

The plausibility effect: Lexical priming or sentential processing?

JANET E. RATCLIFF

Deakin University, Victoria, Australia

Five experiments are reported that investigated whether the plausibility effect is caused by lexical priming resulting from the higher proportion of related words in plausible than in implausible sentences. In Experiment 1, a plausibility effect was demonstrated that was entirely attributable to the way in which lexical items were combined rather than to the properties of individual lexical items. In Experiment 2, the content words from the sentences used in Experiment 1 were shown to produce a similar reaction-time difference in a task in which syntactic processing was disrupted, supporting a lexical priming explanation of the plausibility effect. However, Experiments 3 and 4 demonstrated that, in another task less prone to task-specific strategies but sensitive to plausibility, the disruption of syntactic processing eliminated the effect. In Experiment 5, it was shown that when lexical priming was eliminated, a plausibility effect still occurred. Thus, two separate lines of evidence suggested that the plausibility effect cannot be fully explained in terms of lexical priming.

There has long been debate in the psycholinguistic literature about the extent to which there is free flow of information within the language processor. There are many interactive models of language comprehension (e.g., McClelland & Rumelhart, 1981; Marslen-Wilson & Tyler, 1980; Morton, 1969) that propose unrestricted information flow within the language system and between the language system and other cognitive domains. Such models assume that contextual information of any sort can be used to facilitate any stage of the comprehension process.

In contrast, modular theories of language comprehension (e.g., Cairns & Kamerman, 1980; Forster, 1979) propose that the language processor is composed of separate subsystems for dealing with different types of information. These subsystems are functionally autonomous, or modular, in that they receive input only from the subsystem immediately below them in the chain of processing, and their output is unaffected by the processing of other subsystems or by information from other cognitive domains. Thus, modular models restrict information flow to "bottom-up" processing, whereas interactive systems allow "top-down" information flow as well.

Fodor (1983) also proposed a modular theory of language processing, but whereas Forster (1979) and Cairns and Kamerman (1980) concentrated on modular subsystems within the language processor, Fodor dealt with the modularity of the language processor itself—that is,

its autonomy as a processing unit from other cognitive processing. Fodor suggested that this modularity is a characteristic that is shared by the language processor and the perceptual systems and that this is a necessary result of their shared role of "represent[ing] the world [so] as to make it accessible to thought" (p. 40). Because of this shared role, he described them together under the term *input system*.

Fodor (1983) suggested that input systems need to be modular because speed is important in producing representations of the world. He argued that if input systems were not modular, but had access to all the information in memory, then processing would be slowed down not only by the time involved in memory search in accessing all relevant information, but also by decision making about the relevance of each piece of information in memory to the current identification. This necessity for speed is justified for the perceptual systems in terms of allowing fast responses to stimuli in the environment, which is particularly critical if the stimulus is dangerous. This argument clearly extends to language processing because verbal warnings about dangerous stimuli also require quick action. Thus, Fodor (1983) argued that, for both perception and language, speed of producing a representation is more important than considering all available information to produce the most accurate response. This speed is made possible by modularity, but at the price of intelligence of response.

Fodor (1983) cited an example of such fast but unintelligent processing from studies of the effects of sentence context on word recognition. Several researchers have demonstrated that the presence of an appropriate sentence context can facilitate response latencies in a lexical decision task (e.g., Schuberth & Eimas, 1977; West & Stanovich, 1978), which has been argued to be evidence

The experiments reported are part of a doctoral dissertation submitted to the Monash University Department of Psychology. I would like to express my gratitude to Ken Forster for his help and guidance throughout this study and his helpful suggestions after reading a first draft of this paper. Requests for reprints should be sent to Janet E. Ratcliff, Division of Cognitive Science & Psychology, Deakin University, Victoria 3217, Australia.

for top-down processing that is not consistent with the modularity hypothesis (e.g., West & Stanovich, 1978). However, Swinney (1979) demonstrated that such facilitation can occur, not only from a sentence context as a whole, but also from the meaning of a homographic prime that is inappropriate in the sentence context in which it occurs. What this clearly illustrated is that the priming process involved in this experiment was not an intelligent use of the sentence context to facilitate lexical access, as was originally suggested, but was in fact a highly unintelligent process involving associations between lexical items, which is perfectly consistent with the modularity hypothesis.

Swinney did obtain selective priming by the contextually appropriate sense in this task but only with a delay between prime and target (Onifer & Swinney, 1981; Swinney, 1979). This suggests that sentence context does not influence lexical access directly but is used to select between alternative senses at a later stage of processing. However, Glucksberg, Kruez, and Rho (1986) challenged Swinney's evidence. Using a variation of Swinney's technique involving nonword interference rather than priming, they produced evidence of selective priming by a sentence context even with no delay between target and prime. They suggested, therefore, that Swinney's results were merely an artifact of the backward-priming paradigm used. However, on inspection of their experimental items, it is apparent that their sentence contexts nearly always contained individual lexical items that were closely related to the misspelled target word (e.g., *door* and *keay*, *hot* and *coule*). This raises the possibility that their effect was simply due to intralexical priming rather than top-down sentence priming. Thus, this result does not provide a challenge to Swinney's suggestion of an unintelligent associative priming process that is consistent with the modularity hypothesis.

Fodor (1983) suggested that such priming effects reflect a network of associative connections between items in the lexicon, mirroring what is known about connectedness in the world, and that this may exist specifically to mimic the effects that an intelligent prediction based on sentence context might produce, but without entailing the processing cost associated with a genuine prediction of this type. That is, the language processor gains some of the advantages of intelligence cheaply by building knowledge of the world into the structure of the input system, yet retaining the advantages of modularity.

Semantic Priming: Lexical or Conceptual?

Recent evidence suggests that such an association process may also be the source of the lexical priming effect. Lexical priming, or semantic priming as it is sometimes called, refers to the well-documented finding that a target word such as *doctor* is more quickly responded to in a lexical decision task or naming task when it is presented after a semantically related prime word such as *nurse* than an unrelated word such as *butter* (e.g.,

Becker & Killion, 1977; Fischler, 1977; Meyer & Schvaneveldt, 1971; Meyer, Schvaneveldt, & Ruddy, 1975; Neely, 1976, 1977). It is widely accepted that such priming effects can be divided into two separate types: effects on lexical access, and strategic effects on decision processes that occur after lexical access is complete. In addition, although the lexical decision task is particularly prone to such conscious strategy effects, use of the naming task appears to be a good way to isolate preaccess effects (e.g., Balota & Lorch, 1986; Fischler & Bloom, 1980; Forster, 1981; Koriat, 1981; Lupker, 1984; Seidenberg, Waters, Sanders, & Langer, 1984; Stanovich & West, 1983b; Tanenhaus & Lucas, 1987).

Lupker (1984) recently reported that this semantic priming effect appears to be restricted to highly associated words (e.g., *bread* and *butter*) in the naming task, with little effect of semantic relatedness alone, although semantic relatedness (i.e., a sharing of semantic features as in *pelican* and *sparrow*) does produce priming in the lexical decision task (Fischler, 1977; Lupker, 1984). In light of the finding, mentioned above, that the naming task is a better indicator of effects on lexical access, Lupker's research appears to provide evidence that the only type of semantic priming to affect lexical access is that produced by words that are highly associated. This confirms the view put forward by Fodor (1983) that semantic priming is due to intralexical associative links that result from the common co-occurrence of words in the world.

Recent evidence reported by Seidenberg (1984) and Tanenhaus and Lucas (1987), however, suggests that there is an effect of pure semantic relatedness in the naming task. Tanenhaus and Lucas considered this as evidence against the idea of purely intralexical priming, such as Fodor (1983) proposed, since they assumed it reflects the involvement of the conceptual system. This assumption does not seem adequately justified. Although the priming of unassociated but semantically related words is clearly not the result of associations reflecting knowledge about connectedness in the world of the sort Fodor proposed, there is no evidence to suggest that there are not also lexical connections between semantically related words that could account for priming based on pure semantic relatedness without involving conceptual mediation.

In fact, Meyer and Schvaneveldt's (1971) original interpretation of the semantic priming effect, in terms of Collins and Loftus's (1975) semantic network/spreading activation model, could be interpreted in this way. In this model, each lexical entry, consisting of the name of a concept, is linked to a corresponding node representing that concept in a semantic network. These concept nodes have further links connecting them to nodes representing properties of that concept, and through these to other concepts with semantically related properties. According to this model, when a lexical entry is accessed, activation spreads out along the links of the network, raising the activation levels of neighboring nodes. Thus, priming by semantically related words is explained by this model in terms

of spreading activation through the semantic network, which exists to represent meaning relationships, and thus provides close links between semantically related words.

Lupker (1984) suggested this to be a different type of explanation from Fodor's (1983) intralexical associative network since it involves "higher level cognitive structures like those representing meaning relationships" (p. 710). This conclusion is reasonable if it is assumed that activation of the semantic network does involve higher level cognitive structures, as would be the case if the semantic network were part of a general conceptual level used for general cognitive processing. This was true in Collins and Loftus's (1975) original formulation, but is not necessarily so. A viable alternative view is that the semantic network is a way of representing word meaning and that this is a purely linguistic property. This seems reasonable given that word meaning would appear to be an essential component of any linguistic representation of verbal material and so should be incorporated within the language processor in some way. If this alternative view were accepted, then the semantic network could be validly considered part of the lexical level and therefore part of the language processor. This would then make priming via the semantic network a variation of the operation of intralexical association of the type that Fodor proposed.

Tanenhaus and Lucas (1987) shared Lupker's (1984) assumption that the finding of priming resulting from pure semantic relatedness implicates a general conceptual level and cannot be explained in terms of intralexical associations. Thus, they explicitly assumed that the semantic network is involved at a general conceptual level without considering the alternative, purely lexical level role proposed above. They further suggested that such conceptually mediated priming would also include priming by concepts that are activated as a result of combining word meanings during comprehension. For example, phrases such as *large dog* should prime the word *collie* more than the phrase *dog large*.

If such a result were found, it clearly would implicate a top-down activation process such as Tanenhaus and Lucas (1987) assumed, which would clearly violate the modularity hypothesis because it cannot be explained in terms of intralexical associations. However, as we have seen above, the finding of priming resulting from pure semantic relatedness is not necessarily evidence of conceptually mediated priming in this sense, because it could be explained in terms of an intralexical process. Until direct evidence of high-level conceptual involvement is found, there seems to be no need to consider semantic priming to be inconsistent with the modularity hypothesis.

That such evidence will be forthcoming seems unlikely given the example of sentence-context effects described earlier, in which the meaning of a homograph inappropriate to the sentence context in which it occurred produced a significant priming effect. This seems to suggest that it is not the concepts activated by the whole sentence meaning that produce the priming, as would be expected

from Tanenhaus and Lucas's (1987) view of conceptually mediated priming, but rather that Fodor's (1983) view of a "stupid" intralexical association process is more appropriate. Although it is suggestive, this evidence is clearly not conclusive on this issue, however, since it would be possible for both associative priming and conceptually mediated priming to be contributing to the effects. What is needed to clear up this issue is evidence demonstrating whether or not priming can be produced by concepts activated by combinations of words, as Tanenhaus and Lucas suggested. Thus, the most relevant area to look for such evidence would appear to be studies of sentence-context effects on word recognition, rather than the single-word studies of semantic priming that we have been considering.

Sentence-context Effects: Lexical or Sentential?

Stanovich and West (1983a, 1983b) recently made a suggestion similar to Fodor's (1983) that sentence-context effects are due to an automatic spreading activation process within semantic memory, although they did not address the issue of whether or not this semantic memory is a part of the language processor. They did, however, suggest that it could result from "semantic states induced by combinations of words" (Stanovich & West, 1983b, p. 30). They further suggested that all semantic context effects, including semantic priming, can be explained in spreading activation terms. Thus, they considered sentence-context effects and lexical priming to reflect the same underlying phenomenon. They proposed this as an explanation for Forster's (1981) failure to find facilitation effects of sentence contexts in an experiment in which semantic relationships and associations between words were avoided, in contrast to most studies of sentence context.

Interestingly, a corollary of this argument would appear to be that sentence-context effects are entirely the result of semantic relationships and associations between words, rather than of concepts activated by combinations of the words within the sentence. In fact, Forster's (1981) failure to find facilitation effects of appropriate sentence contexts that do not contain words semantically related to the target word would appear to provide negative evidence on whether priming can result from concepts associated with word combinations, which we have seen to be the critical issue in discriminating between intralexical associative priming and conceptually mediated priming in the sense that Tanenhaus and Lucas (1987) used it.

This example highlights a problem that has become increasingly apparent in the discussion of semantic context effects. Recently, a number of authors (e.g., Forster, 1981; Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; Stanovich & West, 1983b; Tanenhaus & Lucas, 1987) pointed out apparent contradictions in evidence relating to semantic context effects. These contradictions have resulted from a failure to distinguish between differ-

ent types of semantic context, particularly between appropriateness and/or predictability of sentence contexts and the presence of semantic relationships and associations, with the words making up the sentence context.

As we have discussed, this is a particular problem in attempting to evaluate the evidence relating to the modularity of the language processor. It is probably premature to draw firm conclusions from the evidence discussed above for semantic context effects, but there does seem to be general agreement that lexical priming and sentence-context effects result from a common automatic spreading activation process of the Collins and Loftus (1975) variety. Whether this spreading activation process can be explained in terms of intralexical associations consistent with the modularity hypothesis is a more contentious issue, but none of the evidence discussed above was inconsistent with this proposal.

For the rest of this paper, the term *lexical priming* will be used only to refer to such an intralexical associative priming process, be it associative or semantic, rather than in the more general sense that includes conceptually mediated priming. This is merely for convenience of discussion; it should be kept in mind that the possibility of conceptually mediated priming has not been conclusively ruled out.

The Plausibility Effect: Lexical or Sentential?

The study of semantic context effects is not the only area of research in which potential confusion exists between the effects of lexical association and sentential meaning. An obvious candidate for reconsideration in these items is the *plausibility effect*, which is the well-established finding that plausible sentences, which describe very ordinary likely events such as THE DOG CHEWED THE BONE, are more quickly and accurately processed than implausible sentences, which describe bizarre or unusual situations such as THE OCTOPUS ATE THE REFRIGERATOR. This effect has been found by many experiments in a wide variety of tasks. The explanations for the effect have varied enormously, but have always had in common the assumption that some aspect of sentence meaning is responsible for the difference in processing time or accuracy between plausible and implausible sentences.

This effect was first reported in shadowing and recall tasks (Marks & Miller, 1964; Miller & Isard, 1963; Rosenberg, 1968, 1969; Rosenberg & Jarvella, 1970); however, as Rosenberg (1969) acknowledged, plausibility effects in these tasks may be due to easier reconstruction of plausible sentences from partial memory traces at the time of recall and may have nothing to do with normal sentence processing. A plausibility effect has also been found on accuracy of recall in the RSVP task (Forster & Ryder, 1971), even when a two-alternative forced-choice procedure is used to eliminate the possibility of any guessing strategies or criterion shift. This was explained in terms of faster semantic interpretation of plausible sentences leading to more of the sentence being successfully encoded in memory before the end of the very

brief display (Forster, 1974). However, Murray (1982) suggested that the effect could reflect memory processes rather than perceptual processes and that the way to overcome this problem is to measure reaction time to clearly presented and freely available stimuli rather than to measure accuracy of report under degraded stimulus presentation conditions.

The effect of plausibility has been demonstrated in a number of reaction-time (RT) tasks, even in some for which processing of meaning would appear to be completely irrelevant. For example, Forster and Olbrei (1972) and Watson (1976) found that plausible sentences are classified as grammatical more quickly than implausible sentences (the grammaticality task), and Murray (1982) showed that sentences are classified as physically identical faster if their meanings are plausible (the sentence matching task). Like Forster (1974), Watson and Murray both offered explanations in terms of faster processing of the meanings of plausible than of implausible sentences, with the additional suggestion that the processing of meaning was obligatory (Watson, 1976) or faster than lower level processing (Murray, 1982) to account for its occurrence in those tasks for which meaning processing would appear to be irrelevant.

Thus, all the varied explanations for the plausibility effect have shared the assumption that some property of sentential meaning is responsible for the effect. A quick glance at the examples of plausible and implausible sentences used in such experiments make it clear, however, that the plausibility of sentential meaning is not the only variable by which researchers' plausible and implausible sentences differed. For example, among the small number of plausible sentences used as examples were THE DOCTOR CURED THE PATIENT (Rosenberg, 1969), THE CONDUCTOR ASKED THE VIOLINS TO PLAY LOUDER (Forster, 1974), and THE DENTIST EXTRACTED THE TOOTH (Murray, 1982). The implausible counterparts of these sentences were THE DOCTOR SHOOK THE AUTHOR (Rosenberg, 1969), THE CUSTOMERS ASKED THE VIOLINS TO LEAN STEADILY (Forster, 1974), and THE CLERGY ADVERTISED THE BONES (Murray, 1982). It is obvious that these sentences, like the appropriate sentence contexts of the semantic context experiments, not only are varied in sentential plausibility, but also in the number of semantically related and associated words they contain. This suggests the possibility that the plausibility effect is another case in which apparently high-level sentential processing effects may be more simply explained in terms of a low-level associative process, as appears to be the case for sentence-context effects on word recognition.

In fact, the same intralexical associative process leading to faster lexical access of the semantically related words within the plausible sentences could be sufficient to account for all the effects of plausibility that have been found. Such priming would obviously account for faster overall processing times in the RT tasks as a direct result of speeding the lexical access stage of processing. It would

also account for the greater recall of plausible sentences found in the RSVP task, because faster processing of the related words would enable more words to be encoded before the end of the very brief display, in a similar way to the advantage given by faster semantic interpretation in Forster's (1974) explanation. For the shadowing experiments a similar account could be given, whereas for the recall experiments even the priming explanation is unnecessary because it is well documented that recall is higher when related words are involved. Thus, intralexical associative priming provides a viable explanation for the plausibility effect that needs to be investigated.

It should be noted that Forster's (1974) higher level explanation for the plausibility effect is not inconsistent with the modularity hypotheses, as a genuine sentential context effect on lexical access would be, because it is suggested that the effect is occurring entirely at the stage of message-level processing. Thus, each of the competing explanations involves an effect within a single subsystem of the language processor, but the two explanations differ in which subsystem they suggest to be involved—whether the message level or the lexical level.

Lexical Priming Explanations of Plausibility: Empirical Evidence

Murray (1982) addressed the issue of whether the plausibility effect could be explained in terms of lexical priming by taking plausible and implausible sentences that had been shown to produce a plausibility effect and scrambling them, thus producing strings such as BEEN HAD WOUNDED THE SOLDIER. Murray argued that if the plausibility effect were purely lexical, it would be unaffected by such scrambling, because the relationships between the words remained unaltered, but an effect on sentence meaning would be destroyed. Using the sentence matching task, he found no difference in reaction time between scrambled plausible and scrambled implausible sentences, suggesting that the effect must have been due to sentence meaning rather than to intralexical priming.

One problem with Murray's (1982) experiment is that, since lexical priming can be drastically reduced by the presence of intervening words (Bodi, 1982; Gough, Alford, & Holley-Wilcox, 1981), the effect may have been destroyed by the separation of related words rather than the disruption of sentence meaning. Murray argued that this was unlikely because the related words were sometimes closer after scrambling than before, but clearly control of the distance between related words would be preferable.

In an unpublished study, K. I. Forster (personal communication, July, 1977) very neatly eliminated the existence of related words as the sole cause of the plausibility effect by showing that implausible sentences constructed by reversing plausible sentences, such as THE PATIENT CURED THE DOCTOR, were significantly slower than their plausible counterpart, which for this example was

THE DOCTOR CURED THE PATIENT. The task Forster used for this experiment was another RT task, the multiple lexical decision task, in which subjects are asked to decide whether all the words are proper English words, or whether any nonsense words are present. He found that with this task also, plausible sentences were responded to more quickly than implausible sentences. In this experiment, since both plausible and implausible sentences contained the same words at the same distances from each other, the fact that this plausibility effect still occurred is strong evidence against lexical priming as its sole cause.

However, there remains one possible criticism of this conclusion, which is that the order in which the related words occur might be critical, because it is possible that the associations producing the priming may be unidirectional, and hence be effectively unrelated in the opposite order. For example, *doctor* might prime *cure* but not vice versa. This would be reasonably likely since, by Fodor's (1983) account, words become associated by often occurring together in the world; presumably priming will be stronger in the order in which they usually occur than in the reverse direction. That is *bread* will prime *butter* better than vice versa because the words usually occur in that order in the phrase *bread and butter*. If the co-occurrence of words used in plausible sentences, such as THE DOCTOR CURED THE PATIENT, is usually in that same order, which seems likely given the nature of plausibility, then it would seem quite reasonable to suggest that in general the subject of a plausible sentence (in this example, *doctor*) will be a better prime for the verb (in this example, *cured*) and the verb will be a better prime for the object (in this example, *patient*) than in the reverse order. Thus, Forster's evidence cannot quite rule out the possibility of an associative priming explanation of the plausibility effect. This explanation seems particularly unlikely in the face of evidence that backward priming of unidirectional associates occurs in the lexical decision task (Koriat, 1981; Seidenberg et al., 1984). However, a more direct investigation of this unidirectional priming explanation will be undertaken in this paper.

Bodi (1982) found that the existence of related words in sentences had no effect on reaction times in the grammaticality task, although the same sentences did produce a relatedness effect in another task more sensitive to lexical effects. This would appear to support the idea that the grammaticality task is not influenced by lexical priming; therefore, the effect of plausibility found in such tasks is unlikely to be the result of lexical priming. However, Bodi did find a nonsignificant trend toward an effect of lexical relatedness that was of roughly the same magnitude as the significant effect in the more sensitive task. Thus, it is possible that lexical priming may influence the grammaticality task, but that the task is less sensitive to effects of such small magnitudes. Conditions that might produce a larger priming effect or increase the statistical power of the comparison might well result in a signifi-

cant priming effect in the grammaticality task, so this evidence cannot rule out lexical priming as the source of the plausibility effect.

Thus, the evidence from Murray (1982) and Forster (1977) suggests that purely intralexical associative priming is probably not the cause of the plausibility effect, but cannot completely rule it out as a possibility. The purpose of this paper is to provide more conclusive evidence.

EXPERIMENT 1

The first question we addressed was whether a plausibility effect would occur in the absence of syntactic processing, even when the relative position of the content words remained unaltered. If it would not, as Murray's (1982) and Bodi's (1982) results suggested would be the case, this would provide evidence against purely intralexical priming's being the source of the plausibility effect, since such lexical priming should be unaffected by the disruption of syntactic processing. However, before the effect of the disruption of syntax could be tested, it was necessary to start with a set of items for which a plausibility effect was known to occur. In Experiment 1, such a set of items was constructed and tested using the grammaticality task.

In constructing plausible and implausible sentences to test for the effect of plausibility, it is obviously important to control for lexical factors, such as word length and frequency. One difficulty with doing this is that, despite the most careful matching in terms of word frequency norms, a consistent bias toward the use of less natural words in implausible sentences tends to occur. This is because words such as *damsel*, which seem natural only in a specific historical context, reduce the overall plausibility of any sentence in which they occur, and so will tend to be rejected in the construction of plausible sentences, although they may appear in implausible sentences. For example, the sentence THE BOY KISSED THE DAMSEL would be rejected as a plausible sentence because it does not sound very natural, but the sentence THE AARDVARK SADDLED THE DAMSEL would make an acceptable implausible sentence. Clearly, however, the implausibility of this sentence may be partly the result of the unnaturalness of its component words, which may be a lexical property rather than a property of sentence meaning.

Controlling word frequency does not eliminate the possibility of such a bias. Gernsbacher (1984) showed that words of the same low frequency in counts of printed frequency, such as those of Kučera and Francis (1967), do differ in naturalness, or "experimental familiarity" as she calls it. She demonstrated this to be a problem sufficient to have caused inconsistent results in different studies of interactions between the effects of word frequency and other variables, such as orthographic regularity, semantic concreteness, and polysemy. Thus, a systematic bias in naturalness or experimental familiarity could occur between plausible and implausible items in which frequency of usage has been controlled.

In this experiment, the possibility of such a bias was eliminated by constructing the plausible and implausible sentences from the same subset of words and using a counterbalanced design to minimize repetition. Thus, any difference found between the plausible and implausible items in this experiment must be due to the combination in which the words occur: either lexical priming or a sentential effect of plausibility. It was decided that if an effect were found in these sentences, a further experiment would be attempted to decide which of these two possible word combination effects is the source of the plausibility effect.

Method

Items. Thirty plausible and implausible items were constructed in pairs that differed by only one word. For example, the plausible sentence THE USHER TOOK THE TICKETS had an implausible counterpart THE TAILOR TOOK THE TICKETS, which differed by only one word. The words that differed between the two sentences also occurred in another pair of sentences in the same condition, but with the opposite plausibility value. For example, THE TAILOR MADE THE UNIFORMS was a plausible sentence, and THE USHER MADE THE UNIFORMS was its implausible counterpart. Thus, the set of words making up the plausible and implausible sentences were identical, and the syntactic structure was identical for each plausible and implausible pair, so any plausibility effect that may occur in this experiment cannot be attributed to variations in syntax or lexical properties.

Sixty other grammatical sentences of a similar type were also included to test separate hypotheses that are not discussed in this paper.

Since the subjects' task was to decide whether or not the items formed grammatical sentences, distractor items that were ungrammatical and to which the response was "no" were also constructed. Ninety such distractor items were included so that the subjects made an equal number of randomly interspersed "yes" and "no" decisions, and so that they were unable to predict what the next response would be. These items were based on sentences of the same syntactic structure as the test items but with different content words. They were converted to ungrammatical items by swapping the positions of two words (e.g., THE FIGHTER HIT MAT THE). Half of these items were constructed from plausible sentences, and half from implausible, to mimic the experimental items. Because each role-word (i.e., the syntactic subject of the sentence) in the experimental items was repeated once for each subject, half of the distractors also contained a role-word repeated from an experimental item so that subjects would be unable to determine the correct response from the fact that repetition had occurred. Thus, each role-word occurred three times: once in a plausible experimental item, once in an implausible experimental item, and once in a distractor item. There were also 10 practice items, which were a random assortment of items of the above types.

Ratings. Ten raters assessed the plausibility of each of the experimental items. They were asked to assign a value between 1 and 7 to each sentence, where 7 corresponded to *very plausible*, 1 corresponded to *very implausible*, and the other scores represented gradings on the scale at equal intervals. Plausibility was defined as

the ordinarieness or likelihood of the event described by the sentence. Thus a plausible sentence will describe a very ordinary event which has a high probability of occurring in everyday life, whereas an implausible sentence will describe a very bizarre or unexpected event which is not very likely to occur.

Raters were also given examples of plausible and implausible sentences.

The plausible sentences produced a mean rating of 6.6, and the implausible sentences had a mean rating of 2.9. A *t* test showed this difference to be significant [$t(29) = 17.3, p < .001$], which confirmed that plausibility had been manipulated as intended.

Subjects. Forty Monash University students took part in the experiment and were paid for their participation.

Design. An incomplete within-subjects design (Shaughnessy & Zechmeister, 1985) was used, involving two groups of 20 subjects, to counterbalance the effects of different item sets while minimizing repetition of words for each subject. Each group saw the plausible member of half of the plausible/implausible item pairs (Set 1) and the implausible member of the other half of the pairs (Set 2). The other group saw the opposite member of each pair. For example, Group 1 saw THE TAILOR MADE THE UNIFORMS (Plausible Set 1) and THE TAILOR TOOK THE TICKETS (Implausible Set 2), whereas Group 2 saw THE USHER TOOK THE TICKETS (Plausible Set 2) and THE USHER MADE THE UNIFORMS (Implausible Set 1). Each subject, therefore, saw only one version of each item, but over the two groups each item appeared in both versions. Each role-word was repeated within a group, however, and as mentioned above, distractors also containing repeated words were included so that repetition could not be used as the basis for making decisions. The same practice items were seen by both groups of subjects, and the distractors were also the same except for the repeated role-words, which were different for the two groups.

Procedure. In this and subsequent experiments, each subject was presented with written instructions, seated in a booth looking at a video screen controlled by a PDP-11 computer, and then tested individually. Items consisting of 7×9 dot-matrix characters were presented on the screen, always in uppercase. The subjects responded to the items by pressing a blue button with the index finger of the preferred hand if the item was grammatical or a red button with the nonpreferred hand if the item was ungrammatical. They were asked to respond as quickly and accurately as possible and were given feedback after each trial. The subject brought up the next item by pressing a foot pedal, thus controlling the inter-trial interval. A summary of the instructions was presented again on the screen, then practice items were presented, followed by a mixture of experimental items and distractors in a different pseudo-random order for each subject. The subjects were informed when they were halfway through the experiment.

In this experiment, each sentence was presented for 3 sec. The subjects were instructed to press the blue button if the string of words formed a grammatical English sentence or the red button if it was ungrammatical, and examples were given for each type of response.

Results

To minimize the problems associated with the positive skew of RT distributions in this and subsequent experiments, outlying scores were adjusted to cutoff values two standard deviations from the mean RT for each subject. Furthermore, any scores below 200 msec were omitted altogether. In addition, subjects having an error rate of 25% or more (over all items in the experiment) would have been excluded from the analysis. No subjects in this experiment were excluded for this reason.

The mean plausibility ratings, RTs, error percentages, and plausibility differences for each of the experimental conditions (which contained only grammatical sentences) are shown in Table 1. The ungrammatical distractor items were not analyzed. Two analyses of variance were performed, one based on the mean scores for each subject and one based on the mean scores for each item. A treatment × blocks design was used (Keppel, 1982) to eliminate sources of variance associated with the different subject groups in the subject analysis and the different item

Table 1
Mean Reaction Times (RT in msec) and Percentage Errors (%E) in the Grammaticality Task for Plausible and Implausible Items in Experiment 1, and Plausibility Differences in Reaction Times and Errors

	Example of Items	RT	%E
Plausible	THE USHER TOOK THE TICKETS	1094	2.0
Implausible	THE TAILOR TOOK THE TICKETS	1131	2.3
Difference		37	0.3

sets in the item analysis. The two factors in each 2×2 analysis of variance were groups (groups of subjects for the subject analysis, and sets of items for the item analysis) and plausibility (plausible vs. implausible), with repeated measures on the plausibility factor. A $\min F'$ value was calculated on the basis of this analysis. Following the recommendation of Santa, Miller, and Shaw (1979), a 10% alpha level was used to counteract the conservatism of $\min F'$.

The 37-msec RT difference between plausible and implausible items was significant for both subject and item analyses [$F(1,38) = 12.49, p < .01$; $F(1,28) = 5.67, p < .05$] and was significant by $\min F'$ at the 10% alpha level [$\min F'(1,51) = 3.90, p < .10$]. There was no significant difference in the errors for the plausible and implausible conditions ($F(1) < 1, F(2) < 1$).

Discussion

The results show that the plausible items tested in this experiment produced a consistently faster RT than did the implausible items. This difference cannot be attributed to lexical factors since the same words were used in both conditions. It must therefore be the combination in which the words occurred that produced the effect. These items, therefore, appear to be suitable for testing the contribution of asymmetric lexical priming to this word combination effect.

However, one possible problem with the design of this experiment was the use of repeated subject terms. It could be argued that, rather than being an effect of sentential plausibility on normal sentence processing, the results of this experiment may reflect subjects' unusual processing of the items containing the second presentation of each role-word. This unusual processing would be due to previous experience of that role-word in a sentence of the opposite plausibility or in an ungrammatical distractor item. To show that the plausibility effect does represent an effect on normal sentence processing rather than on an experiment-specific episodic memory effect, it is necessary to show that it occurs in the items in which the subject items are presented for the first time. A post hoc analysis of the results showed that the 12 items (out of the total set of 30) in which the subject term occurred for the first time showed a plausibility effect of comparable size to that of the overall effect. In fact, the effect for these nonrepeated items was slightly larger (43 msec as com-

pared with the original 37 msec). This clearly shows that the repetition of role-words is not responsible for the plausibility effect found in this experiment.

The results of this experiment, then, provide a demonstration of word combination effect that cannot be attributed to the effects of repetition. Thus the items from this experiment were deemed suitable for use in the next experiment in which an attempt was made to decide whether this combination effect is a genuine sentential plausibility effect or whether it can be accounted for in terms of asymmetric lexical priming.

EXPERIMENT 2

As mentioned earlier, Murray (1982) attempted to show that lexical priming could not be the cause of the plausibility effect by demonstrating that when plausible and implausible sentences are scrambled, they no longer produce different RTs. For example, scrambled versions of plausible sentences (e.g., BEEN HAD WOUNDED THE SOLDIER) do not differ in RT in the sentence matching task from the scrambled versions of implausible sentences (e.g., THE STONE THE WORM SWALLOWED). Since such scrambles contain the same related words as the sentences from which they are constructed, the fact that disrupting syntax by scrambling of word order destroys the effect implies that the plausibility effect is sentential rather than the result of lexical priming.

In this experiment, similar logic was used, but the disruption of syntactic processing was achieved without altering the relative positions of the content words. The items in this experiment were modified versions of the sentences from Experiment 1, which were shown to produce a plausibility difference in RT. Instead of scrambling the sentences, as Murray (1982) did, syntax was disrupted by removing the function words (which could not contribute to lexical priming) and replacing them with unrelated three-letter content words. For example, THE USHER TOOK THE TOCKETS became the word string SAG USHER TOOK SIP TICKETS. If the plausibility effect that occurred in Experiment 1 were due to lexical priming, then the same effect should occur when subjects read these word strings, because they contain the same content words in the same positions relative to each other. However, if the plausibility effect were the result of processing the meaning of the sentence, then the effect should not occur for these word strings, because sentential processing should not be possible. It should be noted that this manipulation does not produce a perfect baseline measure of normal processing without syntax, because it could be argued that the replacement of function words by content words may have other effects on processing apart from the disruption of syntax. This issue will be addressed in Experiment 5.

The grammaticality task cannot be used with items such as the word strings described above, because Ratcliff (1983) has shown that such ungrammatical items, for which a "no" decision is appropriate in the grammaticality task, are not sensitive to plausibility if the first few

words of the sentence do not form a possible grammatical sequence. This is not surprising because subjects can make a "no" decision on the basis of a few words without completing sentential processing. That this is the strategy subjects adopt is borne out by the fact that such items produce very short RTs, whereas "no" decisions to items which contain a deviation from grammatical sequencing only toward the end of the sentence have larger RTs than "yes" items and are sensitive to plausibility.

In this experiment, therefore, a decision was made to use the multiple lexical decision task, which Forster (1977) (described above) has shown to be sensitive to plausibility. This task is an extension of the task originally used to demonstrate priming (e.g., Meyer & Schvaneveldt, 1971).

Although the nature of the task and the nonsentential nature of the items should discourage syntactic processing, the stereotyped structure of the items with long content words in the second, third, and fifth serial positions and short content words in the first and fourth positions may alert subjects to the existence of a relationship between the larger content words that formed the main constituents of the original plausible and implausible sentences. To disguise this stereotyped structure, some filler items were included that had quite different structures and no underlying sentential relationships between the words.

Method

Items. The experimental items were constructed from the plausible and implausible items from Experiment 1 by removing the function words and replacing them with three-letter content words. The matched plausible and implausible pairs of word strings contained the same three-letter content words. For example, the plausible word string SAG USHER TOOK SIP TICKETS was matched with the implausible word string SAG TAILOR TOOK SIP TICKETS. To prevent subjects from becoming aware that the word strings were modified sentences, which could lead to undesirable strategies, 15 filler items were constructed containing words completely unrelated to each other, with verbs, if any, in serial positions different from those in the experimental items (e.g., FISH PRIEST BACK FILLED SOLO). Forty-five distractor items were constructed. Thirty were of the same structure as the experimental items but contained a legal nonword (e.g., LAD TOGICIAN DARKENED PEW STATE). Also, as in Experiment 1, 15 of these contained a role-word repeated from the experimental items and 15 were similar to the filler items, but contained a nonword. Nonwords appeared in all serial positions to ensure that subjects had to read the entire string before making a "yes" decision. The 14 practice items were a random assortment of each of the above item structures.

Subjects. The subjects were 40 Monash University students who had not taken part in Experiment 1. They were paid for their participation.

Design. A counterbalanced design involving two groups of 20 subjects was used for the experimental items, as in Experiment 1. The same fillers and practice items were seen by both groups of subjects. The distractors were also the same except for the repeated role-words, which were different for the two groups.

Procedure. The procedure was the same as for Experiment 1, except that the subjects' task was different; they were instructed to push the blue button if the word string was made up entirely of proper English words or the red button if it contained a non-sense word.

Results

The mean RTs and percentage errors for each condition are presented in Table 2. No subjects were excluded from analysis for exceeding the accepted error criterion. The plausibility difference was analyzed as in the previous experiment. The 59-msec RT effect was significant [$\min F'(1,52) = 5.57, p < .05$]. There was no significant error effect [$F(1,38) = 1.06, p > .05; F_2 < 1$].

Discussion

The results showed that there was a clear difference between the plausible and implausible items from Experiment 1, even when the function words were replaced by content words. This suggests that the plausibility effect observed in Experiment 1 may not have been produced by sentential processing at all since the items in the current experiment were designed to prevent sentential processing. The simplest explanation appears to be that lexical priming caused the effect in both experiments. This suggests that the results of Forster (1977) and Murray (1982), which seemed to eliminate lexical priming as the source of the plausibility effect, may have been the result of failure to control the order in which the related words occurred in their experiments.

However, other possibilities need to be considered before this conclusion can be accepted. There were two rather striking differences between the results of this experiment and those of Experiment 1 that need to be accounted for: (1) the size of the RTs obtained, which were about 900 msec longer for the multiple lexical decision task than for the grammaticality task, and (2) the size of the plausibility effect, which was approximately twice the magnitude in Experiment 2 as in Experiment 1.

If subjects were merely relying on lexical analysis of the input, the long RTs would be surprising because the grammaticality task must include the same lexical analysis as the lexical decision task, with the addition of syntactic analysis. The grammaticality task would therefore be expected to produce longer rather than shorter RTs relative to those of the lexical decision task. Some of this difference may be explained in terms of the extra difficulty of accessing the two content words, which were lower in average frequency than were the function words, because Bradley (1978) showed that function words were accessed at approximately the same speed as high-frequency content words. However, a frequency effect for two words certainly could not account for a 900-msec difference.

The most likely reason for the longer RTs in this experiment is that the decision component forms a larger proportion of the RTs in the lexical decision task. A larger decision component in the lexical decision task is likely, because a separate decision about lexical status has to be made for each word in addition to the final overall decision that leads to the response. In the grammaticality task, on the other hand, only a single decision needs to be made about whether a grammatical structure can be assigned. The fact that subjects reported finding the lexical decision task tiring and produced much higher error rates than in the grammaticality task supports the suggestion of a larger decision component.

The problem with all tasks involving a lengthy and difficult decision component is that they provide opportunity for conscious strategies that have nothing to do with normal sentence processing, but are designed to simplify the decision component of reaction time. Such a strategy would be possible in this task because, as can be seen clearly from a quick inspection of some of the items used—such as AWE TAILOR MADE HOP UNIFORMS, HUB VISITOR ADMIRER ROB HOUSE, and USE BURGLARS PICKED CUP LOCK—it is not difficult to pick out the words that make up a sentence, especially because the verbs retain their inflections. Because the decision about individual words is obviously a very difficult one, subjects may actually find it easier and faster to check whether there are words that make sense together as a sentence and hence infer that they must all be proper words than to make a decision about those words individually. If such a strategy were adopted, then a plausibility effect may occur despite the attempts to prevent syntactic processing, rather than a lexical priming effect. The possibility of such a strategy was the reason for including filler items without such an underlying sentence structure, but their presence may not have been sufficient to distract subjects from noticing the structure in the experimental items.

The existence of such a strategy would also account for the larger plausibility effect in this experiment. This would not be expected if the effects were the result of the same lexical priming process in the two experiments, but would make sense if the effect in Experiment 2 were due to a conscious strategy rather than to normal linguistic processing. Thus, before any conclusion could be drawn from this experiment, it was necessary to eliminate the possibility that the effect was the result of task-specific strategy rather than of normal linguistic processing. This was attempted in the next experiment.

EXPERIMENT 3

In this experiment another attempt was made to test whether a plausibility effect can survive the disruption of syntactic processing when the relative order of the content words in the sentence was maintained. Thus, the same items were tested as in the previous experiment. This time

Table 2
Mean Reaction Times (RT in msec) and Percentage Errors (%E)
in the Multiple Lexical Decision Task for the Plausible and
Implausible Items from Experiment 2, and the Plausibility
Difference in Reaction Time and Errors

	Example of Items	RT	%E
Plausible	SAG USHER TOOK SIP TICKETS	1971	9.5
Implausible	SAG TAILOR TOOK SIP TICKETS	2030	8.0
Difference		59	-1.5

a task was used that, unlike the multiple lexical decision task, should not be prone to conscious strategies that use relatedness to simplify the decision processes involved in a way quite unrelated to normal linguistic processing. The problem with such strategies is that they may produce an effect of lexical relatedness that has nothing to do with normal lexical priming.

The task chosen for this experiment was the sentence matching task in which subjects decide whether two strings of words are physically identical. As discussed earlier, Murray (1982) showed that this task, like the grammaticality task, is sensitive to plausibility. He also pointed out that it is unlikely to be influenced by any sort of conscious strategy, because a decision about physical identity cannot be made on the basis of whether any lexical or sentential properties exist.

In this experiment, then, if a difference is found between the plausible and implausible word strings, which contain the content words from the stereotyped plausible and implausible items from Experiment 1 in the same order, but with the function words replaced with content words to prevent any sentential processing from occurring, then conscious strategies will not be a possible explanation. Rather, such a result would suggest that lexical priming is the source of the plausibility effect. If, on the other hand, no effect occurs in this experiment, then it will suggest that the content words in the plausible and implausible sentences in Experiment 1 are not sufficiently related to produce a lexical priming effect, and therefore that lexical priming is not the cause of the plausibility effect.

Method

Items. The experimental items were the same as in the previous experiment, but were presented simultaneously as identical pairs, one above the other, for example,

SAG USHER TOOK SIP TICKETS
SAG USHER TOOK SIP TICKETS

There were 30 distractors, which were modeled on the experimental items but contained a word in the second of the pair of sentences that differed by a few letters from the corresponding word in the first sentence, for example,

TOP SAILORS OPPOSED TEN CAPTAIN
TOP SAILORS OPPOSED TAN CAPTAIN

As in the two previous experiments, half the distractors contained role-words repeated from the experimental items. Differences occurred in each of the five serial positions to ensure that subjects had to read all the words to make a "yes" decision. No fillers were included because they were unnecessary for this task. The 14 practice items were once again a random selection of the above types.

Subjects. The subjects were 40 Monash University students who had not taken part in either of the previous experiments. They were paid for their participation.

Design. The design was identical to that of Experiment 2, except that there were no filler items.

Procedure. The procedure was the same as that for the two previous experiments, except that the subjects' task was different; they were instructed to push the blue button if the two word strings were identical or the red button if they were not.

Results

The mean RTs and percentage errors for each condition are presented in Table 3. No subjects were excluded from analysis for exceeding the accepted error criterion. The results were analyzed as in the previous experiment. There was no significant RT difference ($F1 < 1, F2 < 1$) or error difference [$F1(1,38) = 3.09, p > .05; F2(1,28) = 3.51, p > .05$].

Discussion

The failure to find an effect for the items from Experiment 2 with the sentence matching task, which is unlikely to share the conscious strategies possible in the multiple lexical decision task, suggests that the result in the previous experiment may have been due to such a task-specific strategy. This, then, would mean that the results of Experiment 2 do not implicate lexical priming as the source of the plausibility effect in Experiment 1. In fact, if lexical priming had been the source of the effect in Experiment 1, it should also have occurred in this experiment, because the task has been shown to be sensitive to a number of lexical properties (e.g., Chambers & Forster, 1975) and sentential properties (Freedman, 1982; Murray, 1982) and therefore must involve at least lexical access. Also, Forster (1979) reported that lexical priming does occur with this task.

However, the possibility should be considered that the null result in this experiment could be the result of insufficient sensitivity of the task. Murray (1982) has shown that the task does produce plausibility effects, but it is still possible that the present manipulation of plausibility (or relatedness), which barely produced a significant effect in Experiment 1, may be insufficient to produce a significant effect with the sentence matching task. Before any conclusion can be drawn about the possible existence of lexical priming for the content words in Experiment 1, it is necessary to check that the sentence matching task is sufficiently sensitive to detect such an effect.

It should be noted that the long RTs in this experiment (200 msec longer even than those for the lexical decision experiment) should not necessarily be taken as a sign of extensive decision processes. The sentence matching task involves the analysis and comparison of two word strings, and thus the long RTs in this task do not suggest the likelihood of strategies as in the previous experiment, where only one string of five words needed to be analyzed.

Table 3
Mean Reaction Times (RT in msec) and Percentage Errors (%E) in the Sentence Matching Tasks for the Plausible and Implausible Items in Experiment 3, and the Plausibility Differences in Reaction Times and Errors

	Example of Items	RT	%E
Plausible	SAG USHER TOOK SIP TICKETS SAG USHER TOOK SIP TICKETS	2221	4.0
Implausible	SAG TAILOR TOOK SIP TICKETS SAG TAILOR TOOK SIP TICKETS	2230	6.5
Difference		9	2.5

EXPERIMENT 4

To check the sensitivity of the sentence matching task, the plausible and implausible items from Experiment 1, on which the items from Experiments 2 and 3 were based, were used in this experiment with the sentence matching task. If an effect is found in this experiment, it would rule out insensitivity of the task as the reason for the null result in Experiment 3.

Method

Items. The experimental items were the plausible and implausible items from Experiment 1, but were presented simultaneously in identical pairs, one above the other, for example,

THE USHER TOOK THE TICKETS
THE USHER TOOK THE TICKETS

Thirty distractors were included that were similar to the experimental items but that contained a word in the second sentence of the pair that differed by a few letters from the corresponding word in the first sentence, for example,

THE FIGHTER HIT THE MAT
THE LIGHTER HIT THE MAT

As in the three previous experiments, half of these distractors contained a role-word repeated from the experimental items, and differences occurred in each of the five serial positions. The 10 practice items were a random selection of items of the above types.

Subjects. The subjects were 40 Monash University students who had not taken part in any of the previous experiments. They were paid for their participation.

Design. The design was identical to that of Experiment 3.

Procedure. The procedure was identical to that of Experiment 3 except for a slight modification in the wording of the instructions due to the use of sentences rather than word strings.

Results

The mean RTs and errors for each condition are presented in Table 4. Three subjects were excluded from analysis for exceeding the accepted error criterion. The results were analyzed as in the previous experiment. The 47-msec effect of plausibility was marginally significant [$F(1,38) = 6.85, p < .05$; $F(2,1,28) = 5.90, p < .05$; $\min F'(1,61) = 2.98, p > .1$]. There was no error effect ($F(1 < 1, F(2 < 1)$).

Discussion

The results of this experiment are very similar to those of Experiment 1: The RT difference is a little larger, which is to be expected because the processing of a second sentence is involved, but the level of significance is quite similar; the RT differences for both tasks were significant by both $F(1)$ and $F(2)$; the grammaticality effect was significant by $\min F'$ at the 10% level, but the effect for the sentence matching task was not quite significant. Thus, the grammaticality and sentence matching tasks appear to be comparable in sensitivity, so the null result in Experiment 3 for word strings does not appear to have been due to insensitivity of the task. It seems likely, therefore, that the apparent lexical priming effect in Experiment 2 was due to a task-specific strategy, and that the plausibil-

Table 4
Mean Reaction Times (RT in msec) and Percentage Errors (%E) in the Sentence Matching Task for the Plausible and Implausible Items in Experiment 4, and the Plausibility Differences in Reaction Times and Errors

	Example of Items	RT	%E
Plausible	THE USHER TOOK THE TICKETS THE USHER TOOK THE TICKETS	1516	4.8
Implausible	THE TAILOR TOOK THE TICKETS THE TAILOR TOOK THE TICKETS	1563	4.8
Difference		47	0.0

ity effects in Experiments 1 and 4 were not caused by lexical priming.

However, one further criticism of Experiment 3 remains. Bodi (1982) found that related words that produce priming when presented in a list separated by two function words (not in an appropriate order for sentential processing) do not produce priming when separated by two content words. For example, a word string containing a related pair separated by function words (e.g., AS PIPE OFF BY TOBACCO) was responded to more quickly than a matched unrelated word string (e.g., ON LIEUTENANT TO OF TOBACCO), whereas the same related pair in a context of other content words (e.g., DAY PIPE HAT OWE TOBACCO) was not faster than its unrelated counterpart (i.e., LOG LIEUTENANT OWN MOB TOBACCO). Bodi points out that this could be explained in two different ways: (1) in terms of greater salience of the related words in the context of function words, which could enable adaptive strategies to occur, or (2) in terms of a disruption of automatic priming effects by the intervening content words, which does not occur with function words because they do not share the same semantic network system. This second explanation could be due either to the number of content words intervening or to the rate of content word access, exceeding some critical limit on priming. Bodi attempted to decide between the strategic and automatic processing explanations experimentally, but obtained indeterminate results.

The possibility of such a disruption of automatic priming by content words but not function words is critical to the interpretation of the present experiments, because the difference between the sentences of Experiments 1 and 4 and the word strings of Experiment 3 involves both a content word between potentially related words and a higher rate of content word access in the word strings. Thus, if such a limit on priming existed, it could account for the pattern of results obtained in these experiments.

It should be noted that such an explanation is less likely to apply in these experiments than in Bodi's (1982), because in the present experiments the first two content words of the sentence remained adjacent to each other in the word strings, so that any priming they produced would only be prevented in the word strings by the existence of a single three-letter word prior to them, which intuitively seems unlikely to be able to produce such a drastic

reduction in priming. This intuition is confirmed by the fact that Schvaneveldt and Meyer (1973) found a large priming effect in strings of three words in which only the second two are related. Thus, any priming effect produced by the first two content words in the sentences from Experiment 1 should have survived in the word strings. For example, in the word string PEG PORTER CARRIED RIM LUGGAGE, which was constructed from the sentence THE PORTER CARRIED THE LUGGAGE, any priming produced by the words *porter carried* was unlikely to be destroyed by the occurrence of the word *peg* prior to them.

Priming in these first two content words is likely to be a major component of any priming effect that exists. The second and third content words (the verb and object of the sentence) could not contribute to the effect in these experiments because they are held constant for the plausible and implausible items. Thus, the only two possible sources of priming are the subject/verb pairs and the subject/object pairs. By Bodi's (1982) argument, the second of these effects might be destroyed by the unrelated content words due to the increased rate of content word access, but, as we have seen above, the first should survive the addition of the content words. Thus disruption of lexical priming by content words could account for a reduction of the effect in the word strings (perhaps by half), but not a complete elimination of it as was found. Unfortunately, however, because the full effect was barely significant, any reduction of the effect would be likely to result in a failure to reach significance. Thus, this argument rests on the assumption that the 9-msec nonsignificant trend was smaller than the nonsignificant effect, which would be expected if the effect were halved (i.e., approximately 24 msec). Clearly, this is not a strong basis from which to draw conclusions.

Thus, the results of the first four experiments can only be considered weak evidence against lexical priming as the cause of the plausibility effect. Therefore, in the next experiment, another attempt was made to eliminate lexical priming as the source of the plausibility effect, this time using a completely different approach.

EXPERIMENT 5

As mentioned earlier, Forster (1977) found a plausibility effect using the lexical decision task between sentences that contained the same content words, but in reverse order. For example, THE DOCTOR CURED THE PATIENT was responded to more quickly than THE PATIENT CURED THE DOCTOR. This finding eliminates the existence of related words alone as the source of the effect. However, it could be argued that there may still be a residual effect due to the different order in which the words occurred, because it is possible that relatedness effects are asymmetric and that plausible sentences tend to contain related words in the right order to produce priming.

In this experiment, using the sentence matching task as in Experiment 3 and 4, this possibility was examined by also testing the passive versions of the same sentences that, when their content words are in the same order as in the active sentences, have the opposite plausibility value. For example, THE PATIENT WAS CURED BY THE DOCTOR is a plausible sentence, with the same content words in the same order as in the implausible sentence above, and THE DOCTOR WAS CURED BY THE PATIENT is an implausible sentence with the same content words in the same order as in the plausible sentence above. Thus, if both the active and passive versions of the same sentences can produce a plausibility effect between reversed item pairs, the effect cannot be attributed to asymmetric priming, because it should be entirely dependent upon word order and therefore produce opposite effects of plausibility for the active and the passive sentences. Therefore, if a plausibility effect is found in this experiment for both the active and the passive versions of such sentences, the effect must be due to sentential processing, not lexical priming.

Method

Items. The experimental items were 60 quadruplets of sentences containing the same content words, but with word orders producing both plausible and implausible meanings in both active and passive voice, for example, THE LANDLORD EVICTED THE TENANT (active plausible), THE TENANT EVICTED THE LANDLORD (active implausible), THE TENANT WAS EVICTED BY THE LANDLORD (passive plausible), and THE LANDLORD WAS EVICTED BY THE TENANT (passive implausible). Sixty distractor items were also constructed, which were made up from sentence pairs modeled on the experimental items but were made different by changing several letters in the second sentence of each pair. The 15 practice items were a random assortment of all of the above types of items.

Ratings. The active versions of the items were rated for plausibility on a 7-point scale by 10 subjects to ensure the plausibility had been successfully manipulated. The plausible items had a mean rating of 6.79 and a range from 6.2 to 7.0. The implausible items had a mean rating of 2.99 and a range from 1.3 to 5.6.

Subjects. The subjects were 60 Monash University students who were paid for their participation.

Design. A counterbalanced design was used, like that of the previous experiments, but involving four groups of 15 subjects. Each group saw one-fourth of the items in the active plausible form, another one-fourth in passive plausible form, and so on, but each group saw each item in a different form from that seen by the other groups. For example, Group 1 saw THE LANDLORD EVICTED THE TENANT (active plausible), THE EMPLOYEE SACKED HIS BOSS (active implausible), THE MOUSE WAS CAUGHT BY THE CAT (passive plausible), and THE VET WAS EXAMINED BY THE DOG (passive implausible). Group 2 saw THE BOSS SACKED HIS EMPLOYEE (active plausible), THE MOUSE CAUGHT THE CAT (active implausible), THE DOG WAS EXAMINED BY THE VET (passive plausible), and THE LANDLORD WAS EVICTED BY THE TENANT (passive implausible). Group 3 saw THE CAT CAUGHT THE MOUSE (active plausible), THE DOG EXAMINED THE VET (active implausible), THE TENANT WAS EVICTED BY THE LANDLORD (passive plausible), and THE BOSS WAS SACKED BY HIS EMPLOYEE (passive implausible). Finally, group 4 saw THE VET EXAMINED

THE DOG (active plausible), THE TENANT EVICTED THE LANDLORD (active implausible), THE EMPLOYEE WAS SACKED BY HIS BOSS (passive plausible), and THE CAT WAS CAUGHT BY THE MOUSE (passive implausible). Thus, each group saw each item in only one version, but over all groups, each item was seen in each of its versions. The same distractors and practice items were seen by each group.

Procedure. The procedure was identical to that of Experiment 1.

Results

The mean RTs and errors for each condition are presented in Table 5. Five subjects were excluded from analysis for exceeding the accepted error criterion. A $4 \times 2 \times 2$ analysis of variance of RTs (with groups, voice, and plausibility as the three factors) showed a significant main effect of plausibility [$\min F'(1, 108) = 4.41, p < .05$] and no interaction between plausibility and voice [$F1(1, 56) = 1.12, p > .05$; $F2(1, 56) = 1.31, p > .05$]. A corresponding analysis of errors showed no main effect of plausibility [$F1(1, 56) = 2.17, p > .05$; $F2(1, 56) = 1.90, p > .05$] and no interaction ($F1 < 1, F2 < 1$).

Discussion

The results of this experiment show clearly that the plausibility effect is not dependent upon the relative order of the content words in the plausible and implausible sentences, as it would be if it were due to an asymmetrical lexical priming effect. There are positive, although weak, effects of plausibility for both active and passive sentences, and a significant main effect of plausibility in the overall analysis, with no interaction. In fact, the trend toward an interaction is in the direction opposite to that expected if asymmetrical associative priming were the source of the plausibility effect. These results conclusively eliminate priming as the source of the plausibility effect, because they show that the effect is found between sentences containing exactly the same related words, regardless of their order of occurrence.

GENERAL CONCLUSIONS

The experiments in this paper provide two separate lines of evidence against lexical priming as the source of the

plausibility effect. Experiments 3 and 4 provided weak evidence that the plausibility effect in Experiment 1 and the apparent priming effect in Experiment 2 were not attributable to lexical priming because no plausibility effect occurred when syntactic processing was disrupted in a task in which plausibility effects normally occur. Experiment 5 showed that even when lexical priming was eliminated, a plausibility effect still occurred. Therefore, the plausibility effect must involve sentential processing; it cannot be explained in terms of purely lexical effects such as the intralexical associative priming that Fodor (1983) proposed.

This is not to say that lexical priming does not ever contribute to the plausibility effect at all. The results of Experiment 4 showed that for the particular set of sentences used in Experiments 1-4, no significant effect of lexical priming occurred, but this does not eliminate the possibility that other sentences containing more highly related words may have their processing speeded by lexical priming. In fact, Bodi (1982) showed that lexical priming effects can occur when sentences contain related words adjacent to each other, or when the relatedness of the words is highly salient to the subjects reading the sentence. Furthermore, Foss (1982), who used a phoneme monitoring task, reported evidence of priming within sentences for words as far as 12 words apart, although Bodi (personal communication, November, 1985) pointed out that Foss's result is more likely to be due to the predictability of the sentences in which the related words occurred, rather than to the relatedness itself.

It does seem certain from these results that a major component of the plausibility effect is the result of the way in which word combinations contribute to sentence meanings and that this effect is dependent on the syntactic roles the words take in the sentence. This, then, is a sentential effect rather than a lexical effect. Thus, whereas purely intralexical processes appear to be adequate to explain the effects of sentence context on lexical access as well as the lexical priming effect, these processes cannot entirely account for the plausibility effect. It is not simply another example that apparent high-level processing is mimicked by an intralexical associative priming process, but it does

Table 5
Mean Reaction Times (RT in msec) and Percentage Errors (%E) in the Sentence Matching Task for the Plausible and Implausible Versions of the Active and Passive Items in Experiment 5, and the Plausibility Differences in Reaction Times and Errors

	Example of Items	RT	%E
Active			
Plausible	THE LANDLORD EVICTED THE TENANT	2317	3.1
Implausible	THE TENANT EVICTED THE LANDLORD	2355	2.8
Difference		38	0.3
Passive			
Plausible	THE TENANT WAS EVICTED BY THE LANDLORD	2713	2.9
Implausible	THE LANDLORD WAS EVICTED BY THE TENANT	2789	4.8
Difference		85	1.9

appear to implicate a genuine higher order sentential processing explanation as was originally suggested.

This raises the question of why sentential processing should be involved in such experiments. None of the tasks in which a plausibility effect was demonstrated would appear to require processing of sentential meaning at all. For example, the sentence matching task requires only a low-level physical matching process. The fact that sentential processing is involved, then, could be taken as evidence that there is feedback from higher level processing to the lower levels at which such task decisions are made. That is, the existence of a plausibility effect could be taken as evidence against the modularity hypothesis. If this were the case, however, then the top-down processing responsible for the plausibility effect should also be evident in experiments investigating the effects of sentence context on lexical access. But, as we have seen, because sentence-context effects do not appear to occur except as a result of purely lexical associations, this explanation does not seem likely.

Forster (1979) provided an alternative explanation for the influence of sentential processing on apparently low-level tasks in terms of a multilevel race model based on Chambers and Forster's (1975) explanation of word familiarity effects in a word matching task. He suggested that apparently unnecessary higher level processing will speed responses in such tasks as the sentence matching task because it produces a more structured representation. This makes the matching process faster, because once the input has been "chunked" into higher order units, there are fewer units to be compared. This has been used to explain an entire class of effects in which greater structuring of the input has been found to lead to faster matching. For example, it explains why words are matched more quickly than are random letter strings of the same length, because the lexical representations can be compared in one process, whereas letter strings involve a separate comparison process for each letter. Similarly, sentences can be matched more quickly than ungrammatical sequences of words because comparison is needed only for a small number of syntactic units instead of for each word separately. (For further development of this argument, see Freedman & Forster, 1985; Murray, 1982). It is also implied that the plausibility effect is another example of higher order structuring's leading to faster RTs and thus can be explained under the same model.

However, it is not clear that the plausible sentences are more structured at any level than are implausible sentences. Syntactic processing clearly assigns structure to sentences, but it seems unlikely that message-level processing does. Even if it did, it is unlikely that plausible sentences would have more structured representations than would implausible sentences. Thus, plausibility effects do not fit neatly into the multilevel race model.

This leaves sentential plausibility effects not satisfactorily accounted for. Explanation in terms of top-down processing is not consistent with the results of the

sentence-context experiments, whereas the multilevel race model, although it explains the existence of higher level processing in low-level matching tasks, does not readily explain plausibility effects. Thus, although the experiments in this paper have been successful in resolving the issue of the level of processing responsible for the plausibility effect, the fact that sentential processing is implicated opens up a whole new field of inquiry into the causes of this effect. An investigation of the processes underlying this sentential plausibility effect appears to be critical to our understanding of the relationship between the different levels of processing involved in the comprehension of sentences, and this promises to be an important new direction for further research.

REFERENCES

- BALOTA, D. A., & LORCH, R. F. (1986). Depth of automatic spreading activation: Mediated priming effects in pronunciation but not in lexical decision. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *12*, 336-345.
- BECKER, C. A., & KILLION, T. H. (1977). Interaction of visual and cognitive effects in word recognition. *Journal of Experimental Psychology: Human Perception & Performance*, *3*, 389-401.
- BODI, A. (1982). *Lexical priming in sentences*. Unpublished doctoral dissertation, Monash University, Clayton, Victoria, Australia.
- BRADLEY, D. C. (1978). *Computational distinctions of vocabulary type*. Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.
- CAIRNS, H. S., & KAMERMAN, J. (1980). Lexical information processing during sentence comprehension. *Journal of Verbal Learning & Verbal Behavior*, *14*, 170-179.
- CHAMBERS, S. M., & FORSTER, K. I. (1975). Evidence for lexical access in a simultaneous matching task. *Memory & Cognition*, *3*, 549-559.
- COLLINS, A. M., & LOFTUS, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*, 407-428.
- FISCHLER, I. (1977). Associative facilitation without expectancy in a lexical decision task. *Journal of Experimental Psychology: Human Perception & Performance*, *3*, 18-26.
- FISCHLER, I., & BLOOM, P. A. (1980). Rapid processing of the meaning of sentences. *Memory & Cognition*, *8*, 216-225.
- FODOR, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.
- FORSTER, K. I. (1974). The role of semantic hypothesis in sentence processing. In F. Bresson & J. Mehler (Eds.), *Current problems in psycholinguistics*. Paris: Editions du CNRS.
- FORSTER, K. I. (1979). Levels of processing and the structure of the language processor. In W. E. Cooper & E. Walker (Eds.), *Sentence processing: Psycholinguistics studies presented to Merrill Garret*. Hillsdale, NJ: Erlbaum.
- FORSTER, K. I. (1981). Priming and the effects of sentence and lexical contexts in naming time: Evidence for autonomous lexical processing. *Quarterly Journal of Psychology*, *33*, 465-495.
- FORSTER, K. I., & OLBREI, I. (1972). Semantic heuristics and syntactic analysis. *Cognition*, *2*, 319-347.
- FORSTER, K. I., & RYDER, L. A. (1971). Perceiving the structure and meaning of sentences. *Journal of Verbal Learning & Verbal Behavior*, *10*, 285-296.
- FOSS, D. J. (1982). A discourse on semantic priming. *Cognitive Psychology*, *14*, 590-607.
- FREEDMAN, S. A. (1982). *Behavioural reflexes of constraints on transformations*. Unpublished doctoral dissertation, Monash University, Clayton, Victoria, Australia.
- FREEDMAN, S. A., & FORSTER, K. I. (1985). The psychological status of overgenerated sentences. *Cognition*, *19*, 101-131.

- GERNSBACHER, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness and polysemy. *Journal of Experimental Psychology: General*, **113**, 256-281.
- GLUCKSBERG, S., KREUZ, R. J., & RHO, S. H. (1986). Context can constrain lexical access: Implications for models for language comprehension. *Journal of Experimental Psychology: Learning, Memory & Cognition*, **12**, 323-335.
- 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.
- KEPPEL, G. (1982). *Design and analysis: A researcher's handbook* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- KORIAT, A. (1981). Semantic facilitation in lexical decision as a function of prime-target association. *Memory & Cognition*, **9**, 587-598.
- KUČERA, H., & FRANCIS, W. (1967). *Computational analysis of present day American English*. Providence, RI: Brown University Press.
- LUPKER, S. J. (1984). Semantic priming without association: A second look. *Journal of Verbal Learning & Verbal Behavior*, **23**, 709-733.
- MARKS, L. E., & MILLER, G. A. (1964). The role of semantic and syntactic constraints in the memorization of English sentences. *Journal of Verbal Learning & Verbal Behavior*, **3**, 1-5.
- MARSLÉN-WILSON, W., & TYLER, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, **8**, 1-71.
- MCCLELLAND, J. L., & RUMELHART, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, **83**, 375-405.
- MEYER, D. E., & SCHVANEVELDT, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, **90**, 227-234.
- MEYER, D. E., SCHVANEVELDT, R. W., & RUDDY, M. G. (1975). Loci of contextual effects on word recognition. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance*. New York: Academic Press.
- MILLER, G. A., & ISARD, S. (1963). Some perceptual consequences of linguistic rules. *Journal of Verbal Learning & Verbal Behavior*, **2**, 217-228.
- MORTON, J. (1969). The interaction of information in word recognition. *Psychological Review*, **76**, 165-178.
- MURRAY, W. S. (1982). *Sentence matching: The influence of meaning and structure*. Unpublished doctoral dissertation, Monash University, Clayton, Victoria, Australia.
- NEELY, J. H. (1976). Semantic priming and retrieval from lexical memory: Evidence for facilitatory and inhibitory processes. *Memory & Cognition*, **4**, 648-654.
- NEELY, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, **106**, 226-254.
- ONIFER, W., & SWINNEY, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory & Cognition*, **9**, 225-236.
- RATCLIFF, J. E. (1983). *Inference processes in the early stages of sentence comprehension: A study of the plausibility effect*. Unpublished doctoral dissertation, Monash University, Clayton, Victoria, Australia.
- ROSENBERG, S. (1968). Associates and phrase structure in sentence recall. *Journal of Verbal Learning & Verbal Behavior*, **7**, 1077-1081.
- ROSENBERG, S. (1969). The recall of verbal material accompanying semantically well-integrated and semantically poorly-integrated sentences. *Journal of Verbal Learning & Verbal Behavior*, **8**, 732-736.
- ROSENBERG, S., & JARVELLA, R. J. (1970). Semantic integration and sentence perception. *Journal of Verbal Learning & Verbal Behavior*, **9**, 548-553.
- SANTA, J. L., MILLER, J. J., & SHAW, M. L. (1979). Using Quasi F to prevent alpha inflation due to stimulus variation. *Psychological Bulletin*, **86**, 37-46.
- SCHUBERTH, R. E., & EIMAS, P. D. (1977). Effects of context on the classification of words and non-words. *Journal of Experimental Psychology: Human Perception & Performance*, **3**, 27-36.
- SCHVANEVELDT, R. W., & MEYER, D. E. (1973). Retrieval and comparison processes in semantic memory. In S. Kornblum (Ed.), *Attention and Performance IV*. New York: Academic Press.
- SEIDENBERG, M. S. (1984). Constraining models of word recognition. *Cognition*, **14**, 169-190.
- SEIDENBERG, M. S., TANENHAUS, M. F., LEIMAN, J. M., & BIENKOWSKI, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge toward processing. *Cognitive Psychology*, **14**, 489-537.
- SEIDENBERG, M. S., WATERS, G. S., SANDERS, M., & LANGER, P. (1984). Pre- and postlexical loci of contextual effects on word recognition. *Memory & Cognition*, **12**, 315-328.
- SHAUGHNESSY, J. J., & ZECHMEISTER, E. B. (1985). *Research methods in psychology*. New York: Knopf.
- STANOVICH, K. E., & WEST, R. F. (1983a). The generalizability of context effects on word recognition: A reconsideration of the roles of parafoveal priming and sentence context. *Memory & Cognition*, **11**, 49-58.
- STANOVICH, K. E., & WEST, R. F. (1983b). On priming by a sentence context. *Journal of Experimental Psychology: General*, **112**, 1-36.
- SWINNEY, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning & Verbal Behavior*, **18**, 645-659.
- TANENHAUS, M. F., & LUCAS, M. M. (1987). Context effects in lexical processing. *Cognition*, **25**, 198-219.
- WATSON, I. J. (1976). *The processing of implausible sentences*. Unpublished doctoral dissertation, Monash University, Clayton, Victoria, Australia.
- WEST, R. F., & STANOVICH, K. E. (1978). Automatic contextual facilitation in readers of three ages. *Child Development*, **49**, 717-727.

(Manuscript received April 4, 1986;
revision accepted for publication January 8, 1987.)