SYNTACTIC CONSTRUCTIONS AS RETRIEVAL PLANS

By GERARD KEMPEN

Experimental Psychology Unit, University of Nijmegen, Nijmegen, The Netherlands

Four probe latency experiments show that the 'constituent boundary effect' (transitions between constituents are more difficult than within constituents) is a retrieval and not a storage phenomenon. The experimental logic used is called paraphrastic reproduction: after verbatim memorization of some sentences, subjects were instructed to reproduce them both in their original wording and in the form of sentences that, whilst preserving the original meaning, embodied different syntactic constructions.

Syntactic constructions are defined as pairs which consist of a pattern of conceptual information and a syntactic scheme, i.e. a sequence of syntactic word categories and function words. For example, the sequence noun + finite intransitive main verb ('John runs') expresses a conceptual actor–action relationship. It is proposed that for each overlearned and simple syntactic construction there exists a retrieval plan which does the following. It searches through the long-term memory information that has been designated as the conceptual content of the utterance(s) to be produced, looking for a token of its conceptual pattern. The retrieved information is then cast into the format of its syntactic scheme. The organization of such plans is held responsible for the constituent boundary effect.

It is generally assumed that long-term memory representations of sentences or text are conceptual in nature unless the text has been memorized by subjects who anticipated a verbatim recall test. Then, in addition, the literal wording of the sentences is stored. Indeed, several phenomena of sentence recall are easy to explain by appealing to some sort of verbatim storage. One of these is the 'constituent boundary effect': word-to-word transitions *between* syntactic constituents (of a certain level) are more difficult than such transitions *within* constituents (of the same level). A convenient after-learning measure of transitional difficulty is probe reaction time; subjects are presented with one word (probe) of a memorized sentence and respond as fast as possible by saying the next word of the same sentence (Suci *et al.* 1967; Kennedy & Wilkes, 1968).

The constituent boundary effect – this term is intended to be descriptive, not interpretative – has invariably been explained in terms of storage, i.e. by postulating a type of memory storage that favours the combined access to both members of certain word pairs (e.g. two words belonging to the same syntactic constituent) but not of other pairs (e.g. two words from different constituents). A well-known version of this principle includes the notion of hierarchical chunking or grouping (Johnson, 1965, 1970; Wilkes & Kennedy, 1969; see also next section). However, the constituent boundary effect and, generalizing to non-sentence sequences of words or letters, grouping effects do not preclude alternative interpretations in terms of retrieval processes. On this type of account, transitional difficulties are attributed to the organization of a retrieval plan which at certain stages of its execution engages in more elaborate processing activity for locating and outputting the next word of a learned sequence, than at other stages. Johnson (1970), writing in the context of experiments on learning grouped letter strings, remarks that these alternative explanations are empirically indistinguishable because the usual experimental logic for assessing the relative importance of storage and retrieval does not apply.

G. KEMPEN

This logic, if applied to profiles of transition difficulties, would consist of having the subjects retrieve the memorized sequence by means of a retrieval plan that presumably imposes a different profile on to the sequence than the profile to be expected from hypothetical groups or chunks in the memory store. The extent to which the empirical profile obtained under such a retrieval plan instruction conforms to the profile expected from the organization of this plan, determines the plausibility of the retrieval interpretation. In the case of non-sentence strings this method is obviously inapplicable due to the absence of alternative retrieval plans.

However, this limitation does not necessarily exist for sentences. The present paper presents evidence which, on one hand, disconfirms storage interpretations of the constituent boundary effect and, on the other hand, supports a specific retrieval interpretation. The experiments to be reported indicate that, closely related to syntactic constructions, there exist retrieval plans which are capable of searching through conceptual representations of sentences and casting the retrieved information into a specific output format for reproduction purposes. The existence of these syntactic retrieval plans is relevant in two respects. First, they embody a principle of memory search and retrieval that has not yet been described. Secondly, they allow a realistic view on the memory processes during verbatim learning and recall of sentences and text. In current explanations of the constituent boundary effect, constituents are treated on a par with item groups in non-sentence strings. Apart from leaving open the question why, in the case of sentences, subjects often adopt groupings which correspond to syntactic constituents, such explanations fail to recognize the possibility that subjects derive their responses from a memory representation of the sentences' meanings. The to-be-memorized sentences typically are easy to understand and it seems realistic to allow for storage of conceptual representations and for their use during verbatim recall tasks.

SYNTACTIC RETRIEVAL PLANS

Suppose the subjects have committed to memory a few sentences of the type (1) *Two Finns wrote texts* or (2) *Three Greeks read books*. A storage theory of the constituent boundary effect might hypothesize that the long-term memory representation of (1) is the following hierarchy of chunks or registers:

(3) (S (NP (Q *two*) (N *Finns*)) (VP (V *wrote*) (N *texts*)))

The topmost level of the hierarchy is occupied by the register S which contains the symbol sequence 'NP VP'. NP itself is the name of a register with the sequence 'Q N' as its content. At the lowest level there are registers (e.g. Q and N) which always contain one single *terminal symbol* such as *two* and *Finns*.

There are several possibilities for devising a procedure which will reconstruct the memorized sentence from such hierarchies. One of those, which is very similar to the mechanism described by Johnson (1970), will be outlined here. The operations to be performed are controlled by the list of symbols called the pushdown stack, and by the symbol in an auxiliary location called Aux. The sole accessible symbol of the stack is its topmost element, T; the only permissible stack alterations are placing a new symbol on top of the stack and removing T from the stack. Of course, both types of alterations result in a new T. The procedure has been labelled Reprod and

Table 1. *Essentials of a possible procedure for reconstructing
a learned sentence or word string from a hierarchy of labelled chunks or registers*

(For explanation see text.)

Reprod	Fillstack
(1) If T is a terminal symbol then pronounce T else copy T into Aux	(1) If reg(Aux) is not empty then (*a*) copy the last symbol of reg(Aux) on to stack (*b*) remove this symbol from reg(Aux) else Reprod
(2) Remove T from stack	(2) Fillstack
(3) Fillstack	

it includes the subroutine Fillstack; an occurrence of any of these labels in a program statement means that the associated (sub)procedure is called. Table 1 contains the essential part of Reprod.

Reprod starts with the label of the highest-order register (that is, S, in case of hierarchy (3)) on the stack. The assumption has been made that once a symbol has been copied from a register onto the stack it is removed from the register (statement 1*b* of Fillstack). This obviously irrealistic but inessential and simplifying feature is easily amendable. Since T is not a terminal symbol it is copied into Aux and control is transferred to Fillstack. Fillstack puts the symbols of the register which is indicated by Aux (reg(Aux)) onto the Stack in such a way that the first register symbol is T at the moment control is returned to Reprod. With NP figuring as T, now, Reprod calls Fillstack again after having written 'NP' into Aux. Fillstack moves the NP content to the stack, etc. Only after *two* has been placed on top of the stack Reprod will utter the first response word. It is easy to see that if two consecutive responses belong to the same constituent, the amount of symbol manipulation carried out is smaller than if they belong to different constituents: the basis of the constituent boundary effect.

Sentences (1) and (2) are instances of the syntactic construction Quantifier + Noun + Finite main verb + Noun, with the first noun serving as head of the subject phrase, the second noun as object, and the sequence Main verb + Noun as predicate phrase. Now consider strings such as (4) *Two Finns*, (5) *Finns wrote* and (6) *wrote texts*. Each of them, too, is an instance of a grammatical construction and expresses part of the meaning of (1). For convenience, and somewhat stretching the meaning of the term, let us call (4)–(6) *partial paraphrases* of (1). A syntactic construction may be defined as a sequence of syntactic categories and function words which expresses one or more conceptual case of modifier relationships. For short, a syntactic construction is a pair consisting of a syntactic scheme and a conceptual pattern. For example, the Noun + Finite intransitive main verb construction (*Finns wrote*) expresses agentive case relationships between an unmodified actor and an unmodified action. From this definition it is not difficult to see how syntactic constructions could work, in principle, as devices for retrieving information from memory representations and map it into syntactic schemes as output formats. One needs to assume that, for each syntactic construction that is heavily overlearned and relatively simple, the subjects have available a procedure which searches through a designated area of long-term memory – the area which contains the conceptual content of the utterance(s) to be produced. There it looks for a token of its

152   G. KEMPEN

conceptual pattern and transforms the collected information into its output format (consulting, whenever necessary, the internal lexicon).

How can these syntactic retrieval plans account for the constituent boundary effect? Suppose a retrieval plan has located a piece of information which fits into its conceptual pattern. At this moment, the retrieval plan initiates its second task: assembling a word sequence according to the specifications of its syntactic scheme. We will assume that the procedure which performs this task consists of a hierarchy of subprocedures. Take hierarchy (3), for instance. Now, the non-terminal symbols are to be interpreted as labels of subprocedures which assemble a specific type of constituent, and no longer as names of already filled registers or chunks. The terminal symbols indicate the results obtained by the subprocedures. The program of Table 1 governs the order in which the various subprocedures are called and, again, will produce a constituent boundary effect.

Returning to sentence (5), the procedure which fabricates constructions of this type might be represented by $(S'(N\text{—}) (V\text{—}))$, with dashes referring to open places of the output format. Apart from the identity of the non-terminal symbols, this hierarchy is identical to the NP and VP parts of (3).

The probe latency technique entails two deviations from the usual way of sentence reproduction: the reproduction sequence does not necessarily start with the first word of the sentence, and it consists of one word only. Without working out the details it will be assumed that presentation of a probe word sets the current retrieval plan to the state which during normal sentence reproduction would have been present at the moment of pronouncing that word. Furthermore, it is to be supposed that the subjects, in response to the task instructions given to them, have set up a control mechanism which forces the retrieval plan to halt after the first response word has been uttered.

PARAPHRASTIC REPRODUCTION: EXPERIMENTAL PREDICTIONS

In order to apply the previously mentioned experimental strategy for solving storage vs. retrieval problems (see introductory section) to sentence memory one can exploit the availability of partial and full paraphrases of the memorized sentences. If the retrieval interpretation is right and the subjects derive responses from conceptual representations of the learned sentences, then, during appropriately devised recall tests, they must be able to use the retrieval plans which correspond to (partial) paraphrases of the learned sentences. Interpretation of profiles of transition difficulties yielded by this *paraphrastic reproduction* method proceeds in accordance with the general storage vs. retrieval logic. In sentences (1) and (2), the middle transition requires a greater amount of retrieval plan activity than both neighbouring transitions and is expected, other things being equal, to take more time. Now take an experiment where the subjects, after having learned four sentences of type (1) or (2), are told that during the subsequent probe reaction time trials only probe and response words from the four partial paraphrases such as (5) will occur. This represents a situation where the subjects can profitably apply the retrieval routine associated with a familiar construction which is simpler than the construction of the learned sentences. This consideration, however, generalizes to the transitions in partial paraphrases (4) and (6) if their difficulties would be measured in the analogous way.

In this new type of task a general decrease in average latency is to be expected because of the smaller number of possible probe and response items. Since none of the constructions exemplified by (4), (5) and (6) contain compound constituents (of word length two or more), the retrieval hypothesis would expect the mean latency for the middle transition to decrease more sharply than the other transitions when comparing the usual way of measuring probe latencies to this new way. (One cannot make the stronger prediction that the three decrease to the same level of difficulty since the *ceteris paribus* condition is not necessarily true: the words flanking the three transitions are different, and the three subroutines might take different amounts of time.)

As for terminology, the usual method of probe reaction time measurement, where probes are taken at random from the various transitions, will be called 'sentence-wise'. The alternative paradigm with many successive probes belonging to the same transition, will be called 'pairwise'.

The interaction predicted by the retrieval hypothesis (the latency decrease as a consequence of shifting from the sentencewise to the pairwise paradigm is greater for between-constituent transitions than for within-constituent transitions) must not appear, however, in sentences in which the between-constituent transition is not surrounded by a word pair making up a syntactic construction. In the Dutch language such a condition can easily be arranged without altering the words from which sentences like (1) and (2) have been composed.

Word order within the Dutch predicate constituent is different in main and subordinate clauses. In main clauses the finite verb precedes other predicate parts such as object, adverb and prepositional phrase; in subordinate clauses this order is reversed. For instance, if the Dutch equivalent of (1) is embedded as a subordinate clause within a main clause of the type *He says that...*, the result would be – in a word-to-word translation – (6) *He says that two Finns texts wrote.* Here, the between-constituent transition is no longer flanked by a partial paraphrase because (7) *Finns-texts* is not a syntactic construction. Hence, the above interaction between paradigm and transition should not appear.

In this set of predictions the interaction for type (1) sentences is critical with regard to the storage vs. retrieval issue. The storage assumption predicts that the constituent boundary effect will remain constant through variations in the measurement paradigm. In the following experiments these predictions were put to test, together with a control that allowed elimination of a trivial short-term memory interpretation of the outcomes.

EXPERIMENT I

The first experiment is basically a replication, with two modifications, of the traditional constituent boundary effect as measured with the sentencewise probe latency technique. Many of the procedural details also apply to Expts. II–IV.

The first modification is the simultaneous measurement of probe latencies for *more than one* sentence. In earlier studies, sentences were probed one-by-one and subjects had the opportunity to concentrate on one sentence only. The pairwise paradigm, however, presupposes that several sentences are simultaneously kept in memory. In order to equalize the sentencewise and the pairwise paradigms in all non-critical

154   G. KEMPEN

aspects it was decided that in the sentencewise condition, too, a set of four sentences should be memorized and probed concurrently. A pilot experiment with the sentence-wise paradigm pointed out that if the subjects have advance knowledge of the sentence which is about to be probed, the pattern of results is not affected, except for a general decrease in mean latency.

The second modification consists of the introduction of backward responses, i.e. of a condition where the subjects have to name the predecessor, not the successor, of the probe. The purpose of this modification is not merely to add a few data points: it yields the opportunity to test the hypothesis that the constituent boundary effect should be more pronounced in the backward than in the forward reactions. This hypothesis follows from the conception of the retrieval system described above together with the reasonable assumption that any *ad hoc* change in a previously acquired plan slows down the reactions that necessitate that change.

In order to change the program of Table 1 in such a way that it will reproduce a string in reversed order, the only amendment needed is to replace the word 'last' in statement 1*a* of Fillstack by 'first'. In case of a within-constituent transition this statement is executed less often than in case of a between-constituent transition. Taking the hierarchy of (3) for instance, statement 1*a* is encountered only once between responses *two* and *Finns* but three times in between *Finns* and *wrote*. This means that a between-constituent transition, which is already difficult by itself, will suffer more from the requirement of direction reversal.

Method

Subjects were 20 students at the University of Nijmegen who took part in individual sessions. Each subject learned a different set of four Dutch sentences which had been composed from the following lexical material. The first word of all four sentences was *Die* (Those). The second position was occupied by the numerals *twee, drie, vier* and *vijf* (two, three, four, five). The third words were *Belgen, Finnen, Grieken* and *Ieren* (Belgians, Finns, Greeks, Irish). In fourth position occurred the past tense verbs *lazen, leerden, schreven* and *typten* (read, learned, wrote, typed). They belong to the category of middle verbs which can serve both transitive and intransitive function. The final words were selected from four different categories, one word from each category. (The reason for this complication has to do with a theoretical issue that will be touched on in the footnote only.) The categories were Objects (*boeken, teksten, woorden, zinnen*; books, texts, words, sentences), Adjectival Adverbs (*keurig, matig, moeizaam, prettig*; nicely, moderately, laboriously, pleasantly), Time Adverbs (*'s avonds, 's morgens, 's winters, 's zomers*; in the evening, in the morning, in the winter, in the summer), and Frequency Adverbs (*altijd, dikwijls, nimmer, zelden*; always, often, never, seldom). Each word sequence that arises by arbitrary selection of one word from each of the five position groups constitutes a grammatical sentence.

The 20 four-sentence sets were systematically made up in a way that guaranteed maximal between-set variation in the concrete to-be-learned sentences. Learning material of one subject could be, for example, the set *Die drie Finnen typten moeizaam, Die twee Grieken leerden boeken, Die vijf Ieren schreven 's avonds, Die vier Belgen lazen dikwijls*.

The sentences were learned from a memory drum that exposed a new line of the learning material every four seconds, by means of a study–test method. The recall stimuli were the first two words of the sentences, e.g. *Die twee, Die drie*. Memorization went on till the criterion of word perfect recall and, in addition, there was 100 per cent overlearning.

Probe reaction times were measured by a voice-key that started when the probe word appeared in the window of a memory drum and stopped at the onset of subject's pronouncing the response word. The interval between consecutive probes was 6 sec. Reaction times of 3000 msec or longer

Table 2. *Average latencies (arithmetic means) for the two main transitions in Expts. I–IV*

| | Transition and direction | | | |
| | Between-constituents | | Within-predicate | |
Experiment	Forward	Backward	Forward	Backward
I	1725	2061	1561	1591
II	1339	1454	1378	1382
III	1446	1469	1335	1376
IV	1368	1654	1258	1301

were discarded as errors. The instruction required the subject to name, as quickly as possible upon the presentation of a word from one of the memorized sentences, avoiding errors as much as possible, the word that immediately followed the probe word in the sentence concerned (or preceded it, in the case of backward reactions). If the probe was the last word of a sentence subjects had to reply with 'punt' ('stop'); in the backward condition, the pronoun *Die* (Those) served as the analogue of 'punt'. Each subject took part in both forward and backward conditions, the order of these conditions being systematically varied to the effect of controlling for practice influences. The list of probes for a subject contained four occurrences of each of the 16 possible probes in random order, and there were separate lists for the backward and the forward conditions. Both started with four arbitrary training probes. Each subject produced a maximum of $4 \times 16 \times 2 = 128$ latency scores. During the learning stage of the experimental sessions subjects were unaware of the fact that a probe latency task would follow.

Results

Due to the following sources of erroneous responses, not all intended observations could be obtained: incorrect answers; no response within 3000 msec; answers given without sufficient voice volume to stop the timer; and premature stopping of the timer as a result of irrelevant noises made by subject, apparatus or environment. In this paper the data analysis will be confined to the middle and the right-hand transitions, that is, to the between-constituent transition and the transition within the predicate. The left-hand transition (such as *Twee Grieken*) will not be considered because the quantifiers served as recall stimulus during learning. The transitions with respectively 'stop' and 'Those' as responses are even more atypical, for obvious reasons.

Table 2 contains the mean reaction times for the two main transitions and for both forward and backward reactions. The percentage of errors here was 14; they were excluded from further analysis. The constituent boundary effect amounted to 164 msec in the forward direction; it was much stronger in the backward direction: 470 msec. The upper pair of dotted lines in Fig. 1 represents the same data. An analysis of variance confirmed that the Transition by Direction interaction which is at stake here was highly significant ($F = 74 \cdot 5$; d.f. $= 1, 19$; $P < 0 \cdot 0001$). Application of the t test for *a priori* comparisons to the 164 msec difference between the two forward means yielded a value of $t = 4 \cdot 65$ (d.f. $= 19$; $P < 0 \cdot 0005$).

The conclusion is that the procedure of this experiment favours the appearance of the constituent boundary effect, especially in the backward condition. The next experiment concentrates on the pairwise paradigm.
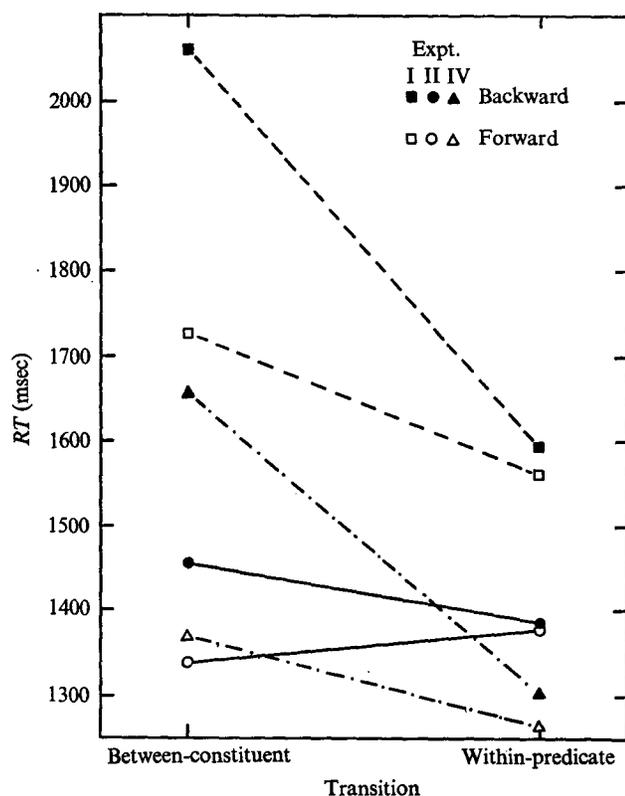
11-2

Fig. 1. Mean latencies for Expts. I, II and IV. Expt. I: sentence wiseparadigm. Expt. II: pair-wise paradigm with partial paraphrase at between-constituent transition. Expt. IV: pairwise paradigm but no partial paraphrase at between-constituent transition.

EXPERIMENT II

Measurement of probe latencies according to the pairwise paradigm in which con-secutive probes come from one type of transition creates a situation where the sub-jects have advance knowledge of the transition being tested, and affords them the opportunity to mobilize a syntactic retrieval plan specifically tuned to that transi-tion, at least if the word pairs surrounding that transition are instances of a syntactic construction. The latter condition is fulfilled by the sentence material of Expt. I because all three transitions are flanked by tokens of syntactic constructions, e.g. *two–Finns*, *Finns–wrote* and *wrote–texts*.

The pairwise paradigm may be expected to entail a general increase in the speed of reactions since it curtails the number of possible probes and responses substantially as compared with the sentencewise paradigm. But if the subjects bring into play specific retrieval routines for the construction exemplified by *Finns wrote*, then this transition no longer suffers from the disadvantage of being a transition between compound constituents. Hence, the reaction times for the between-constituent transition are predicted to improve more than those within constituents.

As in Expt. I, both backward and forward measures were taken, though the two

types were intermingled. That is, not only *Finns* but also *wrote* served as a probe within the same series of trials, and the subjects were instructed to respond with the other member of the pair.

Method

Apart from the paradigm of probe latency measurement the procedure of the present experiment was identical to the previous one. As a matter of fact, the two were run concurrently and 40 subjects were alternately assigned to either of them. The list of probes was divided into four parts, one for each transition measured, namely the between-constituent transition, the transition within the predicate phrase and two transitions within pairs of non-adjacent words, e.g. *Finns–texts*. For each transition there were eight possible probes: the left-hand and the right-hand members of the four surrounding word pairs; they required forward and backward responding, respectively. Each probe word occurred four times on the list. The total number of trials was 4 (transitions) × 8 (probes per transition) × 4 (repetitions of probe words) = 128, as in Expt. I.

The instruction for each transition was communicated to the subject in terms of word pairs occupying certain serial positions in the memorized sentences. For example, 'Each of the following 32 probes is either the third or the fourth word of a sentence; if you see the third word, respond with the fourth word of the same sentence, and vice versa.' In order to control for practice effects the presentation order of the transitions was systematically varied over subjects. They heard about the probe latency measurement only after the learning stage.

Results

The following data do not include 5 per cent of erroneous responses. The main latencies as shown in Table 2 and in the continuous lines of Fig. 1 make amply clear that the between-constituent transition took more advantage of the shift to the pairwise paradigm than the within-constituent transition. The constituent boundary effect has almost disappeared: only in backward direction a weak effect of 72 msec has remained ($t = 2 \cdot 00$; d.f. $= 19$; $P < 0 \cdot 05$; t test for *a priori* comparisons), sharply contrasting with the 470 msec effect of Expt. I. The slight 'negative' constituent boundary effect for the forward reactions is not statistically reliable.

DISCUSSION

Although these results accord with the predictions, they are open to a trivial interpretation. One could hypothesize that during the empty intervals between a response and the next probe the subjects engaged in rehearsing the four word pairs from which the probes were selected and thus were responding on the basis of information kept in short-term or working memory and not by interrogating the long-term memory representations of the learned sentences. If this were true, one would no longer expect the between-constituent transitions to take more time than the other ones. This calls for a control experiment which prevents the subjects from implicit rehearsal during empty intervals.

EXPERIMENT III

In order to ascertain that, at the moment of presentation of a probe, the subject's working memory does not contain the four word pairs which are being tested during a given series of probes, a filler task was introduced that consisted of reading aloud,

158   G. KEMPEN

at high speed, a list of consonant letters. The subjects started this activity as soon as they had produced a response or, in case of no response, upon a signal that the maximum allowed latency of 3000 msec had elapsed. They went on reading until an acoustic signal warned them to look back to the position of the memory drum window where the next probe would appear. The letters were visible in the said memory drum window, to the right of the probe position. The warning signal sounded 700 msec before the next probe was due. There were eight new subjects from the same general population as before. Further procedural details were identical to those in Expt. II, except that all subjects learned the same set of sentences.

The pattern of results (cf. Table 2; 9 per cent errors were discarded) was in good agreement with that of the previous experiment. The means of the backward reactions were even nearly identical to their counterparts in Expt. II. As for the forward reactions, the slight negative constituent boundary effect in Expt. II had changed to a positive effect of 111 msec which, however, was not significant ($t = 1\cdot46$; d.f. $= 7$; t test as above). This might be interpreted as discomforting to the retrieval plan hypothesis because the effect was not very much stronger in the sentencewise paradigm (164 msec, Expt. I). However, since the profile of the backward latencies has been left untouched by the short-term memory control and since, after all, the forward constituent boundary effect is smaller, though not significantly so, than in Expt. I, it is concluded that the short-term memory explanation is not valid. The reader will recall the theoretical and empirical arguments (Expt. I) that backward reactions are liable to stronger constituent boundary effects than forward ones. This justifies more weight to be attached to backward than to forward reactions. The question why the filler task does affect the latter but not the former reactions remains open. (The results of Expt. IV present further arguments against a straight-forward short-term memory interpretation; cf. footnote.)

EXPERIMENT IV

It remains to be demonstrated that in the pairwise paradigm the pattern of outcomes typical for the constituent boundary effect will reappear when the word pairs at both sides of the between-constituent transition do not make up a partial paraphrase of the memorized sentences (cf. the section entitled Paraphrastic Reproduction). One possible way of generating sentences fulfilling this requirement is simply to convert the sentences of the previous experiments into subordinate clauses, for instance, by prefacing them with the conjunction *because*. This yields, in a word-to-word Dutch to English translation, clauses such as *Because those two Finns texts wrote...*, with the pair *Finns–texts* at the boundary between constituents.

One further modification of the sentences was needed. In Expts. I and II, the words following the subject nouns belonged to one syntactic category (verbs), but the words in final position were members of varying parts of speech (nouns, adverbs). Reversing the predicate word order would move the latter to the position following the subject nouns. In order to enhance the comparability of the present experiment with Expt. II it was decided to use only one type of predicate construction, namely Object + Main Verb, so that all words following the main constituent boundary belonged to one syntactic category (see Expt. I for a listing of the object nouns).

Further procedural details were exactly the same as in Expt. III, although there was no filler task. Eight new subjects had been recruited from the same population.

The main results (16 per cent errors were discarded) are shown in Table 1 and by the lower pairs of dotted lines in Fig. 1. It is obvious that the general pattern of the constituent boundary effect has been restored. In the backward direction the effect was again very strong. In the forward direction it is significant ($t = 2\cdot23$; d.f. $= 7$; $P < 0\cdot05$, t test as above), although the between-constituent transition took only 110 msec more than the transition within the predicate phrase. It is also instructive to look at the latencies obtained for the non-adjacent word pairs. In this study, as in Expts. II and III, reactions were measured over the pairs formed by subject nouns and final words. In the previous studies the non-adjacent pairs always produced longer latencies than the adjacent pairs, the difference typically being 150 msec or more (averaged over forward and backward responses). Such an effect would be predicted by the syntactic retrieval hypothesis because non-adjacent pairs like *Finns–texts* or *Greeks–always* do not comprise syntactic constructions, let alone partial paraphrases. In the present experiment, however, the non-adjacent pairs do constitute partial paraphrases (*Finns wrote*, etc.) and one would expect considerable facilitation as a result of this. Indeed, the non-adjacent pairs now take 1350 msec on the average, which is 160 msec below the average of the adjacent between-constituent transition. The conclusion must be that the latency profiles behave in good agreement with the predictions derived from the notion of syntactic constructions as retrieval plans.*

GENERAL DISCUSSION

The first implication of this series of paraphrastic reproduction experiments is that the frequently observed constituent boundary effect in verbatim sentence recall comes into existence at retrieval time. Hierarchically organized retrieval plans provide a sufficient explanation and it need not be assumed that the sentence representations which are accessed by these mechanisms are divided into parts corresponding to syntactic constituents. Plausible candidates for this type of sentence representation are conceptual representations in the form of predicates over arguments or cases, or of imagery codes. This contrasts with the dominant view that verbatim recall largely reflects properties of the phonetic or surface

* Two remarks on the interpretation of Expt. IV. First, it should be noticed that the results of Expt. IV refute the straightforward STM explanation outlined in Expt. III, but they remain compatible with a more limited STM interpretation, namely that the subjects were rehearsing word pairs that constitute syntactic constructions (and no other word pairs). This version is also contradicted by the results of Expt. III, but, if it were true, it would have to add some account of why syntactic constructions are so easy to rehearse!

Second, assuming that sentences receive memory codes in the form of conceptual networks (e.g. *Finns*–actor–*write*–object–*texts*), one might argue that the distance between *Finns* and *wrote* is smaller than between *Finns* and *texts*. This might provide an alternative explanation of the different reaction time profiles of Expts. II and IV. However, this line of reasoning runs counter to another aspect of the data that has not been mentioned yet. If distance over nodes of a conceptual network were an important determinant of reaction times, one would predict substantial effects of the various grammatical relations in the predicate phrase. Remember that, apart from (*a*) object nouns, there were (*b*) adjectival adverbs (verb modifiers), (*c*) time adverbs, and (*d*) frequency adverbs. In conceptual network representations, (*a*) and (*b*) would be more closely connected to the verb than (*c*) and (*d*), that serve as sentence modifiers. But in none of the three experiments did type of grammatical relation produce effects that were statistically reliable. Moreover, the effects changed direction over experiments.

160   G. KEMPEN

structure representations which the subjects supposedly store while memorizing. The
present results indicate that rote learning of sentences gives rise to easily accessible
and retrievable memory representations of their conceptual contents, perhaps in
addition to phonetic representations. Other proponents of retrieval interpretations
of sentence recall phenomena are Loosen (1972) and James *et al.* (1973). The latter
group of investigators take this stand in the context of another so-called syntactic
effect in sentence memory: the 'tendency towards the kernel' as found in the recall
of active, passive, negative, etc., sentences (e.g. Mehler, 1963).

Second and more important, the results point at the existence of long-term
memory retrieval routines closely allied with (highly overlearned and simple) syn-
tactic constructions; i.e. routines which scan sentence representations, read out only
a specific pattern of conceptual information and cast this into a specific syntactic
output format. This contrasts with the view that sentence production proceeds by
first deciding on the content of the utterance, e.g. by somehow retrieving conceptual
information from memory and, subsequently, mapping this information into a
syntactic scheme.

It is suggested that syntactic factors play a role already during the first stage of
sentence production, namely as one of the mechanisms for guiding the memory
retrieval of conceptual content.

REFERENCES

FILLENBAUM, S. (1973). *Syntactic Factors in Memory?* The Hague: Mouton.
JAMES, C. T., THOMPSON, J. G. & BALDWIN, J. M. (1973). The reconstructive process in sentence
    memory. *J. verb. Learn. verb. Behav.* **12**, 51–63.
JOHNSON, N. F. (1965). The psychological reality of phrase-structure rules. *J. verb. Learn. verb.
    Behav.* **4**, 469–475.
JOHNSON, N. F. (1970). The role of chunking and organization in the process of recall. In G. H.
    Bower (ed.), *The Psychology of Learning and Motivation*, vol. 4. New York: Academic Press.
KENNEDY, R. A. & WILKES, A. L. (1968). Response-times at different positions within a sen-
    tence. *Q. Jl exp. Psychol.* **20**, 390–394.
LOOSEN, F. (1972). Cognitieve organisatie van zinnen in het geheugen. (Dissertation, Louvain.)
MEHLER, J. (1963). Some effects of grammatical transformation on the recall of English sentences.
    J. verb. Learn. verb. Behav. **2**, 346–351.
SUCI, G. J., AMMON, P. & GAMLIN, P. (1967). The validity of the probe latency technique for
    assessing structure in language. *Lang. Speech* **10**, 69–80.
WILKES, A. L. & KENNEDY, R. A. (1969). Relationships between pausing and retrieval latency
    in sentences of varying grammatical form. *J. exp. Psychol.* **79**, 241–245.