Q. Jl. exp. Psychol. (1969) **21**, 280–290

SYNTACTIC STRUCTURE MODIFIES ATTENTION DURING SPEECH PERCEPTION AND RECOGNITION

KENNETH ABRAMS
University of Michigan
AND
THOMAS G. BEVER
Rockefeller University

The present paper demonstrates the interaction of syntactic structure and speech perception with a response task which minimizes the effects of memory: reaction time (RT) to clicks during sentences. (1) In 12-word *unfamiliar sentences* each with two clauses, RT is relatively slow overall to clicks located at the end of the first clause but decreases as a function of clause length. Clicks at the beginning of the second clause are not affected by length of the preceding clause. (2) In *familiar sentences*, RT is relatively fast to clicks located at the end of a clause while RT to clicks at the beginning of clauses is relatively unaffected by familiarity. (3) RT is *not* fastest overall to clicks located between clauses either in novel or familiar sentences. (4) As in previous studies, the subject's subsequent judgement of the location of the click tone are towards the clause break. (5) We could find no systematic interaction between RT and subjective click location. Findings (1) to (3) are consistent with the view that perceptual processing alternates between attending to all external stimuli and developing an internal representation of the stimuli. Finding (3) is in conflict with an "information channel" view of immediate attention to speech, which would predict high sensory attention to non-speech stimuli between clauses. However, findings (4) and (5) indicate that the channel view of perception may be correct for that perceptual processing which occurs after the immediate organization of the speech stimulus into major segments.

Introduction

Recent studies have shown that subjects make systematic errors in locating a non-speech stimulus superimposed on speech. For example, when subjects hear a click during a sentence their subsequent estimates of the actual click location are biased toward the point between clauses. This finding has been interpreted as evidence for the view that listeners actively use grammar to impose clause structure on the speech stimulus as they hear it. In this paper we demonstrate this view with a response task that minimizes the possible influence of memory strategies and response biases. We also show that a "channel model" of attention applies to speech processing only after the speech signal has been segmented into perceptual major units.

Grammatical analysis divides sentences into major segments. For example, in sentence (1) there are two clauses,

(1) Peter ran quite fast but we caught him in the narrow valley.

"(Peter ran quite fast)" and "(but we caught him in the narrow alley)." The behavioural relevance of this kind of formal segmentation has been demonstrated in a number of experiments. Using a technique developed by Ladefoged and Broadbent (1960), Fodor and Bever (1965) found that in the dichotic presentation of a click and a spoken sentence, subjects recalled the location of the click as displaced toward the major phrase boundaries between clauses. For example, a click actually located in "fast" or in "but" in (1) was reported as occurring between "fast" and "but." These results have been extended using sentence materials which investigate intonational effects (Garrett, Bever and Fodor, 1966) the effect of sequence probability (Bever, Lackner and Stolz, 1969) and the relative effect of surface and underlying syntactic structure (Bever, Lackner and Kirk, 1969). The conclusion from these studies is that the immediate segmentation of stimulus sentences occurs at points between groups of phrases which are bound together by basic syntactic relations (e.g. the relations "actor, action, object, modifier").

In the click displacement experiments, subjects report the location of the click shortly *after* hearing (and writing down) the sentence. This allows not only instantaneous attentional strategies, but also delayed encoding strategies, and subsequent response strategies to affect the reported click location. Therefore, the previous studies show that the formal syntactic structure of a sentence plays a psychological role in immediate attention, in perceptual analysis, in organizing the response, or in some combination of these three processes.

Our concern in the present research was to investigate how the formal segmentation of speech affects *immediate* attention as measured by reaction time to clicks presented during the sentences. Reaction time to clicks is quite fast (100 to 300 msec.) so that any systematic effects of syntactic phrase structure on reaction time would occur within a syllable of the actual click presentation. Therefore, any systematic effects on reaction time could not be due to slightly delayed operations in the analysis of the sentence or to post-stimulus response strategies. The claim in previous accounts of click location errors has been that the points between clauses represent points of minimum attention to speech and maximum attention to other stimuli. Thus our primary hypothesis was that clicks objectively occurring *in* clause breaks should receive faster reaction times than clicks in any other location.

Experiment 1

Method

Nine 12-word sentences with two clauses each were recorded with subdued (but normal) intonation. These sentences had a clause break after the fifth, sixth or seventh word (see Appendix). Three tape copies of the original sentences were spliced into three different experimental orders such that sentences which were among the first third of one experimental order were among the middle and last third of the other two orders. In each sentence a single pulse was located on the second track of a stereophonic tape in one of three positions relative to the major clause break; in one copy it was placed in the word *before* the break, in one copy *after* the clause break and in one copy it was placed *in* the clause break. (The two words preceding and following the break were monosyllables.) Across the three experimental orders, a given sentence had pulses located in each of three positions relative to the clause break. In each experimental order there were three sentences with pulses before, three sentences with pulses in, and three sentences with pulses after the break; the sentences were presented so that each pulse position was represented once in each third of each experimental order. Thus, the experimental design was balanced across the three variables of clause break position, pulse position relative to the clause break and experimental order. Three normally-spoken 12-word practice sentences with a pulse preceded each experimental order of nine.

The pulse in each stimulus triggered a 25 msec. capacitor discharge click. The click was approximately equal in intensity to the loudest speech sound in the stimulus sentences. Triggering the click also started an electronic timer which was stopped when the subject closed a telegraph key.

Subjects were 12 M.I.T. undergraduates who volunteered their time. The subject was instructed to hold the fingers of his right hand on the key and to close it as soon as he heard the click. After responding to 20 isolated practice clicks, the subject then heard one of the three groups of stimulus sentences. He was told that he would hear a sentence and one click during each sentence. To insure high attention to the reaction time task, the subject was paid 25 cents for fast reaction time responses. (A "fast response" was defined without the subject's knowledge as any response faster than the fastest trial on the five preceding trials.) To insure attention to the sentence, the subject had to write out each sentence immediately after hearing it. The sentence was presented to the left ear and the click to the right ear.

After presentation of the first order of 12 sentences (three practice and nine experimental sentences, hereafter called "presentation 1") ,the subject read a book silently for 5 min. He was then presented with another order of the 12 sentences ("presentation 2"). After another 5 min. of silent reading, the subject was presented with the third order of the 12 sentences ("presentation 3"). Subjects were run individually, and the six possible orders of the three experimental groups of sentences were evenly balanced. Thus, every subject responded to every sentence three times, each time with the click in a different location relative to the clause break.

Results

There are two findings:

(1) The main hypothesis was not confirmed; clicks located in major clause breaks did *not* receive the fastest reaction times.

(2) Clicks located before clause breaks had longer reaction times than those located after clause break.

TABLE I

Mean reaction times to clicks before, in and after points between clauses, in first, second and third presentation of sentences (Experiment I sentences spoken naturally)

	1st presentation	2nd presentation	3rd presentation
Before-clause break	243	202	172
In-clause break	230	184	169
After-clause break	216	189	179

Table I presents the mean reaction time (RT) for each of the click positions in each of the presentation groups. The prediction that in-break clicks would be responded to fastest was not confirmed overall. This was tested first by comparing the average RT for each subject of *in-break* with *before-break* and with *after-break* clicks for all presentation groups together ($P < 0{\cdot}7$ one-tail by Wilcoxon matched-pairs signed ranks test on 12 subjects for each comparison. In this paper all significance measures are by a Wilcoxon matched-pairs signed ranks test across subjects or sentences, unless otherwise indicated). For the separate presentation groups the main prediction was not confirmed except for the difference between before-break clicks and in-break clicks in presentation 2 ($P < 0{\cdot}03$ one-tail).

Clicks preceding breaks were responded to more slowly than clicks following breaks. This was confirmed over all presentation groups considered together ($P < 0{\cdot}05$ two-tail). There was an experience effect: the slowness of RT's to before-break was confirmed for presentation groups 1 ($P < 0{\cdot}05$) and 2 ($P < 0{\cdot}025$) and was not confirmed for presentation 3.

Subjects' RT might have become faster as a function of how much of the sentence had passed; they might have noticed that the sentences were all about the same length, and could have adopted a strategy of increasing expectation for the click as they heard the sentence. If this were true, the relative slowness of RT to before-break clicks might not have been due to the syntactic structure itself, but rather to the fact that before-break clicks on the average appeared one word earlier than after-break clicks. To test this we calculated a separate average RT for each subject for each presentation order, of the responses to the sentences with the clause break after the fifth, sixth and seventh word. None of the differences were significant (by a Wilcoxon test on those subjects who contributed scorable responses in all syntactic positions in every presentation group).

If RT is affected by ongoing perceptual analysis, the results of this experiment suggest that the original view of the nature of the displacement of clicks in speech was not correct. Although before-break clicks receive slower RT than after-break clicks, in-break clicks do not receive the fastest RT. This implies that it is not the case that sensory attention to non-speech material is highest between clauses.

The failure to confirm our main hypothesis and the relative slowness of before-break clicks in Experiment I might be due to acoustic cues of structure such as intonation and pauses rather than due to the active use of structure in sentence perception, since the sentences in Experiment I were spoken with normal intonation. Experiment II was designed to eliminate the possibility of this kind of artifact. Subjects were also asked to locate the clicks after hearing each sentence so that we could compare the processes of click location and reaction to clicks.

Experiment II

Method

The three groups of 12 sentences used in Experiment I were reconstructed artificially. We first recorded the individual words separately in a random list. The appropriate words were then spliced together to produce sentences free from any systematic prosody cues. The placement of the pulses and constitution of the three experimental groups were identical to Experiment I.

Thirty right-handed, undergraduate native speakers of English were run in individual sessions. The subject responded by releasing a telegraph key (instead of closing it as in Experiment I) to 20 trials of a click heard in the right ear. As further practice, the subject then responded to clicks in the right ear with five normally-spoken 12-word sentences presented to the left ear. After this, the experimental sequence was exactly the same as in Experiment I: each subject received three groups of 12 sentences separated by 5 min. of silent reading. Through the experiment, the subject wrote down the sentences after responding to the click, and indicated where in the sentence the click had occurred. The subject was not given a reward for relatively fast responses.

Results

There are three findings (see Table II):

(1) The main hypothesis was not confirmed. In-break clicks did not receive the fastest RT.

(2) (a) RT to clicks at the ends of clauses was slower than to clicks at the beginning of clauses in presentations 1 and 2 and the opposite effect occurred in presentation 3.

(b) Within presentation 1, RT to clicks at the ends of clauses increased with the length of the clause, while RT to clicks at the beginnings of clauses was unaffected by the length of the preceding clause.

(3) (a) Post-stimulus click location responses replicated previous findings: Subjects tended to locate clicks in the break in all sentences and all presentations.

(b) There was no evidence that the mechanisms affecting reaction time and click location are the same.

TABLE II

Mean reaction times to clicks before, in and after points between clauses, in first, second and third presentation of sentences (Experiment II sentences artificially constructed by tape-splicing)

	1st presentation	2nd presentation	3rd presentation
Before clause break	248	207	180
In-clause break	239	201	193
After-clause break	234	194	196

The prediction that in-break clicks should have the fastest RT was not confirmed for any comparison in any presentation order: in-break responses are neither faster than before-break responses, nor are they faster than after-break responses for any presentation order.

As in Experiment I, the presentation order had a considerable effect on the extent to which before-break clicks are responded to more slowly than after-break clicks. In presentation 1, this effect was strongly confirmed ($P < 0{\cdot}016$ one-tail, on 30 subjects); in presentation 2, the effect was weakly confirmed ($P < 0{\cdot}08$ one-tail); and in presentation 3, the effect was not confirmed at all. In fact, in presentation 3, before-break clicks are responded to significantly *faster* than after-break clicks ($P < 0{\cdot}02$ two-tail). The difference between the presentation orders is primarily due to a decrease in reaction time to before-break clicks (71 msec. difference between the first and third presentations in Experiment I, 68 msec, in

Experiment II). Reaction time to after-break clicks is considerably less affected by practice (37 msec. difference for Experiment I, 38 msec. for Experiment II) ($P < 0{\cdot}02$ two-tail across subjects for the relative susceptibility of before-break clicks to familiarity in Experiment II).

The lack of prosodic cues in this experiment led to a higher error rate on correctly recalling the sentence, so each response was categorized according to whether the recall was substantially correct or incorrect. In presentation 1, 34 per cent of the sentences were not recalled correctly so it might be the case that the relative slowness of RT to before-break clicks is arbitrarily related to perceptual processing of the syntactic structure. To check for this, we tested this effect separately on those responses for which the sentence was correctly recalled and on those for which the sentence was incorrectly recalled. The relative slowness of RT to before-break clicks was confirmed across both sets of responses ($P < 0{\cdot}03$, one-tail for correct-recall responses; $P < 0{\cdot}02$ one-tail for incorrect-recall responses).

As in Experiment I, before-break clicks tended objectively to precede after-break clicks by one word. If a systematic decrease in reaction time were correlated with the number of syllables of the sentence that had occurred before the click then the tendency for before-break clicks to receive slow RT's might not be due to the syntactic structure. In Experiment II, there were enough subjects to test this effect for those sentence-syllable positions only in which both before-break and after-break clicks occurred (albeit in different sentences). When only overlapping sentence-syllable positions are included, the effect is confirmed, although less strongly ($P < 0{\cdot}06$, one-tail).

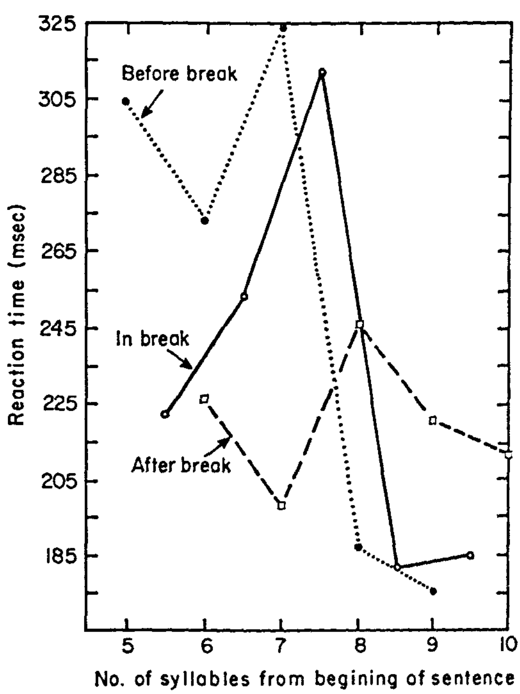

FIGURE 1. The effect of absolute length of preceding sequence on reaction time to clicks in the three different syntactic positions.

However, there was also a general interaction of the number of syllables in the sentence prior to the click with the relative slowness of RT to before-break clicks. This interaction is presented in Figure 1 which displays the RT's according to the number of syllables from the beginning of the sentence in which they occur. (For example, in sentence (1), a before-break click occurred in the fifth syllable, the in-break click between the fifth and sixth syllables and the after-break click in the sixth syllable.) Before-break response times are slower than after-break response times for clicks objectively located in the fifth, sixth or seventh syllables from the beginning of the sentence. That is, RT to before-break clicks are much faster than to after-break clicks for the fifth, sixth and seventh syllable positions ($P <$ 0·001 two-tail grouping the three syllable positions by subject). In contrast, before-break response times are *faster* than after-break response times for clicks objectively located in syllables eight, nine and ten: ($P < 0{\cdot}05$ two-tail by subject). The shift in the relative slowness of RT to before-break clicks appears to be due entirely to an increase in RT to before-break clicks in the later syllable positions (the difference between RT to clicks in syllable positions 5, 6, 7 and in 8, 9, 10 for before-break clicks is significant, $P < 0{\cdot}01$ two-tail by subject); the reaction time to after-break clicks is roughly constant for all syllable positions. The responses to in-break clicks also appear sensitive to the number of syllables from the beginning of the sentence that the click occurs. In-break clicks are responded to significantly faster than after-break clicks, for the clicks in the eighth, ninth and tenth syllable ($P < 0{\cdot}02$ one-tail by subject).

Several results indicate that the psychological mechanism for reacting to clicks is distinct from the mechanism for locating clicks. The pattern of errors in the subjective location of the clicks replicates the results of previous experiments. Seventy-two per cent of click location errors of $\frac{1}{2}$ syllable are toward the major break in presentation 1 and 70 per cent in presentations 2 and 3. However, there is no difference in the confirmation of the click location hypothesis for those clicks in the first five to seven syllables and in the eight to ten syllables in presentation 1. That is, sentence-click combinations which evoke obviously distinct attentional strategies (as revealed by the different effects of syllable positions in presentation 1 and by the different effects of syntax on RT in presentations 1 and 3) confirm the effect of syntax on subjective click location to the same extent.

If the displacement of the reported location of clicks is due to immediate attentional mechanisms, then it might be expected that clicks responded to fastest should either have the least error in reported location or would systematically be reported as preceding clicks which are responded to slowly. The first possibility was tested by correlating separately the accuracy of each subject's location responses with his relative speed on those responses. There was no tendency for positive or negative correlation across subjects for responses in presentations 1 and 2 and a weak positive correlation in presentation 3. In presentation 3, 18 subjects showed a positive correlation, and 12 subjects a negative one ($P < 0{\cdot}07$ by Wilcoxon one-tail on the 30 correlations). This indicates that in familiar sentences, there may be a positive correlation between speed of reaction time and accuracy in subjective location of a click, but that in unfamiliar sentences the mechanisms for reaction time and click location are dissociated. To test the second possibility, the subjective location responses were divided into those which preceded the objective click location, and those that followed the objective location. The reaction time to those clicks subjectively reported as preceding their actual location was not faster than for those reported as following their actual location.

Discussion

The main goal of these experiments was to substantiate the earlier claims that syntactic knowledge is actively utilized in the perception of speech. We found that reaction time to clicks before clause breaks is affected by clause length and by familiarity with the sentence more than the reaction time to clicks after clause breaks. This indicates that syntactic structure does systematically modulate attention during speech perception. The fact that these effects occur in response to artificially constructed stimuli (as in Experiment II) shows that they are not due to any superficial acoustic cues of syntactic structure. Rather, listeners must actively provide the structural analysis which guides their attention as they hear each sentence. The results also indicate that attention to non-speech activity is reduced at the ends of clauses rather than enhanced: before-break clicks receive reactions which actually occur during the clause break, and it is these reactions which are the longest overall; furthermore, reaction times to clicks objectively in the break are not fastest.

These results are not consistent with the view that the points between clauses in a speech "channel" demand the least attention away from a click on a competing click "channel." Several recent experiments support a different view of the interaction of attentional processes (Kahneman, 1967. A similar view is proposed by Wundt in *Physiologischen Studien*, pp. 256 ff.). According to this view, carrying out the perceptual analysis of a stimulus is an *activity* which draws attention away from all other activities. Thus, on this view, competition for immediate attention is not between one sensory information input "channel" and another, but between *all* external sensory inputs and all internal perceptual analysis (see Bever 1968).

In sentences, the clause is a natural unit for internal perceptual analysis: during clauses we listen to the speech *and* non-speech stimuli: at the ends of clauses we encode perceptually what we have just heard. Accordingly, a click at the end of a clause is responded to relatively slowly because it coincides with the point of internal perceptual analysis of the preceding sequence. A click at the beginning of a clause is responded to relatively quickly because it conflicts with relatively little internal perceptual processing.

In both Experiments I and II the relation of the syntax surrounding a click and reaction time reversed for sentences as they became increasingly familiar. The reversal in the effect of syntax on reaction time in familiar sentences reflects a change in the point at which the internal analysis of the clause takes place. In the perception of a novel sentence, the listener organizes the immediately preceding stimulus material near the end of each clause. On the other hand, in the recognition of an already-learned sentence, its internal organization is already present in memory, and the listener's only task is to check which of the previously heard sentences it is. Recent work suggests that the beginning of a constituent predicts the end of that constituent better than vice versa (Johnson, 1966). Thus, a reasonable recognition strategy would be to match internally the beginning of each clause to the set of previously heard clauses to see which of those clauses it is: after comparing the beginning of the clause with the stored representation, the rest of the clause would be relatively predictable. Accordingly, familiarity with a sentence decreases RT to clicks at the end of a clause relatively more than familiarity decreases RT to clicks at the beginning of a clause.

Our interpretation of the effect of internal perceptual analysis on all external attention is also consistent with the fact that clause length decreases RT to end-of-clause clicks but does not affect RT to clicks at beginning of clauses in unfamiliar sentences. The predictability of the end of a clause increases with clause length. Thus, the listener can start his internal perceptual analysis of a long clause before the end of the clause. A click at the end of long clauses would thus occur just after the internal analysis is completed and while the listener is checking the prediction: thus clicks in these positions are responded to relatively fast.

The errors in click location are "attracted" to the break for both unfamiliar and familiar sentences and for sentences with short and long first clauses. Thus the segmental "integrity" of a clause is reflected both in the perception of new sentences and in the recognition of old ones, even though the initial perceptual mechanisms clearly differ. Previous interpretations of the tendency for clicks to be located between clauses (Fodor and Bever, 1965; have followed the proposal of an "n-channel" or "filter-amplitude" theory of auditory attention (Treisman, 1964). On this view, listeners pay attention to the speech except between perceptual units when the attention shifts to the click "channel." The effectiveness of the clause as a unit was demonstrated by the tendency to locate clicks between them allegedly as a function of the ability to switch sensory attention to the click "channel" only between such units.

However, the experiments described in this paper do not support the claim that sensory attention to non-speech stimuli is highest between clauses. This forces us to reconsider the nature of the previous results on subjective click location. If it is not the case that non-speech stimuli receive the highest sensory attention at the points between clauses, why do errors in location of clicks tend to migrate towards the clause break, and why is there a marked tendency for clicks objectively in the clause break to be correctly located?

Many of our results indicate that the attentional system tapped by the reaction time measure is distinct from the behavioural process which produces the systematic errors in click placement. Thus, the systematic nature of errors in click location are due to perceptual processing which follows the initial segmentation of the speech stimulus. We have argued that immediate reaction time interacts with the process of *developing* the internal perceptual organization of the speech. The systematic attraction of click location by clause breaks indicates that the "channel" view of perception is true for speech and clicks once the speech channel has received such initial perceptual organization. That is, listeners first organize the speech into major segments, then they relate the speech and click temporally: it is the latter process that maintains the integrity of the speech units as revealed in the location of clicks.

Certain features of these experiments require investigation with different techniques and study of additional theoretical variables. Reacting to a click in a sentence is itself highly disruptive and may introduce special perceptual properties. Also, the reversal of the relative slowness of RT to before-break clicks for sentences with relatively late clause breaks and for all sentences when they are familiar, indicates that factors such as sentence length and sequence predictability may be as relevant as well as syntactic structure. Nevertheless, these results do demonstrate that syntactic knowledge continuously modulates immediate attention during listening, as well as providing the basis for the subsequent perceptual segmentation of speech into discrete major units.

This work was supported in part by NDEA, AF 19 (628)—5705 to MIT and the U.S. Department of Defence, Advanced Research Projects Agency, Contract No. SD-187, to Harvard University, Center for Cognitive Studies, and in part by the National Institute of Mental Health, Grant No. Pol-MH12623 to Harvard University, Center for Cognitive Studies, by the Harvard Society of Fellows and National Institute of Child Health and Human Development, Grant No. HD 01368–03 to Klaus F. Riegel. The authors are particularly indebted to Dr. Merrill Garrett for his thought and advice. The authors are also indebted to Drs. D. Foss, D. Kahneman, C. Osgood, and P. Carey for reading the manuscript, to Henry Koopmans and Roberta Banay for assisting in the data analysis.

References

ABRAMS, K. (1965). Attention and Speech Perception: AB thesis M.I.T.

BEVER, T. (1968). A survey of some recent work in psycholinguistics, in *Specification and Utilization of a Transformational Grammar*. Final report of Contract No. AF 19 (628)—5127 awarded to W. Plath, IBM, T. J. Watson Research Center, P.O. Box 218, Yorktown Heights, N.Y.

BEVER, T., LACKNER, J. and STOLZ, W. (1969). Transitional probability is not a general mechanism for the segmentation of speech. *J. exp. Psychol.* 79, 387–94.

BEVER, T., LACKNER, J. and KIRK, R. (1969). The underlying structure sentence is the primary unit of immediate speech processing. *Percept. Psychophysics* 5, 225–34.

FODOR, J. and BEVER, T. (1965). The psychological reality of linguistic segments. *J. verb. Learn. verb. Behav.* 4, 414–20.

GARRETT, M., BEVER, T. and FODOR, J. (1966). The active use of grammar in speech perception. *Percept. Psychophysics* 1, 30–2.

GARRETT, M. and BEVER, T. (1969). The perceptual segmentation of sentences. In BEVER, T. and WEKSEL, W. (Eds.), *The Structure and Psychology of Language*. New York: Holt, Rinehart and Winston (in press).

JOHNSON, N. (1965). The psychological reality of phrase structure rules. *J. verb. Learn. verb. Behav.* 4, 469–75.

KAHNEMAN, D. and BEATTY, J. (1966). Pupil diameter and load on memory. *Science, N.Y.* 154, 1583–5.

LADEFOGED, P. and BROADBENT, D. E. (1960). Perception of sequence in auditory events. *Q. Jl exp. Psychol.* 12, 162–70.

TREISMAN, A. (1964). Selective attention in man. *Br. Med. Bull.* 20, 12–6.

WUNDT, W. (1929). *Physiologischen Studien*, Vol. 2. Leipzig: H. Engelmann.

Received 23 May 1969

APPENDIX

The Sentences used in Experiments I, II and III

In addition to his wives, the prince brought the court's only dwarf.
They fought tooth and nail, to get past the huge angry crowd.
By making his plan known, Jim brought out the objections of everybody.
Since she was free that day, her friends asked her to come.
If you did call up Bill, I thank you for your trouble.
We asked the mean old man, to be kind to his dog.
That the matter was dealt with fast, was a surprise to Harry.
After the dry summer of that year, most crops were completely lost.
Any student who is bright but young, would not have seen it.