Journal of Memory and Language **42,** 526–544 (2000)
doi:10.1006/jmla.1999.2694, available online at http://www.idealibrary.com on IDEAL®

The Contribution of Lexical and Situational Knowledge to Resolving Discourse Roles: Bonding and Resolution

Simon Garrod and Melody Terras

Human Communication Research Centre and Department of Psychology, University of Glasgow, Glasgow, Scotland

Resolving links between subsequent referents (e.g., *the car*) and open discourse roles (as in *Keith drove to London yesterday. The car kept overheating*) is crucial for discourse understanding. This article investigates the contribution of lexical semantic factors (e.g., that *drive* implies using a *vehicle*) as compared to more general contextual factors in the on-line resolution of such links. We report an eye-tracking experiment that measures immediate and delayed effects of both kinds of information as readers resolve the reference. The results indicate that lexical information dominates the initial linking process with more general contextual influences emerging later. They are discussed in terms of the distinction between early *bonding* and subsequent *resolution* processes that has been proposed for other kinds of anaphoric interpretation (Sanford, Garrod, Lucas, & Henderson, 1983).

Key Words: discourse; anaphora; verbs; roles.

Understanding connected discourse, as opposed to just understanding the individual sentences it contains, depends upon establishing appropriate links. These links can be of various kinds: anaphoric, when an expression in one sentence points back to some previous expression to which it corefers (Garrod & Sanford, 1977); causal, where an event described in one sentence is taken as the cause or reason for the event described in another (Trabasso & Sperry, 1985; van den Broek, 1988); or in terms of what have sometimes been called discourse roles (Garrod & Sanford, 1981, 1990; Carlson & Tanenhaus, 1988; Tanenhaus & Carlson, 1989). Role-based links occur when something that is mentioned in one sentence is understood to play a particular role in an event that is mentioned in another sentence.

For example, if you were to ask which "car" is being talked about in sentence (2) you would typically get the answer "it's the one Keith took to London" in sentence (1). So the car in (2) is inferred to fill an open (i.e., as yet uninstantiated) role in the event of driving to London described in (1).

(1) Keith drove to London yesterday.
(2) The car kept overheating.

The questions addressed here concern the process by which such discourse role links are established during reading. In particular, is it a top-down expectation-driven process or is it instigated only by encountering the reference itself [e.g., *the car* in (2)]? Are references to discourse roles resolved on the basis of a representation of the complete situation (e.g., Keith driving to London yesterday) or is the process driven by verb-based semantic information independent of the situation (e.g., on the basis of the verb *drove* implying the use of a vehicle)? And, finally, are they resolved on-line when the reference is first encountered or is the process delayed until the whole sentence has been read? The literature is unclear on these points.

This article is organized into three sections. First we consider the current evidence on establishing and resolving discourse roles, reviewing both off-line memory and on-line reading-time studies. Two points emerge from this review: (a) different studies, using different methodologies and materials, lead to very different conclusions about how a reader represents the in-

We acknowledge the helpful comments of Meredyth Daneman, Martin Pickering, Keith Rayner, and two other reviewers on an earlier version of this article. The research reported here was supported by a grant from the UK Economic and Social Research Council to the Human Communication Research Centre.

Address correspondence and reprint requests to Simon Garrod, Human Communication Research Centre, Department of Psychology, University of Glasgow, 56 Hillhead Street, Glasgow G12 9YR, Scotland. E-mail: simon@psy.gla.ac.uk.

0749-596X/00 $35.00

formation in sentences like (1) with conflicting implications about the top-down or bottom-up nature of the role resolution process, and (b) it is not clear from the literature how verb-based lexical knowledge, as opposed to situational knowledge, contributes to establishing the discourse role link.

In the second section, we report an eye-tracking experiment designed both to evaluate the contribution of lexical and situational knowledge and to establish precisely how and when role links are resolved during reading. Finally, we present a two-stage model of the resolution process which takes into account the respective contributions of these two kinds of knowledge in terms of a distinction between bonding and resolution (Sanford & Garrod, 1989; Garrod & Sanford, 1994).

TOP-DOWN OR BOTTOM-UP RESOLUTION OF DISCOURSE ROLES

Early research on discourse roles was motivated by questions about the kind of representation that a reader routinely derives from a text. For example, on encountering (1) does the reader construct a representation that in some way encodes the information that Keith took his car to London? In an influential early study, Johnson, Bransford, and Solomon (1973) showed that participants falsely recognized test items containing an explicit role-filler (e.g., *the man took his car to work*) when they had previously read sentences which had not mentioned that role (e.g., *the man drove to work*). Hence, the authors suggested that the representation did in fact encode the implicit role information because readers could not discriminate between the two in memory. In a similar vein, Paris and Lindauer (1976) demonstrated that implicit roles (e.g., *knife* in the context *The teacher cut into the juicy steak*) were just as effective retrieval cues for sentences describing the bare events as for sentences which also included the role-filler (e.g., *The teacher cut into the juicy steak with a knife*). However, in both cases the authors recognized that their results could reflect inferences made at the time of testing and so might have little bearing on the representational question per se.

The relevance of the Paris and Lindauer (1976) finding was also brought into question by Corbett and Dosher (1978). They replicated the original result but also found that highly predictable role-fillers (e.g., the instrument *knife* for *cutting steak*) were better retrieval cues than less predictable role-fillers even when the original sentence had contained the less predictable item. Thus *knife* was a better cue for the sentence *The teacher cut the steak with a razorblade* than was *razorblade.*

The results of these early studies are therefore equivocal about whether the original representation of sentences like (1) encodes specific information about open discourse roles (see Singer, 1979). In order to avoid some of the problems inherent in the retrieval cue method, McKoon and Ratcliff (1981) applied a different technique. They carried out a series of experiments using materials of the following form:

(3) Bobby got a saw, *hammer,* screwdriver, and square from his toolbox.
(4) Then Bobby *pounded* the boards together with nails.
(5) Then Bobby *stuck* the boards together with glue.

Their main objective was to determine whether the event described in (4), *pounding the boards together with nails,* would activate the potential instrument *hammer,* which had been introduced in (3). They measured activation using the memory recognition probe *hammer,* presented immediately after the subjects had read either sentence (4) or sentence (5) and they found clear evidence for activation following (4) as compared to (5). Follow-up experiments also demonstrated no such activation with less predictable potential instruments [e.g., replacing *mallet* for *hammer* in (3)] and indicated that only the predictable instrument *hammer* became associated in memory with other words in the sentence (e.g., *boards*) after subjects had studied the text.

McKoon and Ratcliff's (1981) results are consistent with the view that open discourse roles, such as those associated with implied instruments, can be encoded into the final representation of a sentence under certain conditions: (i) when the role is strongly associated with the event and (ii) when the role-filler has already been mentioned in the context. This

latter point was reinforced in a more recent study by Lucas, Tanenhaus, and Carlson (1990), who found evidence that an instrument inference is only made when the potential role-filler (e.g., *hammer*) is mentioned in the prior context. However, these findings leave at least two questions unanswered. First, does the strong association between event and role come from knowledge about the situation as a whole, or is it more intimately associated with knowledge of the meaning of a critical verb? Second, to what extent is the linking process part of the on-line resolution of the sentence containing a reference to a discourse role, as it is with the processing of pronouns and other anaphoric references (e.g., Ehrlich & Rayner, 1983; Garrod, Freudenthal, & Boyle, 1994; Marslen-Wilson, Tyler, & Koster, 1993; Rayner & Duffy, 1986)?

VERBS AND DISCOURSE ROLES

There has been a longstanding interest in how syntactic and semantic representations of verbs might encode role-based information (Rumelhart, 1975; Schank, 1972; Schank & Abelson, 1977; Carlson & Tanenhaus, 1988). For example, a verb such as *load* is commonly associated with a THEME thematic role (i.e., the thing being loaded, as in *Harry loaded the truck with furniture*), but this thematic role is often left unfilled (e.g., in *Harry loaded the truck*) and it has been suggested that information about such roles may form part of the syntactic or semantic representation of the verb (Carlson & Tanenhaus, 1988; Mauner, Tanenhaus, & Carlson, 1995). These proposals raise important issues concerning the psychological processes involved in establishing discourse role links.

According to one view, implicit roles reflect our knowledge of the whole situation being described. Hence, if we attribute an instrument, say *hammer,* to the event of *pounding a nail into a board,* this would come from our knowledge of the whole situation and not just knowledge of the meaning of the verb *pound.* On the other hand, it is possible that discourse role links are only routinely made when the role-filler matches some semantic specification on an antecedent verb. In fact, experiments that have examined the automaticity of discourse role resolution indicate that the semantic representation of the verb may be a crucial factor.

The first evidence derives from the conflicting results of Singer (1979) and Garrod and Sanford (1982). Singer (1979) compared reading times for sentences such as (8) below following a sentence that either explicitly mentioned the instrument *shovel* (7) or only presupposed it (6).

(6) The boy cleared the snow from the stairs.
(7) The boy cleared the snow with *a shovel.*
(8) *The shovel* was heavy.

Singer found that the time to read (8) was longer in the context of (6) as compared to (7). This was interpreted in light of Haviland and Clark's (1974) bridging account as indicating that the reader had to compute a bridge between the role-filler *shovel* and the antecedent event *clearing the snow from the stairs,* thereby suggesting that role resolution was not an automatic process. However, Garrod and Sanford (1982) found no evidence of such bridging effects. Using materials like sentences (1) and (2) discussed above and an explicit control context containing the role-filler they found no difference in reading times [i.e., sentence (2) was read just as quickly when preceded by (1) as when preceded by (9)].

(1) Keith drove to London yesterday.
(9) Keith took his *car* to London yesterday.
(2) The *car* kept overheating.

They conjectured that the contrasting results might reflect the degree to which implicit roles were part of the semantic representation of the verbs used in each study. Singer had selected his materials on the basis of a set of event-instrument association norms. His informants were asked what one would use to *clear the snow from the stairs* and 90% responded with *shovel.* In contrast, Garrod and Sanford had selected materials on the basis of verb-role association with informants choosing the implied role given the verb alone. For example, when given *drive* more than 80% of informants generated *car* as the theme. Clearly these two criteria are very different. For example, the semantic association between *clear* and *shovel* is rather weak, even though it may be a preferred

instrument for the event of *clearing snow from the stairs.*

Cotter (1984) explored these differences in greater detail. First, she replicated the original results for both the Singer (1979) and Garrod and Sanford (1982) materials. She then compared the verbs with respect to their dictionary definitions and found that they differed in terms of the probability of having the implied role mentioned as part of the verb's definition. For example, the definition given for *drive* is *to convey in a VEHICLE* but the definition for clear is *to free from obstruction.*[1] For all the verbs used by Garrod and Sanford either the exact role or its superordinate category (e.g., *vehicle* for *car*) was given in the dictionary, whereas this was only true for half of the verbs used by Singer. Interestingly, Cotter also demonstrated that the degree of association between the complete verb phrase (e.g., *clear the snow from the stairs*) and the implied role (e.g., *shovel*) was just as high for the two kinds of material. Hence, Cotter's analysis supports the proposal that open roles may only be directly accessible for processing when there is a strong semantic relationship between verb and role. This conclusion is consistent with the earlier McKoon and Ratcliff (1981) finding that only strongly associated instruments were activated by the occurrence of the subsequent verb.

There is other evidence to suggest that verbs may impose semantic restrictions on their associated role-fillers. In an eye-tracking experiment, Garrod, O'Brien, Morris, and Rayner (1990; see also O'Brien, Shank, Myers, & Rayner, 1988) showed that introducing an instrument such as *weapon* in the context of a verb such as *stab*, which restricts its instrument role to be knife-like, affected the ease of subsequent reference to the weapon as a *knife.* Thus, they found that gaze duration on *the knife* was shorter in the context *stab with a weapon*, which imposes a knife-like restriction on the weapon, than in the context *assault with a weapon*, which imposes no such restriction. One conclusion from this study is that the discourse representation for verbs such as *stab* includes semantic information about their instrument roles that enhances subsequent integration of a matching role-filler.

However, there is one problem with this proposal. Commonly verbs select for very different role-fillers depending upon the context. For example, a verb such as *cut* will strongly select the instrument *knife* in the context of cutting into a steak, but select for *scissors* in the context of cutting someone's hair. The question therefore arises as to whether the strong association between verb and instrument can be mediated by the context in which the verb occurs or results from a purely lexical association between the verb and its role-filler. If the verb alone plays the most important role, we might expect only the most strongly associated or dominant role-fillers to be resolved automatically, as suggested by McKoon and Ratcliff's (1981) probe-recognition study. On the other hand, if the link is established via a representation of the event as a whole, only contextually appropriate role links should be resolved automatically.

The experiment reported here was designed to differentiate between these two possibilities and to establish the precise on-line nature of the resolution of such discourse-role links during normal reading. It arose out of a pilot study (see Terras, 1997) in which we had tracked readers' eye movements while they read materials like those used by Garrod and Sanford (1982). However, for each verb two different role-fillers were selected according to the context in which the verb occurred. Thus a verb such as *write* would appear either in the context of *writing a letter* or *writing an exercise on a blackboard* [see (10) and (11) below]:

(10) The teacher was busy *writing a letter of complaint* to a parent.
(10a) The teacher was busy *writing a letter of complaint with a pen.*

However, she was disturbed by a loud scream from the back of the class and *the pen* dropped on the floor. . . .

(11) The teacher was busy *writing an exercise on the blackboard.*
(11a) The teacher was busy *writing an exercise on the blackboard with chalk.*

However, she was disturbed by a loud scream from the back of the class and *the chalk* dropped on the floor. . . .

[1] *Collins Cobuild English Language Dictionary.*

We compared eye movements as participants read a subsequent clause (shown below each example) containing a target reference to either *the pen* or *the chalk* in contexts which contained either an implicit reference to the instrument through the event described [as in (10) and (11)] or a direct antecedent reference [as in (10a) and (11a)]. Post hoc analysis of the eye-movement data suggested that for each verb there was a dominant role-filler. Readers would take longer overall when interpreting references to implicit roles when the filler was less strongly associated with the verb alone (e.g., *chalk* with *write* versus *pen* with *write*). The presence of a dominance effect suggested a way of contrasting lexical with contextual influences on discourse role resolution and this forms the basis of the present experiment.

For any role-filler pair it is possible to compare appropriate versus inappropriate contexts for both dominant and nondominant verb–role pairings. For example, with the dominant pair *write–pen* you can have appropriate contexts like (10) above and inappropriate contexts like (12) below. Conversely, these same contexts can introduce a nondominant filler *chalk* appropriately in (11) above and inappropriately in (13) below.

(12) The teacher was busy *writing an exercise on the blackboard.*
(12a) The teacher was busy *writing an exercise on the blackboard with a pen.*

However, she was disturbed by a loud scream from the back of the class and *the pen* dropped on the floor. . . .

(13) The teacher was busy *writing a letter of complaint* to a parent.
(13a) The teacher was busy *writing a letter of complaint* with *chalk.*

However, she was disturbed by a loud scream from the back of the class and *the chalk* dropped on the floor. . . .

If the resolution process were driven primarily by the lexical relationship between verb and role-filler, then we would expect early resolution processes to be governed by lexical dominance. Thus, readers should automatically integrate the dominant filler *pen* with the verb *write* irrespective of the context in which it occurs [i.e., in both (10) and (12)], whereas, for the nondominant filler *chalk*, there should be no immediate attempt to integrate the reference even in appropriate contexts such as (11).

Conversely, if resolution were driven by the overall context of introduction, both *the pen* and *the chalk* should be immediately integrated into the appropriate contexts [(10) and (11)] but not the inappropriate ones [(12) and (13)]. Of course it could also be that lexical and contextual factors influence resolution at different points in reading. Thus it is possible to have a two-stage process with the lexical relationship dominating immediate processing and contextual effects emerging later.

The eye-tracking method is ideally suited to uncovering such time-course effects. Early effects will show up in differences in first-pass reading times (i.e., the time spent fixating the region before the eye moves on) when the crucial role-based target references are encountered. Later effects will show up in the second-pass reading times following exposure to the rest of the sentence.

First, let us consider how lexical effects might emerge in the reading-time data. Here the crucial contrast is between conditions in which the role is only implied by the verb [as in (10) and (11)] and conditions in which the role-filler has been explicitly introduced in the context [e.g., as *with a pen* in (10a), or *with chalk* in (11a)]. In line with Singer (1979) and Garrod and Sanford (1982), explicit conditions serve as a baseline against which to examine any effects in implicit conditions that reflect the integration of role-filler with implicit role. Thus to establish early lexical influence we need to look at the increased first-pass reading times for *the chalk dropped* or *the pen dropped* in the *implicit* appropriate context conditions [i.e., (10) and (11)] as compared to the *explicit* appropriate context baseline conditions [i.e., (10a) and (11a)]. The lexical account would predict implicit–explicit differences in first-pass reading times for the nondominant *chalk* but not for the dominant *pen.* We shall refer to this contrast as the *lexical effect.*

To establish early contextual influences we need to make a quite different initial comparison. If context were the primary factor, we would expect to find an increase in first-pass

reading time for *the pen dropped* following the *inappropriate* context *writing on the blackboard* [i.e., (12)] as compared to the *appropriate* context *writing a letter* [i.e., (10)], and a similar increased reading time for *the chalk fell* following the *inappropriate* context *writing a letter* [i.e., (13)], as compared to the *appropriate* context *writing on the blackboard* [i.e., (11)]. This contrast we shall refer to as the *context effect.*

To reiterate, if discourse role resolution is dominated by lexical influences we should expect an early lexical effect for nondominant role-fillers, such as *chalk* in the context of *write,* but no such effect for dominant role-fillers, such as *pen* in the context of *write.* Conversely, if the process is dominated by context we should expect an early context effect for both the dominant and nondominant role-fillers that would be reflected in the contrast between the inappropriate and appropriate implicit contexts.

Finally, the context effect might be mediated by dominance. Thus it could be that only the dominant role-fillers (e.g., *pen*) enable early access to context, in which case we would expect to detect an earlier emergence of the context effect following *pen* than following *chalk.* The experiment was designed to tease apart these various effects over the time course of reading the critical sentences.

METHOD

Participants

Forty-eight students from the University of Glasgow were paid to participate in the experiment. All were native speakers of English. Some of them had previously participated in other eye-tracking studies but none had taken part in any of the pretests.

Materials and Design

Twelve verbs and their 24 context-dependent role-fillers were used to generate the stimulus materials. They were carefully pretested to ensure that each verb–role pairing had a strong association, but with evidence for a clear dominance of one verb–role pairing over the other in the absence of context. This was done in a series of association pretests to establish the memory relationship between verb and role.

TABLE 1

Percentage of Participants Choosing Dominant and Nondominant Role-Fillers Given the Verb (Column 1) or Given the Verb+Object (Column 2) and Percentage of Participants Choosing the Verb Given the Role-Filler (Column 3)

	Verb	Verb+object	Role
Dominant	64.28	91.22	82
Nondominant	5.75	75.67	73.83

Association pretests. The verb–role sets were used in an association pretest for forward (verb-to-role) and backward (role-to-verb) associations. For the verb–role test there were two versions. In one, 30 participants were given the task of deciding, "What do you VERB with?" where "VERB" was replaced with each of the 12 verbs (e.g., "What do you WRITE with?"). In the other, another 30 participants had to establish the degree to which the verb plus its minimal context selected for a special role-filler: the question here was "What do you VERB+OBJECT with?" where "VERB+OBJECT" was replaced with the same verbs but different restrictions (e.g., "What do you WRITE A LETTER with?" or "What do you WRITE ON THE BLACKBOARD with?"). Finally, a third group of 30 participants were given the role–verb association test with the question "What do you do with ROLE?" where "ROLE" was replaced with the set of roles to be used in the experiment (e.g., "What do you do with A PEN?" or "What do you do with A PIECE OF CHALK?"). For each test the participants were allowed to write down as many items as came to mind.

The results from these tests were used to classify fillers as verb dominant as opposed to just context dominant. The verb-dominant ones were taken to be those elicited more often in the default condition (i.e., with the verb alone). The percentage of participants choosing the fillers or verbs for the various association tests is shown in Table 1 sorted by dominance. As can be seen in the table, dominant verb–role pairs produce stronger associations for both the verb–role [$t(11) = 5.82$, $p < .001$) and verb+object–role tests [$t\,(11) = 3.73$, $p < 0.01$]. In fact, for 11 of the 12 verbs the dominant role was mentioned

by at least 30% of participants given the verb alone and in the remaining case the dominant association was with a body part (e.g., *eyes* for *see*), which would not normally be taken as a role-filler (see Dosher & Corbett, 1982).

Finally, the table illustrates that the dominance difference does not carry through to the backward association between role-filler and verb. Despite the slightly greater proportion of participants choosing the verb following dominant fillers, there is no reliable difference across materials [$t\,(11) = 0.89$, $p = 0.39$]. So the basic difference between dominant and nondominant fillers is in the verb–role default association test. Both dominant and nondominant fillers are readily elicited when given the verb plus a minimal context and will strongly elicit the verb when presented alone. Thus the difference is mainly in the forward association between verb and role-filler.

These 24 dominant and nondominant verb–role pairings were incorporated into contexts designed to select for each role. The contexts were also pretested to ensure that they were truly selective for that role-filler.

Context pretest. The final pretest was designed to make sure that the dominant or nondominant contexts would select for their appropriate role-fillers even in the absence of the critical verb. It is important to show that *a teacher doing something with a letter* does not inadvertently suggest the presence of *chalk* or that *a teacher doing something with an exercise on the blackboard* does not inadvertently suggest the presence of a *pen.* If they did so, then it would confound any dominance effect arising purely from verb *write.* To ensure that this was not the case a neutral verb was chosen to replace the critical verb in all the contexts to be used in the main experiment. For example, the two contexts for the verb *write* were: *She was busy writing a letter of complaint to a parent* and *she was busy writing an exercise on the blackboard.* These were changed to: *She was busy reading a letter of complaint from a parent* and *she was busy reading an exercise on the blackboard.* Unlike *write, read* does not require any instrument and should not bias toward either the dominant appropriate role-filler *pen* or the nondominant appropriate role-filler *chalk.*

The modified materials with the neutral verbs were put into the full context used in the main experiment (see Table 2) up to the point where the critical role-filler would have been mentioned. Fifty participants were then required to choose which of the two possible role-fillers (e.g., *pen* or *chalk*) was most likely to occur in that context. For even the least effective context, at least 75% of participants chose the appropriate role-filler for that context (e.g., they chose *pen* in the context of *reading a letter of complaint* and they chose *chalk* in the context of *reading an exercise on the blackboard.*) Furthermore, across all contexts the appropriate role-filler was chosen by 94% of participants on average. Thus the contexts clearly favored their appropriate role-fillers even in the absence of the verb. Hence, any dominance effects could be attributed solely to the verb rather than the context surrounding that verb.

Experimental stimuli. The experimental stimuli were constructed out of these 12 verb–role sets and the pretested contexts. In order to increase the number of experimental materials, each verb–role pair was used twice but in different contexts. An example of one set of stimuli is shown in Table 2 (the complete set of verbs and role-fillers is shown in the Appendix 1 and the complete set of contexts together with the replacement verbs used in the context pretest is shown in Appendix 2).

The stimulus passages all conformed to a standard format. Each had an introductory sentence that established the general context for interpretation. Then the second sentence mentioned the verb together with the role-selecting context and either explicitly stated or implied the target role-filler. The third sentence contained the crucial anaphoric reference to the verb role, which was either the dominant or nondominant role-filler for the verb and was always separated from the sentence containing the verb by at least one intervening clause. There was then a final filler sentence to insure that the passages were coherent.

This produced eight experimental conditions, two within-subjects (and within-items) factors of context type (appropriate or inappropriate context) and explicitness (explicit or implicit introduction of the role) and one between-sub-

TABLE 2

Sample Materials Used in the Experiment

Dominant verb–role pair (WRITE–PEN)	
The teacher worked quietly as the children read their books.	
Appropriate context	
She was busy *writing a letter* of complaint to a parent.	Implicit antecedent
She was busy *writing a letter* of complaint *with a pen.*	Explicit antecedent
Inappropriate context	
She was busy *writing* an exercise *on the blackboard* by the door.	Implicit antecedent
She was busy *writing* an exercise *on the blackboard with a pen.*	Explicit antecedent
Target sentence	
However, she was disturbed by a loud scream from the back of the class and *the pen* dropped on the floor.	
She called for quiet and threatened the class with detention if there was any further disturbance.	
Nondominant verb–role pair (WRITE–CHALK)	
The teacher worked quietly as the children read their books.	
Appropriate context	
She was busy *writing* an exercise *on the blackboard* by the door.	Implicit antecedent
She was busy *writing* an exercise *on the blackboard with chalk.*	Explicit antecedent
Inappropriate context	
She was busy *writing a letter* of complaint to a parent.	Implicit antecedent
She was busy *writing a letter* of complaint *with chalk.*	Explicit antecedent
Target sentence	
However, she was disturbed by a loud scream from the back of the class and *the chalk* dropped on the floor.	
She called for quiet and threatened the class with detention if there was any further disturbance.	

jects (but within-item) factor, target type (dominant or nondominant target).

Eight experimental lists of materials were compiled for presentation, with each list containing six passages in each of the four within-subject experimental conditions. Thus, each passage in each condition was read by six participants. The stimuli were presented in a fixed random order and questions were given at the end of each passage to ensure that participants had read the materials carefully. These 24 materials were then intermixed with a further 36 from another experiment, which acted as fillers.

Apparatus and Procedure

Eye movements were monitored by a Stanford Research Institute Dual Purkinje Generation 5.5 Eye Tracking System made by Fourward Technologies under license to S.R.I. The eye-tracker has an angular resolution of 10′ arc. Viewing was binocular with eye location being recorded from the right eye. The eye-tracking system was interfaced with a Vanilla 386 computer that controlled the presentation of stimuli and recorded the output from the eye-tracking system. The experimental stimuli were presented on a VDU, which also interfaced with the Vanilla. The VDU was located at a distance of 70 cm and the material spanned six to eight lines, with a maximum of 65 characters per line. There were 3.5 characters per degree of visual angle. The position of a participant's eye was sampled every millisecond and analyzed using software that continuously monitored the output in order to establish the sequence of eye fixations and measured their start and finish times to the nearest millisecond.[2] Hence, a continuous record of eye movements, fixation position, and fixation duration was obtained.

[2] This software was developed by Dr. Charles Clifton, whose support we gratefully acknowledge.

Data Analysis

The crucial sentences for testing the lexical and context effects predicted earlier are the sentences containing the target role-filler references (e.g., *the pen dropped* or *the chalk dropped*). So the eye movement record was analyzed in terms of two critical regions of these sentences. First, there was the region containing the role-filler noun (e.g., *pen* or *chalk*). For the dominant materials this region had an average length of 5.7 characters ranging from 4 to 11 characters across the materials. For the nondominant materials its average length was 6.2 characters ranging from 4 to 11 characters across the materials. Analysis from the pilot experiment indicated that the strongest effects were likely to occur in the spillover region beyond the target noun. So a second verb region was defined. The verb region either contained the verb alone (e.g., *dropped*) or, if the next word was an auxiliary or an adverb, the region included that word and the following main verb (e.g., *was put*). This was to ensure that the regions were of sufficient size for analysis. For both the dominant and nondominant materials these regions were identical with an average length 9.2 characters and a range of 5 to 17 characters.

The reading times were analyzed according to the three measures illustrated for the noun region *chalk* in Fig. 1. To detect the earliest effects we used a *first-pass reading time* measure that sums the fixation durations from first fixating the region until the eye moves out of the region either to the left or right (these will be referred to as N1 or V1 times for noun and verb region respectively).[3] To detect early effects including initial repair we used the *regression-path reading-time* measure, which sums all fixation durations from first fixation of the region until the eye goes beyond that region (referred to as N2 and V2). Finally, we recorded *second-pass reading-time* for the noun region as a measure of later processing effects (referred to as N3). This measure includes durations for all fixations on the region which occur after the first-pass reading of the region. In addition we looked at the proportion of first-pass regressions for the verb region in order to check for early repair processes.

[3] For the noun region in which there is only a single word this measure is also referred to as gaze duration.

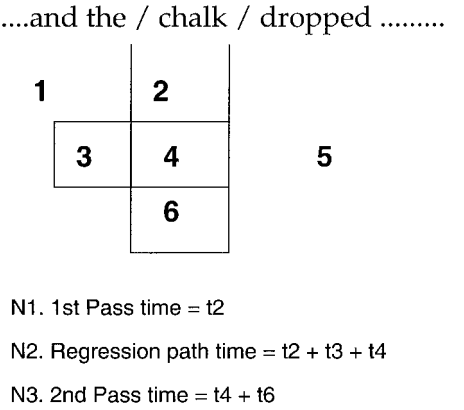

FIG. 1. An illustration of the three reading time measures used in the experiment for the noun region/chalk/. The numbers represent a hypothetical sequence of fixations. The different measures [N1 = First pass (noun region), N2 = Regression path (noun region), N3 = Second pass(noun region)] are shown as summations of durations for those fixations (e.g., $t1+t2\ldots$).

RESULTS AND DISCUSSION

The main predictions concern the earliest point in reading the target sentences at which the lexical and context effects can be detected. As pointed out earlier, the lexical effect relates to the reading-time difference following appropriate implicit contexts (e.g., *writing a letter* for *pen* and *writing on the blackboard* for *chalk*) as compared to appropriate explicit contexts (e.g., *writing a letter with a pen* for *pen* and *writing on the blackboard with chalk* for *chalk*). By contrast the context effect relates to the reading-time difference following inappropriate implicit contexts (e.g., *writing on the blackboard* for *pen* and *writing a letter* for *chalk*) as compared to appropriate implicit contexts (e.g., *writing a letter* for *pen* and *writing on the blackboard* for *chalk*). First we give a brief overview of the main results relating to these two effects and then report the detailed analysis of reading times and pattern of regressions from the critical regions.

Overview of the Main Reading-Time Results

The main results in relation to the lexical and context effects are illustrated in Fig. 2. The

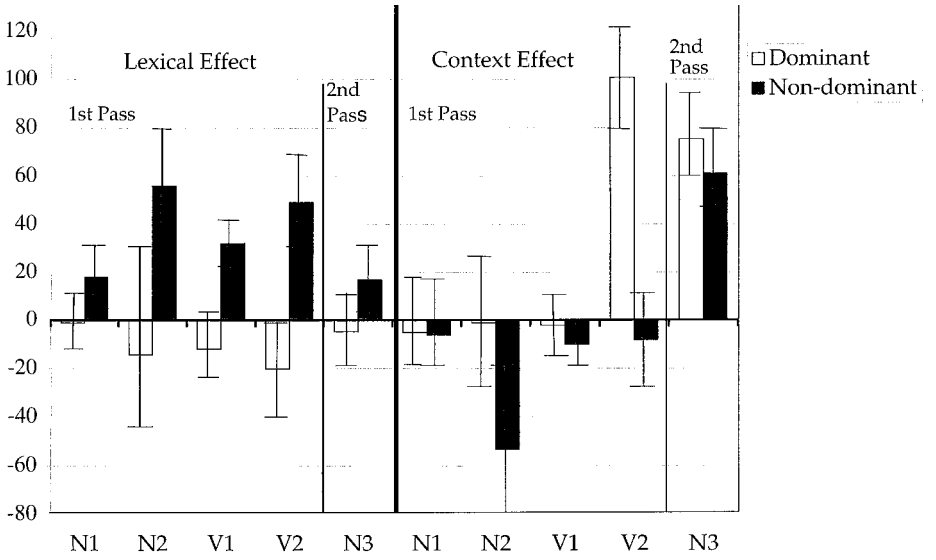

FIG. 2. The lexical and the context effects for the first- and second-pass reading measures in the experiment. The lexical effect is shown as the difference in reading time (in milliseconds) between the implicit appropriate context condition and the explicit appropriate context condition. The context effect is shown as the difference in reading time (in milliseconds) between the inappropriate implicit context condition and the appropriate implicit context condition. N1 corresponds to the first-pass reading-time effect N2 to regression path time effect and N3 to the second-pass reading-time effect for the noun region (see Fig. 1). V1 corresponds to the first-pass reading-time effect and V2 to the regression-path reading-time effect for the verb region (error bars show ± standard error).

left-hand panel shows progressive measures of the lexical effect (represented as the difference in reading time following implicit as opposed to explicit appropriate contexts) for both dominant and nondominant targets. Starting with the first-pass reading of the noun region (i.e., N1 and N2 measures) there is a small and partially reliable lexical effect for nondominant but not for dominant targets. The magnitude of the effect increases and is reliable in both the first-pass (i.e., V1) and regression-path reading times (i.e., V2) for the verb region but is no longer present in the second-pass reading times for the noun region (i.e., N3). This indicates an early lexical influence with nondominant targets exhibiting consistently longer reading times in the implicit conditions.

The right-hand panel shows the same progressive measures of the context effect (represented as the difference in reading time following inappropriate as opposed to appropriate implicit contexts). Here there are no context effects detectable in either the first-pass measures in the noun region or the first pass measures in the verb region (i.e., N1 and V1). However, a strong effect emerges for the dominant targets in the verb region with the regression-path time analysis (i.e., V2). This indicates an early influence of context for the dominant but not the nondominant targets. Finally, there is a strong context effect detectable in the second-pass reading times on the noun (i.e., N3) for both dominant and nondominant targets. These findings are corroborated by the pattern of first-pass regressions following fixation of the verb.

The detailed analysis of these results is given below for the noun and then the verb region.

First-Pass Reading Time: Noun Region

Table 3 shows the first-pass reading time and regression-path times for the noun region averaged across participants and items (i.e., N1 and N2). Only trials with first-pass fixations were included in this and subsequent analyses. The table also shows the probability of first-pass fixation in this region under the different conditions of the experiment.

TABLE 3

First-Pass Reading Times and Regression-Path Times for the Noun Region (in ms)

	Dominant targets		Nondominant targets	
Context:	Appropriate	Inappropriate	Appropriate	Inappropriate
Antecedent				
First pass				
Explicit	255 (54)	252 (57)	246 (66)	263 (67)
Implicit	256 (48)	251 (56)	262 (63)	251 (56)
Regression path				
Explicit	284	299	284	319
Implicit	298	299	340	287

Note. Percentage first pass fixations in parentheses.

The means of the first-pass measures calculated both across participants and across items were entered into 2 × 2 × 2 analysis of variance designs with target type as a between-subject (but within-items) factor and context and explicitness as within-subjects (and within-items) factors. All the analyses we report are based on treating both subjects as a random effect ($F1$) and materials as a random effect ($F2$) and all are reliable at less than the .05 level unless otherwise stated.

As indicated above, the first-pass analyses of the noun region produce only one consistent result, which is a marginally reliable lexical effect for the nondominant targets. There is no evidence for any contextual effect emerging while reading this region.

Thus the overall analysis for first-pass reading times (i.e., N1) revealed no reliable main effects or interactions (For all $F1$s and $F2$s $p > .1$). However, in examining the planned comparison for the lexical effect (i.e., the difference between implicit versus explicit appropriate contexts) there was a marginal effect of 16 *ms* for the nondominant targets [$F1(1,23) = 3.36$, $MS_e = 2377$, $p = .08$; $F2(1,23) = 3.03$, $MS_e = 1686$, $p = .09$].

Given the relatively low probability of first-pass fixation in the noun region it was decided to carry out an additional first-pass analysis extending the region to the left by up to four characters until a fixation was encountered. This fixation was then included as a first-pass reading time on the region. The analyses of the extended region data did not alter the pattern of results. There were no reliable main effects or interactions and only a marginally reliable lexical effect of 20 *ms* between implicit and explicit appropriate context conditions for nondominant targets [$F1(1,23) = 3.31$, $MS_e = 2492$, $p = .08$; $F2(1,23) = 3.56$, $MS_e = 1248$. $p = .07$].

The final analysis we carried out was on the regression-path time data (N2). The data averaged across subjects and items were entered into the same ANOVA designs as used earlier. These produced no reliable main effects or interactions (For all $F1$s and $F2$s $p > 0.1$). However, the lexical effect of 56 *ms* was again marginally reliable for the nondominant targets [$F1(1,23) = 3.58$, $MS_e = 10763$, $p = .07$; $F2(1,23) = 3.87$, $MS_e = 5253$, $p = .06$].

First-Pass Reading Time: Verb Region

The first-pass reading time (V1) and regression-path time (V2) for the verb region are shown in Table 4, averaged across participants and items. The probability of first-pass fixation is also shown. Here there is evidence both of a lexical effect for nondominant targets and an early context effect for dominant targets that appears in the regression-path time analysis.

The first-pass times (V1) were analyzed in the same ANOVA designs as used with the noun region and this produced only one marginally reliable main effect of explicitness (Explicit = 278 *ms*, Implicit = 294 *ms*) in the by-items analysis [$F2$ $(1,23) = 3.58$, $MS_e = 2863$, $p = .07$]. However, there was also a

TABLE 4

First-Pass Reading Times and Regression-Path Times for the Verb Region (in ms)

	Dominant targets		Nondominant targets	
Context:	Appropriate	Inappropriate	Appropriate	Inappropriate
Antecedent				
First pass				
Explicit	285 (80)	266 (80)	284 (79)	275 (77)
Implicit	273 (78)	271 (68)	314 (83)	316 (83)
Regression path				
Explicit	339	343	371	349
Implicit	319	420	419	412

Note. Percentage of first pass fixations is shown in parentheses.

reliable interaction between explicitness and dominance [$F1(1,46) = 4.03$. $MS_e = 4874$; $F2(1,23) = 4.2$, $MS_e = 4249$]. Exploring this interaction further revealed that it was entirely due to the effect of explicitness for the nondominant targets (Explicit = 280 *ms*, Implicit = 315 *ms*; $F1(1,23) = 8.75$, $MS_e = 5298$; $F2(1,23) = 4.89$, $MS_e = 4479$]. No other interactions emerged across either subjects or items (for all $F1$s and $F2$s $p > .1$]. So the first-pass times for the verb region with the nondominant targets show an extension of the earlier lexical effect found in the noun region.

We now turn to the regression-path time data (V2) for the verb region (see Table 4). These were analyzed in the same ANOVA designs as used for the other measures. They revealed a main effect of explicitness [Explicit = 351 *ms*, Implicit = 393 *ms*; $F1(1,46) = 8.16$, $MS_e = 13495$; $F2(1,23) = 4.71$, $MS_e = 11649$] and two interactions, one between explicitness and context [$F1(1,46) = 4.63$, $MS_e = 9252$; $F2(1,23) = 5.85$, $MS_e = 8699$] and another between target type and context, reliable in the by-subjects analysis [$F1(1,46) = 3.8$, $MS_e = 19545$, $p = .05$) but not by items ($F2<1$). To understand this complex pattern of results we need to look separately at the effects associated with the two kinds of target.

First, for the nondominant targets there is again an effect of explicitness (Explicit = 360 *ms*, Implicit = 416 *ms*) which is reliable by subjects [$F1(1,23) = 7.75$, $MS_e = 13884$] and marginally reliable by items [$F2(1,23) = 3.44$, $MS_e = 14198$, $p = .076$]. No other effects approach significance for these targets (All $F1$s and $F2$s < 1). However, when we look at the dominant targets there is a quite different pattern with both an effect of context (Appropriate = 329 *ms*, Inappropriate = 382 *ms*; $F1(1,23) = 4.94$, $MS_e = 13439$; $F2(1,23) = 3.43$, $MS_e = 18959$, $p = .076$] and, more importantly, an interaction between context and explicitness [$F1(1,23) = 4.65$, $MS_e = 12093$; $F2(1,23) = 7.99$, $MS_e = 8238$]. This is due to a reliable context effect as shown in the difference of 101 *ms* between appropriate and inappropriate contexts in the implicit condition [$F1(1,23) = 10.12$, $MS_e = 12093$; $F2(1,23) = 15.88$, $MS_e = 8238$].

So the first-pass reading time and regression-path time analyses of the noun (i.e., N1 and N2) and verb (i.e., V1 and V2) regions confirm the pattern summarized earlier. For the nondominant targets there is some evidence of an early lexical effect in the noun region that shows up as a marginal difference between the implicit and explicit appropriate context conditions (i.e., for *chalk* following *write on a blackboard* as compared to *chalk* following *write on a blackboard with chalk*). This is apparent to some degree in both the N1 and N2 analyses. The effect then increases and is reliable in the V1 and V2 analyses of the verb region. However, for these nondominant targets there is no evidence for an early emergence of a context effect.

The situation for dominant targets is quite different. Here, N1 and N2 analyses of the noun

TABLE 5

Second-Pass Reading Times for the Noun Region (ms)

	Dominant targets		Nondominant targets	
Context:	Appropriate	Inappropriate	Appropriate	Inappropriate
Antecedent				
Second pass				
Explicit	170	195	197	229
Implicit	164	239	214	275

region show neither lexical nor context effects. However, in the verb region there is strong evidence of a context effect in the V2 analysis (i.e., for *pen dropped* following *write on a blackboard* as compared to *pen dropped* following *write a letter*). This suggests an earlier influence of context for the dominant verb–role pairs.

We turn now to the second-pass analyses of the noun region to establish any secondary processing effects associated with the lexical or context effects observed in the first-pass and regression-path times.

Second-Pass Reading Time: Noun Region

The second-pass reading times for the noun region averaged across participants and items are shown in Table 5. The only major result is that of context, for both dominant and nondominant target items.

The data were analyzed in the same ANOVA designs used for the first-pass times. This produced three main effects: target type [Dominant = 193 *ms*, Nondominant = 229 *ms*; $F1$ (1,46) = 5.03, MS_e = 12915, $F2$(1,23) = 7.1, MS_e = 9527], explicitness (Explicit = 198 *ms*, Implicit = 224 *ms*; $F1$ (1,46) = 4.24, MS_e = 7400; $F2$ (1,23) = 4.97, MS_e = 5567], and context [Appropriate = 187 *ms*, Inappropriate = 224 *ms*; $F1$ (1,46) = 22.02, MS_e = 5044; $F2$ (1,23) = 22.66, MS_e = 4928]. The target-type effect is probably simply due to the difference in average region size for the dominant and nondominant targets. The other effects are complicated by a marginally reliable interaction between explicitness and context [$F1$(1,46) = 3.56, MS_e = 5406, p = .06; $F2$ (1,23) = 3.02, MS_e = 5947, p = .09]. This interaction is expected because in the explicit conditions the role-filler has already been encountered in that context in the prior material. Hence, in the explicit inappropriate context condition the reader should have reconstrued their interpretation to accommodate the inappropriate role-filler some time before they encounter the target region.

To understand the pattern of results more clearly we need to analyze the two kinds of targets separately. First, for the dominant targets there is only one reliable main effect, that of context [Appropriate = 167 *ms*, Inappropriate = 217 *ms*; $F1$ (1,23) = 8.34, MS_e = 7206; $F2$ (1,23) = 10.58, MS_e = 5831], which is also present as a context effect in relation to the planned comparison of context across implicit conditions [$F1$ (1,23) = 13.04; $F2$ (1,23) = 9.4]. So this effect extends that uncovered in the regression-path reading-time analysis discussed above.

Turning to the nondominant targets there is a marginal main effect of explicitness [Explicit = 213 *ms*, Implicit = 245 *ms*; $F1$ (1,23) = 3.34, MS_e = 7261, p = .08; $F2$(1,23) = 4.01, MS_e = 5907, p = .06], but this is not associated with a reliable lexical effect in the planned comparison (for the implicit/explicit difference in the appropriate context conditions both $F1$ and $F2$ < 1). This is because there is also a reliable main effect of context [Appropriate = 206 *ms*, Inappropriate = 252 *ms*; $F1$ (1,23) = 17.72, MS_e = 2883; $F2$(1,23) = 11.29, MS_e = 4449] which is primarily associated with the implicit antecedent conditions. Thus, the planned comparison of context across implicit conditions is now reliable for the nondominant targets [$F1$ (1,23) = 7.97; $F2$ (1,23) = 6.87].

The second-pass reading times therefore show that the nondominant targets are now behaving just like the dominant targets in terms of

TABLE 6

Percentage of First-Pass Regressive Saccades from the Verb to an Earlier Region Following the Critical Role Reference

	Dominant targets		Nondominant targets	
Antecedent	Context: Appropriate	Inappropriate	Appropriate	Inappropriate
Explicit	12.7	11.8	11.0	12.5
Implicit	13.4	27.6	15.8	18.6

the emergence of a late context effect and disappearance of the earlier lexical effect. We turn finally to the analysis of the pattern of regressive eye movements following fixation of the verb to check that they confirm the reading-time effects reported above.

First-Pass Regressions from the Verb

The percentage of regressive saccades following first-pass fixation of the verb are shown in Table 6 averaged across participants and items. The overall pattern is consistent with the regression-path reading-time analysis described above (i.e., V2). With dominant targets there are substantially more regressions from the verbs following implicit introduction of the antecedent in inappropriate contexts (e.g., following *the pen fell* in the context of *writing on the blackboard*), as would be expected given the context effect in the V2 data reported above. For the nondominant targets, this pattern does not occur.

To test the reliability of these effects, the data across participants and items were analyzed in the same $2 \times 2 \times 2$ ANOVA designs used for the reading-time data. The analysis confirms our observations. There two main effects, explicitness [Explicit = 12%, Implicit = 18.9%; $F1(1,46) = 6.59$, $MS_e = .338$; $F2(1,23) = 5.76$, $MS_e = 400$] and context (Appropriate = 13.2%, Inappropriate = 17.6%), which is reliable in the items analysis [$F2(1,23) = 4.15$, $MS_e = 536$] but not by subjects ($F1 < 1$). The items analyses also produced a reliable interaction between target type and context [$F2 = 4.457$, $MS_e = 303$] and a reliable three-way interaction between target type, explicitness, and context [$F2(1,23) = 4.75$, $MS_e = 180$].

To interpret this complex pattern of results we need to look at the different target types separately. First, for nondominant targets there is only one marginal effect, which is that of explicitness (Explicit = 12%, Implicit = 17.2%) and it is only marginally reliable in the subjects analysis [$F1(1,23) = 2.95$, $MS_e = 233$, $p = .099$]. No other main effects or interactions emerge for the nondominant target data (for $F1$s and $F2$s all ps $> .1$).

Turning to the dominant targets there is a marginal effect of explicitness [Explicit = 12.3%, Implicit = 20.5%; $F1(1,23) = 3.71$, $MS_e = 427$, $p = .06$; $F2(1,23) = 3.709$, $MS_e = 442$, $p = .06$] and a marginal effect of context in the items analysis [Appropriate = 13.1%, Inappropriate = 19.7%; $F2(1,23) = 3.01$, $MS_e = 562$, $p = .09$]. These main effects are qualified by an interaction between explicitness and context, which is reliable in the items analysis [$F1(1,23) = 9.39$, $MS_e = 232$]. However, the planned comparison for the context effect (i.e., the difference between inappropriate and appropriate contexts in the implicit conditions) is reliable by both subjects and items [$F1(1,23) = 4.29$, $F2(1.23) = 16.62$]. This reflects the 14.2% difference between the percentage of first-pass regressions in inappropriate implicit conditions as compared to appropriate implicit conditions.

So the pattern of first-pass regressions from the verb confirms the earlier analysis of regression path times for the verb region. There is a strong context effect for the dominant targets and only a weak explicitness effect for the nondominant targets.

Summary of the Target Sentence Reading Time and Pattern of Regressive Eye Movements Analyses

The reading-time and regression-pattern analyses point to a clear difference between

processing of the two kinds of target role-filler. The first finding is that with nondominant verb–role pairs (e.g., *write–chalk*) there is evidence of an early lexical effect. Hence, the role-filler is integrated more rapidly following explicit introduction of the antecedent (e.g., *write on a blackboard with chalk*) as compared to implicit introduction (e.g., *write on a blackboard*). However, for dominant verb–role pairs (e.g., *write–pen*) there is no difference between the two forms of antecedent introduction. This result is reminiscent of the earlier findings of Corbett and Dosher (1978) that dominant or highly associated instruments (e.g., *knife* for *cut*) act as superior recall cues even for sentences containing a different instrument (e.g., *the teacher cut the steak with a razorblade*). It is also consistent with the main findings from McKoon and Ratcliff (1981) which showed a systematic advantage for dominant instruments in the probe recognition studies. The second finding is that the dominant verb–role pairs (e.g., *write–pen*) lead to an early context effect with faster V2 reading times following appropriate implicit contexts (e.g., *write a letter* for *pen*) than following inappropriate implicit contexts (e.g., *write on the blackboard* for *pen*). By contrast no such effect occurs with materials containing nondominant targets until the second-pass reading of the noun. This latter finding is also confirmed in the comparison of the percentage of first-pass regressions in which there is a reliable context effect for the dominant role-fillers, but not for the nondominant role-fillers.

Although these results are consistent with the main findings of McKoon and Ratcliff (1981), they do appear to conflict with one of their results. In one of their experiments (Experiment 5) they included a context manipulation in which the antecedent role-filler was introduced in such a way that it could not act as an instrument for the subsequent event. For example a *hammer,* introduced as a possible antecedent instrument for the event *pounding the boards,* was described in the context as a *broken hammer.* In this condition, in contrast to the normal context condition, they found no evidence of a priming effect between *hammer* and *board* in the probe-recognition task. This led them to conclude that the linkage between the instrument and the verb was mediated by the context in which the instrument had been introduced. Their result seems to contradict the present result, which suggests that the *initial* integration of the dominant role-filler happens whatever the contextual appropriateness of the link. The discrepancy between these two findings is probably attributable to the difference in techniques used and, in particular, the point at which the link is being probed. In the present experiment we examined the on-line processing of the role-filler reference and discovered evidence of immediate integration for the dominant role-fillers quickly followed by evidence of contextual evaluation. In the McKoon and Ratcliff study they assessed the final representation of the sentence after it had been understood (following a delay of at least 4.5 s) and in a situation where the instrument had been ruled out on pragmatic grounds. Had they been able to tap into the priming immediately after the critical verb their results might well have been different. We believe that the on-line eye-tracking analysis gives a much clearer picture of the precise time course of the resolution process than is possible with the delayed priming technique.

GENERAL DISCUSSION

The results from the experiment indicate a two-stage process of discourse role resolution. The first stage is consistent with the lexical account and is driven by the lexical link between an antecedent verb and a dominant role-filler. Thus, dominant role-fillers, such as *the pen* for *writing,* are integrated automatically with previous material about writing, whereas nondominant role-fillers, such as *the chalk* for *writing,* are not. Perhaps the most striking finding is that this early integration process is impervious to the influence of the context in which the role was introduced. Thus *writing on a blackboard* is just as effective for initial integration of *the pen* as is *writing a letter* despite the fact that people judge *pen* to be a poor instrument for writing on a blackboard. The experiment also demonstrates that there is a second resolution stage at which context makes an important contribution. Thus measures of subsequent processing difficulty, such as sec-

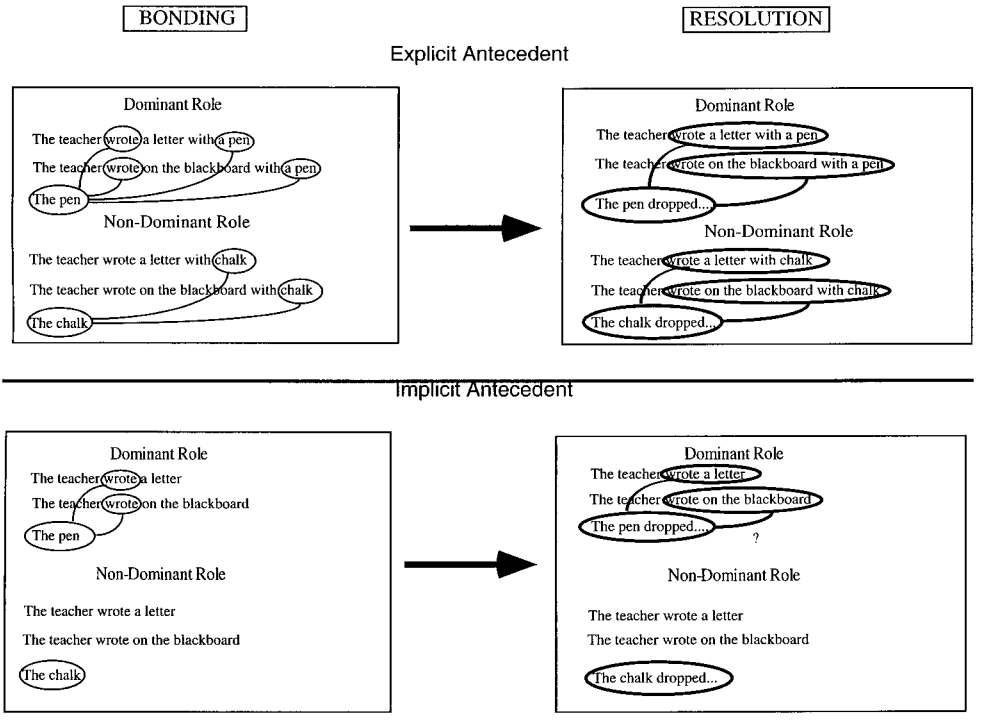

FIG. 3. A schematic representation of the time course of discourse role interpretation in relation to dominant and nondominant verb–role pairs. The left panel illustrates the initial bonding process (thin links), the right panel illustrates subsequent resolution domains (thick links).

ond-pass reading time, show substantial effects of contextual appropriateness. In addition, there is clear evidence that this contextual resolution begins earlier for dominant role-fillers than for nondominant role-fillers. Both the regression-path time analysis for the verb region and the pattern of first-pass regressions from the verb indicate an earlier contextual resolution of the dominant role-fillers.

So the results point to two stages of interpretation with (1) a low-level automatic process associated with establishing some kind of link between the potential role-filler and a previous verb, which we call *bonding,* and (2) a later process which tests and resolves the link with respect to the overall discourse representation, which we call *resolution.* The distinction between low-level bonding processes and high-level resolution processes was first suggested in the context of pronoun interpretation (Sanford, Garrod, Lucas, & Henderson, 1983; Garrod & Sanford, 1990; see also McKoon and Ratcliff, 1989, on a similar distinction which they call recovery versus integration). For pronoun interpretation, it was assumed that readers form a loose superficial attachment between a pronoun and a possible antecedent, bonding, before committing to a full referential interpretation of the pronoun, resolution. For discourse role interpretation, we assume that verbs will only establish bonding links with dominant role-fillers. The verb *write,* for example, would bond with the dominant instrument *pen,* but not the nondominant instrument *chalk.*

To illustrate how this would work, consider Fig. 3 in relation to processing the role references in the materials discussed here. When the critical reference, *the pen* or *the chalk,* is first encountered, the initial bonding takes place (left panel of the figure). For the explicit materials bonds can be set up both with the antecedent noun and, in the case of the dominant *pen,* with the antecedent verb as well (top left panel). However, for the implicit materials only the dominant *pen* will form a bond because the only available bonding site is at the verb and this

only bonds with dominant role-fillers (bottom left panel). On the assumption that bonding signals initial integration of the current material with the context, we would expect failure to bond, as in the case of the nondominant *chalk*, to lead to increased first-pass reading time. This is exactly what happened in the present experiment for the nondominant role-fillers in the implicit conditions.

In the second stage of processing resolution there is also a contrast between the two target types (represented in the right panel of the figure). This is because in the implicit condition the bond for the dominant filler *pen* makes it possible to test the reference earlier against the context and so enables the reader to discover the contextual anomaly at an earlier stage in processing (bottom right panel). Again the pattern of difficulty is consistent with that found in the present experiment: there was an earlier emergence of the context effect for dominant targets (e.g., *pen*) in the regression-path time analysis (i.e., V2) and in the pattern of immediate regressions from the verb. As with interpretation of pronouns, early bonding enables the contextual information to be brought to bear in interpreting the rest of the sentence as soon as it is required. In cases where the role does not fit the most plausible construal of the context situation (e.g., writing a letter with chalk) we must assume that the reader is forced to reconstrue the situation to make it more plausible (e.g., perhaps as a teacher writing the letter on a blackboard as part of a lesson). We suggest that it is this reconstrual process that leads to the increased reading time in the inappropriate context conditions. The discrepancy between the early increase in reading time for the nondominant targets together with evidence of subsequent early contextual disruption for the dominants is what supports the two-stage bonding-resolution account.

At the outset, we raised three unresolved questions about the processing of discourse role links. First, is the process top-down or bottom-up? The results of the experiment are complicated in this respect. On the one hand, it seems that the initial bonding process, as reflected in the lexical effect for nondominant as opposed to dominant role-fillers, relates to a forward association between an antecedent verb and subsequent role-filler. So it could be argued that the effect arises from an intralexical priming process in which the verb *write* speeds up access to the meaning of the dominant role-filler *pen* but not the nondominant role-filler *chalk.* For this to be the case, we would have to assume that the degree and time course of intralexical semantic priming between the verb and its dominant role-filler was equivalent to that of the repetition priming between the antecedent *pen* and the target role-filler *pen* in the explicit condition (i.e., the baseline condition). There are three things that go against this account of the data. First, there is the problem of the distance between the verb and the target role-filler. At least one clause and an average of 12.5 words intervened between the mention of the verb (e.g., *writing*) and the reference to the target role-filler (e.g., *the pen*). It is unlikely that intralexical semantic priming would remain as strong over so much intervening material as the repetition priming in the baseline condition.[4] Second, the lexical effect only emerged weakly in the first-pass reading of the noun itself with the most robust effects coming out in the subsequent verb region as a spillover. As Morris (1994) has demonstrated, such delayed effects in the eye-tracking record reflect postaccess integration rather than intralexical priming. Finally, there is the issue of the absence of early context effects for the nondominant role-fillers. The contexts surrounding the verbs were pretested to ensure that they strongly predicted their appropriate target items. For example, the context for the nondominant item *chalk* contained words such *blackboard* and *exercise,* which would also be expected to prime *chalk.* Yet, there was no evidence of priming effects from the context to the target item in these conditions (e.g., *writing an exercise on the blackboard* does not reduce gaze on *chalk* any more than *writing a letter of complaint*). Hence, we would suggest that the bonding process, like other anaphoric pro-

[4] Instances of long-distance semantic priming (with up to 12 intervening words) have been reported in the literature, but they depend upon retaining the prime word in the focus of attention and probably reflect message level or global context priming (see Foss & Ross, 1983; Hess, Foss, & Carroll, 1995).

cesses, is bottom-up in the sense that it is triggered by the reference to the role-filler itself (See Garrod et al., 1990).

This brings us to the second question raised in the introduction: the extent to which resolution depends on the semantic representation of the antecedent verb as opposed to the overall situation portrayed in the antecedent sentence. The results of the experiment indicate that the initial bonding depends on the semantic representation of the verb, whereas later resolution depends on a representation of the antecedent event as a whole. In light of the discussion above, this raises questions about precisely how the bonding process can operate in such a way that it is triggered by reference to the role-filler but controlled by the representation of the verb. In line with Garrod et al. (1990; see also Gernsbacher, 1989) we suggest that the bonding processes may operate in a similar fashion to retrieval processes in models of memory proposed by Hintzman (1986) and Ratcliff (1978). Initially, the target reference broadcasts to all matching antecedents in the discourse representation in parallel. Potential antecedents then echo or resonate to the extent that they match the semantic features of the referent. In the first instance, both the dominant and nondominant role-fillers will match to some degree the antecedent verb because there is a strong backward association between role-filler and verb irrespective of dominance. However, only in the case of the dominant role-fillers will this be reflected in a reciprocal resonance from the verb to the role-filler because there is only a strong forward association between verbs and their dominant role-fillers. We suggest that it is this combination of backward and forward association that underlies the initial bonding process between the verb and dominant role-filler. What is new in this account is the idea that antecedent verbs as a result of their relationship with role-fillers play a part in anaphoric bonding in much the same way as antecedent nouns.

Finally, we raised the issue of the time course of the interpretation process. The experiment indicates that discourse role interpretation, like that of other kinds of anaphora, is an on-line process initiated soon after encountering the critical reference. However, like pronoun interpretation it occurs in two stages; first, an automatic bonding stage that only takes account of limited lexical information from the prior discourse, and second, a resolution stage in which the link is checked against the broader context.

REFERENCES

Carlson, G. N., & Tanenhaus, M. K. (1988). Thematic roles and language comprehension. In W. Wilkins (Ed.), *Syntax & semantics: Thematic relations* (Vol. 21). London: Academic Press.

Collins Cobuild English Language Dictionary (1987). London: Collins.

Corbett, A. T., & Dosher, B. A. (1978). Instrument inferences in sentence encoding. *Journal of Verbal Learning and Verbal Behaviour,* **17,** 479–491.

Cotter, C. A. (1984). Inferring direct objects in sentences: Some implications for the semantics of Verbs. *Language and Speech,* **27,** 25–45.

Dosher, B. A., & Corbett, A. T. (1982). Instrument inferences and verb schemata. *Memory & Cognition,* **12,** 531–539.

Ehrlich, K., & Rayner, K. (1983). Pronouns assignment and semantic integration during reading: Eye-movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behaviour,* **22,** 75–87.

Foss, D. J., & Ross, J. R. (1983). Great expectations: Context effects during sentence processing. In G. B. Flores d'Arcais & R. J. Jarvella (Eds.), *The process of language understanding* (pp. 169–191). Chichester, UK: Wiley.

Garrod, S. C. (1994). Resolving pronouns and other anaphoric devices: The case for diversity in discourse processing. In C. Clifton, Jr., L. Frazier, & Rayner, K. (Eds.), *Perspectives on sentence processing* (pp. 339–359). Englewood, NJ: LEA.

Garrod, S., Freudenthal, D., & Boyle, E. (1994). The role of different types of anaphor in the on-line resolution of sentences in a discourse. *Journal of Memory and Language,* **33,** 39–68.

Garrod, S., O'Brien, E. J., Morris, R. K., & Rayner, K. (1990). Elaborative inferencing as an active or passive process. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **16,** 250–257.

Garrod, S., & Sanford, A. J. (1977). Interpreting anaphoric relations: the interpretation of semantic information while reading. *Journal of Verbal Learning and Verbal Behavior,* **16,** 77–90.

Garrod S. C., & Sanford A. J. (1981). Bridging inferences in the extended domain of reference. In J. Long & A. Baddeley (Eds.), *Attention & Performance* (Vol. IX, pp. 331–346). Hillsdale, NJ: Elbaum.

Garrod, S. C., & Sanford, A. J. (1990). Referential processing in reading: Focusing on roles and individuals. In D. A. Balota, G. B. Flores d'Arcias, & K. Rayner (Eds.) *Comprehension processes in reading* (pp. 465–486). Hillsdale, NJ: Erlbaum.

Garrod, S., & Sanford, A. J. (1994). Resolving sentences in a discourse context: How discourse representation af-

fects language understanding. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 675–698). New York: Academic Press.

Gernsbacher, M. A. (1989). Mechanisms that improve referential access. *Cognition,* **32,** 99–156.

Haviland, S. E., & Clark, H. H. (1974). What's New? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior,* **13,** 512–521.

Hess, D. J., Foss, D. J., & Carroll, P. (1995). Effects of global and local context on lexical processing during language. *Journal of Experimental Psychology: General,* **124,** 62–82.

Hintzman, D. L. (1986). "Schema abstraction" in a multitrace memory model. *Psychological Review,* **93,** 411–428.

Johnson, M. F., Bransford, J. D., & Solomon, S. K. (1973). Memory for tacit implications of sentences. *Journal of Experimental Psychology,* **98,** 203–205.

Lucas, M. M., Tanenhaus, M. K., & Carlson, G. N. (1990). Levels of representation in the interpretation of anaphoric reference and instrument inference. *Memory & Cognition,* **18,** 611–631.

Marslen-Wilson, W. D., Tyler, L. K., & Koster,C. (1993). Integrative processes in utterance resolution. *Journal of Memory and Language,* **32,** 647–666.

Mauner, G., Tanenhaus, M. K., & Carlson, G. N. (1995). Implicit arguments in sentence processing. *Journal of Memory and Language,* **34,** 357–382.

McKoon, G., & Ratcliff, R. (1981). The comprehension processes and memory structures involved in instrumental inference. *Journal of Verbal Learning and Verbal Behavior,* **20,** 671–682.

Morris, R. K. (1994). Lexical and message-level sentence context effects on fixation times in reading. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **20,** 92–103.

O'Brien, E. J., Shank, D. M., Myers, J. L., & Rayner, K. (1988). Elaborative inferences during reading: Do they occur on-line? *Journal of Experimental Psychology: Learning, Memory and Cognition,* **14,** 410–420.

Paris, S. G., & Lindauer, B. K. (1976). The role of inferences in children's comprehension and memory for sentences. *Cognitive Psychology,* **8,** 217–227.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review,* **85,** 59–108.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects or word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition,* **14,** 191–201.

Rumelhart, D. E. (1975). Notes on a schema for stories. In D. G. Bobrow & A. M. Collins (Eds.), *Representation and understanding: Studies in cognitive science.* New York: Academic Press.

Sanford, A. J., & Garrod, S. (1989). What, when and how: Questions of immediacy in anaphoric reference resolution. *Language and Cognitive Processes,* **4,** 235–263.

Sanford, A. J., Garrod, S. C., Lucas, A., & Henderson, R. (1983). Pronouns without explicit antecedents. *Journal of Semantics,* **2,** 303–318.

Schank, R. C. (1972). Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology,* **3,** 552–631.

Schank, R., & Abelson, A. (1977). *Scripts, plans, goals, and understanding: An enquiry into human knowledge structures.* Hillsdale, NJ: Erlbaum.

Singer, M. (1979). Process of inference during sentence encoding. *Memory & Cognition,* **7,** 192–200.

Tanenhaus, M. K., & Carlson, G. N. (1989). The role of thematic structures in interpretation and parsing. *Language and Cognitive Processes,* **4,** 211–234.

Tanenhaus, M. K., & Carlson, G. N., (1989). Lexical structure and language comprehension. In W. Marslen-Wilson (Ed.), *Lexical representation and process.* Cambridge, MA: MIT Press.

Terras, M. T. (1997). *Discourse roles: A mechanism to establish coherence.* Ph.D. dissertation, University of Glasgow, Glasgow.

Trabasso, T., & Sperry, L. L. (1985). Causal relatedness and importance of story events. *Journal of Memory and Language,* **24,** 612–630.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic disambiguation. *Journal of Memory and Language,* **33,** 258–318.

Van den Broek, P. (1988). The effects of causal relations and hierarchical position on the importance of story statements. *Journal of Memory and Language,* **27,** 1–22.

(Received May 18, 1999)
(Revision received October 28, 1999)