*Journal of Semantics* 7: 221–244 

# The Cognitive Resolution of Anaphoric Noun References

JOCHEN MÜSSELER and GERT RICKHEIT
*University of Bielefeld, Federal Republic of Germany*

## Abstract

The cognitive resolution of references depends on whether an anaphor has previously been mentioned in the antecedent text or whether a relation between the anaphor and the antecedent has to be constructed by means of an inference. This inference process in comparison with a mere concept repetition produces an increase in reading and comprehension time. In the present paper aspects of this processing difference are examined. Recognition data are used in addition to reading and comprehension times in order to compare the mental representations generated by the different processes.

Experiment 1 shows that the differences in processing time not only depend on the repetition of the concept, but are also produced by varying the semantic distance between concepts. Differences in recognition performance were not observed, indicating that the resulting text representations were similar on a semantic level. In Experiment 2 the point of time of the inference process during text processing was determined by a word-by-word presentation. The results show that the inference process starts immediately after the reception of the critical reference concept and ends with the completion of the proposition. Finally, Experiment 3 examines whether inference processes are different after specification or generalization of a previously mentioned concept. Here the inference effect occurs with specification rather than generalisation. The recognition data indicate that it is questionable whether an inference process takes place when a concept is generalised.

## INTRODUCTION

In order to understand a text it is important to relate different parts to larger semantic units. This process is controlled by various cognitive mechanisms in which reference processes within and between sentences play a central part. These references are based on the knowledge of the reader or listener and are responsible for the integration of text information into the cognitive representations (Rickheit, Schnotz and Strohner 1985, Schnotz 1988). In text processing research, different referential relations have been studied. For the cognitive resolution of *anaphoric noun references* it was shown by Haviland and Clark (1974) and Clark (1977) that additional time was needed in processing the so called '*bridging inferences*' (see also Tanenhaus and Seidenberg 1981, Sanford and Garrod 1981, Dell, McKoon and Ratcliff 1983, Müsseler, Rickheit and Strohner 1985):

(1a′) Horace got some *beer* out of the car. (close inference)
(1b) The *beer* was warm.

(1a″) Horace got some *picnic supplies* out of the car. (distant inference)
(1b) The *beer* was warm.

The referential distance between the first and the second sentence varies only through a specification that is expressed in the concepts 'picnic supplies' vs. 'beer'. In the second example there is no immediate relationship between these concepts. The relationship has to be constructed in the knowledge of the reader who has to identify 'beer' as a 'picnic supply'. This assumed inference process is not necessary in the close inference condition because of the repetition of the word 'beer'. The postulated inference is empirically supported by comprehension and reading time differences. In the distant inference version the processing times of the second sentence differ by about 180 msec (Haviland and Clark 1974) and 270 msec (Müsseler, Rickheit and Strohner 1985) compared with the close inference condition.

Although it is plausible to characterise these processing time differences as differences in inference processes there are several open questions: contrary to the example mentioned above, a variation in the *semantic distance* between the concepts to be linked seems to be more informative (see also Garrod and Sanford 1977), especially to exclude a mere word repetition effect caused by the close inference condition. For example, when substituting 'beer' by 'drinks' in sentence (1a′), the semantic distance between 'drinks' and 'beer' still seems to be smaller than between 'picnic supplies' and 'beer'. The time difference should be preserved. Additionally it has to be checked that the sentences in both conditions are combined and integrated in an appropriate way. Only then should the reception of the close and distant inference condition lead to comparable representations for the text as a whole. In the following Experiment 1 we therefore combined on-line (i.e. the reading and comprehension times) with off-line measurements (i.e. a recognition test).

Another question is at what point is the inference process performed during the reception phase. Is it already initiated and completed when reading the concept 'beer' in the second sentence or does the inference process lag behind the encoding of this concept? Experiment 2 is concerned with this question. In the last experiment different text situations are examined in which the inference process becomes necessary after specification or generalisation of an antecedent.

## EXPERIMENT 1

In this experiment the inference effect is replicated under controlled conditions in order to eliminate the following possible explanations of the reading and comprehension time differences.

One can say that the relevant concepts in the distant inference condition are not combined in a suitable way so that time differences are derived from the two concepts existing side by side or 'in parallel' in which case they are not necessarily connected. The prolonged processing time is then due to a change of topic (Just and Carpenter 1980, Haberlandt 1984) and/or an *intended* inference process, which is not necessarily performed. The characteristic *integration* could be missing.

Even if one assumes the integration process to have taken place, it is not clear whether the text representations due to the different inference processes are comparable. Reading and comprehension time differences can only be of interest if one can assume that comparable text representations are produced after encountering both text versions. That is to say, on reception of two totally different texts one would never think of comparing reading and comprehension time differences.

Furthermore, if we assume that encountering both text versions leads to comparable text representations, it has to be clarified that shorter comprehension time is not due to the repetition of an anaphor (see also Haviland and Clark 1974, Haberlandt and Graesser 1985). Accordingly, comprehension time differences are not necessarily due to a prolonged inference process but rather to a word repetition effect. In other words, short processing times would then result because of word repetitions in the close inference condition.

The following experiment tries to exclude these possible explanations. By varying the semantic distance between concepts, the word repetition effect is controlled (as done by Garrod and Sanford 1977). To make sure that the integration process has taken place and the resulting comparable text representations have been attained, additional data were collected from a recognition test. However, in order to guarantee that the relevant inference process did not take place during the recognition phase, subjects were only confronted with the recognition tests after the presentation of several experimental text blocks. It should then be impossible for the subjects to remember semantic or syntactic details of the text surface. Additionally, in this experiment, we examined whether the inference process only affects the local processing time of the critical sentences or whether it is also carried over to the following sentences.

## *Method*

### Subjects

Thirteen female students and 17 male students of different faculties of the University of Bielefeld took part. They were paid for participating. Their mother tongue was German and their average age was 23.23 years.

## Reception Texts

Sixty four-sentence texts were constructed. For each text there was a distant and close inference version (altogether 120 texts). The sentences 1, 3 and 4 of the two text versions were exactly the same; only the second sentence was modified in such a way that the inference to the third sentence was either more difficult or more easy to perform. This was achieved by using a concept which was immediately related to the word, or a concept which belonged to the presented scenario. In contrast to the experiment of Haviland and Clark (1974), the concept repetitions are not inherent for the close inference condition. The following text examples (translated from German) make this clearer:

close inference version:
(2a) Karin drove with her car through the countryside.
(2b′) Karin glanced across to the *field*.
(2c) The *grain* was ripe and was being harvested by the farmers.
(2d) The people were busy in the fields.

distant inference version:
(2a) Karin drove with her car through the countryside.
(2b″) Karin glanced across to the *mountain range*.
(2c) The *grain* was ripe and was being harvested by the farmers.
(2d) The people were busy in the fields.

The first sentence is the setting. The concept 'field' in the second sentence of the close inference version (2b′) should be related easily to the concept 'grain' of the third sentence (2c); this is not the case with the concept 'mountain range' in the distant inference version (2b″). The definite article 'the' of the sentences (2c) marks the anaphoric relation to the previous sentence.

## Recognition Sentences

For each of the 120 texts three different types of recognition sentences were constructed from the second sentence: *identical sentences*, *performed inferences* and *elaborations*.

*Identical sentences* represented a mere repetition of the second sentence:

(2b-1) Karin glanced across to the field. (identical sentence)

or for the distant text version:

(2b-2) Karin glanced at the mountain range. (identical sentence)

Alternatively, the *performed inferences* represented a text content probably resulting from the inference process between the second and third sentence. If

'Karin glanced across the field' (or 'the mountain range') and it was stated that 'The grain was ripe', then this implies that

(2b-3) Karin glanced at the grain. (performed inference)

The third type of recognition sentences contained aspects which were clearly *elaborative*. With such sentences it should be tested whether the subjects were ready to relate text concepts within the recognition phase. For example, in the text it is not mentioned (although possible) that 'Karin, leaving the car, glanced at the field'. This makes it possible that

(2b-4) Karin glanced at the car. (elaboration)

In addition to these three types of recognition sentences, distractor sentences were used. They were taken from other parts of the texts (sentences (2a) or (2d)). All recognition sentences were true with respect to the text.

## Design

One-half of the subjects were given 30 distant inference texts and 30 close inference texts. The other half of the subjects were given the complementary versions. This results in a two-way factorial design with the factors text version (close vs. distant) and the processing times of the third and fourth sentences following the critical second sentence. The dependent variable was the reading time for each sentence.

To each text two recognition sentences were given: an identical sentence, a performed inference, an elaboration or a distractor. Altogether 60 recognition sentences and 60 distractors were presented. Half of these were mere repetitions of the text, the other half had been modified to perform inferences or elaborations.

Accordingly, the recognition phase was based on a two-way factorial design with the factors text reception (distant vs. close version) and type of recognition sentence (identical, performed inference, elaboration). Each cell of the design was filled with 10 recognition sentences. The dependent variable was the recognition frequency.

## Apparatus and Procedure

The texts were displayed on a personal computer. The first sentence was presented in the middle of the screen. As soon as the subject had read the text, he or she had to press a button for the presentation of the following sentence. This appeared below the first sentence but without the previous sentence disappearing. In this way all four sentences were on display at the end of each text.

The subjects were not informed that reading times were measured. They were instructed to read as 'normally' as possible and to avoid repeated reading.

The texts were presented in six blocks with ten four-sentence texts respectively. After each text there was a pause of 4 seconds before the next text appeared. After a text block the recognition items were given in random order. For this task the subjects had to press two other buttons (recognised vs. not recognised). They received no feedback on the correctness of their answers.

## *Results*

### Reception Phase

Extreme reaction times (M $\pm$ 2*SD) were replaced by the average reaction time of the subject per cell of the design. In accordance with Clark (1973), reading and comprehension times were dependent variables in two ANOVAs. Text version and sentence position were 'fixed effects'. 'Random effects' were in one ANOVA the subjects, in the other the texts. The resulting F-values were transformed to min F'.

Figure 1 shows the mean reading times. The analyses with subjects ($F_S(1,29) = 4.52$, $MS_E = 1408308.43$, $p < 0.05$) and with texts ($F_T(1,118) = 64.01$, $MS_E = 198925.22$, $p < 0.001$) show a significant increase in reading times between the third and fourth sentence; accordingly, the min $F'(1,33) = 4.22$, $p < 0.05$ was

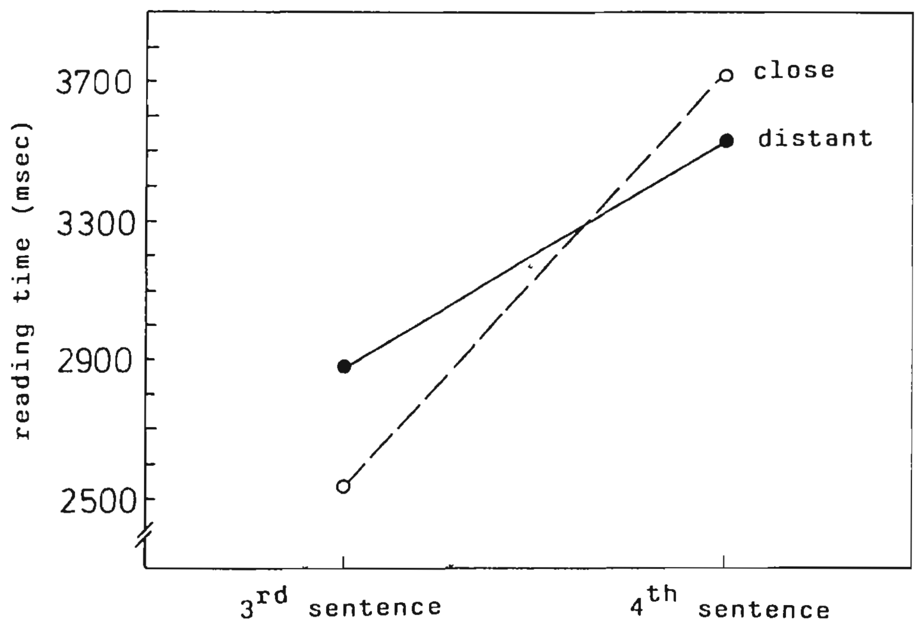


**Figure 1** Average reading times of the critical third and the following fourth sentence in the distant and close inference condition.

also significant. The main factor text version had no effect on the results ($F_S(1,29) = 1.32$, $MS_E = 35667.55$, n.s.; $F_T(1,118) = 0.47$, $MS_E = 197870.31$, n.s. and min $F'(1,65) = 0.17$, n.s.).

The interaction of the factors proved to be significant ($F_S(1,29) = 15.14$, $MS_E = 35771.30$, $p < 0.001$; $F_T(1,118) = 5.45$, $MS_E = 198925.22$, $p < 0.05$); min $F'(1,143) = 4.01$, $p < 0.05$). This is due to the increased reading times of the third sentence and the shorter reading times of the fourth sentence in the distant text version in comparison with the reading times of the close text version. The Scheffe Test shows a hybrid interaction: the two inference versions of the third sentence differ ($p < 0.01$ with subjects, $p < 0.05$ with texts), whereas in the fourth sentence only a tendency can be noted ($p < 0.10$ with subjects, n.s. with texts).

## Recognition Phase

The average recognition frequencies were dependent variables in two ANOVAs one with subjects, the other with sentences. The results are shown in Figure 2.

An effect can be noted with the different types of recognition sentences ($F_S(2,58) = 353.13$, $MS_E = 0.018$, $p < 0.001$; $F_T(2,354) = 275.66$, $MS_E = 0.047$, $p < 0.001$ and min $F'(2,238) = 154.18$, $p < 0.001$). Identical sentences were more easily recognised than performed inferences and the latter were better than elaborations. As in the reception phase, the factor text version had no effect

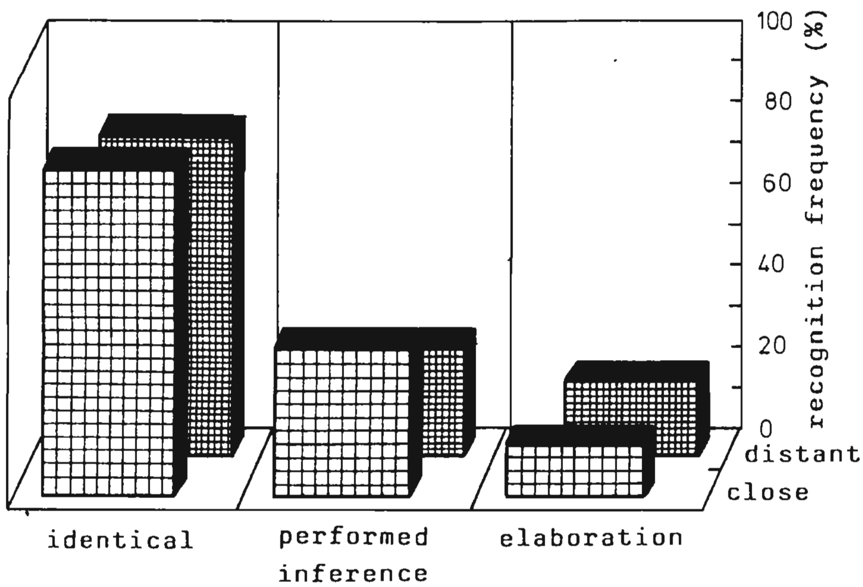


**Figure 2** Mean frequencies of the recognised sentences determined by the presented texts (close vs. distant inference condition) and the three different types of recognition sentences (identical, performed inference and elaboration).

($F_S(1,29) = 0.73$, $MS_E = 0.022$, n.s.; $F_T(1,354) = 0.68$, $MS_E = 0.047$, n.s. and min $F'(1,114) = 0.35$, n.s.).

Again, interaction between the factors is critical. Even though the 5 per cent level of significance was exceeded in both ANOVAs ($F_S(2,58) = 4.02$, $MS_E = 0.023$, $p < 0.05$ and $F_T(2,354) = 3.89$, $MS_E = 0.047$, $p < 0.05$), there is no significant effect with the min $F'$ value (min $F'(2,104) = 1.98$, n.s.). With reference to this result, the recognition frequencies of the performed inferences and the elaborations do not differ in the two inference versions close and distant.

## *Discussion*

The critical reading time difference of approximately 175 msec equals the results of Haviland and Clark (1974). Apart from this, the time difference is not only due to the concept repetition but is also realised when the conceptual distance is varied. This basically replicates in German an earlier finding by Garrod and Sanford (1977) using English texts. Secondly, it was shown that the inference effect was carried over only to the third sentence and could not be found in the fourth. The general increase in reading time of this sentence is probably due to the end of an episode, i.e. the end of the text (Haberlandt 1980, Haberlandt, Berian and Sandson 1980). Thirdly, if the recognition phase can actually be judged as a criterion of mental representations, it can be concluded that an integration process has taken place, resulting in similar representations after the reception of distant and close text versions.

However, this last interpretation has to be restricted: it is only acceptable if the inference process is not ascribed to the recognition phase (see above). In addition, it can only concern the retrieval of semantic memory traces. Nevertheless, differences may exist on a text surface level as indicated by the difference between the recognition of identical sentences and the performed inferences. Finally, there is a marginally significant interaction between the type of recognition sentence and the close vs. distant text reception. Although not reliable by min $F'$, this could indicate a weak difference in representation as a function of presentation condition. In the following experiments we will keep this interaction in mind.

## EXPERIMENT 2

One of the results of the first experiment was that the inference process measured in reading and comprehension times is local and does not extend to the following sentence. The question is still open as to at what point of the critical third sentence the inference is performed.

Just & Carpenter (1980) interpret their results from eye movement experiments in relation to a multiple processing levels model. The information is encoded, referred to semantic contents and stored in memory. Firstly, they assume that the processing of a word starts *immediately* on all levels after reception (*immediacy assumption*). Secondly, it is postulated that the processing time of a word is reflected in its fixation time (*eye-mind assumption*). The eye fixation of a word is as long as one needs to process it. Even the reference process to other previously read parts of the text are reflected in the fixation time.

These assumptions have been empirically and theoretically criticised. Ehrlich and Rayner (1983) postulate that eye movement merely serves to supply information to a memory which can then be referred to during complex text processing. In this way the semantic processing of a word and the resolution of reference can lag behind the encoding of a word (cognitive lag, Rayner 1978).

With regard to the inference process, it can be postulated that following the first assumption the inference process starts immediately after reading the critical concept and is also finished at that point. If we follow the second assumption this inference process takes place later.

These opposing hypotheses were tested with a word-by-word presentation called the moving-window technique. The text presentation was continuous and stops at critical text positions. It is assumed that the reaction times at the point of restart after a stop should reflect the necessary processing times.

## *Method*

### Subjects

In this experiment 10 male and 20 female students of the University of Bielefeld participated and were paid afterwards.

### Reception Texts and Recognition Sentences

Exactly the same texts as in Experiment 1 were used, except that 30 distractor texts were interspersed. The distractors had the same structure as the other texts and served to prevent the subjects from predicting the stop in the critical third sentence (see below). Accordingly, the distractor texts never stopped in this sentence.

The experimental texts stopped in three controlled positions in the third sentence: (1) after the noun phrase (position 1); (2) after the verb phrase (position 2); and (3) at the end of the sentence (position 3). Each subject only encountered one of the three stops in this sentence. Additionally all texts were stopped unsystematically up to four times in different places per text. The following example shows the position of stops in the critical third sentence (marked with *).

(2c) The grain* was ripe* and was being harvested by the farmers*.

## Design

The position of the three stops in the third sentence was varied in the close and distant inference versions. All subjects were randomly confronted with all combinations of the two (2 × 3–) factors.

The design of the recognition phase was extended by one additional factor resulting from the different positions of stops in the critical third sentence. This turned out to be a three-way factorial design with the factors text version (close vs. distant), the position of the three stops (after the noun phrase, after the verb phrase, and at the end of the sentence) and the type of recognition sentences (identical, performed inference, elaboration).

## Procedure

The texts were presented word-by-word with the moving-window technique on a computer screen. Each letter of a word was masked with a '+' to relieve eye fixation. One word was presented for 300 msec[1] and subsequently replaced by its mask. Immediately following was the presentation of the next word. As mentioned above, the text stopped at critical positions. At the point of stopping the subject had to press a button to continue with the reading of the text. The reaction time for this task was registered.

The texts were presented in six blocks containing 10 experimental texts and five distractors. The recognition test was between each block as in Experiment 1.

# *Results*

## Reception Phase

Again the extreme reaction times were replaced by their average reaction time. The ANOVA shows a significant result for the factor inference ($F_S(1,29) = 8.33$, $MS_E = 17066.41$, $p < 0.01$; $F_T(1,59) = 8.08$, $MS_E = 34969.32$, $p < 0.01$ and min $F'(1,79) = 4.11$, $p < 0.05$). With the text versions of the type distant inference the reaction times are slower than with the close inference (see Figure 3).

The position factor shows only a tendency in the analyses for subjects ($F_S(2,58) = 2.41$, $MS_E = 17797.09$, $p < 0.10$) and for sentences a statistical

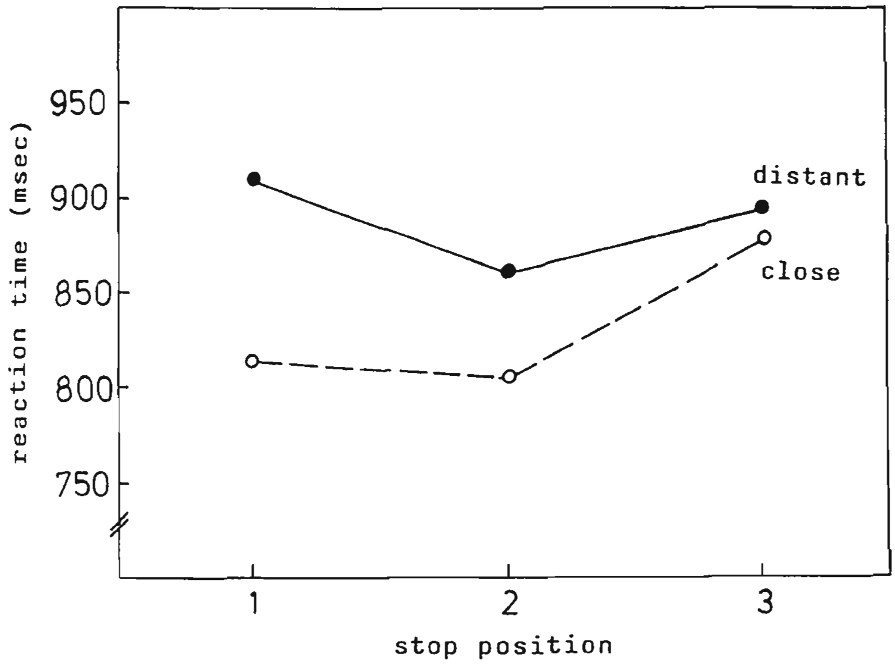


**Figure 3** Mean reaction times of the critical stops (position 1: noun phrase, position 2: verb phrase and position 3: at the end of the sentence) determined by the two versions of close and distant inference.

significance ($F_T(2,118) = 3.81$, $MS_E = 22499.94$, $p < 0.05$). The min F′ value is not significant (min $F'(2,129) = 1.48$, n.s.). A similar result shows the interaction of the two factors ($F_S(2,58) = 3.47$, $MS_E = 7011.27$, $p < 0.05$; $F_T(2,118) = 2.91$, $MS_E = 16605.93$, $p < 0.10$ and min $F'(2,164) = 1.58$, $p < 0.25$). However, Scheffe Tests for the stopping positions show significant results in position 1 (for subjects and texts $p < 0.01$) and 2 (for subjects and texts $p < 0.05$) with the distant and close inference condition. Position 3, however, does not.

## Recognition Phase

The frequency data in this experiment were based on only three to four recognition sentences per cell and subject, because we used the same recognition sentences as in Experiment 1 despite one additional factor (i.e. three positions of stops). For this reason the mean recognition frequency of about 100 recognition presentations of all subjects and sentences served as estimated values in an ANOVA which used only one observation unit per cell (see Bortz 1977, p. 396). This statistical procedure uses as error variance the residual variation in which the interaction variance and error variance of the second order are confounded. With the use of the additivity test of Tukey (1949) it was examined whether there was any interaction at all. This seems to be the case

($F(1,3) = 89.54$, $MS_{Balance} = 0.28$, $p < 0.001$), although interpretations on the main factors and any interaction effects should be more conservative.

However, the different types of recognition sentences show a significant effect as in Experiment 1 ($F(2,4) = 513.50$, $MS_{Residual} = 6.47$, $p < 0.001$, see Figure 4). Additionally, the tendency to recognise sentences from the close inference condition rather than distant inference condition is higher than chance ($F(1,4) = 8.79$, $MS_R = 6.47$, $p < 0.05$). These main effects, however, should be interpreted with their interaction ($F(2,4) = 13.71$, $MS_R = 6.47$, $p < 0.05$). As in Experiment 1, there is a tendency to recognise elaborative sentences of the distant inference version rather than those of the close inference. Surprisingly, this effect is reversed with the identical sentences. For the critical performed inferences there seems to be no difference in the two conditions.

Additionally, the interaction of the position factor and type of recognition sentence shows significance ($F(4,4) = 11.82$, $MS_R = 6.47$, $p < 0.05$, see Figure 5). With the recognition of the identical sentences there is no evidence that the stop position is in any way relevant, whereas with the performed inference it seems that they are recognised less when stopped at the second position and hardly ever recognised when stopped at the end of the third sentence. On the other hand, this tendency is reversed with the elaborations. There is no difference in recognition frequencies when stopped at the third position with both the performed inferences and elaborations.

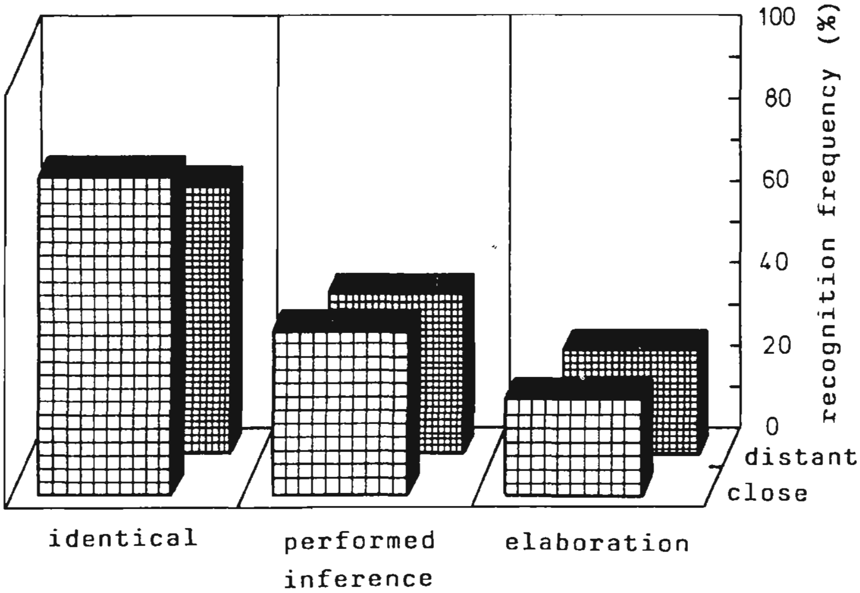


**Figure 4** Mean frequencies of the recognised sentences determined by presented texts (close vs. distant inference condition) and the three different types of recognition sentences (identical, performed inference and elaboration). The factor 'position of stops' was averaged.

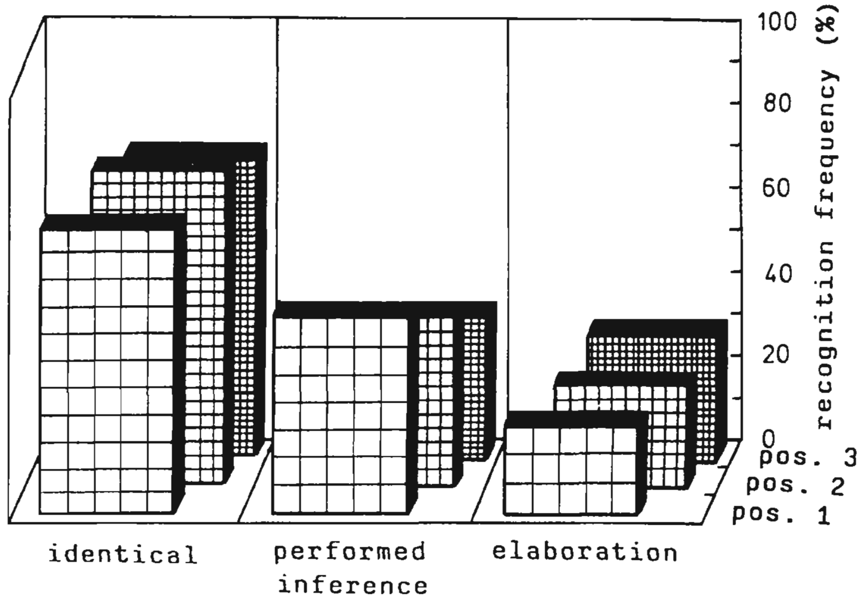


**Figure 5** Mean frequencies of the recognised sentences determined by the position of stops (positions 1–3) and the three different types of recognition sentence (identical, performed inference, elaboration). The factor 'inference version' was averaged.

## *Discussion*

Looking at the results of the reception phase, the inference is performed immediately after reading the critical concept (position 1) as well as after the verb phrase (position 2). The reaction time differences of about 97 msec and 56 msec at these two positions added together give almost the time that is needed for reading the whole sentence as seen in Experiment 1. This result complies with the tendency of the recognition phase to perform the inference when stopped in the first positions of the sentence. This implies that the inference is only performed by the subject when the necessary processing time is allowed. With our means of presentation this only seems to be the case when the text is stopped at the appropriate place. Accordingly, one can conclude that the process of inference is performed immediately after the reception of the critical word concept and finishes with the construction of the first semantic proposition.

However, one problem with the procedure is that readers in a free situation may spend different amounts of time fixating words in different parts of the sentence. This is not taken into account with our procedure. Word-by-word presentation can possibly produce artificial effects, making it uncertain whether positions 1 and 2 reflect the time of the inference in normal reading. Indeed, results from experiments based on artificial reading methods have always to be treated cautiously (Günther 1989). On the other hand, different studies show that a word-by-word presentation is comparable with normal reading (Just, Carpenter and Woolley 1982, Ward and Juola 1982, Juola, Ward and McNamara 1982). Additionally, we demonstrated in a previous study that a presentation similar to the one used here led to comparable text representations on a propositional level (Müsseler and Nattkemper 1986).

## EXPERIMENT 3

In this experiment texts are introduced which make inferential references of different qualities necessary. One can vary the semantic distance of concepts when the antecedent is *specified* differently. The texts of Experiments 1 and 2 were principally constructed in this way. On the other hand, concepts can be *generalised* differently. This gives rise to the question of how manipulations of this sort induce different information processing. Garrod and Sanford (1977, Exp. 1) followed up this variation with the following pairs of sentences:

(4a) A *robin* would sometimes wander into the house.
(4b) The *bird* was attracted by the larder.

(5a) A *bird* would sometimes wander into the house.
(5b) The *robin* was attracted by the larder.

They found that sentence (4b) was read faster than sentence (5b). This would seem to indicate a processing advantage for the anaphoric generalisation. However, the two sentences (4b) and (5b) differ so that the reading times may not be comparable at all.

One could proceed with the variation in a different way: when a concept is specified (in the following 'zooming-in' version), for example, 'animal-fish-shark', the semantic distance of 'fish' and 'shark' is shorter than that of 'animal' to 'shark'.[2] On the word level there are a number of experiments and theoretical discussions which have dealt with such semantic distance effects (Wilkens 1971, Meyer, Schvanefeldt and Ruddy 1975). For the present experiment examples of the following sort are taken:

'zooming-in', word level:
(6a′) The schoolgirl admired *the fish*. (close inference)
(6b) *The shark* was very aggressive.

(6a″) The schoolgirl admired *the animal*. (distant inference)
(6b) *The shark* was very aggressive.

Processing time differences are to be expected because the unspecific concept animal can be split into many subcategories. The number of subcategories in the close inference version is clearly more restricted.

If one reverses the concepts and proceeds by going from the specific to the general category 'shark–fish–animal' (in the following 'zooming-out' version), it is realised on a textual basis in the following way:

'zooming-out', word level:
(7a′) The schoolgirl admired *the fish*. (close inference)
(7b) *The animal* was very aggressive.

(7a″) The schoolgirl admired *the shark*. (distant inference)
(7b) *The animal* was very aggressive.

A processing time difference can be expected only when the next highest category is activated; in this case 'fish' is associated with 'animal' and 'shark' only with 'fish'. However, because the number of main categories is restricted in generalisation as opposed to when they are specified, one does not necessarily anticipate processing time differences. With the concept 'animal' no new information is introduced at the point of text reception which the processing system does not already know. In comparison with the specification, the text does not introduce something new because it is already in the world knowledge of the reader. The concept 'fish' semantically implies 'animal'. In other words, the definite article and the concept 'the animal' are a signal for focus maintenance in the 'zooming-out' version and do not announce anything new, whereas in the 'zooming-in' version a mental *focus tracking* (e.g. Sidner 1983, Schnotz 1986) is probably necessary because of specification.

To control this type of influence an additional variation is introduced that is neutral in this respect: the 'zooming-in' or 'zooming-out' in a spatial scenario ('mountain range–field–grain' vs. 'grain–field–mountain range'):[3]

'zooming-in', spatial scenario:
(8a′) Karin glanced at *the field*. (close inference)
(8b) *The grain* in the field had grown.

(8a″) Karin glanced at *the mountain range*. (distant inference)
(8b) *The grain* in the field had grown.

'zooming-out', spatial scenario:
(9a′) Karin glanced at *the field*. (close inference)
(9b) *The mountain range* was impressive.

(9a″) Karin glanced at *the grain*. (distant inference)
(9b) *The mountain range* was impressive.

Comparing the examples (8) and (9), enough new information in the second sentences is introduced in order for the reader to make a spatial focus change of the following discourse or theme necessary. However, this focus change should be performed more easily in the close inference versions independent of the zooming condition.

## *Method*

### Subjects

Forty students of the University of Bielefeld took part in this experiment. They were paid for participating.

### Reception Texts

From Experiment 1 and 2, 40 texts were modified to correspond to the example texts of (6) to (9). Mainly the second and third sentences were changed. The first and last sentences remained. In addition, 20 distractor texts were introduced.

### Recognition Sentences

For each of the 40 experimental texts recognition sentences of the type 'performed inference' were constructed. In addition 60 distractors were used.

### Design

Half of the subjects were given the 'zooming-in' versions, the other half the 'zooming-out' versions. Within subjects the inferential factor ('close' vs. 'distant') and the referential context factor ('word level' corresponding to examples (6) and (7) vs. 'scenario' corresponding to examples (8) and (9)) were varied.

### Procedure

The texts were read as in Experiment 1 sentence-by-sentence. The same procedure was used except for the fact that the 60 four-sentence texts were divided in four experimental blocks of 15 texts.

## *Results*

### Reception Phase

All extreme reading times were excluded from the ANOVA. As in Experiments 1 and 2 the inference effect between the close and distant condition is evident in the 'zooming-in' version (Figure 6). In the 'zooming-out' version it can only be found in the spatial scenario condition.

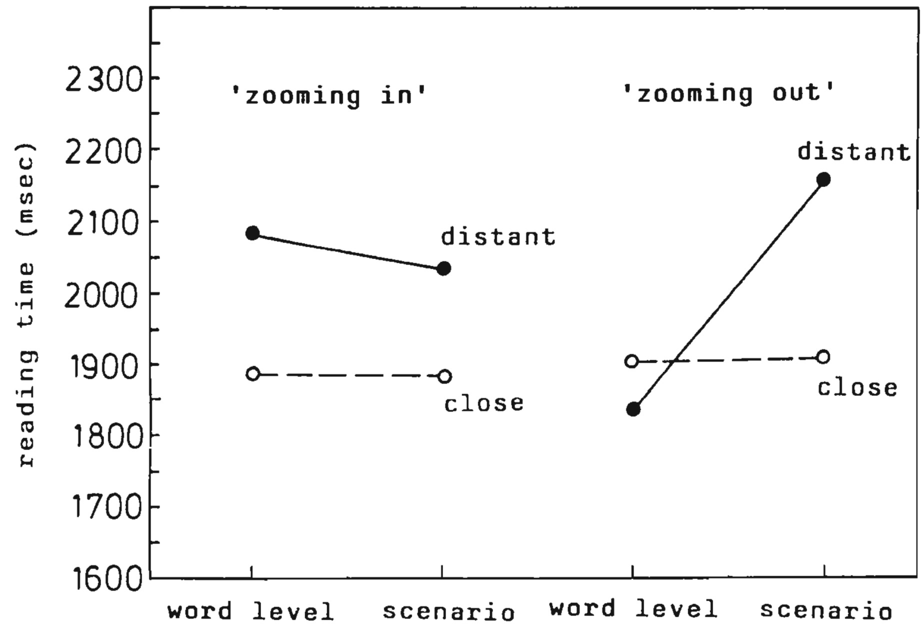


**Figure 6** Mean reading times of the critical third sentence determined by the inference conditions (close vs. distant), the referential context (word level vs. spatial scenario), and the 'zooming' factor (in vs. out).

Statistically one can expect an interaction of all three factors, but the results only show a weak tendency ($F_S(1,38) = 2.71$, $MS_E = 116690.43$, $p < 0.15$; $F_S(1,38) = 3.82$, $MS_E = 60997.27$, $p < 0.10$ and min $F'(1,78) = 1.59$, n.s.). Instead there is an interaction between the 'zooming' factor and the referential context in the ANOVA of subjects ($F_S(1,38) = 6.10$, $MS_E = 60116.86$, $p < 0.05$) and of texts ($F_T(1,38) = 3.89$, $MS_E = 120331.25$, $p = 0.056$), but not for the min $F'$ value (min $F'(1,76) = 2.38$, n.s.). The main factors close and distant inference were tendentiously significant ($F_S(1,38) = 3.93$, $MS_E = 179470.63$, $p = 0.055$ and $F_T(1,38) = 15.75$, $MS_E = 53587.00$, $p < 0.001$ yielding a min $F'(1,59) = 3.11$, $p < 0.10$).

## Recognition Phase

As in the preceding experiments the recognition frequency of the two inference conditions hardly show a difference in the 'zooming-in' versions (Figure 7). There are differences, however, in the 'zooming-out' text versions. Here, for the word level more sentences were recognised after the reception of the close inference condition. This effect is reversed in the scenario condition. The ANOVA in effect shows a significant interaction of the three factors ($F_S(1,38) = 5.76$, $MS_E = 210.20$, $p < 0.05$, $F_T(1,38) = 5.30$, $MS_E = 228.36$, $p < 0.05$), but only a tendency with the min $F'$-value (min $F'(1,80) = 2.76$, $p < 0.15$). Additionally,

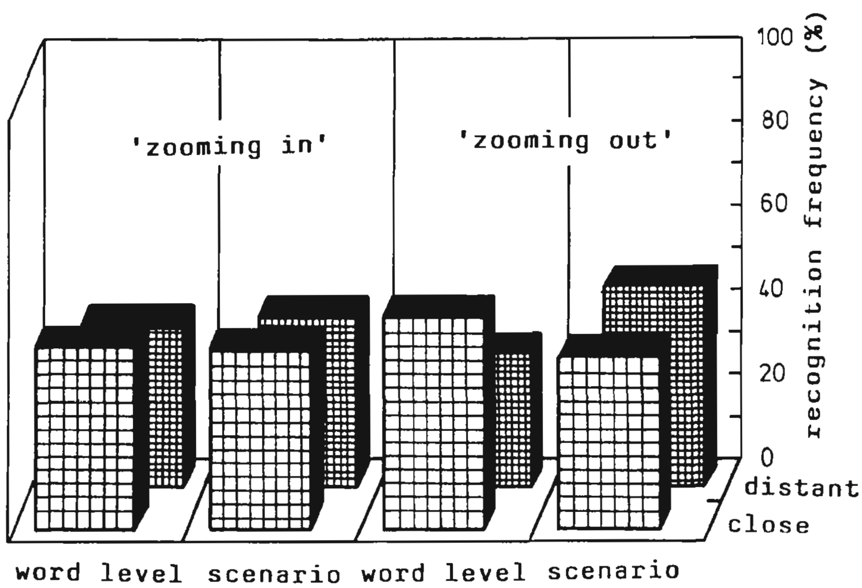


**Figure 7** Mean frequencies of the recognised sentences (type: performed inference) determined by the referential context (word level vs. scenario), the 'zooming' factor (in vs. out), and the inference factor (close vs. distant).

the interaction between the inference factor and the referential context is significant ($F_S(1,38) = 10.00$, $MS_E = 210.20$, $p < 0.01$; $F_T(1,38) = 8.37$, $MS_E = 251.25$, $p < 0.01$ and min $F'(1,89) = 4.56$, $p < 0.05$).

## *Discussion*

The reading time differences in the reception phase prove the expected results. Garrod and Sanford's (1977, Exp. 1) findings show that concepts are processed by about 200 msec faster when generalised than when they are further specified. This is in accordance with our results on the word level.

However, the findings of the recognition phase show results which lead to the assumption that the text representations after the reception of the 'zooming-out' version are different. Let us have a closer look at the critical recognition sentences of the performed inference type: they arise through the exchange of the last-named concept of the second sentence with the first-named concept of the third sentence.

With respect to the 'zooming-out' version on the *word level*, this exchange went more unnoticed when they were lined up closer in the hierarchy of concept categories. On the other hand, the concepts in the distant inference condition which did not border on the hierarchy of word categories had a greater tendency to be rejected. This result would generally have been expected if the subjects had not performed any inference process particularly in the distant inference condition.

Experiments 1 and 2 and the 'zooming-in' versions show that the contrary is true due to the fact that reading and comprehension time decreases and no important differences were found in the recognition phase.[4] In the 'zooming-out' version there is no difference in the reading and comprehension time on the word level but there is a difference in the recognition phase. This result points to the fact that the inference had not necessarily been performed in the reception phase so that different mental text representation results, and hence the subjects can distinguish between the recognition sentences. In other words, if the inference is not performed or there is no necessity for it in the texts used, as could be the case in the 'zooming-out' versions on the word level, then different text representations are still available. Consequently the inference process leads to a qualitative change in the text representations, at least in those that are retrieved in the recognition phase.

For the *spatial scenarios* the findings are different: in this case the inference variation leads to increased reading and comprehension times in the 'zooming-out' as well as the 'zooming-in' versions. Performed inferences of the 'zooming-out' versions are better recognised when the spatial focus is extended further as is the case in the distant inference version. With the 'zooming-in' version this did not happen. A speculative interpretation would have to be that the 'zooming-out' version requires a qualitative change of the mental scenario representation when the spatial mental model is extended. This representation would not equal the antecedent one. With the spatial 'zooming-in' this change is not necessary because the new concept incorporated into the spatial mental model is only more detailed information. In both cases the reading and comprehension times can increase. Nevertheless, their mental text representations cannot necessarily be compared.

## GENERAL DISCUSSION AND CONCLUSION

In summary, the above experiments give the following results. Firstly, the inference effect measured by reading and comprehension times is not bound to any concept repetition, but can also be found when the semantic distance between concepts is varied. This basically replicates an earlier finding in English (Garrod and Sanford 1977), but with slightly different form of semantic distance manipulation.

Secondly, the recognition test proved not to be different between the close and distant inference condition (with one exception, see below). So the text representation constructed by the close and distant inference process seems to be comparable for the text as a whole, at least on a semantic level. This is a necessary assumption to secure the integration process combining the two critical sentences.

Thirdly, the inference process starts immediately after the reception of a critical concept and ends with the construction of the first propositional unit. This result is consistent with a *weak* version of Just and Carpenter's (1980) immediacy hypothesis. In this way, the identification of the antecedent in the present study seems to be a similar recursive mechanism to that used, for example, in pronominal resolution (for differences between noun and pronoun resolution, see Cloitre and Bever 1988). With the encoding of a referent a search-and-match procedure is initiated which establishes the reference's relation to the preceding discourse and identifies the object it refers to (e.g. Carpenter and Just 1977, Ehrlich and Rayner 1983, but see also, for an alternative view concerning pronominal resolution, Müsseler and Terhorst 1990).

Fourthly, with the specification ('zooming-in') of a concept on the word level the inference effect occurs more easily than with generalisation ('zooming-out'). With the latter the inference process as such is questionable when looking at the recognition performance; at least there are differences in the semantic memory traces.

Finally, both the spatial 'zooming-in' and the spatial 'zooming-out' lead to increased reading and comprehension times which reflect the inference effect; again, as far as the recognition performance is concerned, only the specification seems to lead to similar mental text representations. This indicates that the 'zooming-out' version requires a qualitative change of the mental scenario representation when the spatial mental model is extended. With the spatial 'zooming-in' this change is not necessary, because the new concept incorporated into the spatial mental model is only more detailed information.

Nevertheless, the question remains as to whether the increased reading and comprehension time reflects a greater processing expense of the same process or whether it reflects an additional process. In this regard, previous assumptions from Clark and Haviland (1977) have to be made more clear: Clark and Haviland assumed a syntactic parser which dissects the present sentence into 'new' and 'given' components. The 'given' component signalled by the definite article connects the concept to the antecedent which should be no problem when the concept is repeated. The actual inference process is only accomplished in their distant inference condition because here the antecedent can only be identified when it is connected with the world knowledge of the recipient (Rickheit, Schnotz and Strohner 1985). With regard to the first example, 'beer', it should be known to be a 'picnic supply'. The 'bridging' process is then an additional process, which is not necessary in their close inference condition or at least is processed in a different manner.

In the present texts this inference process is accomplished in both cases: the antecedent is not more easily identified in the close inference condition by concept repetition but rather by its semantic distance. Accordingly, one can assume a gradual change in the information process. The different processing times are then due to semantically closer distances to an already constructed mental model, whereas the processing cost becomes greater when one has to relate it to a semantically more distant mental representation. Garrod (1985) called the first type of inference a pseudo-inference which tends to be processed easily and in general automatically, in contrast to true inferences which are derived from a propositional representation and need intensive processing time.

It is, however, questionable whether one need assume increased processing; the effect could also be interpreted as a reduction in processing time with the close inference condition. Results from experiments on the word level show that by varying the semantic distance, different processing times can be found (e.g. Wilkens 1971, Meyer, Schvaneveldt and Ruddy 1975). A semantic facilitation or a reduction in processing time in contrast to a neutral condition can be stated even during the resolution of word ambiguities (Swinney 1979, 1982). In contrast to these findings the inference effect mentioned in the above experiments could be due to a processing reduction or extension, because one cannot decide empirically between these two alternatives. Reduction or extension can only be decided with respect to neutral conditions. Whatever these look like, one can only postulate theoretically.

The findings of the last experiment in particular allow us to assume that reference also depends on structural features of the critical concepts. It obviously makes a difference whether in the following sentence a concept is specified or generalised. At least the inference effect could only be seen after specification on the word level. As indicated above, one could interpret these findings with hierarchically constructed network models (see also Collins and Quillian 1972). Here one could assume that the processing time is determined by the amount of subcategories. In the specification case the number of subcategories is increased, whereas with generalisation their numbers decrease.

## Acknowledgement

The research reported here was supported by grants from the Deutsche Forschungsgemeinschaft (Ri 314/9). Thanks are due to Julie Carson and Andrea Reuther for many insightful stylistic suggestions. Requests for reprints should be sent to Jochen Müsseler and Gert Rickheit, Fakultät für Linguistik und Literaturwissenschaft, Arbeitseinheit Psycholinguistik, Universität Bielefeld, Postfach 8640, D-4800 Bielefeld 1, Federal Republic of Germany.

JOCHEN MÜSSELER AND GERT RICKHEIT
*University of Bielefeld*
*Postfach 8640*
*D-4800 Bielefeld 1*
*Federal Republic of Germany*

## NOTES

1 The mean fixation time in normal reading is about 250 msec (e.g. Rayner 1978) depending on text difficulty and reading ability. In a pilot experiment we increased the presentation time until a comfortable reading was achieved.

2 Note that in German 'a shark' or 'a fish' is taken as 'an animal' even in some non-technical contexts.

3 In the spatial 'zooming-out' and 'zooming-in' an obvious part-of-relation is established with respect to the spatial mental model. On the word level only a hierarchical class relation was realised. For example, the 'grain' is part of the 'field', but a 'fish' is not a part of an 'animal' but rather a membership of the class 'animal'.

4 But note that in Experiment 1 a marginally significant interaction (though not reliable in min F′) between the recognition test type and the close vs. distant text version was reported. This may be caused by the confounding of 'zooming-in' and 'zooming-out' in these texts.

## REFERENCES


Bortz, J. (1977), *Lehrbuch der Statistik für Sozialwissenschaftler* (Textbook of statistics for the social sciences), Berlin, Springer.

Carpenter, P. A. & Just, M. A. (1977), 'Reading comprehension as eyes see it', in P. A. Just & M. A. Carpenter, (eds), *Cognitive Processes in Comprehension*, Hillsdale, NJ, Erlbaum, 109–39.

Clark, H. H. (1973), 'The language-as-fixed-effect fallacy: a critique of language statistics in psychological research', *Journal of Verbal Learning and Verbal Behavior*, **12**, 335–59.

Clark, H. H. (1977), 'Inferences in comprehension', in D. LaBerge & S. J. Samuels (eds), *Basic Processes in Reading: Perception and Comprehension*, Hillsdale, NJ, Erlbaum, 243–63.

Clark, H. H. & Haviland, S. E. (1977), 'Comprehension and the given-new contract', in R. O. Freedle (ed.), *Discourse Production and Comprehension*, Norwood, NJ, Ablex, 1–40.

Cloitre, M. & Bever, T. G. (1988), 'Linguistic anaphor, levels of representation, and discourse', *Language and Cognitive Processes*, **3**, 293–322.

Collins, A. M. & Quillian, M. R. (1972), 'Experiments on semantic memory and language comprehension', in L. W. Gregg, (ed.), *Cognition in Learning and Memory*, New York, Wiley, 117–37.

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**, 121–32.

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**, 45–87.

Garrod, S. C. (1985), 'Incremental pragmatic interpretation versus occasional inferencing during fluent reading', in G. Rickheit & H. Strohner (eds), *Inferences in Text Processing*, Amsterdam, Elsevier, 161–81.

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

Günther, U. (1989), 'Lesen im Experiment' (Reading in experiments), *Linguistische Berichte*, **122**, 283–320.

Haberlandt, K. (1980), 'Story grammar and reading time of story constituents', *Poetics* **9**, 99–116.

Haberlandt, K. (1984), 'Components of sentence and word reading times', in D. E. Kiereas & M. A. Just (eds), *New Methods in Reading Comprehension Research*, Hillsdale, NJ, Erlbaum, 219–51.

Haberlandt, K., Berian, C. & Sandson, J. (1980), 'The episode schema in story comprehension', *Journal of Verbal Learning and Verbal Behavior*, **19**, 635–50.

Haberlandt, K. & Graesser, A. C. (1985), 'Component processes in text comprehension and some of their interactions', *Journal of Experimental Psychology: General*, **114**, 357–74.

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**, 512–21.

Juola, J. F., Ward, N. J. & McNamara, T. (1982), 'Visual search and reading of rapid serial presentations of letter strings, words, and texts', *Journal of Experimental Psychology: General*, **111**, 208–27.

Just, M. A. & Carpenter, P. A. (1980), 'A theory in reading: from eye fixations to comprehension', *Psychological Review*, **87**, 329–54.

Just, M. A., Carpenter, P. A. & Woolley, J. (1982), 'Paradigms and processes in reading comprehension', *Journal of Experimental Psychology: General*, **111**, 228–38.

Meyer, D. E., Schvanefeldt, R. W. & Ruddy, M. G. (1975), 'Loci of contextual effects on visual word-recognition', in P. M. A. Rabbit & S. Dornic (eds), *Attention and Performance*, New York, Academic Press, Vol. V, 98–118.

Müsseler, J. & Nattkemper, D. (1986), 'Visuelle sequentielle Darbietung und "normales" Lesen: ein Vergleich zweier experimenteller Darbietungsarten' (Visual sequential presentation and 'normal' reading: a comparison of two modes of presenting texts), *Psychologisches Beiträge*, **28**, 107–19.

Müsseler, J., Rickheit, G. & Strohner, H. (1985), 'Influences of modality, text difficulty, and processing control on inferences in text processing', in Rickheit, G. & Strohner, H. (eds), *Inferences in Text Processing*, Amsterdam, Elsevier, 247–71.

Müsseler, J. & Terhorst, E. (1990), 'Pronominale Besetzung: ein alternativer Mechanismus neben der rekursiven Auflösung?' (Pronominal occupations: an alternative mechanism in addition to the recursive resolution?), *Sprache und Kognition*, **9**, 37–49.

Rayner, K. (1978), 'Eye movements in reading and information processing', *Psychological Bulletin*, **85**, 618–60.

Rickheit, G., Schnotz, W. & Strohner, H. (1985), 'The concept of inference in discourse comprehension', in G. Rickheit & H. Strohner (eds), *Inferences in Text Processing*, Amsterdam, Elsevier, 3–49.

Sanford, A. J. & Garrod, S. C. (1981), *Understanding Written Language*, Chichester, Wiley.

Schnotz, W. (1986), *Kohärenzbildung beim Aufbau von Wissensstrukturen mit Hilfe von Lehrtexten* (Coherence in the construction of knowledge on the basis of textbooks), Tübingen, University, Institut für Fernstudien, Forschungsbericht Nr. 39.

Schnotz, W. (1988), 'Textverstehen als Aufbau mentaler Modelle' (Text understanding as construction of mental models), in H. Mandl & H. Spada (eds), *Wissenspsychologie*, Munich, Psychologie Verlags Union, 299–330.

Sidner, C. L. (1983), 'Focusing and discourse', *Discourse Processes*, **6**, 107–30.

Swinney, D. A. (1979), 'Lexical access during sentence comprehension: (re)consideration of context effects', *Journal of Verbal Learning and Verbal Behavior*, **18**, 645–59.

Swinney, D. A. (1982), 'The structure and time course of information interaction during speech comprehension: lexical segmentation, access, and interpretation', in J. Mehler, E. C. T. Walker & M. F. Garrett (eds), *Perspectives on Mental Representation*, Hillsdale, NJ, Erlbaum, 151–69.

Tanenhaus, M. K. & Seidenberg, M. S. (1981), 'Discourse context and sentence perception', *Discourse Processes*, **4**, 197–220.

Tukey, J. W. (1949), 'One degree of freedom for nonadditivity', *Biometrics*, **5**, 523–40.

Ward, N. J. & Juola, J. F. (1982), 'Reading with and without eye movements: reply to Just, Carpenter, and Woolley, *Journal of Experimental Psychology: General*, **111**, 239–41.

Wilkens, A. (1971), 'Conjoint frequency, category size, and categorization time', *Journal of Verbal Learning and Verbal Behavior*, **10**, 382–5.