Lexical processing and the mechanism of context effects in text comprehension

Amanda J.C. Sharkey
and
Noel E. Sharkey

Centre for Connection Science,
University of Exeter.

Two models of context effects on lexical processing are described; (a) the Construction-integration (CI) model (Kintsch, 1988) and (b) the Lexical Distance (LD) model (Sharkey, N.E. in press; 1989). Both can account for some well known effects in the priming and ambiguity literature. However the mechanisms by which they operate differ. The CI model presumes connections between related items in the lexicon. In this model, it is assumed that, during the initial stages of processing, the associates of a word in the lexicon are *always* activated in a context-independent manner. It is postulated that textual priming effects can only occur after this phase of sense activation. Lexical priming effects and textual priming effects are the result of the operation of different processes. In the LD model, on the other hand, there are no associations between items in the lexicon. Words are represented in the lexicon as vectors of microfeatures. Context effects are conceptualised as a measure of network distance from an initial state to a target state. Both lexical priming effects and textual effects exert their influence in the same way; reducing the time taken to move to the target state. An experiment is reported in which textual priming effects are examined in an attempt to test the predictions made by the two models. In this experiment, a facilitatory influence of knowledge-based texts on a lexical decision task was demonstrated, despite short SOAs and the absence of an immediately preceding associatively related word. According to the CI model the SOAs involved were too short to allow anything but associative priming; therefore the results of this experiment favour the LD model. This research clears up a conflict in the literature between two sets of experimental findings (Kintsch & Mross, 1985; Sharkey & Mitchell, 1985). It also illustrates the use of psychological experimentation in allowing a principled choice to be made between two models.

Two kinds of context effects on lexical processing can be identified; lexical priming effects and textual priming effects (Foss, 1982; Kintsch & Mross, 1985; Keenan et al, in press; Sharkey, A.J.C., 1989; Sharkey & Mitchell, 1985). Lexical priming effects stem from the presence of an immediately preceding associatively-related word (e.g if the word NURSE is immediately preceded by the word DOCTOR it will be recognised faster than if it is preceded by a row of Xs). Lexical priming effects, as they occur in word lists have been extensively studied (e.g. Gough et al, 1981; Meyer & Schvaneveldt, 1971). Textual priming effects are those that are due to the relationship of the processed word to the preceding text, rather than the immediately preceding word (e.g the word MENU might be processed faster if it occurred in a story about a restaurant, than in a story about visiting the dentist). These effects have been less extensively studied. In this paper we shall attempt an empirical evaluation of the accounts given of both kinds of context effects in two models; Kintsch's (1988) Construction-Integration (CI) model and Sharkey's Lexical Distance (LD) model (Sharkey, N.E. in press; 1989).

These two models were selected for comparison because, unlike other models of word recognition (i) they are computationally specified, and (ii) they deal with lexical processing in terms of a global view of text comprehension - in both, lexical processing takes place in a model which also accounts for the inferences made at the propositional level. These models are similar in certain other respects; for instance they both posit the existence of a propositional and a lexical level of representation. However, quite different contextual mechanisms are proposed. As will be discussed later, these processing differences have empirically testable consequences.

Lexical priming effects: A major difference between the CI and LD models is whether or not activation is assumed to spread between items in the lexicon. In Kintsch's CI model, words are represented in the lexicon such that there are connections between related items. By means of the mechanism of spreading activation (Meyer & Schvaneveldt, 1971), the level of activation in one unit in the lexicon can affect the level of activation of other units. In Sharkey's LD model there are no connections between related lexical nodes. Instead words have a distributed representation in the lexicon made up of three types of microfeatures: semantic, situational and graphemic. Access to the lexicon is through graphemic microfeature activation; the semantic microfeatures are similar to semantic features such as is-male, has wings etc; and the situational microfeatures provide information about the contextual setting (e.g. the situational microfeatures for NURSE would contain information about hospitals). Each microfeature may be shared by several concepts. For example, DOCTOR is likely to share many situational microfeatures with NURSE because of overlapping job roles, and place of work. Each microfeature has an activation value; therefore a lexical entry can be characterised as a vector of microfeatures. Each vector of microfeatural activations can be identified as a point in n-dimensional space (where n is the number of microfeatures in the lexicon).

Because the nature of the lexical representation assumed by the two models differs, so too does the way in which lexical priming effects are accounted for. In the CI model, lexical priming effects are assumed to be the result of spreading activation in the lexicon: '..right after a word is perceived, it activates its whole associative neighbourhood in a context-independent way with the consequence that strong associates of a word are likely to be represented in working memory and hence will be primed in a lexical decision task..' (Kintsch, 1988 pg 172). Therefore, if the word 'nurse' is presented to the system just after the word 'doctor', its recognition will be speeded.

In the LD model, lexical priming is not the result of spreading activation in the lexicon; rather it reflects a measure of network distance from an initial state to a target state. When a prime word is presented to the system it activates its corresponding set of graphemic microfeatures, and sets the system on a downward descent in the energy function until a stable minimum of microfeature activity has been reached. Finding a stable minimum is the equivalent of locating the lexical entry for the prime word. We refer to this as the *initial* state of the system. Now, it is the relationship of this initial state to the representation of the target word which is responsible for the response time predictions. To make this clear, imagine the representation of the prime and target words as two vectors of microfeature activations in the lexical space $\mathbf{L}^n$. Let these vectors represent the starting state of the system $\mathbf{s}$ and the required or target state $\mathbf{r}$. Then the Euclidean distance between the two points $\mathbf{s}$ and $\mathbf{r}$ is given by $\|\mathbf{s} - \mathbf{r}\|^{1/2}$, where length $\|\mathbf{v}\| = (\mathbf{v} \cdot \mathbf{v})^{1/2}$. A major assumption of the model is that the greater the distance from an initial state to a target state, the longer will be the recognition time for a target word. If a target word is related to the prime, it will take less time to process because the two words will share some microfeatures and the system will be closer to the target state than if the target was not related to the prime. Thus both models make the same general predictions about lexical priming effects (although the LD model also accounts for a number of other effects as well such as interactions between context, frequency and stimulus quality).

Textual priming effects: The two models differ in their accounts of textual priming effects. In the LD model, textual priming effects have the same causal factor as lexical priming effects i.e. the time taken to move through n-dimensional lexical space from an initial state (the vector of microfeatures that resulted from the processing of the previous word) to a target state (the vector of microfeatures implied by the word being processed). The difference between textual priming and lexical priming in the LD model is the result of processes occurring in a knowledge-net which is external to the lexicon. This net holds the model's knowledge of the world in the form of a network of massively interconnected propositions. During reading, once a proposition has been constructed, it activates the knowledge net with the result that it relaxes on a stable configuration of situationally related propositions. This stable state is maintained until cues from the text indicate otherwise. These active propositions broadcast activation to the situational microfeatures in the lexicon, holding them constant. This means that incoming words that are related to the knowledge structure will be processed faster, since the activation of the situational microfeatures will ensure that their representations will be closer to the state of the lexical system than the representations of contextually unrelated words; so it will take less time to traverse the intervening distance.

In the CI model, unlike the LD model, it is assumed that textual priming effects result from the operation of a different mechanism from that implicated in lexical priming effects. When a word is encountered in text, in the initial stages of processing, its associates are activated through spreading activation in the lexicon. This activation of associates occurs in a context-independent manner. Textual priming effects exert their influence only as a result of the process of integration in the following phases of sense selection and sense elaboration. The process of construction that results in context-independent activation of concepts at the lexical level is also assumed to operate at the propositional level; when a proposition is constructed, associatively and semantically related propositions are also activated. The result of the construction process is the production of a network expressable as a connectivity matrix. Via the process of integration, activation is then spread through the network until the system stabilises. More specifically, an activation vector representing the activation values of all the nodes in the network is repeatedly postmultiplied with the connectivity matrix. This cycle is repeated until the system stabilises. The end result of this process of integration is that '..positively interconnected items strengthen each other, while unrelated items drop out and inconsistent items become inhibited..' (Kintsch, 1988 pg 171).

Kintsch (1988, pg 164) criticises top-down accounts of lexical processing. He suggests that it is too difficult to '..make a system smart enough so that it will make the right decisions, yet keep it flexible enough so that it will perform well in a variety of situations..'(Kintsch, 1988, pg 164). In other words, a system that expects particular lexical items is unlikely to perform well (a point with which we agree). He seeks to avoid the problem of inflexibility through the use of the combined processes of construction and integration. In his system, the process of construction results in the activation of a number of elements which are selected amongst during the operation of the integration process. The idea is that the correct element is likely to be amongst those that are generated in this haphazard manner. Nonetheless, it is claimed here that even though the LD model admits a top-down influence on the lexicon, it does not suffer from the sort of inflexibility that Kintsch discusses. This is because the system does not expect particular lexical items. Instead, it is predisposed towards the reception of a *class* of lexical items - those which share situational microfeatures with the active knowledge structure.

Kintsch (1988) illustrates his account of textual context effects in terms of the processing of lexically ambiguous words. When an ambiguous word is processed, it activates its associates in a context-independent manner (e.g. the word 'mint' will activate both 'candy' and 'money'). Then the process of integration will result in the selection of one of these alternatives, and the deactivation of the other. This account fits with available data on lexical ambiguity. If the interval between the presentation of the ambiguous word and the subsequent presentation of an associatively related word is short, then even the processing of context-inappropriate associates will be facilitated. However, if the interval is longer then the process of integration will have been called into operation, and only the context-appropriate associates of the ambiguous word will be primed.

This explanation fits much of the available data on lexical ambiguity resolution (e.g Kintsch & Mross, 1985; Seidenberg et al, 1982; Swinney, 1979), although there are some exceptions (Blutner & Sommer, 1988; Glucksberg, Kreuz & Rho, 1986; Tabossi, 1988). An explanation of these findings can also be couched in terms of the LD model, through use of the distance metric as in a related connectionist model by Kawamoto (in press). Kawamoto also discusses the structure of his lexical network in terms of points in energy space. He shows that homophones are close together in the energy landscape and separated by a high energy ridge. Before the system settles on the final meaning of a homophone, the system may move along the ridge between the alternative meanings. Therefore, until the system has stabilised, both sets of microfeatures will be equally available and words sharing either set will be primed. Once the system has stabilised, only words that share microfeatures with the stable configuration will be primed.

In summary, the LD and CI models are fairly equivalent in their ability to account for effects reported in the literature on lexical priming. Where they differ is in their explanations of the effects of textual context on unambiguous words. Sharkey and Mitchell (1985) demonstrated facilitation of lexical decision responses to words related to knowledge-based contexts (e.g MENU in a restaurant context). It is clear how such an effect would be explained in the LD model. As the restaurant story is processed, a related set of propositions would be assembled. These propositions would then hold active in the lexicon the situational microfeatures associated with restaurants. Since the word MENU is associated with restaurants it shares these microfeatures, and therefore it will take less time for the lexical processor to move from its initial state when the word is perceived to the target state which represents its meaning.

It is less clear how such an effect would be explained in the CI model. The CI model is primarily designed to select among alternatives generated when the previous word is processed. Unless, MENU was an associate of the preceding word, it is not clear how its processing would be facilitated. Kintsch is more specific about *when* textual context exerts its influence on lexical processing, than he is about *how* this influence is exerted. He argues that '..if the target word closely follows the priming word, so that the processing of the prime is still in its initial stages ..... context-appropriate inferences that are not associatively related to the priming word are not responded to any faster than unrelated control words..' (Kintsch, 1988 pg 171). On the other hand, if more time is allowed, '..context-appropriate associates are still primed, but inappropriate associates no longer are, whereas context-appropriate inferences now become strongly primed..' (Kintsch, 1988 pg 171).

In the CI model, the interval of time between the target word and the immediately preceding word is crucial in determining the effects that will be obtained. Kintsch (1988) refers to another study (Till, Mross & Kintsch, 1988) which suggests that context-inappropriate associates are not deactivated until 300 milliseconds have passed. In an earlier paper, Kintsch & Mross, (1985) suggest that Sharkey and Mitchell's (1985) facilitation results are due to the self-paced method of presentation they used. Kintsch and Mross's argument is that the self-paced presentation of the sentences in the knowledge-based contexts allowed enough time for evidence of textual context effects to become detectable. When Kintsch and Mross (1985) used an experimenter-paced visual presentation of knowledge-based texts (each word appeared on the screen for 150 milliseconds), they found no clear evidence of any knowledge-based facilitation of lexical decisions. In fact, they failed to find any clear evidence of knowledge-based facilitation, even when they used a self-paced or a delayed presentation of the knowledge-based texts. This finding supports the CI model, but means that there is a discrepancy in the literature between the results obtained by Kintsch and Mross (1985) and those obtained by Sharkey and Mitchell (1985). One reason might be that Sharkey and Mitchell's results were in some way the result of the task they employed. A more likely explanation is that the discrepancy was due to a difference in the materials used. Kintsch and Mross (1985) based their materials on Galambos's (1982) norms. However, unlike the norms used by Sharkey and Mitchell (1985), the Galambos norms were first generated by Galambos, and subsequently rated by subjects. The subject-generated norms of Sharkey and Mitchell may therefore have provided a better reflection of the subjects' knowledge.

In the experiment reported in this paper, an attempt was made to resolve this inconsistency by looking at the effect of textual priming effects using the same task as that employed by Kintsch and Mross (1985), but with materials based on the subject-generated norms of Sharkey and Mitchell (1985). Knowledge-based texts were presented visually on a screen at a rate of one word per 200 milliseconds. This control over the rate of presentation made it possible to test predictions derived from the two models. According to the CI model, at short delays only associates of the preceding word should be primed. Kintsch is very specific about this; following short SOAs, '..the discourse context is actually irrelevant to the priming effect. What matters is merely the associative relation between the prime word and the target word..' (Kintsch, 1988 pg 171). However, in the experiment, care was taken to ensure that the word which preceded the lexical decision target was not associatively related to it. Therefore, according to the CI model, the processing of the target word should not be facilitated. According to the LD model, evidence of priming should be detected under these circumstances. In the LD model, as long as enough time has been allowed to construct propositions on the basis of the text, then words which share situational microfeatures with the text will be facilitated. No delay between the immediately preceding word and the target is needed for this textual facilitation to occur. According to the LD model, lexical priming does not contribute to the process of textual priming. Therefore, the LD model predicts that evidence of facilitation should be obtained in the proposed experiment despite the absence of an immediately preceding associatively related word.

AN EMPIRICAL STUDY OF THE ONSET OF TEXTUAL PRIMING

Texts based on Sharkey and Mitchell's (1985) norms were presented to subjects using a task similar to that employed by Kintsch and Mross (1985) (see Example 1). Lexical decision targets interrupted the text at certain points. A crucial factor in this experiment is that the knowledge-based stories were written so as to ensure that the word that immediately preceded the lexical decision target had not appeared in the subject-generated norms, and was not associatively related to the target. The lexical priming effects that are obtained in word lists have been shown to disappear if even one word intervenes between prime and target (e.g Gough et al, 1981). Therefore, if any facilitation is evident in this experiment it must be the result of the influence of textual context, not associative relations between words.

As stated above, the predictions of the two models about the outcome of this experiment are clear. According to the CI model, priming should not occur. The delay that intervenes between the preceding word and the lexical decision target is too short to implicate anything but context-independent activation of associates, and the word that immediately preceded the lexical decision target was not associatively related to it. In contrast, the LD model predicts the occurrence of priming. According to the LD model, once related propositions have been constructed, (by reading introductory sentences such as those shown in Example 1), the influence of the propositional level will predispose the system towards words which share situational microfeatures with the relevant knowledge structure. Therefore priming should occur despite the relatively short SOAs and the absence of an immediately preceding associatively related word.

Materials and Design: 24 subjects took part in this experiment, and each read 40 stories. Twenty of these stories were knowledge-based passages based on Sharkey and Mitchell's norms (Sharkey & Mitchell, 1985). Twenty were distractor stories (containing a mixture of word and nonword targets). The stories were presented on a computer screen, one word at a time, at a rate of one word every 200 milliseconds. The task was self-paced in one respect; subjects initiated the presentation of each sentence. This ensured that they had time to understand one sentence before moving on to the next.

At certain points in each story, a lexical decision target was presented. Subjects were required to indicate as quickly as possible whether or not this target was a word. Care was taken to ensure that the word that immediately preceded the lexical decision target was not associatively related to it. There were 6 conditions in the experiment (as shown in Example 1 below). The targets were either related or unrelated to the surrounding text and they were presented in one of three sentence positions; in an initial sentence position, at the end of a main clause or at the end of a subordinate clause.

<u>Example 1</u>

Introductory sentences
Toby was going to his little friend Edward's birthday party. Toby's mother took him along to the party. Toby rushed into the room to see everyone else.
Sentences containing related targets
a) The balloon GAMES he grabbed belonged to someone else.
b) He grabbed hold of a balloon GAMES and waved it above his head.
c) Whether he grabbed hold of a balloon GAMES or not he certainly had one in his hand.
Two Fillers: The children were all very excited. Edward was quite red in the face.
Sentences containing unrelated targets
a) The sofa FIRMS was fun to jump on, Toby decided.
b) Toby jumped on the sofa FIRMS and was told off by the adults.
c) Given that Toby jumped on the sofa FIRMS its not surprising he got told off by the adults.
Concluding sentences
He sobered up a little. The adults tried to get the children under control. By the end of the party they were completely shattered.

Results and Discussion: The differences between Unrelated and Related targets were statistically significant for all conditions (see Table 1 below). The magnitude of the differences were as follows:

Initial portion of sentence - MinF(1,61) = 10.33, $p < .005$); End of subordinate clause - $F_1(1,23)$ = 4.97, $p < .036$, $F_2(1,38)$ = 3.56 $p < 0.067$; End of main clause MinF(1,60) = 5.72, $p < .025$. These results clearly favour the LD model over the CI model. The CI model cannot account for textual context effects which occur following short delays. The LD model predicts the occurrence of textual context effects even at short delays and when the immediately preceding word is not associatively related.

Table 1

SENTENCE POSITION

	Within-clause	End-main-clause	End-sub-clause
Related	810.03	851.73	806.94
E%	0.48	0.35	0.0
Unrelated	902.94	910.84	896.72
E%	0.71	0.48	0.83
Neutral baseline (words in isolation) =		723.51	
E%		0.47	

Overall the findings suggest that (a) a difference in materials is likely to have been responsible for the difference in the results obtained by Sharkey and Mitchell (1985) and Kintsch and Mross (1985), and (b) because the word that immediately preceded the target was not related to it, the knowledge-based priming results were not due to the confounding influence of lexical priming from an immediately preceding word. Nonetheless, one way in which an attempt might be made to weaken the current experimental results would be to suggest that the effects obtained were in fact due to the confounding influence of lexical priming from earlier parts of the text (even though an attempt was made to avoid this possibility when writing the materials). Keenan et al (in press) raise the possibility that lexical priming effects in text sustain for longer than they do in word lists. They in fact explicitly suggest that Sharkey and Mitchell's (1985) results might be due to the confounded influence of lexical priming. Keenan et al (in press), and Ratcliff (1987) both argue that it is important to control for the potentially confounding influence of lexical priming. Our argument is that the reverse may also be true, and apparent examples of lexical priming may actually be textual context effects. If subjects are allowed time to interpret a text, how can it be demonstrated that apparent lexical priming effects are really due to context-independent sense activation, and not to the influence of the interpreted text? Foss and Ross (1983) make a similar point when they suggest that apparent instances of lexical priming may be deceiving since, '... the operative relation is that between the semantically interpreted phrase and the next word..' (Foss & Ross, 1983). In any case, two studies reported by Sharkey, A.J.C. (1989) suggest that lexical priming effects are unlikely to have been responsible for the apparent textual priming effects reported above.

In the first of these two experiments, associative prime words were presented in sentence contexts as shown in Example 2.

Example 2
(the ^ sign indicates the various positions at which the target could be presented)

Associative
The boy ^ is sometimes ^ unkind and ^ thoughtless.
The boy unkind and sometimes is ^ thoughtless.
Unrelated
The law ^ is sometimes ^ unkind and ^ uncompromising.
The law unkind and sometimes is ^ uncompromising.
Target: GIRL

Like the experiment reported above, the words in the sentence were presented visually at a rate of one word per 200 milliseconds. The prime words were followed by lexical decision targets that were either related or unrelated to the prime. The lexical decision targets followed the prime words after an interval of either 0,2 or 4 words in sentences, or after an interval of 4 words in a scrambled sentence, as illustrated in Example 2 below.

In this experiment, no evidence of priming was found for associative words, even at the beginning of sentence. That is, there were no statistically significant difference between the Unrelated and Related in any of the sentence positions (F1 and F2 < 1 in all cases except lag 2 where $F_2(1,29) = 1.40$, $p > .24$).

The second experiment was conducted to check that evidence of lexical priming effects could be obtained, under some circumstances, using the same task. Essentially the same experiment was run with the same materials, but the sentences were presented in scrambled form, as shown in Example 3. Scrambled sentences are the equivalent of word lists. The results obtained were those that would be expected when associative primes are presented in word lists. Clear evidence of priming was obtained when there were no intervening words between prime and target ($MinF(1,44) = 6.05$, $p < 0.025$). No priming was obtained when the interval between prime and target was filled with intervening words.

Example 3

Associative
Thoughtless boy ^ unkind the ^ sometimes is ^ and.
Unrelated
Thoughtless law ^ unkind the ^ sometimes is ^ and.
Target: GIRL

These findings add to others in the literature, (e.g. Gough et al, 1981; Meyer & Schvaneveldt, 1971;) by demonstrating that lexical priming effects are less likely to occur in texts than in word lists. Not only do lexical priming effects in normal texts seem not to sustain, they do not even seem to occur. Since the textual priming experiment reported here used the same task and experimental settings it is clear that lexical priming effects could not be responsible for the evidence of facilitation that was obtained.

CONCLUSIONS

The experimental results reported here favour the Lexical Distance model (Sharkey, in press; 1989) as a more accurate account of lexical processing in humans than the Construction-Integration model (Kintsch, 1988). In the textual priming study reported here, evidence of the effect of knowledge-based contexts on lexical processing was obtained even though the targets were not preceded by related words, and the SOAs involved were short. These findings run counter to predictions arising from the CI model. Kintsch (1988) proposes that the effects of knowledge-based contexts will be detected only if an SOA greater than 400 milliseconds is used, and as we have seen this delay is crucial to the operation of the model in its ability to explain effects in the ambiguity literature. In contrast, the LD model accurately predicts the results of the current experiment. Sharkey (in press; 1989) proposes that propositions activate situational microfeatures at the lexical level, thus holding the lexical processor close to the target state of incoming contextually related words until contextual cues indicate a change of context. Potential criticisms of the experiment were allayed by presenting data which demonstrated that in the same task lexical priming effects occurred only in word lists and not in sentences. These findings show that our textual priming results were not due to the confounding influence of lexical priming.

Although the results of the textual priming study favour the LD model over the CI model, it cannot be conclusively stated that textual context exerts an influence on the lexicon prior to lexical access. As argued elsewhere, (Sharkey A.J.C., 1989), current techniques do not permit an unassailable demonstration of non-modularity in language processing. It is possible that the results reported here could be accounted for in some other incarnation of a modular system. However, as yet such a system has not been developed in sufficient computational detail to allow the derivation of precise empirically testable predictions. It is therefore concluded that the minimal top down account offered by the LD model fits the data better than its alternatives.

REFERENCES

Blutner, R. & Sommer, R. (1988) Sentence processing and lexical access: The influence of the focus identifying task. *Journal of Memory and Language*, 27, 359-367.

Foss, D.J. (1982) A discourse on semantic priming. *Cognitive Psychology*, 14, 590-607.

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*. New York: John Wiley.

Galambos, J.A. (1982) Normative studies of six characteristics of our knowledge of common activities. Cognitive Science Tech. Rep. 14, Yale University

Glucksberg, S., Kreutz, R.J., & Rho, S. (1986) Context can constrain lexical access: Implications for models of language comprehension. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 3, 323-333.

Gough, P.B., Alford, J.A. & Holley-Wilcox, P. (1981) Words and Contexts. In O.J.L. Tzeng & H. Singer (Eds) *Perception of print: Reading research in experimental psychology*. Hillsdale, NJ: Erlbaum

Kawamoto, A.H. (in press) Distributed representations of ambiguous words and their resolution in a connectionist network. In S.L. Small, G. W. Cottrell, and M.K. Tanenhaus (Eds) *Lexical ambiguity resolution in the comprehension of human language*.

Keenan, J.M., Golding, J.M., Potts, G.R., Jennings, T.M. & Aman, C.J. (in press) Methodological issues in evaluating the occurrence of inferences. In A. Graesser and G.H. Bower, (Eds) *Learning and Motivation, Vol 24*: Academic Press.

Kintsch, W. (1988) The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95, 163-182.

Kintsch, W. & Mross, E.F. (1985) Context effects in word identification. *Journal of Memory and Language*, 24, 336-349.

Meyer, D.E. & Schvaneveldt, R.W. (1971) Facilitation in recognising pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*. 90, 227-35.

Ratcliff, J. (1987) The plausibility effect: Lexical priming or sentential processing? *Memory and Cognition* 15, 6, 482-496.

Seidenberg, M.S., Tanenhaus, M.K., Leiman, J.M. & Bienkowski, M. (1982) Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology*, 14, 489-537.

Sharkey, A.J.C. (1989) Contextual mechanisms of text comprehension. Unpublished PhD dissertation.

Sharkey, N.E. (in press) A connectionist model of text comprehension. To appear in D.Balota, G.B. Flores d'Arcais, and K.Rayner (Eds) Comprehension processes in reading.

Sharkey, N.E. (1989) The Lexical Distance model and word priming. Proceedings of the 11th Annual Conference of the Cognitive Science Society.

Sharkey, N.E. & Mitchell, D.C. (1985) Word recognition in a functional context: The use of scripts in reading. *Journal of Memory and Language*, 24, 253-270.

Swinney, D.A. (1979) Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, 18, 645-659.

Tabossi, P. (1988) Accessing Lexical Ambiguity in different types of sentential contexts. *Journal of Memory and Language*, 27, 324-340.

Acknowledgements: We would like to thank Economic and Social Research Council (C08250015) and Leverhulme Trust (A/87/153) for supporting this research.