effects were found. Again, the measure was not on-line, since probe recognition latencies were taken only at sentence end. Thus, the immediate effect of verb causality across sentence boundaries in anaphor resolution is not clear. The following experiment can help clarify the results of the above two studies by studying both S-E and E-S verbs with and without gender contrast within an on-line crossmodal probe recognition framework.

Method

Subjects. Subjects were 96 fluent speakers of English, none of whom had participated in the other experiments.

Stimuli. The stimuli sentences were identical to those in Experiment 1, except that the first clause was given sentential status, the word "because" was eliminated, and the anaphoric pronoun was the first word of a new sentence. Thus, all stimuli consisted of two sentences of the form:

Name1 Verbed Name2 Adverb phrase. Pronoun was so Adjective

e.g. Gary amazed Alan time after time. He was so talented

The recording of the first clause used in Experiment 1 was used as the first sentence in these materials. However, because of the fluid nature of speech, it was impossible to delete the "because" in the recordings from Experiment 1 and have natural sounding second sentences. Therefore, the second sentences were recorded anew by the same native speaker as in Experiment 1. The stimuli were then formed by splicing the newly recorded second sentence in after old recordings of the first sentence. Filler stimuli were also altered to fit this pattern.

In order to have time delays equivalent to those in Experiment 1, the same marking point was used in the first sentence, and a mark was placed into the newly recorded second sentences immediately after the pronoun. Thus, the four time delays were as in Experiment 1: Delay 1 was 100 msec after the second name in the first sentence; Delay 2 was immediately after the pronoun in the second sentence; Delay 3 was 200 msec downstream from the pronoun, and Delay 4 was at the end of the second sentence.

All other details of design and procedure were identical to Experiment 1. Four subjects with more than 3 replaced or trimmed latencies, or an average correct reaction time greater than 2 seconds on target sentences were replaced. Data trimming resulted in replacing 2.8% of the probe reaction time data in this experiment.

Results

Probe reaction time. As in the previous experiments, there were effects due to concurrent processing load and the first mention advantage. Concurrent processing load was evident in the main effect of delay ($F(3,285)=14.8, p<.001$; $F(3,90)=13.1, p<.001$), with faster reaction times occurring at later delays, particularly at the end of the sentence (Delay 1: 1135 msec; Delay 2: 1088 msec; Delay 3: 1106 msec; Delay 4: 984 msec). The significant

effect of probe ($F(1,95)=38.1, p<.001$; $F(1,30)=21.4, p<.001$), reflected a first mention advantage of 88 msec (first name: 1034 msec; second name: 1122 msec). There was also a significant main effect of gender ($F(1,95)=17.7, p<.001$; $F(1,30)=16.4, p<.001$), with faster latencies occurring with different-gender (1047 msec) than same-gender participants (1110 msec). These main effects were all qualified by interactions.

There was a two-way interaction between verb type and gender ($F(1,95)=8.7, p<.005$; $F(1,30)=4.1, p=.05$). The effect of gender was weaker for S-E verbs (different: 1072 msec vs. same: 1103 msec) than for E-S verbs (different: 1022 msec vs. same: 1116 msec). Recall that, in Experiment 1, weak main effects were found for both verb type and gender such that responses were faster to sentences with E-S verbs and to sentences with a gender contrast. In the current experiment these two factors interacted producing the fastest times for E-S verbs with a gender contrast. A second interaction, significant in the subjects analysis and marginal in the items, occurred between verb type and probe ($F(1,95)=5.4, p<.05$; $F(1,30)=3.0, p<.10$), with a stronger first participant advantage for S-E verbs (name 1: 1027 msec vs. name 2: 1148 msec) than for E-S verbs (name 1: 1041 msec vs. name 2: 1096 msec). The fact that there was not a further interaction with delay, as was the case in Experiment 1, indicates that this effect of verb semantics on participant availability generally held across all time delays. This is shown in Figure 4, where the verb type by probe interaction is further broken down by delay. For S-E verbs (Figure 4a), there is a reaction time advantage for the first name at all time delays; it is especially strong at delays 3 and 4. For E-S verbs (Figure 4b), latencies to first and second participants are nearly equal at delays 1 and 2, while the first participant advantage starts to assert itself at later delays. Thus, at every delay, the differential between the first and second name is larger for the S-E verbs than E-S verbs. Experiment 2 rules out the explanation of this difference in relative activation as strictly priming from verb semantics apart from pronoun resolution. The fact that this effect also held at the first time delay, before the pronoun was supposedly heard,

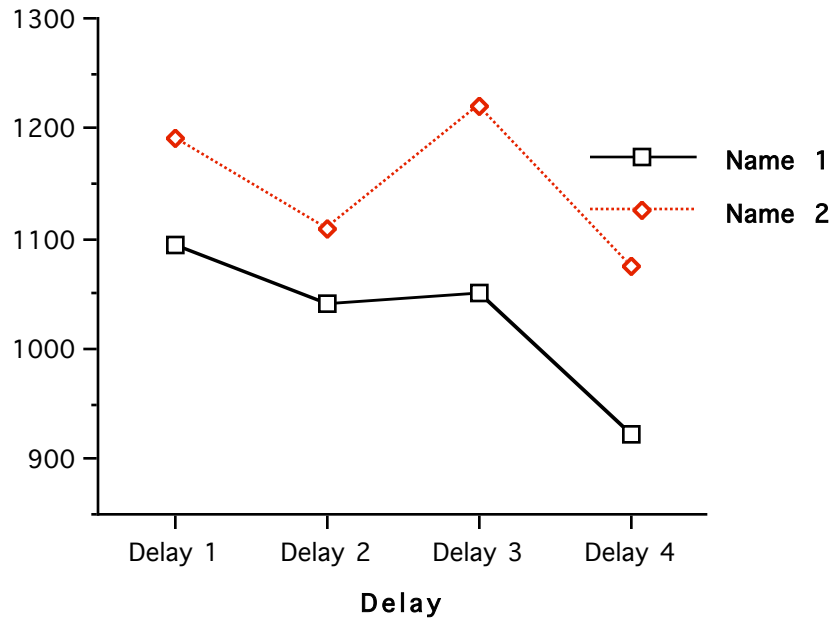
could be due to strategic use of verb semantic information for the pronoun they knew was upcoming. However, it is also possible that subjects actually heard the pronoun before they completed their response at the first time delay. The elimination of the word “because” between the first clause and the pronoun caused the pronoun to occur earlier relative to this first time delay.

A two-way interaction between probe and delay ($F(3,285)=3.2, p<.05$; $F(3,90)=1.8, n.s.$) was qualified by a three-way interaction between gender, probe, and delay ($F(3,285)=2.8, p<.05$; $F(3,90)=2.9, p<.05$), which is shown in Figure 5. Examination of the sentences with same-gender participants shows that, whereas the first mentioned advantage is present at delays 1, 3, and 4, this tendency is not present at delay 2 (see Figure 5a). Thus, at the pronoun, both names that agree in gender with this pronoun are equally activated. This loss of the first mention advantage at the pronoun indicates that listeners are trying to pay equal attention to both gender-matching names as possible candidates for coreference. Examination of the sentences with different-gender participants shows a slight first mentioned advantage at delay 1, with increasing advantage at the later delays (see Figure 5b). For these gender contrast sentences there is no evidence for the promotion of the name agreeing in gender with the pronoun.

An examination of individual cell means shows how the equal availability of names occurred for the gender same condition for the S-E and E-S verbs. For S-E verbs the accessibility of the second name is strongly increased at the pronoun (delay 1: 1222 msec; delay 2: 1077 msec); a short time later it is suppressed to re-establish the first participant advantage (delay 3: 1273 msec). For E-S verbs the accessibility of the first name is decreased at the pronoun (delay 1: 1156 msec; delay 2: 1212 msec); it is then later enhanced to re-establish the first participant advantage (delay 3: 1058 msec). Thus, in both cases, the

first participant advantage is erased at the pronoun by changing the availability of the experiencer.

STIMULUS-EXPERIENCER VERBS



EXPERIENCER-STIMULUS VERBS

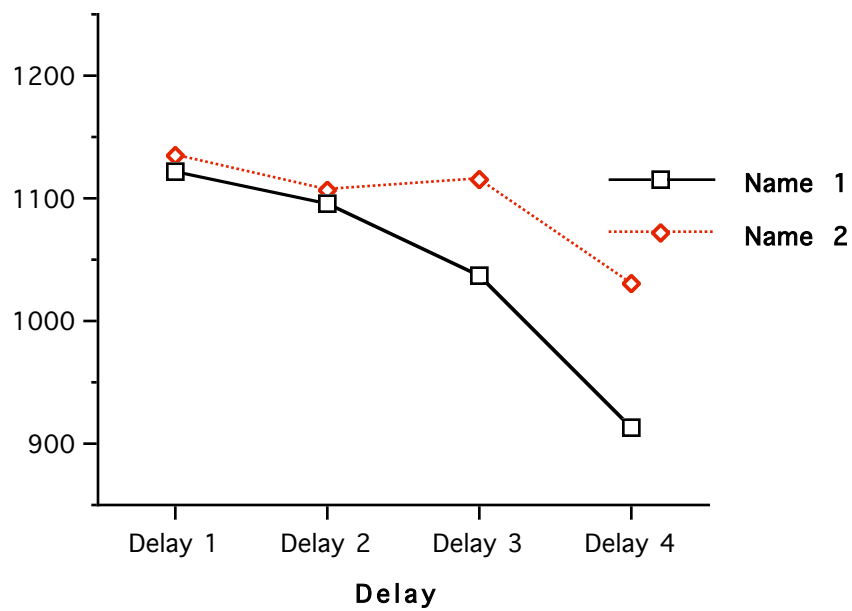


Figure 4: Experiment 3: Interaction between verb type and probe, further broken down by the different delays a) S-E verbs; b) E-S verbs

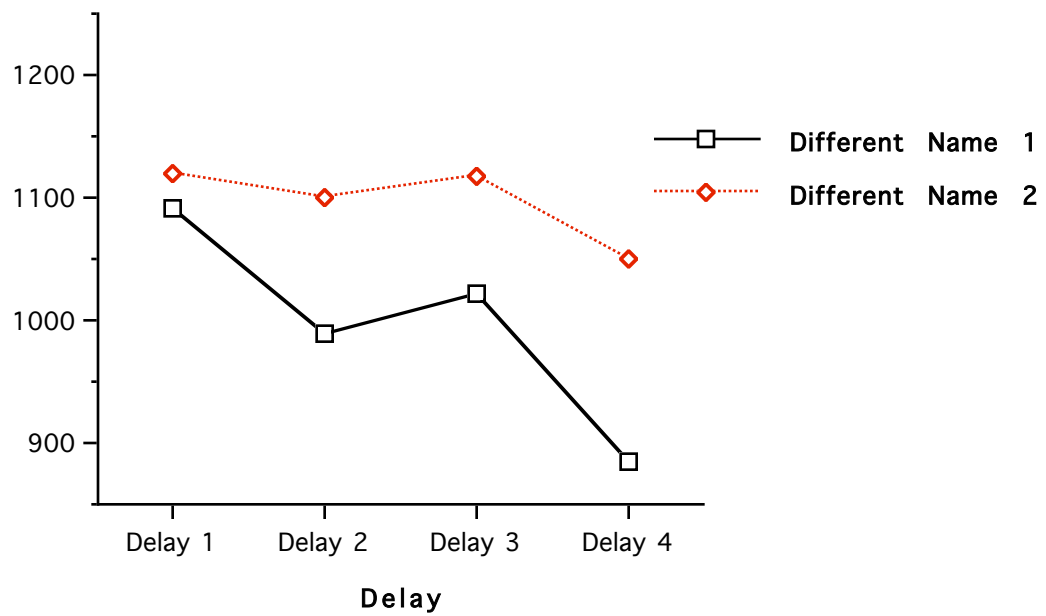
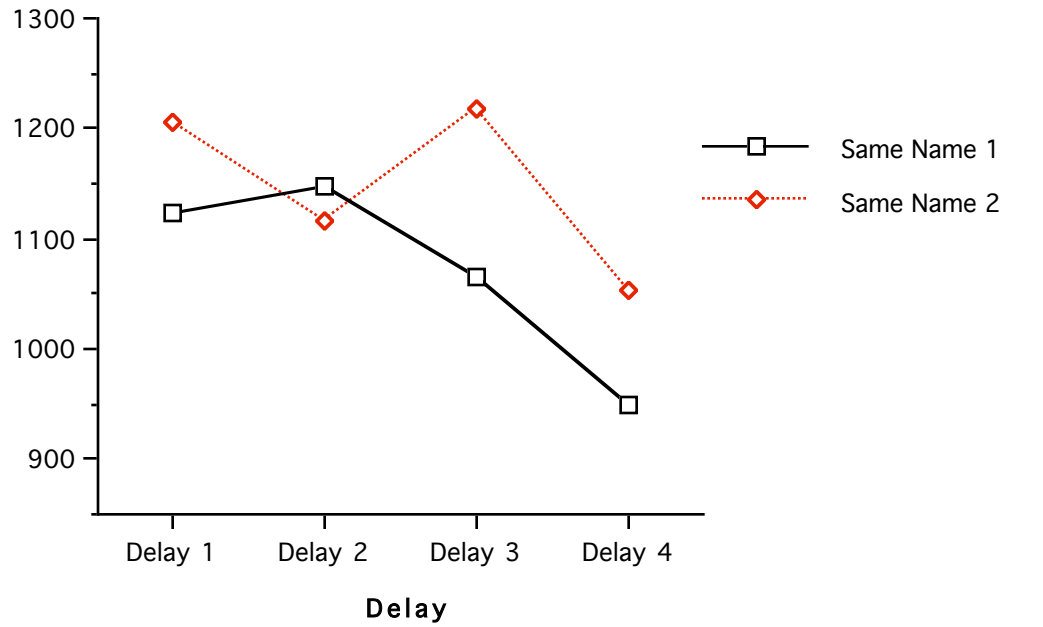


Figure 5. Experiment 3: Interaction of gender, probe and delay a) same gender; b)
different gender

The results of Experiment 3 refine those of Hudson et al. (1986) and McKoon et al. (1993). These two studies, which had the pronoun in a separate sentence from the implicit causality verb found evidence for the effect of verb causality on anaphor resolution for S-E verbs (Hudson et al., 1986) and E-S verbs (McKoon et al., 1993) using off-line measures. This study finds evidence for use of verb causality with an on-line measure. In addition, while neither of these studies found evidence for gender effects, the current study found that when both participants matched the pronoun in gender, both were equally available. When only one participant matched the pronoun in gender, the first participant, not necessarily the gender-matching participant, was more available.

Probe judgment error rate. Error rates were low on the probe judgment task, averaging about 1.3%. A significant interaction occurred between gender, probe and delay in the subjects analysis ($F(3,285)=3.1, p<.05$; $F(3,90)=2.6, p<.06$). There were generally lower error rates to the first than second participant, although this tendency was sometimes reversed for different delays across the gender conditions. Again, because of the low error rate, we do not wish to overly interpret this interaction.

Comprehension question error rate. Error rates to the comprehension questions were again rather high, averaging around 15%. Verb type was significant in the subjects analysis and marginal in the items ($F(1,95)=7.5, p<.01$; $F(1,30)=3.6, p<.07$), with a slightly higher error rate to S-E verbs (17%) than to E-S verbs (14%). Gender was significant ($F(1,95)=73.3, p<.001$; $F(1,30)=98.6, p<.001$), with more errors when there was not a gender contrast (22%) than when such a contrast was present (8%).

There were three two-way interactions with verb type. Verb type interacted with gender ($F(1,95)=12.1, p<.001$; $F(1,30)=9.3, p<.005$), with the highest error rates occurring for same-gender S-E verbs, indicating that, without gender information to help determine the referent, people had some tendency to give the second noun as the referent in sentences with same-gender nouns and S-E verbs. Verb type interacted with probe ($F(1,95)=25.9, p<.001$; $F(1,30)=30.6, p<.001$) indicating that subjects made fewer errors when the probe was the correct answer -- this was name 1 for S-E verbs and name 2 for E-S verbs. Finally, verb type also interacted with delay ($F(3,285)=2.7, p<.05$; $F(3,90)=2.9, p<.05$). Error rates were higher for S-E verbs than E-S verbs at delay 1; this tendency diminished over time, and finally reversed at delay 4. Thus, the earlier probe interruptions wreaked more havoc for S-E than E-S verbs, indicating the greater difficulty people had with the former verb type.

Summary. In this experiment we again saw the general effects on participant activation of processing load and the first participant advantage. In addition, there was a general effect of implicit verb causality across all time delays such that there is a larger first participant advantage for S-E verbs than for E-S verbs. The effect of implicit causality effectively cancelled the first mention advantage at the first two time delays for E-S verbs, at later delays it just dampened this advantage relative to the S-E verbs. Again, as in experiment 1, implicit causality information served at best to cancel the first mention advantage; it never was strong enough to make the actual referent of E-S verbs (the second participant) more accessible than the first.

While effects of implicit verb causality were found both here and in Experiment 1, it is important to note the difference in time course between the two. In Experiment 1 implicit causality specifically affected latencies at the pronoun and at the end of the sentence, where it cancelled out the first mention advantage. In the current experiment verb causality affected participant activation across all time delays, and was strong enough to cancel out the first

mention advantage around the time the pronoun was encountered. (As argued above, the effect probably occurred at delay 1 as well as delay 2 due to the pronoun being encountered before delay 1 responses were complete.) Thus, the elimination of "because" and the introduction of a sentence boundary in Experiment 3 did not eliminate the effect of implicit causality information, it just changed its pattern across time. In Experiment 1 it was focused at the pronoun, and then suspended until end of sentence wrap-up. In Experiment 3, it had an on-going effect, but did not show the sharp end of sentence effect. Thus, the lack of "because" or the introduction of a sentence boundary in Experiment 3 allowed the use of implicit causality information to continue across the sentence, rather than suspending it to the end.

Relative activation levels for candidate antecedents was also affected by gender in Experiment 3. If both candidates agreed in gender with the pronoun, the accessibility of the experiencer was changed so that both candidates were equally available, overcoming the general first participant advantage. This effect occurred at the pronoun, and was short-lived. There was no evidence for use of gender contrast in anaphor resolution.

Discussion

The activation of candidate referents is determined by many factors. Some factors, such as processing load affect both candidates equally. For example, in each of our three studies, latency to probe recognition decreased systematically across the four time delays. As listeners hear more of a sentence, there are fewer remaining unresolved relations and they can respond more quickly to probes. Other factors, such as the first participant advantage, cause one candidate (in this case the first) to be more accessible than another. The first participant advantage is a general accessibility advantage and occurs whether or not anaphor resolution processes are involved. It is pervasive and is not localized to a particular point in sentence processing. When the first participant advantage does occur in the context of

anaphor resolution it may help promote the first name as the default referent. Across sentence types and across discourse structures, the first noun of the sentence is always the most likely candidate referent for the pronoun. Betting on the first noun as the referent is a relatively safe bet, since promotion of the first noun involves a very uniform and quick bias which is often right and which can always be reversed when needed.

Other factors are more specifically involved in anaphor resolution processes. We investigated two of these here--implicit verb causality and gender. As for the first factor, the results of these three experiments show that anaphor resolution processes can make immediate use of implicit verb information--both Experiments 1 and 3 showed the cancellation of the first mention effect for E-S verbs at the pronoun. However when a pronoun is not present, verb causality does not automatically promote accessibility of the causer (Experiment 2). Because the referent of the pronoun was always queried by the comprehension question following each sentence in Experiments 1 and 3, it is possible that subjects elected to devote extra attentional resources to pronominal processing. Therefore, it is not certain from these experiments whether implicit verb causality information is used automatically in anaphor resolution, or is a function of strategies under subject control. However, evidence from other studies where such strategies were made less likely by variations in the materials still showed effects of verb causality (Garnham et al., 1992), indicating that such information may be used automatically. There were also differences in the time course of the use of implicit causality information after the pronoun occurred. In Experiment 1 the influence was suspended a bit downstream from the pronoun and returned at sentence end; in Experiment 3 the effects occurred across all time delays. These differences in time course must be due to differences in sentence structure between the two experiments.

We also found effects of the second factor, gender, on anaphor resolution processes. In Experiment 1, gender contrast information increased the availability of the gender-matched participant. However, this effect was only significant in the subjects analyses and not in the item analyses. This effect did not interact significantly with delay, showing no localized effects of gender. However, an examination of the cell means did show some evidence for an immediate effect of gender at the pronoun with gender-mismatched participants being suppressed at the pronoun. The lack of significance of this trend fits in with previous research that has failed to find an immediate effect of gender on anaphor resolution. One possible explanation for the lack of significant gender effects across many studies is the possibility that such effects are under strategic control, and depending on the subjects or the experimental task demands, the subject may or may not make immediate use of gender information (Garnham et al., 1992).

We did find immediate effects of gender at the pronoun in Experiment 3; however the nature of this effect was not that gender mismatched participants were suppressed. Rather, in same-gender sentences, both nouns were equally activated at the pronoun, thereby wiping out the first participant advantage. That is, the effect of gender information found in Experiment 3 was focused on the equal availability of candidate referents at the pronoun in the same-gender condition, rather than exclusive promotion of the sole gender agreeing referent in the different-gender condition. Thus, the results of Experiment 3 do not necessarily contradict the findings of other studies which have failed to show the immediate promotion of the gender agreeing referent in different-gender conditions (Gernsbacher, 1989; MacDonald & MacWhinney, 1990; Tyler & Marslen-Wilson, 1982); however, these results do establish that gender agreement can eliminate the first participant advantage bias in same-gender conditions. The fact that this gender effect is only found in Experiment 3 and not Experiment 1 indicates that the use of gender information may be affected by sentence structure as well as subjects' strategies.

The time course of information use in anaphor resolution

Given the above results and those in the previous literature, what can we conclude about the time course of anaphor resolution? First, it appears that the degree to which anaphor resolution processes are triggered immediately at the pronoun is subject to task demands and sentence structure (Greene et al., 1992; McKoon et al., 1993). There are cases when anaphor resolution processing can be demonstrated immediately at the pronoun (Experiments 1 and 3). However, even when resolution processes begin immediately at the pronoun, full resolution may not be completed until some time later (Ehrlich & Rayner, 1983; Sanford & Garrod, 1989; Stevenson, 1986). This is because resolution processes are data-sensitive--often needing to wait for relevant input from the results of other sentence comprehension processes before they can be completed. Since the output of these other comprehension processes may be affected by such variables as task demands or sentence structure, these effects can, in turn, affect the completion of anaphor resolution. Let us consider how this analysis helps to explain the results of both the current study and earlier research.

Implicit verb causality. Verb causality information is inherent in the semantics of the verb itself. In experiments with single sentences or pairs of sentences, no larger text inferencing processes are necessary to get at the semantics of the verb. Thus, when anaphor resolution processes are triggered by the presence of a pronoun after a implicit causality verb, this information should be quickly available and used to promote the relative availability of the participant consonant with the causality. This is exactly what we found. In both Experiments 1 and 3 implicit causality information was able to cancel out the first mention advantage right at the pronoun. The subsequent time course of the use of this information was influenced by sentence structure.

Other types of semantic information. If the type of semantic information needed to help resolve an anaphor is not readily available at the pronoun, the impact of this information will not be immediately detectable. In particular, if the information involves inferences over the text, anaphor resolution will have to await the output of these inferential processes before it can use this information. For example, Gernsbacher (1989) used stimuli of the following type to provide unambiguous semantic information about the referent of the pronoun before the anaphor occurred.

Bill lost the tennis match to John.

(1) Accepting the defeat, he walked slowly toward the showers.

(2) Enjoying the victory, he walked slowly toward the showers.

In completion 1, “he” obviously refers to “Bill,” while in completion 2, “he” refers to “John”. However, no effect of this semantic information was found on the activation of the two names immediately after the pronoun (Gernsbacher, 1989), indicating this type of information was not immediately used to assign an antecedent. Note, however, that the type of inference necessary to use the information at the beginning of either of these completions involves reference back to the previous sentence. It seems plausible that such inference processes would take an appreciable amount of time, and might not be completed by the time the pronoun is encountered. Information from such inference processes would become available somewhere after the pronoun, and its influence seen there (Gernsbacher, 1989; Tyler & Marslen-Wilson, 1982).

Gender information. In order to use gender information in anaphor resolution, the gender of each potential antecedent must be readily available. Gender is a semantic feature of most proper names and thus, should be quickly available. However, unlike verb causality information, gender information is often unable, by itself, to identify a unique referent, since the various competing referents may share the same gender. There is no easy way to know

in advance whether an attempt to match on the basis of gender agreement alone will be successful. Depending on the task and materials, subjects may choose not to make immediate use of gender information (Garnham et al., 1992). Instead of trying to anticipate the gender of a pronoun, it may be more profitable to resolve pronominal anaphors by using gender along with other cues as they become available. This delayed use of gender information is in accord with what has typically been reported (Gernsbacher, 1989; MacDonald & MacWhinney, 1990; Tyler & Marslen-Wilson, 1982). This is not to say that gender could not be used immediately if this is a beneficial strategy for certain situations (Garnham et al., 1992), just that gender cues do not automatically induce anaphor resolution processes.

Sentence structure effects. Changes in sentence structure can cause changes in how strong information availability will be, and the timing of this availability. For example, the presence of a causal conjunction such as "because" may cause implicit causality information to be particularly important to end of sentence integration processes; the absence of this conjunction may diffuse this effect. Sentence boundaries can cause some information to become less available (e.g., verbatim information), while other information becomes more available (e.g., higher level abstractions). Thus, the occurrence of a sentence boundary may hurt some aspects of the data-sensitive anaphor resolution process while aiding others. For example, in the current work, sentence boundaries were found to effect the use of gender information. In the gender same condition both candidates were equally available at the pronoun when the pronoun occurred after a sentence boundary (Experiment 3); this was not the case when only a clause boundary intervened (Experiment 1). Thus, the abstraction of gender information may be promoted by sentence wrap-up, and may therefore be subsequently more available to any anaphor resolution processes. This does not mean that gender information would have to be used after a sentence boundary--we found no

promotion of the gender matched candidate in the gender different condition after a sentence boundary.

Interestingly, two other studies which indicated immediate resolution of pronominal anaphor both had the pronoun occurring at the end of a sentence (Leiman, 1982; Cloitre & Bever, 1988). Although these studies used a different methodology than the current experiment (i.e., they used lexical decision to semantically related words rather than a probe recognition task, and in addition did not systematically manipulate gender), the results nonetheless suggest that some types of information potentially useful to anaphor resolution are more readily available after a sentence boundary.

Syntax. Information from the syntactic constraints in a sentence can also be used in anaphor resolution. For example, Nicol and Swinney (1989) report an experiment in which syntactic binding constraints made only one of several possibilities a legal antecedent. They found immediate activation at the pronoun only for the antecedent which fit the syntactic constraint. Thus, this type of structural information seems to be immediately available to anaphor resolution processes.

In summary, the results of our experiments have led us to adopt a model of anaphor resolution in which resolution can be initiated immediately upon encountering the pronoun. At all points, the baseline candidate for resolution is the first noun of the sentence, but this baseline is easily cancelled when additional information becomes available. Depending on the type of information and the sentence structure involved, a full resolution of the anaphor may sometimes have to wait on the output of other sentential and strategic processes. Thus, some studies may find that anaphor resolution begins immediately at the pronoun, while others may find no effects until some distance downstream from the pronoun because these

studies are measuring the effect of different types of information, or using different types of sentence structure.

In general, when the information is strong and readily available at the pronoun, resolution processes should begin immediately. Such readily available information includes implicit verb causality and syntactic binding constraints. It may also include some aspects of gender information across sentence boundaries. If the information is weak, or not available until another sentential process is complete, the onset of anaphor resolution should be delayed. Such information may include semantic information involving inferences across the text, and gender information within a clause or sentence. Note that the same type of information which is readily available in one type of sentence structure may not be available in another. The wrap-up processes that occur at clause and sentence boundaries may weaken the effects of verbatim information, but may promote the abstraction of other types of information. Thus, not only the type of information, but also the structure of the passage will determine the time course for the resolution of the anaphor.

Title Note

We thank Lenore Frigo, Young-Ran Lee, Bruce Tran, Nita Wadhvani, George Warren and Katherine Williams for help in running the subjects, Bruce Chapman, Watee Klongnugrob and Won Chan Jung for writing the data collection program, and Debbie Shaibe for general help. Thanks also go to several anonymous reviewers, and Greg Murphy and Ed Shoben for editorial advice. The preliminary results of the first experiment were reported at Psychonomics, 1990. Requests for reprints should be directed to Janet McDonald, Psychology Department, Louisiana State University, Baton Rouge, LA 70803.

References

- Au, T. K. (1986). A verb is worth a thousand words: The causes and consequences of interpersonal events implicit in language. Journal of Memory and Language, *25*, 104-122.
- Brown, R. & Fish, D. (1983). The psychological causality implicit in language. Cognition, *14*, 237-273.
- Caramazza, A., Grober, E., Garvey, C., & Yates, J. (1977). Comprehension of anaphoric pronouns. Journal of Verbal Learning and Verbal Behavior, *16*, 601-609.
- Caramazza, A. & Gupta, S. (1979). The roles of topicalization, parallel function and verb semantics in the interpretation of pronouns. Linguistics, *3*, 497-518.
- Chafe, W. L. (1972). Discourse structure and human knowledge. In R. O. Freedle & J. B. Carroll (Eds.), Language Comprehension and the Acquisition of Knowledge. Washington, D. C.: V. H. Winston & Sons.
- Chang, F. R. (1980). Active memory processes in visual sentence comprehension: Clause effects and pronominal reference. Memory & Cognition, *8*, 58-64.

- Clark H. H. & Sengul, C. J. (1979). In search of referents for nouns and pronouns. Memory & Cognition, 7, 35-41.
- Cloitre, M. & Bever, T. G. (1988). Linguistic anaphors, levels of representation, and discourse. Language and Cognitive Processes, 3, 293-322.
- Corbett, A. T. & Chang, F. (1983). Pronoun disambiguation: Accessing potential antecedents. Memory & Cognition, 11, 283-294.
- Daneman, M. & Carpenter, P. A. (1980). Individual differences in working memory and reading. Journal of Verbal Learning & Verbal Behavior, 19, 450-466.
- Daneman, M. & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. Journal of Experimental Psychology: Learning, Memory & Cognition, 9, 561-584.
- Dell, G. S., McKoon, G. & Ratcliff, R. (1983). The activation of antecedent information during the processing of anaphoric reference in reading. Journal of Verbal Learning and Verbal Behavior, 22, 121-132.
- Ehrlich, K. (1980). Comprehension of pronouns. Quarterly Journal of Experimental Psychology, 32, 247-255.
- Ehrlich, K. & Rayner, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. Journal of Verbal Learning and Verbal Behavior, 22, 75-87.
- Garnham, A. & Oakhill, J. (1985). On-line resolution of anaphoric pronouns: Effects of inference making and verb semantics. British Journal of Psychology, 76, 385-393.
- Garnham, A., Oakhill, J. & Cruttenden, H. (1992). The role of implicit causality and gender cue in the interpretation of pronouns. Language and Cognitive Processes, 7, 231-255.
- Garvey, C. & Caramazza, A. (1974). Implicit causality in verbs. Linguistic Theory, 5, 459-464.
- Gernsbacher, M. A. (1989). Mechanisms that improve referential access. Cognition, 32, 99-156.

- Gernsbacher, M. A. (1990). Language Comprehension as Structure Building. Hillsdale, NJ: Lawrence Erlbaum.
- Gernsbacher, M. A. & Hargreaves, D. J. (1988). Accessing sentence participants: The advantage of first mention. Journal of Memory & Language, 27, 699-717.
- Gernsbacher, M. A., Hargreaves, D. J. & Beeman, M. (1989). Building and accessing clausal representations: The advantage of first mention versus the advantage of clause recency. Journal of Memory & Language, 28, 735-755.
- Greene, S. B., McKoon, G. & Ratcliff, R. (1992). Pronoun resolution and discourse models. Journal of Experimental Psychology: Learning, Memory & Cognition, 18, 266-283.
- Hudson, S. B., Tanenhaus, M. K. & Dell, G. S. (1986). The effect of the discourse center on the local coherence of a discourse. In the Program of the Eighth Annual Conference of the Cognitive Science Society. Hillsdale, NJ: Lawrence Erlbaum.
- Jarvella, R. (1971). Syntactic processing of connected speech. Journal of Verbal Learning & Verbal Behavior, 10, 409-416.
- Just, M. A. & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. Psychological Review, 87, 329-354.
- Leiman, J. M. (1982). A chronometric analysis of referent assignment to pronouns. Unpublished Doctoral Dissertation, Wayne State University.
- MacDonald, M. C. (1986). Priming during sentence processing: Facilitation of responses to a noun from a coreferential pronoun. Unpublished doctoral dissertation, University of California, Los Angeles.
- MacDonald, M. C. & MacWhinney, B. (1990). Measuring inhibition and facilitation from pronouns. Journal of Memory & Language, 29, 469-492.
- MacWhinney, B. (1977). Starting points. Language, 53, 152-168.

- MacWhinney, B. (1985). Grammatical devices for sharing points. In R. Schiefelbusch (Ed.), Communicative competence: Acquisition and intervention. Baltimore: University Park Press.
- MacWhinney, B., & Pléh, C. (1988). The processing of restrictive relative clauses in Hungarian. Cognition, 29, 95-141.
- McKoon, G., Greene, S. B., & Ratcliff, R. (1993). Discourse models, pronoun resolution, and the implicit causality of verbs. Journal of Experimental Psychology: Learning, Memory & Cognition, 19, 1040-1052.
- Millis, K. K. & Just, M. A. (1994). The influence of connectives on sentence comprehension. Journal of Memory & Language, 33, 128-147.
- Nicol, J. & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. Journal of Psycholinguistic Research, 18, 5-19.
- Sanford, A. J. & Garrod, S. C. (1989). What, when and how?: Questions of immediacy in anaphoric reference resolution. Language and Cognitive Processes, 4, 235-262.
- Smyth, R. H. (1986). Cognitive aspects of anaphora judgment and resolution. Bloomington, IN: Indiana University Linguistics Club.
- Stevenson, R. J. (1986). The time course of pronoun comprehension. In the Program of the Eighth Annual Conference of the Cognitive Science Society. Hillsdale, NJ: Lawrence Erlbaum.
- Stevenson, R. J. & Vitkovitch, M. (1986). The comprehension of anaphoric relations. Language and Speech, 29, 335-360.
- Tyler, L. K. & Marslen-Wilson, W. (1982). The resolution of discourse anaphors: Some on-line studies. Text, 2, 263-291.
- Vonk, W. (1984). Pronoun comprehension. In A. G. Gale & F. Johnson (Eds.), Theoretical and Applied Aspects of Eye Movement Research. Amsterdam: North-Holland.

Vonk, W. (1985a). The immediacy of inferences in the understanding of pronouns. In G. Rickheit & H. Strohner (Eds.), Inferences in Text Processing. Amsterdam: North-Holland.

Vonk, W. (1985b). On the purpose of reading and the immediacy of processing pronouns. In R. Groner, G. McConkie & C. Menz (Eds.), Eye Movements and Human Information Processing. Amsterdam: North-Holland.

Appendix

Experiment 1 conjoined the main clause to the phrase containing the pronoun with "because." Experiment 2 used the prepositional phrase rather than the "because" clause as a completion. Experiment 3 eliminated the "because," and the phrase containing the pronoun was presented as a separate sentence.

Stimulus-Experiencer verbs

Gary amazed Alan/Ellen time after time (because) he was so talented at the juggling competition

Liz amazed Rose/Rob this morning (because) she was so cheerful during the coffee break

Cindy amused Laura/Larry around the clock (because) she was so entertaining during the power outage

Carl amused Lyle/Lynne frequently (because) he was so humorous during the comic presentation

Peter annoyed Dennis/Debbie occasionally (because) he was so talkative during the opera broadcast

Nan annoyed Bess/Ben yesterday (because) she was so pushy during the evening news

Linda bored Mary/Martin to tears (because) she was so repetitious during the discussion group

Sam bored George/Joyce beyond belief (because) he was so monotonous during the lengthy presentation

Tim deceived Jake/Jane repeatedly (because) he was so underhanded during the police investigation

Sharon deceived Michelle/Michael on purpose (because) she was so cunning at the murder scene

Beth disappointed Pam/Paul bitterly (because) she was so hard-hearted at the anniversary party

Matthew disappointed Henry/Helen last week (because) he was so unmotivated at the board meeting

Ted frightened Mark/Marge to death (because) he was so violent at the Halloween party

Sheila frightened Bonnie/Barry at times (because) she was so unpredictable during the gruesome story

Gail inspired Joan/Jim tremendously (because) she was so enthusiastic at the pep assembly

Brian inspired Richard /Rita last night (because) he was so optimistic at the religious revival

Experiencer-Stimulus verbs

Karen/Kevin admired Becky openly (because) she was so intelligent at the tennis tournament

Ross/Ruth admired Craig quite a bit (because) he was so courageous during the debate championship

Andy/Amy detested Walter to the extreme (because) he was so obnoxious after the cheating scandal

Kate/Keith detested Peg intensely (because) she was so condescending during the trial testimony

Julie/Jerry envied Kathy sometimes (because) she was so artistic after the well-deserved promotion

Dan/Dawn envied Frank in secret (because) he was so good-looking during the awards banquet

Charlie/Shirley feared David all the time (because) he was so aggressive during the assault trial

Jan/John feared Val extremely (because) she was so vengeful after the violent attack

Barb/Bob hated Ann a great deal (because) she was so manipulative after the damaging betrayal

Daniel/Donna hated Jason with a passion (because) he was so bossy after the bitter falling out.

Jack/Joy noticed Doug immediately (because) he was so overdressed at the Japanese restaurant

Betty/Billy noticed Lisa at the party (because) she was so striking at the governor's mansion

Jill/Jeff pitied Claire much of the time (because) she was so unsuccessful during the custody case

Steven/Stella pitied Harry deeply (because) he was so unfortunate during the funeral service

Ken/Kay trusted Tom completely (because) he was so sincere during the hypnosis session

Kathryn/Calvin trusted Brenda at once (because) she was so forthright at the psychotherapy workshop