The Regeneration of Syntax in Short Term Memory

LINDA LOMBARDI
University of Maryland

AND

MARY C. POTTER
Massachusetts Institute of Technology

Immediate recall of a sentence is normally highly accurate. Previous work (Potter & Lombardi, 1990) has suggested that this accuracy is not due to a surface ("verbatim") representation of the string of words, but to regeneration of the sentence from a conceptual representation, using activated lexical entries. The present study tests an extension of that hypothesis: that the surface syntax of the to-be-recalled sentence is not directly represented in memory, but is regenerated using normal mechanisms of sentence production. In Experiments 1 and 2, induced intrusions of verbs were not prevented by incompatibility between the surface syntax of the original sentence and the subcategorization requirements of the intruded verb. In Experiments 3 and 4 subjects judged whether a probe verb could replace a verb in the to-be-recalled sentence; again, subjects had difficulty rejecting verbs on the basis of surface syntactic incompatibilities. In all four experiments subjects who recalled a sentence with a syntactically mismatched verb usually made syntactic changes that restored grammaticality. The results support an extension of the regeneration hypothesis: In regenerating a sentence from its conceptual representation the selection of the verb determines the syntactic structure of the sentence. When more than one structure is compatible with the conceptual representation and with the chosen verb, a structure that has been recently activated is likely to be reused. © 1992 Academic Press, Inc.

The present study tests the hypothesis that the surface syntax of a perceived sentence, although used in comprehension, is not directly represented in memory. What is directly represented in short-term memory is the sentence’s conceptual message. In recall, normal mechanisms of sentence production are used to express this message, drawing on currently active lexical representations (Potter & Lombardi, 1990). The syntax of the sentence is normally reproduced correctly because the original verb has been selected, and it dictates the appropriate syntax. In this way, the perceiver succeeds in repeating the original sentence with perfect or near-perfect accuracy.

Because short-term recall of a sentence is so accurate, it is ordinarily uninformative about the memory representation(s) being used. However, a method to induce representation-dependent errors in immediate recall was developed by Potter and Lombardi (1990), who asked subjects to read or listen to a sentence and then (after a brief distractor task) to recall it. When a plausible synonym for one of the words in the sentence was included in the distractor task (e.g., castle for palace), intrusion of that word in recall was frequent. Moreover, the
probability of an intrusion depended on the meaning of the whole sentence, not simply on the two words in question. Overall recall of the sentence, apart from the induced intrusions, showed the high accuracy typical of short-term sentence memory.

These findings led us to propose that accurate short-term memory for a sentence depends on two main factors: a conceptual representation of the sentence and the recent activation of a set of lexical units. In recall, the speaker expresses the conceptual representation in a manner similar to normal speech production (e.g., Bock, 1982; Bock & Brewer, 1974; Greenbaum, 1970; Levelt, 1989), except that lexical selections are made preferentially from currently active lexical units. Activating an extraneous but plausible competing word can therefore induce an intrusion.

An implication of this hypothesis is that there is no direct representation of the syntax of a sentence in short-term memory once the sentence has been parsed and understood. Rather, the syntax is regenerated when the subject uses the activated lexical items to express the conceptual message of the sentence. However, the materials of Potter and Lombardi's (1990) experiments were not designed to test the question of memory for syntax explicitly. In the sentences used, if the subject had a conceptual representation of the sentence (together with activated lexical items) available as recall begins, then donate might intrude in (1b) as well as in (1a). The subject would then generate a syntactic structure compatible with the verb, such as (2a), rather than the more literally correct response, the ungrammatical (2b).

(1a) The rich widow is going to give a million dollars to the university.

(1b) The rich widow is going to give the university a million dollars.

Verbs like give can alternate between these syntactic forms without changing the conceptual content of the sentence, but other dative verbs like donate are restricted to the prepositional-phrase (NP-PP) form (2a); the double noun phrase form (NP-NP) is ungrammatical (2b).

(2a) The rich widow is going to donate a million dollars to the university.

(2b) *The rich widow is going to donate the university a million dollars.

Consider what might be expected if the subject reads either (1a) or (1b) and then is presented with donate on an incidental distractor task before recalling the sentence (in the paradigm of Potter & Lombardi, 1990). If, during recall, subjects have the surface structure in mind when they select the verb, they would intrude donate in place of give only in (1a), not in (1b). On the other hand, if only a conceptual representation of the sentence (together with activated lexical items) is available as recall begins, then donate might intrude in (1b) as well as in (1a). The subject would then generate a syntactic structure compatible with the verb, such as (2a), rather than the more literally correct response, the ungrammatical (2a).

When two or more structures are compatible with the chosen verb (correct or intruded), the regeneration hypothesis as just described makes no prediction about which structure will be produced in recall. Presumably such factors as recency and frequency of use would determine the choice. In studies of speech production, recently encountered syntactic structures influence the choice of structure in the production of a new sentence (e.g., Bock, 1986). Crucially, such a preference for reusing a recent structure is found even when the priming sentence is unrelated in conceptual content to the sentence being produced, suggesting that the primed structure is represented autonomously. Such priming ef-
fects are consistent with the regeneration hypothesis: there is a bias to reuse the original structure as long as it is compatible with the selected lexical items, in particular the verb. Such a priming effect would parallel the effects of lexical activation on lexical choice: other things being equal, recently activated words and syntactic structures will be chosen when the sentence is regenerated from a conceptual representation.

In the present experiments we used dative verbs and also other verbs with differing complement structures to test the hypothesis that surface syntactic structure is not directly represented in short-term memory, but that it is regenerated during recall. The main predictions are that verb intrusions in recall will not be prevented by incompatibility between an intruder verb and the original surface syntax and that when such intrusions occur (but rarely otherwise) the syntactic structure of the recalled sentence will differ from that of the original, as required for grammaticality. The first two experiments used the same method as Potter and Lombardi (1990); the third and fourth experiments employed a complementary method.

**Experiment 1**

As in the method used by Potter and Lombardi (1990), on each trial the subject read a sentence, performed a brief distractor task, and then recalled the sentence. The distractor task consisted of the presentation of five verbs in sequence, followed by a capitalized test verb that had just appeared on the list—on half the trials—or was a new word. The subject had to indicate whether the test word had been on the list (a simple task) and then recall the sentence aloud. Unknown to the subject, the distractor list often included a lure word, a synonym of a verb in the sentence; this verb was never the one tested in the distractor task.

As in the earlier study with noun lures, we expected that the lure verb would intrude in place of the sentence verb on some trials. In critical cases, the lure verb required a different surface structure than the one in the sentence to express the same grammatical relations. The question was whether this incompatibility would prevent intrusions, showing that recall is based not only on conceptual representations and activated words, but also on explicit memory for the syntactic structure of the sentence.

Two types of verbs were investigated in Experiment 1: dative verbs that do or do not alternate (as in (1) and (2) above) and verbs that are synonymous except that one requires a prepositional phrase (PP) complement (e.g., *think about*) and the other a noun phrase (NP) complement (e.g., *ponder*).

**Method**

**Subjects.** The subjects were 28 college students who had volunteered to participate and were paid.

**Materials and design.** Of a total of 50 sentences, 16 were experimental sentences with lures (10 datives and 6 NP versus PP complements), 6 were dative control sentences (without a lure), 6 were control sentences with a compatible verb synonym as a lure, and 22 were filler sentences. Each sentence was followed by a list of five verbs. Except for 12 of the filler sentences and the 6 dative controls, there was always one verb on the list that was a potential substitute or lure for the sentence verb. The remaining verbs on the list were chosen to be similar in length but unrelated in meaning to the lure verb or sentence. In addition, there were 10 practice sentences and lists; none included potential word substitutes on the list.¹

The 10 dative sentences used verbs that participate in the dative alternation, while the lure required the NP-PP form. Five of the sentences in a given version of the experiment were written in the NP-PP form (as in (1a)), and the other five in the NP-NP

¹ A list of stimulus materials used in all the present experiments is available from either author.
(double object) form (as in (1b)). In the complementary version the forms were switched. The verb pairs were (with the alternator first): give/donate; show/display; send/transmit; send/submit/tell/mention; bring/convey; read/preach; mail/address; read/say; offer/suggest.

The six NP/PP complement sentences could include either of two synonymous verbs, but one required a prepositional phrase as the complement, whereas the other required a direct object. The verb pairs (with the relevant preposition shown in brackets) were: request/ask [for]; discuss/talk [about]; want/wish [for]; ponder/think [about]; examine/look [at]; protest/object [to]. In each version three of the sentences used the NP complement verb and three the PP complement verb, counterbalanced between versions; the other verb of the pair appeared on the distractor list. The preposition was included with the relevant verb when that verb was in the sentence (3a), but not when that verb was in the list (3b).

(3a) The committee has to talk about the issue before it makes a decision. *DISCUSS
(3b) The committee has to discuss the issue before it makes a decision. *TALK

The six dative control sentences used dative alternating verbs but did not have a potential substitute in the list. The sentences were presented in the NP-PP form, half in the NP-NP form, counterbalanced across the two versions of the experiment. These verbs were lend; write; teach; serve; sell; take.

In the six synonym control sentences, straightforward verb synonyms appeared in sentence and list; they were exchanged between versions. The sentences were in a form appropriate for both verbs. The pairs were complete/finish; purchase/buy; choose/select; answer/reply; grow/increase.

The 22 filler sentences were included to increase the diversity of the types of verbs used. Twelve of these sentences had no lures on the distractor list; the other 10 included lure verbs that would have required substantial changes had they been intruded (e.g., not succeed/fail; put [that question to x]/ask [x that question]), and in fact intrusions were rare.

The 50 experimental and filler sentences were randomized. There were two versions of the experimental materials, counterbalancing the form of the experimental sentences.

Procedure. The procedure was the same as that used in Potter and Lombardi (1990). Sentences and distractor lists were shown using rapid serial visual presentation (RSVP), in which each word appears in sequence in the same location. Each trial began with a row of three asterisks, which remained in view for 300 ms, followed by a blank of 350 ms and then the sentence, at 200 ms per word. The first letter of the first word was capitalized and the other letters were in lower case. A mask of percentage signs was displayed for 517 ms immediately following the last word of the sentence and then the five-word list was presented for 250 ms/word (to 10 subjects) or 183 ms/word (to the remaining 18 subjects). (The two rates of presentation of the distractor list did not have a significant differential effect on any of the results and will not be discussed further.) At the end of the list the mask reappeared for 250 (or 183) ms and then the capitalized test word appeared for 500 ms. The subject responded to the test word by pressing a key for Yes or a key for No and then recalled the sentence aloud. The test word was on the list on half the trials. On the 16 experimental trials a possible substitute (lure) verb appeared on the list; it was never the word tested and it was never the last word on the list.

Results

Overall, the sentences were recalled fairly accurately; leaving out the critical verbs, 88.5% of the words were recalled. This level of accuracy for 9–14 word sentences is equivalent to the level of recall
TABLE 1

EXPERIMENT 1: PERCENTAGE OF RECALLED SENTENCES WITH AND WITHOUT AN INTRUDED LURE VERB AND (GIVEN THAT RESPONSE) THE PERCENTAGE OF SPONTANEOUS CHANGES IN THE SYNTAX OF THE VERB PHRASE

<table>
<thead>
<tr>
<th>Lure type</th>
<th>Sentence form presented</th>
<th>Intrusion</th>
<th>% Resp.</th>
<th>% Change</th>
<th>No intrusion</th>
<th>% Resp.</th>
<th>% Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonalternator</td>
<td>NP-PP</td>
<td>11</td>
<td>7</td>
<td>89</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*Nonalternator</td>
<td>NP-NP</td>
<td>7</td>
<td>90</td>
<td>93</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No-lure control</td>
<td>NP-PP</td>
<td>—</td>
<td>—</td>
<td>100</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No-lure control</td>
<td>NP-NP</td>
<td>—</td>
<td>—</td>
<td>100</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*PP</td>
<td>NP</td>
<td>6</td>
<td>100b</td>
<td>94</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*NP</td>
<td>PP</td>
<td>21</td>
<td>100c</td>
<td>79</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. The asterisk (*) indicates the conditions in which an intrusion of the lure would create an ungrammatical sentence unless other changes were made.

a The verb in the sentence was an alternator.
b Addition of the preposition.
c Deletion of the preposition.

achieved in comparable studies without distractor tasks, and hence qualifies as “verbatim” recall. Accuracy in the distractor task was also high: 95% of the test words were responded to correctly.

Datives. Table 1 shows the main results. The first question was whether the nonalternator lure word intruded in place of the alternator verb, in recall—it did, on 8.9% of the recall trials—and the second question was whether intrusions of the lure verb were confined to the sentences with compatible surface syntax—they were not. While the overall results indicate that the lure was slightly more likely to intrude when the sentence was in the form NP-PP than when it was in the form NP-NP (11% versus 7%), the difference was not reliable by a sign test, n = 12, x = 4. (Throughout the paper, analyses of variance and t tests are used when possible, but when n’s are small or other assumptions of parametric statistical tests are not met, nonparametric tests are substituted.)

The third question was whether, when a syntactically incompatible lure intruded in recall, the syntax of the recalled sentence changed (if needed) to accommodate it. Table 1 shows the percentage of spontaneous changes in the verb phrase made during recall of the dative sentences, conditional on whether an intrusion was made. The percentage of changes in recall of the dative controls that had no potential substitute in the list is also shown. For the NP-PP sentences, changes in syntax were very rare, whether or not there was an intrusion. For the NP-NP sentences, however, in 9 of the 10 cases in which subjects intruded the lure verb, the syntax was changed so that the recalled sentence was grammatical with the new verb (the sentence was changed to NP-PP or else the indirect object noun phrase was omitted). When the lure verb was not intruded the syntax of the verb phrase was much less likely to change (18% vs. 90%), \( \chi^2(1, N = 140) = 22.6, p < .001 \). The dative controls, without synonym lures and with no spontaneous intrusions of other verbs, produced 7% changes for NP-PP sentences and 15% for NP-NP sentences, \( \chi^2(1, N = 168) = 2.14, n.s. \) Thus, to sum up, changes in the surface structure of the

2 The chi-square test is based on the total number of trials in this condition, even though the trials came from a smaller number of subjects and, so, the assumption of independence was violated. Thus, the result must be interpreted with caution.
verb phrase were relatively rare, except in
the critical case: NP-NP sentences into
which a nonalternating verb had been in-
truded, when a change was required to
avoid producing an ungrammatical sentence.

**PP versus NP complements.** The other
set of experimental sentences used PP com-
plement verbs and their transitive syn-
onyms. The intrusion results are shown in
the lower part of Table 1. When the sen-
tence had an NP complement (transitive)
verb like *ponder* and the lure verb was *think*
(which would be synonymous only with the
addition of a prepositional complement
"about . . ." ) there were only 6% intru-
sions of "think." But when the PP comple-
ment verb was in the sentence and the lure
was the NP complement verb, there were
21% intrusions. This difference was signif-
icient by a sign test, $n = 18, x = 3, p < .01,$
two-tailed. Note that in the former case the
lure is not a synonym for the sentence verb
unless the missing preposition is supplied—
*think* is not a synonym for transitive *pon-
der*—so it was not as obvious a potential
substitute as in the reverse case.

As with the dative sentences, when there
was an intrusion the syntactic structure of
the verb phrase was changed to restore
grammaticality to the recalled sentence, by
adding or deleting the preposition. This
happened 100% of the time with the PP and
NP complement sentences, and (not sur-
prisingly) there were no spontaneous addi-
tions or deletions of the preposition when
the original verb in the sentence was cor-
rectly recalled.

**Synonym controls.** The sentences with
lure verbs that were syntactically compati-
bile with the target verbs produced 14% in-
trusions.

**Discussion**

The main result of Experiment I was
that, in recall of a recently read sentence,
intrusion of a synonymous verb presented
on a distractor list was not prevented by an
incompatible surface structure in the to-be-
recalled sentence. When an intrusion oc-
curred in the incompatible structure, the
correct structure for the selected verb was
generated in recall. Note that since the sub-
ject was instructed to recall the sentence
verbatim, there was no incentive to change
the structure in this fashion. For the dative
nonalternator lures with alternator verbs in
the sentence, no significant difference was
found in the intrusion rate as a function of
the compatibility of the surface structure.
Still, intrusions were somewhat fewer in the
incompatible cases, suggesting a small
incompatibility effect.

PP complement verbs such as *think about*
were frequently replaced with transi-
tive synonyms such as *ponder*. Thus they
show the same behavior as the dative
verbs: If "ponder" is activated and can ex-
press the conceptual message of the sen-
tence it may be used in recall, despite
the fact that the sentence produced by the sub-
ject will not have the same surface syntax
as the original sentence. Having the transi-
tive synonym (e.g., *ponder*) in the sentence
and the PP complement verb (without the
preposition—e.g., *think*) on the list mark-
edly reduced the number of intrusions, be-
cause the verb alone does not readily activ-
ate its PP complement sense. (We con-
sider the PP complement verbs in more
detail in the general discussion.)

Both in dative sentences with the NP-NP
construction and in PP/NP complement
sentences, a word-for-word substitution is
not possible—a preposition must be deleted
or word order must change or both. But, as
expected under the hypothesis that recall is
generated from the conceptual message
with preferential selection of activated lex-
ical items, this does not form a significant
barrier to synonym substitution because
subjects are not representing word order or
prepositional phrases as such. Subjects are
not changing remembered word order, but
generating word order anew.

**Experiment 2**

Experiment 2 was carried out to replicate
Experiment 1 and to include a new dative
condition in which both the verb in the sentence and the lure verb were alternators; as before, NP-NP and NP-PP structures were counterbalanced. The purpose was to determine whether the propensity to intrude even a syntactically appropriate verb would be affected by the sentence structure and to assess the baseline tendency to change the syntax when a verb is intruded, even when such a change is not required to ensure grammaticality. In most other respects Experiment 2 was identical to Experiment 1.

Method

Subjects. There were 40 subjects in Experiment 2 from the same pool as the subjects in Experiment 1.

Materials and design. Minor changes in materials were made, including the substitution of provide for suggest (dative nonalternators); also show and tell replaced lend and take as dative control verbs.

Five of the six PP- and NP-complement pairs were used (the sixth was omitted by mistake) and also the 12 filler sentences without lures. Some filler items were changed or deleted. Otherwise all materials already described were retained.

Sixteen new dative sentences used eight pairs of verbs, both of which could take the NP-NP construction. Two sentences were written for each verb pair. A given subject had verb 1 in one sentence and verb 2 in the other sentence; the lures were verb 2 and verb 1, respectively. The two sentences were widely separated in the experiment. In each case half the sentences were in the NP-PP form and half in the NP-NP form. The verb pairs were: take/bring; send/mail; award/give; lend/loan; fix/make [a meal]; get/buy; reserve/save; bake/make.

Each subject saw 31 experimental sentences, 6 dative controls, and 15 filler sentences. Three there were 10 practice sentences.

3 Four new fillers were dative sentences with non-alternator verbs and non-alternator lures (e.g., allocate/distribute). Only the NP-PP form of the sentence was used. These sentences produced 23% intrusions. The materials were counterbalanced over four versions.

Procedure. The procedure was the same as that of Experiment 1, except that Group 1 (N = 20) saw the sentences at 200 ms/word (as in Experiment 1) and Group 2 (N = 20) at 150 ms/word. The list was presented at 183 ms/word. Subjects were debriefed at the end of the experiment and almost none of them realized that the dative construction was being tested.

Results and Discussion

As in Experiment 1, overall sentence recall was fairly accurate. Apart from the critical verbs, Group 1 recalled 90% of the words and Group 2 recalled 86%. Although Group 2 (who viewed the sentences at 150 ms/word) made more intrusions—17% overall—than the 10% made by Group 1 (who viewed the sentences at 200 ms/word), t (38) = 3.01, p < .01, two-tailed, the pattern was similar in the two groups and their results are reported together.

Dative verbs. The percentage of verb intrusions is shown in Table 2. For the set of dative sentences which were carried over from Experiment 1 the results were similar to those of Experiment 1. For the critical case of non-alternator lure and an alternator verb in the sentence, the lure was more likely to intrude when the sentence was in the form NP-PP (14%) than in the form NP-NP (8%), sign test (n = 16, x = 4), p < .05, one-tailed. This result is consistent with the hypothesis that lexical insertion is at least partially constrained by the presence of an incompatible sentence structure in the to-be-remembered sentence. However, the results for alternator verbs both in the sentence and as lures suggests another explanation of the different rates of intrusion. For these sentences more intrusions were also made when the sentence was in the NP-PP form (17%) than in the NP-NP form (13%)—although the difference was not significant.

An analysis of variance was carried out on the percentage of lure intrusions when
### Table 2

<table>
<thead>
<tr>
<th>Lure type</th>
<th>Sentence form presented</th>
<th>Intrusion</th>
<th>No intrusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% Resp.</td>
<td>% Change</td>
<td>% Resp.</td>
</tr>
<tr>
<td>Nonalternator</td>
<td>NP-PP</td>
<td>14</td>
<td>26</td>
</tr>
<tr>
<td>*Nonalternator</td>
<td>NP-NP</td>
<td>8</td>
<td>71</td>
</tr>
<tr>
<td>Alternator</td>
<td>NP-PP</td>
<td>17</td>
<td>25</td>
</tr>
<tr>
<td>Alternator</td>
<td>NP-NP</td>
<td>13</td>
<td>17</td>
</tr>
<tr>
<td>No-lure control</td>
<td>NP-PP</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>No-lure control</td>
<td>NP-NP</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

**Dative verbs**

- NP versus PP complements
  - *PP* NP 3 100* 97 0
  - *NP* PP 15 100 85 0

*Note.* The asterisk (*) indicates the conditions in which an intrusion of the lure would create an ungrammatical sentence unless other changes were made.

- The verb in the sentence was an alternator.
- Addition of the preposition.
- Deletion of the preposition.

the sentence was in the NP-PP versus NP-NP form and when the lure was an alternator versus a nonalternator, with group (rate of presentation) as a between-subjects variable. While there was a main effect of group, with the faster rate of presentation producing more intrusions, *F*(1,38) = 6.03, *p* < .02, there were no significant interactions of group with the other variables. Likewise, there was a marginal main effect of the type of verb, with alternators producing more intrusions than nonalternators, *F*(1,38) = 3.85, *p* < .06. Most important, NP-PP sentences were associated with more intrusions than NP-NP sentences, *F*(1,38) = 4.62, *p* < .04, but there was no interaction with type of verb: *F*(1,38) = 0.11. Thus, the higher rate of intrusions of nonalternators in NP-PP sentences than in NP-NP sentences does not indicate that the surface structure exerted a constraint on the intrusion of a nonalternator, but seems instead to indicate that NP-NP sentences are generally more resistant to intrusions that the same sentences in NP-PP form. The dative controls with no lure again showed no spontaneous intrusions of synonymous verbs.

**Changes in syntax.** The percentage of spontaneous changes in the syntax of the verb during recall is also shown in Table 2. When the sentence would have been grammatical without a syntactic change, changes were made on 7% to 26% of the trials. The one case in which a change was necessary to avoid an ungrammatical sentence was when a non-alternator intruded in place of an alternator in an NP-NP sentence. In the 14 instances of that kind of intrusion, 10 (71%) were recalled with a change in the syntax, and all of these resulted in a grammatically acceptable sentence. When NP-NP sentences were recalled without an intrusion, syntactic changes occurred on only 14% of the trials, *χ*²(1,180) = 24.9, *p* < .001, two tailed. (In the NP-PP case, also, intrusions were associated with an increase in syntactic changes (*p* < .05), but this tendency was not as marked as in the NP-NP condition, *χ*²(1,41) = 8.57, *p* < .01, two-tailed.)

**NP- and PP-complement verbs.** Sen-
tences with these verbs produced 3% and 15% intrusions, respectively, a difference significant by a sign test, n = 12, x = 2, p < .05, two-tailed. This pattern was like that in Experiment 1. When the PP-complement verb in a sentence was replaced by the NP-complement verb, the preposition was always deleted.

Conclusions. Experiment 2 confirmed the main conclusion from Experiment 1: spontaneous intrusion of a lure verb during recall is not prevented by a mismatch between the surface structure of the target sentence and that required by the intruding verb. The mismatch is resolved in most cases by making appropriate changes in the sentence during recall, so that the resulting sentence is grammatical, even though the exact wording is compromised. Not only the dative sentences, but also the sentences with PP- and NP-complement verbs produced intrusions and appropriate corrections to the surface structure, in this case the deletion or addition of a preposition. Both the occurrence of such syntax-violating verb intrusions and the changes then made during recall support the hypothesis that the surface syntax of the verb phrase is not directly represented in short-term memory for a sentence, but is regenerated once the verb has been selected. The lower rate of intrusions of nonalternators into the NP-NP form of the sentence appears to result from the fact that the NP-NP form is less susceptible to intrusions, although we cannot rule out the possibility than an incompatible surface structure exerts some constraint on intrusions. Finally, in both experiments the surface structure of the sentence was usually recalled accurately when it was compatible with the verb (intruded or correct). This result is consistent with syntactic priming, as noted earlier (see also the general discussion).

Experiment 3

In Experiment 3 a more direct method was used. Subjects were asked to judge whether a single probe verb could fit into the sentence they had just read. After making that decision the subject recalled the sentence, substituting the verb if it was judged to fit. If people have access to the surface structure of a sentence in short-term memory, they should have no difficulty in rejecting a probe verb that does not fit that structure. If they do not have direct access to the surface structure, then an error pattern similar to that observed in Experiments 1 and 2 should be found.

Method

Subjects. There were 16 subjects from the same pool as Experiments 1 and 2. Materials and design. As in Experiments 1 and 2, two groups of verbs were studied: datives and PP versus NP complement verbs. There were eight dative verb sets, each a triple consisting of two alternators and one nonalternator. For one of the verbs in the triple, an alternator, two sentences with different content were written such that either of the two remaining probe verbs (one an alternator and one not) provided an appropriate paraphrase of the sentence verb in each of the sentences. Thus, for example, the two basic sentences for the send: mail/transmit triple were (4) and (5), each shown in the NP-PP and NP-NP forms:

(4a) The agent will send his report to the government when his mission ends. MAIL or TRANSMIT

(4b) The agent will send the government his report when his mission ends. MAIL or TRANSMIT

(5a) The ambassador's staff will send our message to the embassy before his visit. MAIL or TRANSMIT

(5b) The ambassador's staff will send the embassy our message before his visit. MAIL or TRANSMIT

Each of the two basic sentences appeared in every version of the materials, but the form of the sentence (NP-PP or NP-NP)
and the probe word (alternator or nonalternator) were counterbalanced over versions. Of the four counterbalanced conditions, only one was such that a replacement by the probe verb would result in an ungrammatical sentence: when the sentence was in form (b) (NP-NP) and the probe word was a nonalternator ("TRANSMIT"). The verbs were: promise: give/donate; send: mail/transmit; offer: send/submit; get: buy/purchase; make: build/construct; read: teach/explain; task: bring/carry; read: tell/relocate. Subjects saw a total of 16 dative sentences, so there were four trials of each sentence-probe word type per subject.

The second group of verbs used in Experiment 3 were PP/NP complement verbs like those in Experiments 1 and 2—for example, think about ponder. There were two versions of these sentences, one with the PP complement verb and one with the NP complement verb; the other verb served as the probe word. The preposition was not included when the PP complement verb was the probe word.

Ten filler sentences and probes were intended to be very easy "no" trials: both the syntax and the meaning of the probe word were highly incompatible with the sentence. Ten other fillers used sentence and probe verbs with similar meaning but incompatible syntax: for example, advise/recommend.

Each subject saw 22 experimental sentences, 20 filler sentences, and 10 practice sentences.

Procedure. Each trial began with a row of three asterisks presented for 300 ms, a 200-ms blank, and the words of the sentence for 183 ms each. After a blank of 317 ms, the capitalized probe word appeared for 333 ms. Subjects were instructed as follows: "When you see the test word, you should decide as quickly as possible whether that word could replace a word in the sentence without having to change the rest of the sentence. The meaning may change somewhat but the word order and the rest of the words in the sentence should stay the same." They were instructed to respond by pressing a right-hand key for yes or a left-hand key for no; reaction time (RT) was measured. The instructions continued, "If you said yes, repeat the sentence as it would be with the test word inserted; if you said no, repeat the original sentence."

Results and Discussion

Overall, recall of the sentence was again good; of all words other than the critical verb, subjects recalled 89%. A standard procedure for truncating unusually long response times (RTs) was followed: RTs longer than 5 s were first truncated to 5 s (in Experiment 3 there were none); then the mean and standard deviation of each subject’s RTs were calculated, separately for yes and no responses, and RTs longer than the mean plus 2 SD were truncated to that number (5% of the trials). The percentage of correct and incorrect responses, the mean RTs, and the percentage of grammatical changes in syntax are shown in Table 3.

Datives. Consider first the case in which both the sentence verb and the probe verb were alternators, and therefore the probe word (MAIL, in the examples above) was acceptable whether the sentence structure was NP-PP or NP-NP. Subjects correctly accepted the probe word 88% of the time, with an overall mean RT of 1447 ms. and there was no significant difference between the two sentence structures in either errors or RTs. On the 12% of trials when subjects said "no" incorrectly, RTs were long.

Now consider the critical case, when the probe word was a nonalternator (TRANSMIT, in the examples above). When the sentence structure was NP-PP (as in 4a and 5a), subjects again had no difficulty accepting the probe verb (91% correct, mean RT 1379 ms). However, when the sentence structure was NP-NP (as in 4b and 5b) (and the correct answer to the probe word was therefore "no"), subjects correctly rejected the probe word on only 33% of the trials, with a mean RT of 1829 ms. Both
TABLE 3
EXPERIMENT 3: PERCENTAGE OF CORRECT AND INCORRECT RESPONSES, MEAN RT, AND PERCENTAGE OF
GRAMMATICAL CHANGES IN SYNTAX

<table>
<thead>
<tr>
<th>Probe type</th>
<th>Sentence form</th>
<th>Correct</th>
<th>Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>% Resp.</td>
<td>RT</td>
</tr>
<tr>
<td>Alt.</td>
<td>NP-PP</td>
<td>(Yes) 88</td>
<td>1460</td>
</tr>
<tr>
<td>Alt.</td>
<td>NP-NP</td>
<td>(Yes) 88</td>
<td>1434</td>
</tr>
<tr>
<td>Nonalt.</td>
<td>NP-PP</td>
<td>(Yes) 91</td>
<td>1379</td>
</tr>
<tr>
<td>*Nonalt.</td>
<td>NP-NP</td>
<td>(No) 33</td>
<td>1829</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Probe type</th>
<th>Sentence form</th>
<th>Correct</th>
<th>Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*PP</td>
<td>NP</td>
<td>(No) 71</td>
<td>1870</td>
</tr>
<tr>
<td>*NP</td>
<td>PP</td>
<td>(No) 23</td>
<td>1856</td>
</tr>
</tbody>
</table>

**Dative verbs**

**NP versus PP complements**

<table>
<thead>
<tr>
<th>Probe type</th>
<th>Sentence form</th>
<th>Correct</th>
<th>Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>% Resp.</td>
<td>RT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. The asterisk (*) indicates the conditions in which inserting the probe verb would create an ungrammatical sentence unless other changes were made.

*The verb in the sentence was an alternator.

Addition of the preposition.

Deletion of the preposition.

Accuracy level and reaction time were significantly different in the two conditions: for accuracy, $F(1,15) = 66.0, p < .001$, and for RTs, $F(1,10) = 5.03, p < .05$ (five subjects with one or more empty cells were omitted from the RT analysis).

A comparison of "yes" responses to nonalternator probes in the two sentence conditions indicated a significant difference in the number of "yes" responses (91% in the NP-PP case and 67% in the NP-NP case), $F(1,15) = 6.04, p < .03$, and a marginally significant difference in RTs, $F(1,10) = 4.59, p < .06$, suggesting that subjects were at least somewhat sensitive to the unacceptability of the nonalternator verb in the NP-NP construction.

**Syntactic changes in datives.** The percentage of changes in the syntax of the verb phrase is also shown in Table 3. Changes ranged from 0% to 34% except in the one critical case, when subjects did mistakenly say yes to the nonalternator verb when the sentence was in the NP-NP form. Then, as they recalled the sentence with the test verb replacing the original verb, 67% of the time a change was made to restore grammaticality. But when subjects gave the correct answer, "no," to the same sentence-probe combination, only 19% of recalls included a change in syntax, $\chi^2(1, N = 64) = 11.36, p < .001$, two-tailed.

**NP versus PP complement verbs.** The correct answer to these sentences was "no" in both cases, because a preposition would have to be added or deleted to make the new sentence grammatical. The lower part of Table 3 shows the main results. When the sentence used the NP verb, the PP verb was correctly rejected in 71% of trials. When the sentence used the PP verb, subjects were much less likely to correctly reject the NP verb: only 23% of the trials were correct noes. The difference in accuracy between the conditions was significant, $F(1,15) = 24.9, p < .001$. Thus, as in Experiments 1 and 2, there were more erroneous substitutions when a preposition would have to be deleted (think [about], ponder) than when it had to be supplied. A comparison of the false "yes" RTs in the two conditions showed a marginally significant difference, $t (6) = 2.37, p < .06$, two-tailed: the condition that produced more false yeses also produced faster responses.

Almost all mistaken inclusions of the
probe verb did lead to addition or deletion of the preposition to restore grammaticality, whereas the preposition was never added or deleted when the original verb was correctly retained.

**Filler sentences.** In the 10 filler trials where both syntax and semantics were inappropriate there were only 3% false yes answers, and in no case did the subject actually try to insert the verb, so subjects clearly can tell an impossible fit when they see it. On the other hand, when the probe verb was conceptually appropriate but syntactically inappropriate (e.g., *advisesel recommend, or intransitive eatconsume*), there were 56% incorrect yeses, again with grammaticality-preserving changes in recall of the sentence in most cases.

**Summary and Conclusions: Experiment 3**

For datives, when the meaning of a substitute verb was appropriate subjects had difficulty rejecting it, even if using the verb in that syntactic structure would result in an ungrammatical sentence. When incorrectly inserting such verbs they usually changed the syntax of the recalled sentence to make it grammatical. Subjects were not entirely blind to the surface syntax, however, because their propensity to respond *Yes* and their speed in doing so tended to be greater when "yes" was the correct answer.

For the NP- versus PP-complement verbs (e.g., *ponder/think [about]*), verb-preposition pairs such as "think about" apparently function as a conceptual unit, so that "ponder" seemed to be an appropriate substitute; however, "think" as a probe apparently did not readily activate the conceptual unit "think about" and so subjects were less likely to make the mistake of accepting it as a replacement for "ponder."

The pattern of results in Experiment 3’s more explicit task was similar to that in Experiments 1 and 2’s sentence recall task with an implicit lure: subjects make numerous mistaken verb insertions when the surface structure of the sentence is inconsistent with the verb, although they are able to make the correct discrimination some of the time. Thus both the present and the earlier experiments provide support for the regeneration hypothesis. Subjects were explicitly instructed that they should be able to make the insertion without making any other changes in the sentence. But they appear to make the decision on the basis of similarity of meaning and the potential to support the arguments of the original verb, while having difficulty noting an incompatibility with the surface syntax: for example, they accept nonalternators when the sentence is in the NP-NP form. This is evidence that the syntax of a sentence in short-term memory is not readily available for reflection.

Then, when the subject recalls the sentence with the verb inserted, s/he repeats it with the syntax appropriate to that verb, although this results in a change from the original form of the sentence. The subject usually does not produce the ungrammatical sentence which would result from simply removing the original verb and inserting the new verb into that slot, as we would expect if a verbatim representation of the sentence were available. Rather, the subject produces a grammatical utterance using the verb and other elements of the sentence. This is what we expect if the subject uses his usual grammatical mechanisms to regenerate the sentence.

**Experiment 4**

The three experiments described so far support the hypothesis that the surface syntactic structure of a sentence is not generally part of its short-term memory representation, even when that memory representation is complete enough for near-verbatim recall. Were such a syntactic representation available, it should consistently prevent the intrusions and insertions of syntactically incompatible verbs that we observed. The frequency of these errors supports the claim of the regeneration hypothesis that short-term memory for sen-
sentences is essentially a conceptual store. In Experiment 4 we investigated properties of the conceptual representation itself—what internal structure it might have.

The method was the same as that used in Experiment 3, except that in this experiment we used sets of dative verbs that were non-synonymous. All dative verbs by definition take the same abstract argument structure: the noun phrases associated with them must fulfill the thematic roles of Agent, Theme, and Goal, and at an abstract level the verb expresses a relationship of transfer of possession (Gruber, 1965; Jackendoff, 1972). The alternators allow two different syntactic structures that express these arguments, and a nonalternator like donate only allows the first syntactic structure to express its arguments:

The author gave a letter to the library
The author wrote a letter to the library
The author donated a letter to the library
AGENT THEME GOAL

The author gave the library a letter
The author wrote the library a letter
AGENT GOAL THEME

Although write and give are not synonyms, they share important elements of meaning at the level of argument structure: the abstract relationships among the noun phrases are the same. Our hypothesis was that the conceptual representation of the sentence includes this abstract level at which all dative verbs have similar properties. This level is important in structuring the meaning of the sentence, and so it seems likely that it would be important in the structure that represents the meaning of the sentence in memory.

If subjects have encoded the relationships among the noun arguments in this way, it should be possible for them to replace the verb with a non-synonymous verb that allows the same relationships; they would be replacing only one element in the conceptual representation without making other changes—in particular, no changes in the structure of the representation need to be made. In addition, we predict that the results should show the same pattern as in the previous experiments—that surface syntactic structure should not prevent the insertion, since the task engages a representational level at which differences in surface syntax are not represented.

Method

Subjects. There were 16 subjects from the same pool as the previous experiments.

Materials and design. Dative sentences were constructed in the same manner as in Experiment 3 except that the verb sets were chosen so that the probe verb was not synonymous with the verb in the sentence. All the critical verbs made equal sense in the sentence context, apart from the difference in syntax, as in the following example.

(6a) The famous author will write a letter to the library. GIVE or DONATE
(6b) The famous author will write the library a letter. GIVE or *DONATE

The sets were: write: donate/give; sell: relate/tell; make: purchase/buy; give: carry/bring; buy: construct/build; read: submit/send; lend: transmit/mail; assign: explain/teach. The probe words in each set (for example, donate/give) were approximately synonymous in an effort to make both of them about equally semantically “distant” from the word in the sentence. Two different sentences were composed using each verb. The forms of the sentences (NP-PP or NP-NP) and the two probe words (alternator or nonalternator) were counterbalanced within and between subjects; each subject had four trials in each of the four conditions.

A mistake in the counterbalancing reduced the n in one condition from 64 to 60 trials.
TABLE 4
Expenrrrve 4: PenceNrece on Connecr eNo INconnecr ResroNses, Meex RT, eNo Psncpxrece or Gnruurrrcr.L CtrlNcrs rN Synrex

<table>
<thead>
<tr>
<th>Probe type</th>
<th>Sentence form</th>
<th>Correct</th>
<th>Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>% Resp.</td>
<td>RT</td>
</tr>
<tr>
<td>Alt. NP-PP</td>
<td>(Yes) 92</td>
<td>1133</td>
<td>1</td>
</tr>
<tr>
<td>Alt. NP-NP</td>
<td>(Yes) 94</td>
<td>1156</td>
<td>1</td>
</tr>
<tr>
<td>Nonalt. NP-PP</td>
<td>(Yes) 92</td>
<td>1221</td>
<td>0</td>
</tr>
<tr>
<td>*Nonalt. NP-NP</td>
<td>(No) 38</td>
<td>1800</td>
<td>17</td>
</tr>
<tr>
<td>*PP NP</td>
<td>(Yes) 33</td>
<td>1635</td>
<td>0</td>
</tr>
<tr>
<td>PP NP</td>
<td>(Yes) 73</td>
<td>1267</td>
<td>0</td>
</tr>
<tr>
<td>*NP PP</td>
<td>(No) 32</td>
<td>1526</td>
<td>0</td>
</tr>
<tr>
<td>NP NP</td>
<td>(Yes) 90</td>
<td>1189</td>
<td>0</td>
</tr>
</tbody>
</table>

**Note.** The asterisk (*) indicates the conditions in which inserting the probe verb would produce an ungrammatical sentence unless other changes were made.

*The verb in the sentence was an alternator.

*For the remaining 43% of the wrong yeses (16 trials) the recalled sentence was ungrammatical by our criterion, but perhaps not by the subject's criterion; see text.

*For 73% of the cases (11 trials) recall was ungrammatical by our criterion.

A set of NP- versus PP-complement sentences was also constructed using nonsynonymous probe words. The test sentence had two versions which were similar in meaning, using either an NP or a PP verb, and there were two probe words, one an NP verb and one a PP verb. For example, the two sentences in one pair used the verbs *examine* and *look* + [at]; the probe words were *arrive* [at] and *visit*. Note that *arrive* could replace *look* in (7b) without changing the preposition *at*.

(7a) The historians will examine the site today. *VISIT* or *ARRIVE*

(7b) The historians will look at the site today. *ARRIVE* or *VISIT*

(8a) Some tourists may examine the king’s tomb in the afternoon. *VISIT* or *ARRIVE*

(8b) Some tourists may look at the king’s tomb in the afternoon. *ARRIVE* or *VISIT*

Sets of four verbs were used. The two sentence verbs, one with a preposition, are listed first, followed by the two probe verbs. One of the probe verbs required a preposition in the context of the sentence, as shown in brackets. The sets were: *ponder/think about: explain/read [about]; request/ask for: attend/register [for]; discuss/talk about: read/wonder [about]; examine/look at: visit/arrive [at]; protest/object to: rewrite/add [to]; demand/insist on: approve/vote [on]. Two sentences were written for each of six verb sets, so that each subject saw 12 sentences of this type; the sentence verbs and probe verbs were counterbalanced between subjects.6

The 10 "no" fillers were the "easy no" fillers used in Experiment 3.

**Procedure.** The procedure was the same as that of Experiment 3.

**Results and Discussion**

Overall recall accuracy was high: 95% of the words (other than the critical verbs) were recalled. This relatively high level of accuracy may reflect the fact that sentences in Experiment 4 were 2.5 words shorter on

6 A mistake in the counterbalancing reduced the n in one condition from 48 to 40.
average than those in Experiment 3. Before analysis, RTs were truncated as in Experiment 3: 5% of the RTs were truncated to that subject’s mean plus 2 SD.

**Datives.** The results are shown in Table 4. Looking first at the probe words that were alternators, responses were correct yeses an average of 93% of the time, regardless of whether the sentence form was NP-PP or NP-NP. RTs in the two cases were also similar, F < 1.0. When the probe was a nonalternator the subjects made 92% correct yeses to the NP-PP sentences; however, when the sentence form was NP-NP, subjects had difficulty (as in Experiment 3) rejecting the verb: only 38% of the responses were correct rejections, F(1,15) = 40.8, p < .001. The RTs in the latter case were also longer, F(1,11) = 28.3, p < .001. Surprisingly, the RT to say “yes” incorrectly to nonalternator verbs with NP-NP sentences was not significantly longer than to say “yes” correctly to an alternator verb. When the nonalternator was incorrectly incorporated into an NP-NP sentence, 57% of the sentences were altered so that they became grammatical, compared with only 17% spontaneous syntactic changes when subjects correctly rejected the nonalternator, χ²(1, N = 60) = 7.50, p < .01, two-tailed.

The pattern of responses was almost identical to that of Experiment 3, although the probe verbs in Experiment 4 did not have as close a semantic relationship to the verbs in the sentence. In both experiments subjects had trouble rejecting nonalternator probes as substitutes for dative verbs in double-object sentences, and when they did mistakenly accept such verbs they then frequently changed the syntax of the sentence during recall.

Even so, on a substantial percentage of trials when the subject accepted a nonalternator into an NP-NP sentence the syntax was not modified (33% of such cases in Experiment 3 and 43% in Experiment 4). On these trials the subject produced an ungrammatical sentence such as “Jane wants to purchase the children a birthday cake.” We would argue that the pattern of recall is best explained by variation in the acceptability of the NP-NP structure from verb to verb and subject to subject. The subjects in Experiment 3 were asked to fill out a questionnaire after the experiment that asked for grammaticality judgments, including NP-NP structures with some of the experimental verbs. Verbs that we considered alternators were mostly accepted as alternators by the subjects, for example, 93% for buy and 96% for write. But some of the nonalternators were accepted in the NP-NP construction by some subjects: for example, 32% accepted transmit; 14%, relate; and 75%, purchase.

Given these judgments, it seems likely that many or all of the “mistaken” acceptance of nonalternators and the consequent production of “ungrammatical” sentences were cases in which the recalled sentence was grammatical for the subject. This also means that some answers we counted as incorrect “yes” answers may have been correct given that subject’s idiolect. If we set aside all trials in which the subject inserted the verb and made no syntactic changes (on the conservative assumption that subjects made no changes because the utterances were grammatically acceptable to them), the percentage of correct rejections of nonalternator probes, given an NP-NP sentence, rises from 33% to 42% in Experiment 3 and from 38% to 52% in Experiment 4. Note that the corrected percentages are still markedly below both the correct “yes” responses to the same probes with the NP-PP sentence form, and below the high rate of correct rejections of the “easy no” filler sentences (97%, in Experiment 4). Thus, the overall results for the datives support the regeneration hypothesis, with the caveat that on some tri-

---

7 Bock (personal communication, November 7, 1990) points out that the apparent acceptability of “ungrammatical” sentences may have been increased by priming from grammatical sentences in the experiment which used the structure.
could insert a dative verb with a different meaning, *donate*. Despite the fact that it resulted in a change in both meaning and surface syntax, subjects made the same kinds of errors as in Experiment 3, responding with *donate a letter to the library*.

These results suggest that the subjects depend for their decision on a representation which encodes the abstract conceptual relationships among the items—the level at which *write* and *donate* have similar properties. They cannot be carrying out the task by using a representation of the surface syntactic structure, since in that case they would either reject the verb or produce the ungrammatical *donate the library a letter*. But also they cannot be doing the task by judging whether the resulting sentence would have the same meaning as the original, since it does not. Rather, they must be using an abstract representation, like the argument structure of the verb—a level of structure at which different dative verbs can be substituted without making any other changes in the representation at that level.

There must, of course, also be a level at which subjects encode the more subtle details of the meanings of words and sentences; we do know the difference in meaning between writing and donating. But the results of Experiment 4 indicate that the conceptual representation must also contain a more abstract structural level that can be tapped relatively easily.8

* Although we refer to argument structure as an aspect of conceptual representation, this is not crucial to the argument. Some linguists have proposed that argument structure is a level of syntactic representation and lexical conceptual structure is the related level of semantic representation (see, for example, Grimshaw, 1990, and references therein). Our data cannot address the subtleties of any of these theories. They provide support for, and require, the existence of some level of representation separate from phrase structure and separate from the level of meaning which determines synonymy, which encodes more abstract properties of the relationship between a verb and its noun arguments. However, the data do not allow us to make any more specific claims about the properties of this level.

**General Discussion**

The present experiments explored the representation (or nonrepresentation) of surface syntax in short-term memory. In an earlier paper we proposed that the predominant representation used in immediate recall is a conceptual representation (Potter & Lombardi, 1990). The recaller has in mind the meaning of the sentence, which is then expressed using the normal processes of sentence production. An implication of this hypothesis is that the surface syntax of the sentence is not retained in short-term memory, since surface syntax is not part of the conceptual message of the sentence. The present experiments tested the hypothesis that syntax is regenerated rather than remembered by using materials that allowed two possible syntactic structures to express the same message. For some verbs (e.g., dative alternators) both structures were possible, but for others (nonalternators) only one of the two structures was grammatical.

Using the synonym-activation method of Potter and Lombardi (1990), Experiments 1 and 2 showed that an incompatibility with the surface syntax did not prevent intrusions of the lure verb, and when such intrusions occurred the syntax was usually changed to conform with the requirements of the intruded verb. In Experiments 3 and 4 subjects judged whether a test verb could fit into the sentence with no other changes. Thus, we were explicitly asking the subjects to remember the sentence in full detail. Even so, subjects had difficulty rejecting a verb which differed only in surface syntactic possibilities. This is evidence that, as predicted by the regeneration hypothesis, there is little or no direct representation of surface syntactic structure in short-term memory. Conceptual similarities are what constrain intrusions and insertions, not similarities in syntactic subcategorization. The sense of conceptual similarity which is relevant is a fairly abstract one, since subjects made no more errors of com-
mission when the dative verbs were close synonyms (Experiment 3) than they did when the verbs shared only the same argument structure (Experiment 4). (In both experiments the probe verbs in question fit with the pragmatics of the arguments in the target sentence.) Unlike argument structure, whether a dative verb has the NP-NP subcategorization in addition to the NP-PP is a fact which is not related to the meaning of the verb, and this purely syntactic feature does not routinely block intrusions or mistaken insertions of a verb.

**Syntactic priming and the accuracy of recall.** When subjects had not intruded or inserted an incompatible verb they usually did recall the sentence with the same syntax as the original, even when there was an alternative surface structure available. In Experiment 3, for example, when the probe verb for a dative sentence was an alternative, either the NP-PP construction or the NP-NP construction would have been grammatical, yet only 13% of recalls incorporating the probe verb switched to the other form. (This percentage is based on 96 cases in which both arguments of the verb were recalled, omitting 16 trials on which one argument was dropped.) But we argue that this is an implicit priming effect, not the result of explicit memory for the surface structure. Previous studies of memory for sentence syntax have shown that syntactic priming from an immediately preceding sentence can influence the syntactic structure of a novel sentence produced by the subject even when the conceptual and lexical content of the produced sentence is unrelated to the priming sentence (e.g., Bock, 1986; Levelt & Kelter, 1982). Such priming occurs for similar surface structures even when there are substantial conceptual differences in the thematic roles of the constituents in the priming sentence and the produced sentence—for example, a locative to phrase can prime a prepositional dative (Bock & Loebell, 1990). The preposition used in the priming sentence is not critical (to and for prepositional phrases may prime each other, Bock, 1989), but to as an infinitival does not prime the prepositional dative. This shows that the priming phenomenon is truly an effect of syntactic structure and not, say, of priming of closed-class lexical items or rhythmic structure.

Bock and Loebell (1990) argue that what is primed are the mechanisms that create surface constituent structures, not a memory trace of a syntactic tree. If subjects were reusing the exact same structure stripped of its terminal elements, the adding or subtracting of adjectives, determiners, and the like would be expected to influence the amount of priming, but they do not. Bock and Loebell’s hypothesis concurs with the present hypothesis that the surface syntax is regenerated in recall. Their finding allows us to explain the generally accurate repetition of the original syntax even though it is not represented explicitly in the short-term memory trace of that sentence: in regenerating the sentence from a conceptual representation, subjects are likely to use the syntactic structures primed by their reading of the original sentence, as long as these structures are compatible with the verb. When they are not, however—as when the intruded or substituted verb is incompatible with the surface structure of the original sentence—the regenerative process usually overrides the priming effect to produce changes consistent with the new verb.

**PP and NP Complement Verbs**

The results with prepositional complement verbs support the findings from dative verbs. Subjects both intrude synonymous NP complement verbs into PP complement sentences (Experiments 1 and 2) and are often willing to replace a PP complement verb with an NP complement verb in Experiments 3 and 4. They then almost always spontaneously delete the preposition along...
with the verb. For example, “ponder” replaces the synonymous “think about” despite the fact that the surface trees required are different. This suggests that the verb-preposition combination stands in the same conceptual relationship to the NP object of the preposition as the transitive verb to its direct object—that is, that the argument structures are roughly as in (9):

(9) John thought about the question
    John pondered the question

In addition we see that the verb alone (e.g., think) does not point to the verb-preposition combination (think about) with a great deal of success, even when a synonym such as ponder is available (Experiment 3) and even less so when it is not (Experiment 4). This implies that the verb-preposition combination acts as an independent unit in the mental lexicon, not as a subcategorization of the verb “think,” so that providing the probe verb “think” is rarely sufficient to activate the entry for think about.10

Alternative Hypotheses

The regeneration hypothesis has the following properties. Short-term memory consists of a conceptual representation of the sentence. Surface syntax, then, is not represented. Recently encountered words (lexical items) remain activated, and the mechanisms for producing recently encountered syntactic structures are primed.

In recalling the sentence the subject attempts to express the conceptual message, choosing lexical items and generating a syntactic structure to do so. Activated lexical items are used if appropriate. If two lexical items that express the same message are activated—as in Experiments 1 and 2 where a lure verb was included in the distractor task—the subject may select either one. (In the case of Experiments 3 and 4 the process of recall is the same, except that the subject is deliberately choosing to substitute the probe for the original verb.) The syntactic subcategorization possibilities of the chosen verb constrain the syntactic structure of the sentence produced in recall. When two or more structures are compatible with the conceptual representation and the verb, a recently activated structure is likely to be used. Thus, if there is no intrusion or an intrusion of a verb with the same subcategorization possibilities, recall of the rest of the sentence will usually be verbatim, since the syntactic structure primed by the original sentence will be used.

However, if the subject chooses a verb which is incompatible with the primed syntactic structure of the original sentence, a structure appropriate to that verb will be used. This is expected under the regeneration hypothesis since the sentence is generated using the normal mechanisms of language production.

The alternatives to the regeneration hypothesis are two. The first is that short-term memory is a verbatim representation of a string of words. There is surely some memory of this type, since we can remember unstructured lists of random items. However, this type of memory cannot account for results such as those presented here, since subjects cannot remember lists of the length of the present sentences. Memory for sentences, then, must take advantage of structure.

The second alternative is simply that there is a representation of surface syntactic structure in short-term memory, but it is easily forgotten. The errors that our subjects make occur when they forget the original structure and so they do not realize that the verb they are considering is incompatible with it. This hypothesis, however, can-

10 A reviewer questions whether the contrast between NP complement and PP complement verbs is an appropriate case to test our hypothesis, because the latter may involve a reanalyzed structure [think about], NP. However, it seems likely that they nevertheless have at least an initial phrase structure that is different from simple NP complement verbs. For instance, our verb-preposition pairs can be distinguished from verb-particle constructions, which behave syntactically more like single constituents.
not account for the distribution of syntactic changes: they occurred routinely when the chosen verb was incompatible with the original syntax and rarely otherwise. If the changes occur because of forgetting, there should be just as many changes for a given structure in conditions in which the syntax was compatible as when it was incompati-
ble, contrary to what we observed.

Experiments 3 and 4 deserve some special attention in this discussion. In this task, subjects are not making the intrusion inadvertentiy, but are asked to judge whether the insertion is possible and if so, to make it. In this task it proved to be quite difficult for the subject to reject a probe verb on the basis of surface syntax, and when the verb was mistakenly incorporated, the syntax was usually changed, although it was unlikely to be changed otherwise. A possibility suggested by a reviewer is that the subject does have a memory store of the surface syntax which is used in recall but does not use it in the judgment task. If this is the case, it seems likely that when they attempt to repeat the sentence with the wrong verb inserted they should realize their error and make some comment that it was actually impossible to repeat the sentence that way. But this rarely happened.

In summary, although subjects do forget parts of sentences in some trials, forgetting cannot account for the systematic pattern of insensitivity to surface syntax that we have observed in the present experiments. But why, if the regeneration hypothesis is correct, are subjects ever able to detect the incompatibility between a given probe verb and the surface syntax of the sentence? Although subjects did say yes mistakenly at a high rate in the cases we have discussed, the percentage of incorrect yes answers was always lower than that of correct yes answers. For datives in Experiments 3 and 4, correct yes answers ranged from 88-94%, whereas the critical incorrect yeses were 67% and 62%, respectively. For the NP/PP complement verbs in Experiment 4, correct yes answers were 73% and 90% of responses in the PP and NP sentences, whereas incorrect yeses (probed by the wrong type of verb) were 68% and 33%, respectively. Thus, except for NP complement probes of PP complement verbs, some sensitivity to the incompatibility was always evident. This may mean that on some trials subjects were able to retain enough of the surface string to detect the incompatibility as they covertly regenerationed the sentence with the new verb. Nevertheless, the high rate of incorrect accep-
tances in the critical conditions shows that such retention was not the usual case, and the regeneration hypothesis is required to explain the pattern of results observed.

Conclusions

The results of these experiments show that the surface syntax of a sentence that is in "verbatim" short-term memory is not directly represented as a string of words, an acoustic sequence, a phrase structure, or in any other form in which the sequence of words is completely specified. If so, we would not expect to observe spontaneous intrusions of recently activated verb synonyms in recall, particularly when the syntax of the intruder was incompatible with the surface syntax (Experiments 1 and 2). Neither would we expect the many erroneous judgments of whether a probe verb could successfully replace the verb in a test sentence when the surface syntax (but not the semantic structure) was incompatible with the probe (Experiments 3 and 4). These errors were systematic, occurring under conditions in which recall and judgments were generally very accurate.

The restricted pattern of inaccuracies supports the regeneration hypothesis of Potter and Lombardi (1990): Subjects recall a sentence by expressing its message using recently activated words. They generate the appropriate syntax from the verb, using recently activated syntactic structures when they are compatible with the verb and the conceptual representation. The results suggest that there is little or no representa-
tion of the original surface syntax in short-term memory for the sentence, because subcategorization differences which involve only the surface syntactic structure of the sentence (such as the difference between dative alternators and nonalternators) have only small effects on spontaneous intrusions or substitution judgments. What constraining effects there are may mean that surface structure is represented on some occasions, or that other forms of representation such as an unstructured string may be used to detect incompatibility in some instances. Finally, the conceptual structure we hypothesize includes, in addition to the exact meaning of the sentence, a more abstract level of representation of the verb and its noun arguments.

REFERENCES


(Received August 13, 1991)

(Revision received February 20, 1992)