Regeneration in the Short-Term Recall of Sentences

MARY C. POTTER AND LINDA LOMBARDI
Massachusetts Institute of Technology

Verbatim short-term memory for a sentence has been taken as evidence for a surface representation different from the conceptual representation characteristic of longer-term memory. In seven experiments we investigated an alternative hypothesis: that immediate recall involves regeneration of the sentence from a conceptual representation, using words that have been recently activated. A key claim is that the activated lexical items are unordered. To test this hypothesis, a synonym of a word in the sentence was presented in a secondary task before or after the sentence, prior to recall. As predicted, these lure words were intruded frequently (Experiments 1 and 2), but only when supported by the meaning of the whole sentence (Experiments 3 and 4). In Experiment 5 as high an intrusion rate was obtained for sentences read in RSVP at 100 ms per word as at the 200 ms rate of the other experiments. Experiment 6 showed that listeners make even more lure-based intrusions than readers. In Experiment 7 4-year-old children made intrusions similar to those of adults. The results support the hypothesis that a sentence is regenerated in immediate recall from a representation of its meaning, using recently activated words. Only when the right set of words is active will recall be “verbatim.” © 1990 Academic Press, Inc.

Why is immediate recall of a sentence so accurate? The usual answer (e.g., Glanzer, Fischer, & Dorfman, 1984; see von Eckardt & Potter, 1985, for a review) is that short-term memory provides a special verbatim representation of the sentence, whereas in long-term memory only its meaning is retained. We first review briefly the basis for this widely accepted view and then we consider an alternative possibility, which is tested in the experiments that follow.

Short-term memory for unrelated lists of words is verbatim for up to six or seven words, but as the to-be-remembered list increases in length the accuracy of recall declines markedly. Unlike longer-term memory for lists, the verbatim memory span is little influenced by the presence or absence of semantic relations among the words (e.g., Baddeley, 1966, but see Huttenlocher & Newcombe, 1976), suggesting that the short-term representation is different from the meaning-based long-term representation.

In short-term memory for a sentence many more words—12 or more—can be recalled verbatim, but as in the case of lists there is a sharp drop in accuracy as the sentence memory span is exceeded, that is, when more than the most recent sentence is to be recalled. In that case the gist of earlier sentences may be recalled, but not the exact wording (e.g., Jarvella, 1971; Glanzer, Dorfman, & Kaplan, 1981; Marslen-Wilson & Tyler, 1976). Tests of recognition memory for sentences also have shown that memory is essentially verbatim immediately after a sentence has been heard or read, but details of the syntax and vocabulary (although not the meaning) are lost if even one sentence intervenes between pre-
sentation and test (Sachs, 1967, 1974; see also Anderson & Paulson, 1977).

Thus there is evidence from both list and sentence experiments for two stages in memory, one short-term and verbatim and the other longer term and meaning-based. This evidence, in turn, has been taken to suggest that short-term memory includes a short-lasting representation of the surface form of information, such as an acoustic, phonological, orthographic, or articulatory representation. Evidence that including phonologically confusable words in the material to be recalled or as distractors makes comprehension more difficult and immediate recall less accurate is consistent with this view (see Baddeley, 1986, for a review). Although the meaning of words and sentences is also available in short-term memory, meaning does not correlate perfectly with surface form because a given meaning can be expressed in many different ways. Therefore, a surface representation (not simply the meaning) is believed necessary for verbatim recall.

A problem with the standard view just described is that it accounts for the sentence (verbatim) span of up to about 20 words with the same mechanism as the word span of six words. One common way of explaining this disparity in capacity is to invoke some version of chunking (Miller, 1956): just as words are learned chunks of several letters or phonemes, so sentences are chunks of familiar phrases (Wickelgren, 1977). Since short-term memory capacity is counted in chunks rather than letters or phonemes, that might explain the greater capacity for words in sentences. The observation that sentences expressing sensible or familiar ideas in everyday language are easier to repeat than more complex or unusual sentences is consistent with the chunking account, although the details of such an account have not been worked out.1

What is less clear is how the chunking account fits with the standard view that a surface representation—usually understood to mean a perceptual or perhaps motoric representation—is responsible for verbatim short-term memory. If the surface representation is, for example, an ordered string of word-sounds, it has a capacity of only about three to six words, judging by memory-span studies. Even if the surface representation is supplemented by a central executive system with some memory capacity (as in Baddeley’s multiple-components model of working memory, reviewed in Baddeley, 1986), there is no account of how the representation of the surface string is combined with word or phrasal chunks that are represented in a lexical, syntactic, or conceptual system or how these representations are mapped smoothly onto a single output during recall.

A different possibility, the one we propose here, is that immediate recall of a sentence is not based on a surface representation in the usual sense, but—like longer-term recall—is based on a representation of the meaning of the sentence (e.g., a propositional representation). In long-term recall remembered concepts or propositions of a sentence or text are expressed using normal

1 A chunk in Miller’s original formulation was a previously-learned unit, and short-term memory had slots for 7 ± 2 such units. How well-learned an item had to be, to become a single unit, was not specified; nor was the model extended to more loosely interconnected items such as the words in a sentence, among which there might be associations but no identifiable separate chunks. There are two chief ways to model short-term capacity limits for material connected in this way. One is to model the probability that a given grouping of words will be represented as a single chunk versus as separate words (e.g., “sat down” or “the last chapter” or “put on his coat”), with the probability increasing as a given sequence of words becomes more familiar or frequent. The other is to treat short-term memory capacity as a continuous quantity that can augment the strength of existing links so that they reach threshold for immediate recall; when the material to be recalled is already partially interconnected, the added strength provided by short-term memory can be distributed over more links, yielding a larger number of words recalled in a sentence than in a random list. These two types of models might be difficult to distinguish empirically.
speech-production mechanisms. A set of lexical elements is selected to express the conceptual structure and a syntactic frame is constructed accordingly (e.g., Bock, 1982, 1987; Garrett, 1982; and Levelt, 1989). Even when specific sentences have been memorized and verbatim recall is later requested, this proposition-based route seems to be used (Flores d'Arcais, 1974; James, Thompson, & Baldwin, 1973; Kempen, 1977; Levelt & Kempen, 1975).

In longer-term recall paraphrases are common, as this model would predict (Anderson, 1974; Graesser, 1978; but see Peterson & McIntyre, 1974). What, then, accounts for the verbatim nature of immediate memory recall? Our hypothesis is that recently activated lexical entries ("logogens," in Morton's 1970 terminology) are the ones most likely to be selected for expressing the conceptual content of the sentence, so that the very words just used in the sentence have a high probability of being selected. Instead of using a "surface" representation such as a phonemic string or any ordered perceptual representation, recall relies on a conceptual representation of the sentence and on active but unordered lexical entries.

Thus, according to the present view, the only qualitative difference between short-term and longer-term recall of sentences is in the availability of a set of activated lexical entries in the former case. This regeneration hypothesis about short-term sentence recall was tested in the present study by using a new paradigm in which a synonym of a word in the to-be-remembered sentence was presented in an incidental task (either before or after the sentence) on the same trial. For example, the following sentence was presented:

(1) The knight rode around the palace searching for a place to enter.

Immediately after reading the sentence, but before recalling it, the subject read a short list of nouns in the incidental task; in this example, for half the subjects the list included the lure word "castle" (roughly synonymous with "palace"). "Castle" could replace "palace" without altering the meaning of the sentence significantly. We predicted that this incidental presentation would be sufficient to activate the lure word's lexical entry so that at the time of recall it would provide an alternative to the original word and would sometimes intrude in recall. If, contrary to the regeneration hypothesis, the verbatim character of immediate memory depends on an ordered surface representation of the string of words there would be no reason to expect that the synonym lure would intrude.

**Experiment 1**

In Experiment 1 the paradigm just described was used to test the hypothesis that "verbatim" recall of a sentence is based in part on a conceptual representation of the sentence message and in part on (unordered) recent activation in the lexicon. The to-be-recalled sentence was presented and immediately followed by a list of five unrelated nouns and then a capitalized probe word. The subject was instructed to decide whether or not the probe word had appeared on the list (by pressing one key for Yes and another for No), and then to recall the sentence aloud. The list-probe task was described as a distracting task intended to increase the difficulty of recall. On half the trials the (quasi-)synonymous lure word (hereafter, the synonym lure) was included on the list, but was never the word probed.
On the other half of the trials an unrelated control word replaced the lure on the list. The dual task was subjectively fairly difficult but, as will be seen, recall of the sentence was generally very good.

**Method**

**Subjects.** The 12 subjects were college students who had volunteered to participate and were paid. All were native English speakers.

**Materials.** Twenty sentences were written, each including a noun (the "target") which had a synonym or near-synonym (the "lure") that could replace the target without significantly changing the overall meaning of the sentence. Pilot experiments were carried out to develop the materials; the final 20 sentences and lures were selected from a larger set. The sentences were 11 to 15 words in length (mean, 12.5 words). For each sentence a list of five nouns was selected, with the constraints that the words should be similar in length to the target and lure for that sentence, but should be unrelated to the meanings of words in the sentence.

The sentences were written so that, in the experimenters' judgment, the synonym lure was at least as appropriate as the target noun, in the context of the sentence. That is, when one of the two words was judged more appropriate it was designated as the lure. For example, in (1) the lure word "castle" fits the scenario of the sentence somewhat better than the target word "palace." Brewer (1975), who studied synonym substitutions in longer-term cued recall of sentences, found that the extent of spontaneous synonym substitutions in long-term memory was determined by the relative naturalness of the substitute word. By biasing the sentence to favor the lure word as a replacement for the target word, we hoped to obtain a measurable number of spontaneous intrusions even when the lure word was not on the distractor list. The additional effect of having just seen the lure word could than be compared with the spontaneous rate of intrusions. The materials are shown in Appendix A.

**Apparatus.** The experiment was run on a TEK microcomputer; the CRT had a fast-fade phosphor.

**Design and procedure.** Two versions of the materials were used, counterbalancing for whether or not the synonym lure was on the list following a given sentence. Half the sentences in each version were followed by the lure. When the lure was included it replaced one of the five words; it was never the last word on the list and it was never the probe word that followed the list. For half the lists the probe word was positive and for half, negative. The order of the sentences was randomized and that order was used in both versions. There were six practice trials, without lures on the list.

The sentence and list were presented using RSVP (rapid serial visual presentation), with each word centered on the same point on the screen. The rate used for the sentence, five words per second, is close to the normal reading rate for college students. Each trial began when the subject pressed the space bar on the computer keyboard. Three asterisks appeared for 300 ms at the center of the screen where the subsequent words would appear, followed by a blank screen for 350 ms and then the words of the sentence for 200 ms per word. The first word was capitalized, but no punctuation was used. Immediately after the last word of the sentence a row of percentage signs, serving as a visual mask, appeared for 517 ms, followed by the five lure words at 250 ms per word and then a 250 ms row of percentage signs. Finally, the capitalized probe word appeared for 500 ms.

The subjects were told that the study concerned reading and memory for sentences and that the list-probe task was intended to increase the difficulty of the sentence task. They were instructed to respond to the probe word by pressing one of two response keys and then to recall the
sentence aloud as accurately as possible. The experimenter recorded the recall response.

Results and Discussion

Overall recall of the sentences was good: 89% of the words (apart from the target word) were reported correctly. Similarly, the target word was omitted on only 5% of the trials, apart from intrusions of the lure word. Thus, despite the 3 s or so required to present and perform the list-probe task, recall of the sentence was nearly verbatim.

The main question was whether the lure word would intrude at all and—if so—whether it would intrude more frequently when it had been on the word list following the sentence. The lure intruded spontaneously on 9% of trials when it was not on the list and intruded on 27% of the trials when it was on the list, by subjects $t(11) = 3.17$, $p < .01$, by items, $t(19) = 4.10$, $p < .01$, $F'_{min}(1,23) = 6.28$, $p < .05$.

The results strongly support two claims of the conceptual regeneration hypothesis. First, recent lexical activation makes a word preferentially available for use in immediate recall, as shown by the marked increase in synonym intrusions when the synonym was present in the distractor list. Second, the accuracy ordinarily observed in immediate recall of a sentence is, like long-term recall, based in large part on a representation of sentence meaning, not simply on a surface representation of the words. This claim is supported by the substantial tendency for a suitable synonym to intrude even when it was not on the distractor list (9%)—although as just noted, the intrusion rate was much higher when the lure had been presented on the list (27%). In contrast, there were virtually no intrusions of the other words on the list, and overall accuracy in recall of the sentences was high.

Experiment 2

In Experiment 1 recall of the sentence was very good, apart from intrusions of the critical word, showing that the experiment did tap so-called verbatim immediate memory. Nonetheless, the delay imposed by presenting the list after the sentence may mean that recall was not from short-term memory. To investigate that possibility, in Experiment 2 the list was presented before the sentence; only the probe and probe response intervened between the end of the sentence and recall.

Method

Except as noted, the method was identical to that of Experiment 1.

Subjects. The subjects were 16 students from the same pool as those in Experiment 1; none had been in Experiment 1.

Procedure. On each trial the five words of the list immediately followed the fixation asterisks and blank, at 250 ms/word. The list was followed by a 250 ms mask and then the words of the sentence at 200 ms/word. A mask of 517 ms followed the sentence and then the probe word appeared in capital letters for 500 ms. As in Experiment 1, the subject first responded to the probe by pressing one of two keys to indicate whether or not the probe had been on the list and then recalled the sentence aloud.

Results and Discussion

As in Experiment 1 recall of the sentences was generally quite accurate: apart from the target word, 86% of the words were recalled correctly. Aside from lure intrusions there were only 8% errors in the recall of the critical word. When no lure word was present on the list there were 9% spontaneous intrusions of the lure, and when the lure word was present there were 22% intrusions: by subject, $t(15) = 3.32$, $p < .01$, by items, $t(19) = 2.67$, $p < .02$, $F'_{min}(1,34) = 4.33$, $p < .05$.

Although there were somewhat fewer lure intrusions in Experiment 2 than in Experiment 1, a comparison of the two experiments showed a main effect of the lure, $F'_{min}(1,40) = 8.45$, $p < .01$, but no signifi-
cant difference between the experiments and no interaction, all Fs less than 1.54. The elapsed time between the last word of the sentence and the beginning of recall was reduced from about 3 s in Experiment 1 to about 1.5 s in Experiment 2. That put the recall task in Experiment 2 more clearly within the scope of short-term memory, and the similarity of the results in the two experiments suggests that Experiment 1 also involved short-term memory.

The similarity of the results in the list-before and list-after conditions rules out explanations of lure intrusions based solely on their recency relative to the target word in the sentence. Evidently presenting the lure word before the sentence and requiring the subject to remember it well enough to respond appropriately to the probe was sufficient to keep the word active and available as an intrusion.

**Experiment 3**

A possible alternative interpretation of the high rate of intrusions is that the lure and the target word were associated preexperimentally, so that when one was presented it tended to activate the other. During recall the associated word, primed by its presentation of the distractor list, might sometimes displace the target word. This hypothesis would also explain the occurrence of spontaneous intrusions even when the lure was not presented. Brewer (1975) considered this hypothesis in his study of spontaneous synonym substitutions in longer-term cued recall, but rejected it on the grounds that high associates only intrude when they are synonyms—e.g., big for large—and not when they are antonyms (e.g., small) or coordinates. However, a more subtle version of the association hypothesis should be considered: perhaps when processing sentences people set themselves to make only synonym associations (word associations are readily directed by such sets).

To distinguish the word association hypothesis from the present hypothesis, target–lure pairs were presented with two types of sentence. One type of sentence was similar to those used in Experiments 1 and 2 and the other was designed to elicit a conceptual representation of the target word that would be distinct from that of the lure word (cf. Barclay, Bransford, Franks, McCarrell, & Nitsch, 1974). Compare, for example, sentence (2) with sentence (1):

(2) The royal guests danced in the palace to the music of an orchestra.

While both sentences include the word “palace,” and while the lure word “castle” is acceptable in either sentence as a replacement for “palace,” the overall meaning of the sentence makes “castle” a more appropriate intrusion in (1) than in (2). If the word association hypothesis is correct, then intrusions should be as likely in (2) as in (1). The present hypothesis, in contrast, asserts that the overall meaning of the sentence governs the choice of lexical items in recall (with a propensity to use only recently activated items), and so “castle” would be expected to intrude more frequently in (1).

**Method**

Except as noted, the method was like that of Experiment 1: the list and the probe followed the sentence.

**Subjects.** Twenty-four subjects from the same pool as in the previous experiments participated; none had been in the previous experiments.

**Materials and design.** Two sentences were written for each of 20 target–lure word pairs; 18 of these pairs were ones used in the previous experiments. One version (the Match version) was identical to or similar to the sentence used in the earlier experiments: it attempted to make the lure word at least as appropriate as the target word in expressing the overall message of the sentence. The second version (the Mismatch
version) attempted to reduce the appropriate-ness of the lure word by bringing out aspects of the target word’s meaning or connotation that were not shared with the lure. The core meaning of the target word was not changed, however—for example, the chess meaning of “castle” was not used—and the lure remained acceptable in the sentence. The materials are shown in Appendix B.

There were four versions of the experiment, counterbalancing Match and Mismatch sentence versions and the presence or absence of the lure on the list. Thus, each subject saw ten Match and ten Mismatch sentences, half with the lure and half without. The order of the sentences was randomized.

**Results and Discussion**

Overall accuracy of sentence recall was again high; 87.5% of the nontarget words were recalled correctly. Apart from lure intrusions, the target word was recalled incorrectly on 97% of the trials. The percentage of lure intrusions in each condition is shown in Table 1, together with the results of Experiments 1 and 2. In recall of the Match sentences the lure intruded on 38% of the trials (and intruded spontaneously on 13% of the control sentences with no lure). In recall of the Mismatch sentences, the corresponding percentages were 13% and 3%. Analyses of variance were carried out on the number of intrusions in each condition. The main effects of sentence type (Match versus Mismatch) and Lure (present versus absent) were both significant, $F'_{\text{min}}(1,29) = 10.82, p < .01$, and $F'_{\text{min}}(1,41) = 22.33, p < .01$, respectively. The interaction between these variables was significant in the subject analysis, $F_1(1,23) = 7.87, p < .01$, and the item analysis, $F_2(1,19) = 4.70, p < .05$, although $F'_{\text{min}}$ was not significant. Separate tests showed that Match sentences were more likely to produce intrusions than Mismatch sentences both when the lure was present (by subjects, $t(23) = 5.93, p < .01$, by items, $t(19) = 3.38, p < .01$, $F'_{\text{min}}(1,31) = 8.82, p < .01$), and when the lure was not (by subjects, $t(23) = 2.88, p < .01$, by items, $t(19) = 2.46, p < .05$, although $F'_{\text{min}}(1,40) = 3.50, p < .10$).

The results for the Match sentences replicate those of Experiments 1 and 2. The

### Table 1

**Percentage of Errors in Recall of Target Words and Percentage of Nontarget Words Recalled in Experiments 1, 2, 3, 5, and 6**

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Condition</th>
<th>Lure intrusions</th>
<th>Spontaneous intrusions</th>
<th>Other target errors</th>
<th>Nontarget words recalled</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Standard$^a$</td>
<td>27</td>
<td>9</td>
<td>5</td>
<td>89</td>
</tr>
<tr>
<td>2</td>
<td>Lure before</td>
<td>22</td>
<td>9</td>
<td>8</td>
<td>86</td>
</tr>
<tr>
<td>3</td>
<td>Match</td>
<td>38</td>
<td>13</td>
<td>8</td>
<td>87$^b$</td>
</tr>
<tr>
<td>5</td>
<td>Mismatch</td>
<td>13</td>
<td>3</td>
<td>10</td>
<td>86</td>
</tr>
<tr>
<td>6</td>
<td>100 ms/word</td>
<td>31</td>
<td>12</td>
<td>5</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>Visual</td>
<td>33</td>
<td>6</td>
<td>9</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>Auditory</td>
<td>53</td>
<td>13</td>
<td>6</td>
<td>91</td>
</tr>
</tbody>
</table>

*Note. “Lure intrusions” were those occurring when the critical synonym was on the distractor list; “spontaneous intrusions” were intrusions of the lure word on control trials when it was not on the distractor list; “other target errors” were other omissions or substitutions for the target word in experimental and control trials combined.*

$a$ Except as specified, in all conditions including the standard condition rapid serial visual presentation (RSVP) was used, words appeared for 200 ms each, the sentence was followed by the distractor list with the lure or a control word, and the lure matched the sentence (i.e., it was an appropriate paraphrase of the target word).

$b$ For Match and Mismatch sentences combined.
apparent increase in the percentage of lure intrusions in Experiment 3 may reflect either the changes made in many of the sentences, or the relatively low proportion of Match-with-lure sentences—25% versus 50% in the earlier experiments. Having fewer such trials may have reduced the likelihood that subjects would notice the lure–target relationship and be put on their guard against lure intrusions.

As predicted by the conceptual regeneration hypothesis, the Mismatch sentences produced dramatically fewer lure intrusions than did the Match sentences. This result is strong evidence against the word association hypothesis, because the same pairs of words were used as targets and lures in both the Match and Mismatch sentences; what varied was the overall meaning of the sentence in which the target word was embedded. A straightforward associative substitution, such as “castle” for “palace”, should have been equally likely in both sentences. The regeneration hypothesis proposes that recall is driven by a conceptual (for example, propositional) representation of the meaning conveyed by the sentence as a whole, so that only words that express this meaning will be used in recall. The lure word was less expressive of sentence meaning in the Mismatch condition than in the Match condition and so was less likely to be intruded.

There is a possible way to save the word association hypothesis, which is to argue that the Match sentences were more likely to include other words in addition to the target word that are associated with the lure. For example, “knight” might be associated more strongly with the lure “castle” than with the target “palace” (helping to produce intrusions), whereas “royal” or “danced” might be more strongly associated with the target “palace” than with the lure “castle,” thereby decreasing intrusions. It seems implausible, however, that such relatively remote word associations could produce the very large observed difference in intrusion rates between the two types of sentences, given that the word most strongly associated with the lure—the target itself—was present in both versions of the sentence. Furthermore, it was almost invariably the target word that was displaced when the lure intruded, not one of the other “associates” (for example, castle did not displace knight). To make the multiple word association hypothesis work, one must add constraints on which associates actually intrude and where they intrude—and these constraints begin to make the modified association hypothesis equivalent to the regeneration hypothesis.

Experiment 4

In Experiment 4 a different method for assessing the short-term representation of sentences was adopted: the task was to recognize whether or not a given probe word had been in the sentence. False recognition of synonyms has been demonstrated in numerous studies (e.g., Mandler & Worden, 1973; see a review by Kennedy, 1975). In a study in which the meaning of a noun in a sentence was compared with the meaning of the whole phrase in which the noun appeared, Green (1975) showed that subjects had greater difficulty rejecting a probe word that paraphrased the phrase than rejecting a probe that paraphrased just the head noun (but only when the main task was to continue the sentence; the reverse was found when the task was to memorize the sentence). In Green’s experiment the sentence was held constant and the probe changed. In Experiment 4 the target word and probe were held constant and the sentence was changed to make the conceptual representation match or mismatch the lure probe, as in Experiment 3. The question was whether subjects would take longer to reject the identical lure word (or accept it falsely more often) after Match than after Mismatch sentences.

If the immediate representation of a sentence is conceptual, then subjects should have more difficulty rejecting the lure with Match sentences. Compared to unrelated
control words, however, the lure word was globally related to both Match and Mismatch sentences, so the lure word was expected to be substantially harder to reject in both cases than unrelated filler words. Potter, Valian, and Faulconer (1977) showed that subjects are readily and rapidly able to assess the general relevance of a word to the scenario or theme of an immediately previous sentence, even when the word is not a possible substitute for any word or idea in the sentence. Therefore, the prediction was that, while the synonym lure would be substantially more difficult to reject than a wholly unrelated word, Match sentences would produce even more difficulty than Mismatch sentences.

Method

Subjects. The 29 subjects (including one who was replaced) were from the pool previously described; none had participated in any of the earlier experiments.

Materials and design. The 20 sentences of Experiment 3, in their Match and Mismatch versions, were used in Experiment 4, together with 20 filler sentences and six practice sentences. A probe word followed each sentence; no lists were presented. There were four versions of the materials, corresponding to the four versions in Experiment 3: each experimental sentence was either the Match or the Mismatch version, and the sentence was followed by the target word itself (the positive probe) or by the lure word (the related negative probe). The 20 filler sentences consisted of 10 with positive probes of adjectives and verbs (all the experimental sentences were probed with nouns) and 10 with negative probes (four with words related to adjectives or verbs in the sentence and six with unrelated nouns or verbs). The filler sentences were the same in all versions of the experiment. The same random order of the experimental sentences was used as in Experiment 3, and the filler sentences were inserted in a pseudorandom order.

Procedure. The subjects were instructed to read each sentence and then decide as rapidly as possible whether the probe word that followed had appeared in the sentence. Each trial began with the same row of asterisks and blank as in the earlier experiments, and each word of the sentence was presented for 200 ms, followed by a 250 ms mask of percentage signs and then the probe word, capitalized, for 500 ms. The yes or no response was made by pressing one of two keyboard keys; response time was measured from the onset of the probe word.

Results and Discussion

One subject whose mean RT was more than 2 standard deviations longer than that of the group was replaced. RTs longer than 2 s (less than 1%) were regarded as errors. For each subject, separately for yes and no responses to the experimental sentences, RTs more extreme than two standard deviations plus or minus the mean were truncated to that score. The mean correct RTs and error rates for the experimental sentences are shown in Table 2. The overall error rate was 7.3%: the variation was too small to justify statistical analysis.

The main question concerned the relative time to reject lure probes of Match and Mismatch sentences. In the subject analysis the difference of 49 ms was significant and it was nearly so in the item analysis, $F(1,27) = 6.94, p < .02$ and $F(1,19) = 4.15, p < .06$, respectively. Thus matching lures were, as predicted, harder to reject, even though the target–lure pairs were identical in the Match and Mismatch conditions; only the sentence changed. The four filler

<table>
<thead>
<tr>
<th>Condition</th>
<th>Lure (No)</th>
<th>Target (Yes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Match</td>
<td>RT</td>
<td>Errors</td>
</tr>
<tr>
<td></td>
<td>896</td>
<td>.08</td>
</tr>
<tr>
<td>Mismatch</td>
<td>847</td>
<td>.09</td>
</tr>
<tr>
<td>Difference</td>
<td>49</td>
<td>-.01</td>
</tr>
</tbody>
</table>
sentences with noun probes that were entirely unrelated to the sentence meaning took a mean of 753 ms to reject, which was (as expected) significantly faster than rejections of the lure words after both Match (892 ms) and Mismatch (847 ms) sentences, t(27) > 7.0, p < .001, in each case. The error rate was also low (1.8%) for these fillers.

Given the relatively small (although significant) difference between Match and Mismatch lures in relatedness to the sentence (as assessed by the speed of rejecting the lure) it is perhaps surprising that there was such a large difference in intrusion rate in recall, in Experiment 3. While the word-probe method of Experiment 4 appears to be sensitive chiefly to the global relevance of the test word to the sentence meaning, the intrusion paradigm of Experiments 1–3 is more sensitive to the exact fit of the lure to a particular slot in the sentence.

Experiment 5

Experiments 1–4 support the hypothesis that immediate verbatim memory for a sentence, like longer term memory, relies in large part on a conceptual representation of the sentence. Just how rapidly is the conceptual representation established? In previous work using more rapid rates of RSVP than the essentially normal reading rate of 200 ms per word used in the present experiments, there has been evidence that sentences can be understood when read at rates as high as 83 ms per word (e.g., Forster, 1970; Potter, Kroll, & Harris, 1980; Potter, Kroll, Carpenter, Yachzel, & Sherman, 1986). However, it has not been clear how completely readers are able to process sentences at those rates; for one thing, memory for a continuous RSVP paragraph was much poorer at 10 words per second (wps) than at 6.7 wps, which in turn was poorer than at 3.3 wps (Potter et al., 1980).

Experiment 5 provided a test of the rapidity with which a conceptual representation is established in reading a sentence, by presenting the sentences at 100 ms per word. If the conceptual representation of a sentence is less complete immediately after presentation when it has been read at 100 ms rather than at 200 ms per word, then recall in the former case might rely on a surface representation and synonym intrusions would be expected to be less frequent than at the slower rate of presentation.

Method

Subjects. Twelve subjects from the same pool used previously participated in this experiment. The experiment was run shortly after Experiment 1.

Design and procedure. The method of Experiment 5 was identical to that of Experiment 1 except that the sentences were presented at the rate of 100 ms per word rather than 200 ms per word. The list that followed was presented at 250 ms per word, as in Experiment 1. Subjects were told that the sentences would be presented very fast.

Results and Discussion

The main results are shown in Table 1. Overall recall of the sentences was very good: 86% of the nontarget words were recalled correctly. Apart from lure intrusions, only 5% of the trials had errors in recall of the target word. The lure intruded spontaneously on 12% of the trials without the lure in the list and 31% of the trials with a lure, by subjects, t(11) = 4.41, p < .01, and by items, t(19) = 3.52, p < .01, F_min(1,30) = 7.57, p < .01. The comparable percentages in Experiment 1, where the sentences were presented at 200 ms per word, were 9% and 27%. An analysis combining Experiments 1 and 5 showed a significant main effect on lure intrusions of the presence of the lure on the list, F_min(1,39) = 11.89, p < .01, but no effect of presentation rate (between the two experiments) and no interaction between the presence of the lure and rate, all Fs < 1.66.

The fact that synonym substitutions were just as frequent when subjects read at a high rate is consistent with the conclusion that
sentences are immediately and automatically processed to a conceptual level (even when presented very rapidly) and that the conceptual representation plays a role in recall. But is it plausible that sentences presented serially at 100 ms a word could be understood as they are read? Potter et al. (1980) showed that readers could understand RSVP paragraphs read at that rate, even though (as noted earlier) the readers forget much of the detail (see also Masson, 1986). Potter et al. (1986) also provided evidence that RSVP sentences presented at 83 or 100 ms per word are understood as they are read. What the present results show is that understanding of such sentences is not only sufficient for accurate recall, but also for susceptibility to the systematic errors in recall that are predicted by the regeneration model.

**Experiment 6**

In Experiments 1–5 subjects read the sentences serially, which had the important advantage of controlling the amount of processing time given the sentence. Reading serially is like listening, in that the perceiver receives the words one at a time, has no control over the rate, and cannot reread the material. But, unlike listening, serial reading is unfamiliar; further, like ordinary reading, it lacks some important features of spoken speech, such as stress and intonation patterns. Comparisons of reading in which the words are displayed serially with reading of simultaneously displayed text and with listening have not shown major differences in comprehension or memory when overall rate is matched (see Potter, 1984, for a review). Still, it would be important to know whether or not the synonym intrusion phenomenon occurs in recall of a spoken sentence, given that speech is for most people the first and most natural form of language. If the regeneration view of short-term sentence recall is correct, then synonym intrusions should occur regardless of the modality of presentation of the sentence and lure list.

There is another important respect in which RSVP reading is unlike listening: it does not provide a direct auditory trace. Although readers do normally retrieve phonological or sound-based representations of words, this internally generated representation does not seem to be as persistent as an auditory trace (and visual persistence is unlikely to play any significant role when each word appears in the same spot as the previous one). One role that has been proposed for auditory traces is to assist the listener in holding on to the words of a sentence: if so, one might expect a smaller number of synonym substitutions with spoken than with written sentences.

In Experiment 6 auditory and visual sentences, matched for rate, were presented to two groups of subjects.

**Method**

In most respects the method was like that of Experiment 1.

**Subjects.** Subjects were from the same pool described previously; none had been in the earlier experiments. There were 28 subjects, 14 in the auditory group and 14 in the visual (RSVP) group.

**Materials and design.** The sentences were the same as the Match sentences in Experiment 3 except for changes in four sentences, as shown in Appendix B. Some unrelated words in the lists were changed because they sounded like some word in the sentence or were difficult to understand when spoken. A new random order of the sentences was used. The new materials were used in both the auditory and visual conditions. The auditory sentences were recorded by a practiced speaker at a rate that resulted in approximately the same overall time per sentence as the RSVP sentences, which were presented for 200 ms per word. Two copies of the sentence recordings were made and spoken lists were added, with the presence or absence of the lure counterbalanced over sentences between versions. The list began about 500 ms after the last word to the sentence and
was spoken at the rate of 250 ms per word; the probe word followed after a 250 ms pause.

Procedure. For the visual group the procedure was like that of Experiment 1. For the auditory group, the instructions were modified appropriately. In the auditory group subjects listened through headphones. Each trial began when the subject pressed a foot pedal; there was nothing equivalent to a fixation point or to the visual masks that followed the sentence and the list in the visual version. The timing of the sentence, the list, and the intervals between the sentence, list, and probe were matched as closely as possible to those of the visual group.

Results and Discussion

Table 1 shows the main results of Experiment 6. The general level of accuracy of recall was high in both conditions: apart from the target word, 87% of the words were recalled correctly in the visual group and 91% in the auditory group. Aside from intrusions of the critical word, there were 9% recall errors for the target word in the visual group and 6% in the auditory group. For the visual group the synonym lure replaced the target in 33% of the recalled sentences when the lure was on the list and there were 6% spontaneous intrusions when the lure was not on the list. For the auditory group the comparable intrusion rates were 53% and 13%, respectively. Analyses of variance on the number of intrusions showed significant main effects of group (visual versus auditory), $F_{min}(1,45) = 4.53, p < .05$, and of presence or absence of the lure, $F_{min}(1,40) = 41.8, p < .001$. The interaction between group and lure condition was also significant, $F_1(1,26) = 4.92, p < .05, F_2(1,19) = 4.63, p < .05$. Separate analyses showed significant effects of the lure condition in both groups for both subjects and items, all $ps < .001$.

The results of Experiment 6 show that the tendency to intrude an incidentally activated synonym is even greater when the sentence and list are spoken than when they are written. This result might seem surprising, in view of the hypothetical role of a sound-based representation in aiding sentence processing. However, it is quite possible that the modality effect was limited to the list, which was the most recent part of the trial. An auditory trace of the lure word could have made it especially salient during recall, increasing the likelihood of intrusion.

In any event, the main conclusion from Experiment 6 is that the tendency to make synonym substitutions is present for auditory as well as visual sentences, just as the regeneration hypothesis predicts. The possible availability of a lingering sensory trace, in the case of the spoken sentences, did not enhance the accuracy of recall or prevent meaning-based intrusions.

Experiment 7

The purpose of Experiment 7 was to extend the investigation of conceptually-based immediate memory for sentences to young children. Apart from the intrinsic interest of investigating the short-term representation of sentences in children, such a study could shed light on the question of whether the adult results reflect a memory strategy rather than a fundamental characteristic of short-term memory. Memory strategies such as rehearsal are known to develop in the early school years (Flavell, 1977), so pre-school children are unlikely to make use of them. Since pre-school children clearly do understand sentences, if they make synonym intrusions in something the same manner as adults that would suggest that the underlying mechanism is not the result of a strategy acquired in later school years. The failure to find intrusions in the sentence recall of young children would be less informative, because (in addition to the possibility that intrusions result from an acquired strategy) there would be other possible explanations: for example, greater reliance on surface mimicry in recall, or a lack of overlap in the conceptual
representations of the putative synonyms, in the child’s lexicon (see, for example, Clark, 1987; Merriman & Bowman, 1989).

Method

Subjects. The subjects were 16 children (eight girls and eight boys) between the ages of 3:11 and 5:2 (mean age 4:5) who were attending one of two daycare centers. Four additional children (ages 4:3, 4:3, 4:5, and 4:8) were not able or willing to repeat sentences, so they were not included in the experiment.

Materials and design. Sixteen synonym or other related pairs of nouns that we judged to be appropriate for 4-year-olds were selected. For each pair, a sentence of 6 to 9 words (mean 7.6 words) was constructed which was appropriate for either word. For each sentence a four-word list was constructed which included the synonym and three other age-appropriate nouns. The sentences and lists are shown in Appendix C. The synonym was either the first or the last word on the list, counterbalanced over materials and subjects. Which synonym appeared in the sentence and which on the list was also counterbalanced, making a total of four versions of the experiment. (The list always included a synonym.) Four children participated in each of the four versions.

There were seven practice sentences, two without lists, two with three-word lists, and three with four-word lists.

Procedure. The experimenter visited the daycare group and played with the children before beginning the experiment, which was carried out with one child at a time. The child was invited to play a game of repeating sentences: “I’ll say something and you say the same thing.” Children who were shy were coached through the first three-word practice sentence one word at a time, and then the whole sentence. After the first two practice sentences a hand puppet named Max (a pig) was brought out and introduced. The experimenter said to the child: “Max wants to play the game with us. He thinks you are so good at repeating sentences that he is going to try to make it harder. I will say a sentence and at the end of it Max will jump up and say some words. Then, you have to remember my sentence, and don’t let him confuse you.” The experimenter then spoke the next practice sentence in a normal voice, followed immediately by the pig (which had been resting on the table) “jumping up” and moving its mouth while the experimenter spoke the list in a slightly higher pitch with a vigorous list intonation: e.g., “house! car! fish!” Although most of the children were able to repeat just the sentence, if they forgot the practice sentence or repeated the list the experimenter said something like, “You let him trick you! Remember, you have to say my sentence and not let him confuse you. Let’s try that one again.” If the child began to perform the task correctly and confidently, some of the practice sentences were omitted. All children received the two practice sentences without lists and at least one sentence with a three-word and one with a four-word list.

The session was tape-recorded and transcribed later by the experimenter.

Results and Discussion

The 4-year-old subjects were able to repeat the sentences fairly accurately. The mean percentage of words other than the critical target that were reported correctly was 84%: apart from the critical word, 64% of the sentences were repeated entirely accurately. On 52 (20%) of the 256 trials the synonym lure was intruded in recall; it was substituted for the critical target word on 48 of those trials and on four trials both words were included in recall. On three trials (1%) another word from the list was substituted. Synonym intrusions were about twice as frequent when the lure was the first word on the list than when it was the fourth (last) word. All but two of the 16 sentences generated at least one synonym intrusion; 10 sentences generated at least one intrusion in each direction (e.g. bunny was changed
All but one of the children made at least one synonym intrusion; the number of intrusions out of the 16 trials ranged from 0–7, $M = 3.25$, $SD = 2.08$. Neither age nor overall accuracy in recall of the sentences (excluding the critical target word) was correlated with the number of synonym intrusions made (Kendall’s rank order correlation: $S(16) = -8$ and $+2$, respectively).

The pattern of results is thus similar to that for adults: synonym lures were intruded in recall, but other words on the lure list were intruded only rarely. Unlike the experiments with adults, there was no control condition in which no synonym lure was on the list (this was done to keep the experiment short enough to be completed by the young subjects). However, it seems unlikely that the intrusions we observed were largely or entirely spontaneous (i.e., would have occurred had there not been a lure on the list). Although the subjects did make other paraphrases in recall from time to time, none appeared with the consistency of the lure intrusions.

In conclusion, the similarity between 4-year-olds in Experiment 7 and the adults in the earlier experiments indicates that sophisticated adult strategies are not responsible for the prevalence of lure intrusions in recall of sentences: the phenomenon is clearly present in sentence recall by 4-year-old children.

**General Discussion**

To return to our initial question, why is immediate recall of a sentence ordinarily so accurate? Not, we have suggested, because short-term memory includes a verbatim surface representation that can simply be read off. Instead, immediate recall of a sentence is like long-term recall in that it begins with a representation of the meaning of the material to be recalled. Unlike long-term recall, in immediate recall the original lexical items are likely still to be activated and can be selected to express the message, resulting in verbatim or near-verbatim recall.

One implication of this hypothesis is that if there is more than one activated word that is capable of expressing a given concept in the message, the wrong one may be chosen in recall. In the present experiments this prediction was tested by presenting a synonym lure in an incidental task that immediately preceded or followed the sentence that was to be recalled. While the other words of the sentence were recalled with high accuracy, the critical target word was, as predicted, frequently replaced by the lure word (34% of the trials, averaging over the experiments). Because the lure word was chosen to be highly appropriate in the sentence, some of the time (10% of the trials) it also intruded spontaneously when not presented as a lure, consistent with the hypothesis that immediate recall is based on sentence meaning rather than on a representation of the surface string. Activating the word by presenting it on the distractor list increased the intrusion rate dramatically (from 10% to 34%), supporting the claim that recently activated words are selected preferentially in immediate recall.

Ordinarily, the selection of recently activated words to express the meaning of a just-read (or just-heard) sentence guarantees accurate recall, giving rise to the impression that there is a specialized verbatim memory store. The present findings suggest that this impression is mistaken, at least in the case of sentences of 10 words or more (we did not test shorter sentences). As a number of theorists have suggested, the earlier assumption that there was just one form of short-term storage was wrong; instead, there are multiple subsystems that together account for short-term memory (e.g., Baddeley, 1986; Crowder, 1982; Monsell 1984; and Potter, 1983). The present study shows how two of these systems, conceptual and lexical, work together in recall of sentences.

Thus, short-term sentence memory is like long-term memory in being primarily...
SHORT-TERM RECALL OF SENTENCES

Conceptual or propositional rather than literal. In long-term memory the conceptual content is usually all that is retained, so that the particular words, phrases, or syntactic structures of the original communication are lost (e.g., Bartlett, 1932; Bransford & Franks, 1972; Brewer, 1975; Sachs, 1967, 1974). Having a set of recently activated words to draw on is what distinguishes short-term from long-term recall.

Short-term recall of word lists versus sentences. A coherent conceptual representation is not available for a list of words (i.e., a random string of words), but for short lists there is ample evidence that an auditory-like short-term representation is available for visual as well as spoken lists. This representation preserves the order of items and can be refreshed by rehearsal (e.g., the articulatory loop and phonological store postulated by Baddeley, 1986, and his colleagues). However, the capacity of that system is only three to six words, so it can play at best only a supporting role in recall of longer sentences. The acoustic representation of a list, we conjecture, is used to retrieve activated lexical entries (just as conceptual representations cue activated lexical entries in the model we propose for sentence recall). A partially degraded acoustic trace may be sufficient to cue the correct word among those entries that are active. When there is no lexical entry, as with nonsense words, the immediate span is much shorter (Brener, 1940), presumably because the acoustic trace must be complete.

When the words on a list are acoustically similar, a partially degraded acoustic trace may not be able to discriminate between two currently activated words, for example, bud and bug, so that words may be recalled correctly but in the wrong order, producing the familiar negative effect of acoustic similarity on ordered short-term recall (e.g., Conrad, 1965; Schweickert, Guentert, & Hersberger, 1990). Acoustic similarity would not necessarily interfere with the number of words recalled regardless of order (as in free recall). In longer-term recall neither the original lexical items nor an acoustic representation would be active, so recall would have to be based primarily on conceptual information. If the subject remembered that some words rhymed or sounded similar, however, that could be used as a retrieval strategy, leading to a benefit for lists of phonologically similar words. That is just the pattern of results that has been observed in some studies (e.g., Watkins, Watkins, & Crowder, 1974; see Crowder, 1976, for further discussion).

Thus we suggest that immediate memory for a sentence relies on a structured conceptual representation whereas memory for short lists makes use of an ordered acoustic representation, but that recall in each case involves the use of the respective representation to retrieve activated lexical entries. An intermediate case is recall of strings of words that are stochastic approximations to English, ranging between random lists and normal sentences. For example, Marks and Jack (1952) constructed strings in which any two, three, or five adjacent words were constrained to come from a sentence, or the whole string was a sentence. The closer the approximation to English, the longer the string that could be recalled verbatim; complete sentences were markedly better than even fifth-order approximations, with the mean of the longest string a subject could recall being 15.1 words for sentences and 10.0 words for fifth-order approximations, but 7.7 for second-order approximations. This general pattern is what would be expected if lexical entries are activated in all cases, but only fragmentary conceptual structures can be constructed for approximations to a sentence.

Function words. Although the present hypothesis can explain the verbatim recall of content words in a sentence, function words seem to be more of a problem. Some function words are used frequently enough to be almost continuously active in normal discourse. In some cases two such words—
for example, the determiners “the” and “a”—could be used interchangeably or omitted without a major alteration of meaning. Without an ordered surface representation, such errors in recall might then be expected. Inspection of the 9–14% of errors on nontarget words in the present experiments indicates that these kinds of errors do occur. Similarly, Greenbaum (1970, cited by Clark & Clark, 1977) found that when subjects immediately recalled short sentences with an awkwardly placed adverb (e.g., “He badly needed the money”) they often moved the adverb to the more natural position. (Greenbaum’s task required subjects to pluralize the subject pronoun in recalling the sentence, which may have been sufficient to prevent recall from the articulatory loop even though the sentences were very short.) These kinds of errors are consistent with the present hypothesis.

Surface structure. Some forms of paraphrase make only minimal changes in vocabulary: notably, a change from active to passive voice or vice versa. Yet subjects rarely change the voice of a sentence in immediate recall. One possible explanation is that the phrase structure of the sentence remains primed and tends to be reused (Bock & Loebell, 1990: but see Footnote 3). Another possibility is that the conceptual representation includes information about focus and other subtle interpretative differences that are sufficient to determine the voice of the regenerated sentence. (This may also explain why voice changes were more readily recognized than other formal changes, in Sachs’ experiments, 1974.) A related question is why the probability of intrusion of the lure word is almost always lower than the probability of correctly recalling the target word, even though the lure word was often a somewhat better fit to the rest of the sentence (e.g., “palace” versus “castle,” in (1)). Here, two factors were pitted against each other: fit to the rest of the sentence, and the specific conceptual information conveyed by the word actually in the sentence. To the extent that subjects included in their conceptual representation information that was specific to the target word and was not shared by the lure, intrusion of the lure would be less likely than correct recall. In longer-term memory, however, such minor conceptual distinctions as those differentiating active from passive voice or a given word from a near-synonym are likely to be lost, as experiments already cited have shown.

The present experiments do not rule out the possibility that immediate sentence recall is based on a mixture of representations of the target sentence that includes acoustic, orthographic, and syntactic (as well as conceptual and lexical) representations of the target sentence. But a parsimonious interpretation of the present and earlier results is that only conceptual and lexical representations are involved. The conceptual representation is expressed as in normal sentence production; choice of lexical items from those currently activated—in particular, the choice of the verb—constrains the detailed syntactic structure of the sentence so that recall is verbatim, or nearly so.

Lexical activation. The present hypothesis posits that recently perceived words remain activated for some period. This period must be relatively brief—perhaps lasting only a few seconds—or there would be interference from earlier sentences in the recall of the most recent sentence in running text, contrary to evidence for excellent last-sentence recall in experiments such as those of Jarvella (1979). But just how long does lexical activation persist? In studies of repetition priming of words, there is evidence for a long-standing effect of the prior presentation, in some cases persisting for days. Significantly, however, this effect does not readily generalize between different surface forms such as spoken and written words (see Monsell, 1985, for a review), suggesting that the process involved in these long-term effects is perceptual (e.g., recognition of an orthographic pattern)
rather than conceptual or fully lexical. Such long-term perceptual effects generally are not correlated with recall and are therefore irrelevant to the present hypothesis.

Evidence for short-lasting activation of lexical entries comes from paradigms such as the Peterson-Brown method for evaluating passive persistence of words in short-term memory, when the subject is performing a demanding secondary task that prevents rehearsal. In this paradigm, the probability of recall declines sharply over the first 10 or 15 s of intervening activity. A major reason for the decline is that the current to-be-remembered words become confused with the words presented on earlier trials, when all the words are drawn from a single semantic category or are in other ways conceptually related (see Wickens, 1972, for a review). If we consider this loss to be a loss of primary activation, it has about the right time course to account for accurate recall of the most recent sentence using the proposed mechanism of lexical selection. Direct tests of sentence recall by Martin, Roberts, and Collins (1968) and Martin and Walter (1969) using the Brown-Peterson distraction method have shown a marked drop in accuracy between 1 and 10 s of delay (no intermediate delays were used), with further drops at longer intervals.

Speech errors involving anticipations, exchanges, or blends of words have been taken as evidence that, during speech production, the two words were simultaneously active. Such errors almost always involve words in the same sentence (Garrett, 1975). Although such speech errors are superficially similar to those observed in the present experiments in that one activated word replaces or intrudes on another, there is a critical difference: speech errors are by definition mistakes that the speaker can (and usually does) recognize and correct, because the speaker can hear that what was said means something different from what was intended. In the present experiments, in contrast, it was rare for a subject to make an intrusion and then correct it, because the error resulted from an inaccuracy of memory rather than an inaccuracy of speech production. Nonetheless, both phenomena may require the same precondition: persisting activation of potentially competing lexical items.4

Short-term memory and sentence processing. How does the view of memory for sentences that emerges from the present study fit with recent theories of the role of "verbatim memory" in sentence comprehension (see Baddeley, 1986, for a review)? It is generally agreed that at least the current phrase, clause, or perhaps the whole sentence must be available in a complete form in order to parse the sentence, resolve surface anaphors, make bridging inferences, and so on. It has commonly been assumed that this means that a surface representation is available, but we have argued that that assumption may not be justified. Although there are numerous studies showing that sound-based representations of the words of a written sentence are activated during reading, it is not clear what role those representations play in processing (but see Black, Coltheart, & Byng, 1987). The needed level of representation may be lexical in some cases, syntactic in others, and conceptual or propositional in still others (see, for example, the distinction between surface and deep anaphora, reviewed in Garnham, 1987). Further, what-

4 Another example of an error resulting from an inaccuracy in matching present information to long-term memory is the Moses illusion investigated by Erickson and Mattson (1981) and more recently by Reder (1989; Reder & Cleeremans, 1990). In this illusion, a near match between a question such as "How many animals of each kind did Moses take on the Ark?" and information in long-term memory leads subjects to treat the mismatching information (Noah replaced by Moses) as though it were correct. This phenomenon is in some respects a long-term parallel to the recognition task used in Experiment 4. Reder has suggested (personal communication, July 1989) that the factors that affect susceptibility to the Moses illusion will be similar to those that produce intrusions or false recognition in the present study.
ever the needed representation may be, it could either persist in memory or be reconstructed when needed from a more abstract representation. To distinguish between surface persistence of the sentence and the present regeneration hypothesis, one would need to study cases in which the two candidate representations diverged.

Conclusion. The view of sentence memory supported by the present experiments is that immediate memory for a sentence is conceptually based and reconstructive, like long-term recall. Memory is (nearly) verbatim not because of a special short-term representation of the surface sequence, but because the regenerative process of recall makes use of recently activated but unordered entries in the lexicon to express the ideas in the sentence, using the normal mechanisms of sentence production. Thus, if two words are synonymous in a given context and both have been recently activated, either may be chosen in recall. Under ordinary circumstances, however, such a situation would be rare and so recall would be verbatim.

This conceptual view of sentence memory invites a reconsideration of questions about short-term memory in sentence processing, replacing the idea of an ordered surface string that is independent of deeper levels of processing (and can be consulted if things go wrong) with a representation that is based on the deepest, message level. If this new view of short-term memory for sentences is correct, then initial errors of interpretation during perception should be difficult to correct unless they are caught within three or four words (the presumed duration of the surface trace, in this view); otherwise, a listener would need to regenerate the original sentence from its conceptual base (together with activated lexical items), attempting to locate the error during this reconstruction. It remains to be seen whether this new view can account not only for the pertinent psycholinguistic data, but also for evidence from cases of patients who have deficient memory spans but adequate comprehension of sentences.5

Appendix A: Materials in Experiments 1, 2, and 5

Each trial consisted of a sentence and a distractor list with a probe (the capitalized last word). The critical word in the sentence is shown in italics, as is the synonym lure on the distractor list; following the slash is the word that replaced the lure on control trials.

1. The modern art display at the museum has some very unusual paintings.
   month exhibit/floor summer license hotel HOTEL

2. He gave up his chair so that the old woman would not have to stand.
   page metal road seat/goal shore ROAD

3. The history professor waited impatiently for his reply to the unexpected question.
   answer/county energy market island doctor NICKEL

4. The movers rolled up the carpet and threw it into the truck.
   rug/fur boot tub acid lime ACID

5. The gym instructor started off each day with some jumping jacks and sit-ups.
   teacher/umbrella platform aircraft diamond carnival BASEBALL

6. The shrub near the porch needed to be trimmed before the party.
   straw onion flag bush/trunk mist SPEAR

7. The new curtains for the guest room are made of brightly colored material.
   liquor meat pilot fabric/queen steel PILOT

8. The fireman was honored for his courage in saving the helpless children.
   phrase bravery/drill knee convent ocean NIECE

9. They moved into their new home a week after he started his job.
   group board house/power voice school BOARD

10. The canoe skimmed across the water

5 See Saffran (1990) for a compatible approach.
as the men looked for fish. gas lake/land
wheel bear fort GAS
11. The report was checked by four people before a mistake was found. cloud error/guid powder barn fiber BARN
12. The knight rode around the palace searching for a place to enter. turtle recipe booth castle/vowel medal SHOWER
13. They went down to the cellar and rummaged through some boxes of toys. storm football mirror basement/climate pistol PISTOL
14. Bill fixed supper for the family after working at the office. ranch barrel dinner/temple legend parade RANCH
15. The woman scientist used snake poison in her research on brain function. cart jockey venom/shovel altar plumb FERRY
16. The evening was chilly so he put on his coat before leaving for the ballgame. figure amount jacket/society heart music PEANUT
17. The boxers stood in the center of the ring while the referee spoke. window middle/horse income plane staff THIGH
18. The mound of soil grew as they dug up the flower bed. dirt/desk gift horn lunch maid ALBUM
19. They stacked all the cartons in the rear of the closet. pocket dollar chamber back/award screen MANTLE
20. The woman asked the author to sign her copy of his new book. writer/shelter traffic animal league factor TRAFFIC

APPENDIX B: MATERIALS IN EXPERIMENTS 3, 4, AND 6

In Experiments 3 and 4 each sentence had two versions: in (a) the target and lure matched and in (b) they mismatched. Experiment 6 used the (a) versions of the sentences, with a few changes [shown in square brackets] and with a new random order of the sentences. The lure word (the recognition probe in Experiment 4) is italicized, as is the critical word in the sentence. The distractor lists and list probes shown in Appendix A were used with the corresponding lure-target pairs in the present experiments, with only minor changes.

1. exhibit a. The modern art display at the museum has some very unusual paintings. [The art museum has a new display of modern paintings.]
   b. The new window display at the store has some very fancy dresses.
2. seat a. He gave up his chair so that the old woman would not have to stand.
   b. He got another chair so that his uncle could join him on the patio.
3. hazard a. The rocks falling on the highway have become a danger for drivers.
   b. The gang hanging around the mall has become a danger to shoppers.
4. answer a. The history professor waited impatiently for his reply to the unexpected question. [The history professor waited impatiently for the reply to his unexpected question.]
   b. The general waited anxiously for the reply to his urgent message.
5. rug a. She decided to move the carpet from the study to the hall.
   b. She decided to tack the carpet onto the stairs in the hall.
6. teacher a. The gym instructor started off each day with some jumping jacks and sit-ups [...situps.]
   b. The driving instructor started each lesson with some basic rules and helpful tips.
7. material a. The new curtains for the guest room are made of brightly colored material.
   b. The new curtains for the guest room are made of poor quality material.
8. bravery a. The fireman was honored for his courage in saving the helpless children.
   b. The little boy showed his courage by standing up to the bully.
9. house a. They built their home in less than six months.
b. They have owned their home for more than twenty years.

10. middle a. The boxers stood in the center of the ring while the referee spoke.
   b. The astronomer pointed the telescope at the center of the galaxy.

11. lake a. The guide in the canoe paddled across the water toward the island.
   b. The little girl wearing the red swimsuit paddled in the water near the island.

12. error a. The tax return was checked three times before the mistake was found.
   b. The decision to take the job was the biggest mistake of his life.

13. castle a. The knight rode around the palace searching for a place to enter.
   b. The royal guests danced in the palace to the music of an orchestra.

14. basement a. They went down to the cellar and rummaged through some boxes of toys. [They can’t store things in the basement because of the dampness. Lure: cellar]
   b. He went down to the cellar and found a bottle of wine for the evening.

15. dinner a. The butler prepared supper for the ambassador and his wife.
   b. Jimmy ate supper with his brother after playing in the yard.

16. venom a. The woman scientist used snake poison in her research on brain function.
   b. The woman scientist used poison in her research on brain function.

17. jacket a. The evening was chilly so the man put on his coat before leaving for the ballgame.
   b. The weather was cold so the lawyer put on a coat to walk to the office.

18. dirt a. The pile of soil grew as they dug up the flower bed.
   b. The rich soil was plowed and watered to prepare the flower bed.

19. oven a. Jane took the casserole from the stove and served it to the visitors.
   b. Jane took the frying pan from the stove and served it to the visitors.

20. writer a. The author sat down to work on the last chapter of his new book.
   b. The author sat down to sign a copy of his new book.

APPENDIX C: MATERIALS IN EXPERIMENT 7

The pair of words used as both lure and target (in different versions) is shown in italics in each sentence. The word used as a lure either preceded or followed (in different versions) the three other distractor words shown.

1. My friend got a rabbit/bunny for his birthday. boat hand floor

2. Everyone should wear mittens/gloves when it snows. chair pig egg

3. The girl wants to pet the cat/kitty. train song bowl

4. The kids/children can’t play with the stove. clock hat paper

5. Don’t step in the puddle/water with your new shoes. truck head bottle

6. The lady/woman in the store is buying apples. sun color street

7. The teacher will read the story/book after lunch. puzzle mouth bag

8. An adult/grownup should help with the scissors. arm spoon crayon

9. At home/At my house, I have a red balloon. swing duck window

10. He asked his mother for a glass/cup of milk. light button foot

11. The puppy/dog wants to go for a walk. soup door phone

12. The shoes/sneakers are in the closet. farm cake bear

13. Take off your coat/jacket and hang it up. name car finger

14. The toys/blocks are all on the floor. cow night leg

15. The boy rode a pony/horse at the zoo. bath money fish

16. I like to eat vegetables/carrots for dinner. stairs yard paint

REFERENCES

SHORT-TERM RECALL OF SENTENCES


(Received December 12, 1989)

(Revision received April 30, 1990)