JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOUR **15**, 43–51 (1976)

# Effect of Word Transitional Probability on Phoneme Identification

JOHN MORTON AND JOHN LONG

*Medical Research Council, Applied Psychology Unit, Cambridge, England*

The results of an experiment are reported which show that with identical preceding context, the same initial phoneme targets contained in high transitional probability words were responded to significantly faster than those in low transitional probability words. The result argues for the importance of transitional probability as an independent variable in sentence processing. The finding seriously weakens the conclusions of recent studies which have failed to control for transitional probability and whose results have been taken to support a model of sentence comprehension based on fixed capacity mechanisms operating concurrently at all linguistic levels. An alternative model which restricts the operation of limited capacity to processes higher than phoneme and word identification is proposed.

In a series of experiments, Foss and his collaborators have investigated the time taken for subjects to respond to a target phoneme in a presented sentence. Foss and Lynch (1969) showed that the reaction time to a word beginning with /b/ was affected by the complexity of the surface structure surrounding that phoneme. Thus the RT to *broke* was shorter with right-branching sentences such as *The store sold the whisky that intoxicated the rioter that broke the window* than in self-embedded sentences like *The rioter that the whisky that the store sold intoxicated broke the window*. This was accounted for by a time-sharing hypothesis, it being supposed that some limited-capacity decision-making process is implicated at various linguistic levels, for example, phonological, lexical, surface structure, deep structure. In such a model, the difficulty of the decision at any one level would affect the speed and accuracy of decisions at another level. Since the syntax of self-embedded sentences is more complex, *RT*s to words in such a framework would be longer. Foss (1969) then found that the *RT* to a target phoneme was affected by the frequency of occurrence of the preceding word. In addition, *RT*s were shorter when the target occurred later in the sentence. These results were also interpreted in terms of overlapping analysis mechanisms, and were considered adequate to rule out models in which analyzing mechanisms for the identification of phonological and lexical items do not overlap. Cairns and Foss (1971) also looked at the effects of manipulating the frequency of the word preceding the target. They found an effect on *RT* only when the word preceding the target was an adjective. The frequency of occurrence of nouns or verbs did not affect the *RT*s to words following them. This result was also explained in terms of overlapping mechanisms with some modifications to explain the restriction of the effect to adjectives.

There are two ways of identifying a target phoneme: acoustically; or on the basis of the internal response following recognition of the word containing it. Two lines of evidence indicate the latter more likely. Corcoran (1966) showed that when subjects were asked to cross out all occurrences of the letter *e* in a text, they tended to miss those which were not pronounced. Then Savin and Bever (1970) showed that *RT* to nonsense syllables was faster than *RT* to the initial phoneme of the

Requests for reprints should be sent to Dr. John Morton, MRC, Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF, England.

same nonsense syllables in the same test, a result confirmed with words as the basic unit by McNeill and Lindig (1973) and Foss and Swinney (1973). In this later paper, Foss modified his earlier view that phoneme identification occurs directly and argued that phonemes are detected by decomposing words.

This is not a *necessary* conclusion from the results cited. It could be that phonemes are detected independently of the word recognition process but that the results of detection are not as readily available to the response mechanism. Note that even with independent phoneme detection, information from a word recognition system would also be required to establish that a real target had occurred, since stimuli which contain the target phoneme in nontarget positions must also be excluded.

However, if phonemes are detected by decomposing words, it follows that any factor which delays the recognition of a word will delay the detection of the target phoneme by the same amount, even if all the processes of interest were independent. That is to say, while Foss considers the effects of the complexity of the material preceding the target upon phoneme detection, we feel it essential first to consider the effects of such complexity on the recognition of the target word itself. Such effects could be interpreted in terms of Foss's time-sharing model, but as Savin and Bever (1970) observe, an alternative model in which the recognition of words does not compete directly for processing time is also possible. In such a model, the time taken to recognize a word would be affected by the results of higher-order processes, whether limited or not. Specifically, the higher-order processes provide contextual cues which interact with sensory information, whose availability is spread in time to produce word recognition. The more contextual cues are available the less sensory information is required and so the faster a particular word is recognized. And this, as already stated, should be directly reflected in phoneme detection times.

The crucial test of the effects of the predictability of a target word is to compare *RT* to a target phoneme under the conditions in which the stimulus is identical up to and including the target phoneme but the target word itself differs in its probability of occurrence in the context and so in its ease of recognition (Morton, 1964; Tulving, Mandler, & Baumal, 1964; Rubenstein & Pollack, 1963). Any differences in reaction time found under these conditions can only be attributed to the processing of the word containing the target phoneme. As we have observed, such an account does not need to postulate any central analyzing mechanism, nor does it require that syntactic and phonemic analysis compete directly for limited capacity.

## Design and Method

### *Sentences and Targets*

Eighty sentences in all were used, each 15 words long; 40 containing real and 40 containing dummy targets. The real target sentences were taken from Morton's 1967 population norms for sentence completion.[1] Real targets were initial noun phonemes, 20 of which were plosive /b, d, p, t, k/ and 20 nonplosive /h, l, m, s, w/, all phonemes appearing with equal frequency. Each real target sentence was associated with a word having a low and a high probability of occurrence in the context of the sentence, both of these words having the same initial target phoneme and the same overall frequency of occurrence in the language (mean frequencies per million words following Thorndike & Lorge, 1944, were low 81.0 and high 81.6); e.g., *At the sink she washed a plate/pan, thinking of the time when she was younger. Plate* and *pan* both have an A frequency rating, but very different completion scores—*plate* .23, *pan* .02. The mean transitional probabilities were .365 for high and .001 for low targets. Low and high transitional probability target words did not differ significantly in the mean number of

[1] Available on request from A.P.U.

syllables (high 1.3, low 1.4), number of letters (high 4.7, low 5.0) or in spoken duration (high 440 msec, low 438 msec). Some minor modifications of the completion sentences were necessary, for example, substituting a target word not appearing in the norms (counted as zero completion score) and changing sentence length. In a sentence only the target word began with the target phoneme. Target sentences are given in the Appendix.

The 40 dummy sentences were intended to obscure the fact that all real targets were nouns and never appeared in the last third of the sentences. The dummy sentences were taken from novels and were not noticeably different in construction from real target sentences, for example, *Inside the big reception centre, it was full of desks and chairs round the wall* (phoneme target /b/). The targets were nouns, verbs and adjectives, placed at any point in the sentences and were made up of the same plosive-nonplosive sets as the real target sentences.

*Design*

Two groups of 12 subjects were used. Real and dummy target sentences were identical for both groups as were dummy phoneme target words. Within the real target sentences, for any given sentence, one group received the low and the other the high transitional probability phoneme target word, these being allocated equally and randomly between groups.

*Presentation*

A Ferrograph tape recorder was used to record the sentences on one track and a reaction time trigger for the phoneme target on the second track. The sentences were read with normal intonation by a female voice unconnected otherwise with the experiment and unaware of the target words. Presentation was by means of a loud speaker. Reaction times to target phonemes were automatically recorded by a Modular One computer. The timing of the trigger on the second track in relation to the target phoneme was checked by recording both speech and the trigger on a Mingograph chart recorder. This is illustrated in Fig. 1. The correction times were automatically applied by the computer. Spoken recall of each sentence was recorded for later checking.

*Procedure*

Each trial lasted 30 sec and was made up as follows: trial number announcement (1 sec), target phoneme (1 sec) announced both as a letter *B* and as a sound *buh*, pause (2 sec), sentence (5 sec), spoken recall (21 sec). The subjects were tested singly in a separate cubicle, communication with the experimenter being by two-way intercom. Recorded instructions asked the subjects to push a response button when they heard a word which began with the appropriate target phoneme. In addition subjects were asked to recall the sentence verbatim. Subjects were familiarized with the reaction time response procedure and given five practice trials, including recall. If an anticipatory response occurred or no response was made, trials were rerun at the end of the session. This occurred on 2.3% of trials.

The subjects were all female volunteers from the Applied Psychology Unit panel, who were paid for their participation.

## Results

The mean reaction times in msec, measured from word onset, for all words with target phonemes are shown in Table 1. In the results which follow, only combined *F*-ratios are reported. As recommended by Clark (1973), these are based on separate analyses of variance with both subjects and materials considered as random factors. Since only dummy target words were identical for all subjects, this procedure required separate analysis of the data from real and dummy target sentences. In neither case was the difference between groups significant (*dummy*: Group 1 = 475 msec, Group 2 = 523 msec,

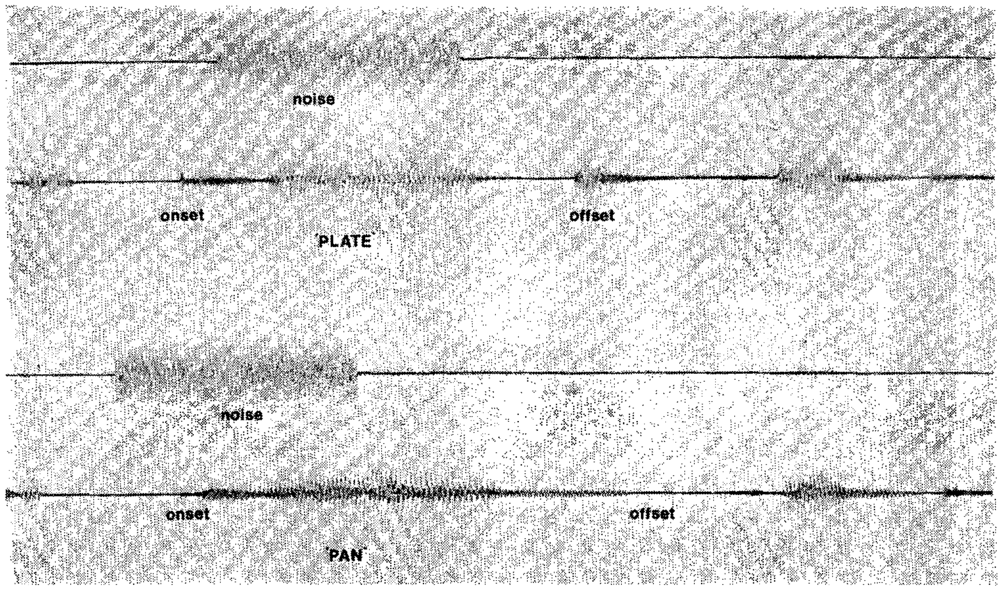


FIG. 1. Mingograph recording of a high (*plate*) and a low (*pan*) transitional probability word containing the same initial target phoneme.

TABLE 1

MEAN REACTION TIMES (MSEC) TO DUMMY AND REAL PHONEME TARGETS

| | Group 1 | | Group 2 | |
|---|---|---|---|---|
| Target | Plosive | Nonplosive | Plosive | Nonplosive |
| Dummy | 453 | 497 | 503 | 543 |
| Real | 423 | 486 | 450 | 500 |

max $F'(1, 33) = 3.48$; *real*: Group 1 = 455 msec, Group 2 = 475msec, max $F'(1, 38) = 1.43$; $p > .05$ in both cases). Plosive targets, however, appear to have been responded to faster than nonplosive targets (*dummy*: plosive = 478 msec, nonplosive = 520 msec, min $F'(1, 52) = 3.34$, $p > .05$, but max $F'(1, 52) = 4.09$, $p < .05$; *real*: plosive = 437 msec, nonplosive = 493 msec, min $F'(1, 52) = 6.53$, $p < .025$). The interaction term failed to reach significance for either class of target. Given the absence of an effect due to groups, any differences due to transitional probability can be interpreted directly.

Since mean duration was shorter for the plosive (411 msec) than the nonplosive (466 msec) target words, and by an amount almost identical with the reaction time differences obtained (duration difference 55 msec; reaction time differences 50 msec), the difference between the two phoneme classes might be best imputed to the duration differences between them and thus between their respective word members. Note that the time measurements were made from the onset of the sound corresponding to the word as shown in Fig. 1 presented earlier. It is arguable that for words beginning with a stop consonant one ought to measure from the beginning of the closure preceding the consonant release. For fluent running speech this would mean, in general, at or just after the end of the preceding word. The justification for such a procedure would be that the silent interval is itself a sensory cue for a stop consonant and thus for a word beginning with a stop consonant. If such a correction is made then the *RT* difference found between plosive and nonplosive targets (also noted by Foss & Swinney, 1973; 58 msec) would be reversed, the mean closure time for our stimuli being 79 msec.[2]

Mean reaction times to real target phonemes only are shown in Table 2. Analysis of variance again showed the difference between groups not to have been significant, max $F' < 1.0$. Again plosive phonemes were responded to significantly more quickly than nonplosive phonemes, min $F'(1, 50) = 7.07$, $p < .01$. Plosive target words were also shorter (plosive, 398 msec; nonplosive, 480 msec). The main finding, however, was that high transitional probability word targets were responded to significantly more quickly than low transitional probability word targets, min $F'(1, 54) = 22.17$, $p < .001$, the difference being 73 msec or an improvement of 15%. Since none of the interactions reached significance, the effect held equally for plosive and nonplosive phoneme target words. All 24 subjects without exception showed the effect. Thirty-five of the 40 sentences containing real targets showed the effect.

When measured from the off-set of target words, the effect of transitional probability was smaller in magnitude. The mean time between the ends of the target words and the response was 5 msec for the high probability targets and 49 msec for the low probability targets. Analysis of variance was equivocal as to the significance of the difference (untransformed data, max $F'(1, 43) = 3.77$, $p > .05$; transformed data, min $F'(1, 30)$, $p < .025$). The direction of the effect, however, was not different from that exhibited by responses measured from word on-set.

TABLE 2

MEAN REACTION TIMES (MSEC) TO PHONEME TARGETS CONTAINED IN LOW AND HIGH TRANSITIONAL PROBABILITY WORDS

| Transitional probability | Group 1 | | Group 2 | |
|---|---|---|---|---|
| | Plosive | Nonplosive | Plosive | Nonplosive |
| Low | 460 | 516 | 491 | 538 |
| High | 386 | 457 | 408 | 462 |

Table 3 shows percentage errors—scored as word omissions, additions, substitutions and transpositions—for real target word sentences. Neither the difference between groups (Mann-Whitney $U$, $p > .05$) nor the difference between low and high transitional probability (Sign test, $p > .05$) target sentences was significant. The reaction time measures can thus be interpreted directly since both groups performed the sentence recall task similarly.

TABLE 3

PERCENTAGE WORD RECALL ERRORS FOR REAL TARGET SENTENCES

| Transitional probability | Groups | |
|---|---|---|
| | 1 | 2 |
| Low | 1.04 | 1.13 |
| High | 1.17 | .83 |

[2] Foss and Swinney report *RT* to /s/ of 481 msec, to /b/ of 431 msec and to /k/ of 415 msec. Their stimuli were lists of words read at .5-sec intervals. Our measurements of such lists reveal systematic variations in onset times which accord well with the discussion (Morton, Marcus, & Frankish, in preparation).

Thus we conclude that high transitional probability target phoneme words were responded to significantly faster than low transitional target phoneme words. The effect did not depend on differences between the two subject groups. The effect was shown for both classes of target phoneme.

## Discussion

Our results show clearly that the time taken to respond to the presence of a target phoneme is affected by the relation between the word containing the target and the preceding context. Our results cannot be explained in terms of the properties of the target words or in terms of differences in ongoing processing at the time of occurrence of the target. It must be, then, that the identification of the target phoneme followed or competed with the word identification which was itself affected by the context. In the introduction, it was stated that the possibility of phoneme identification occurring on the basis of acoustic cues could not be excluded, since increased *RT* to phonemes as opposed to words or syllables might be due to difficulties of access to the results of the identification process. The present results rule out this possibility for a model in which word recognition and phoneme detection are independent. This is not so for a limited capacity model, in which phoneme identification and word recognition compete for time. In the latter case high transitional probability words require less time than low transitional probability words, and would therefore be associated with faster *RT*s.

The results also reveal the predictive nature of speech processing. The mean time between the ends of the target words and the response was only 5 msec for the high probability targets and 49 msec for the low probability targets. It must then be the case that the majority of responses, even to the low probability targets, were initiated before the end of the target word, on the most optimistic estimate of the time between identification and response. And since there is a difference in the response time to targets contained in words of different probability it might well be that these words are identified in the majority of cases, before the acoustic signal has finished. Further, it is necessary that the context has had time to exert its influence before or during the arrival of the target word. A glance at the context sentences in the Appendix will show that in some cases at least the crucial context words occurred not more than three or four syllables before the target. Whatever the mechanism of contextual facilitation of word identification, our results indicate that the entire process must be completed in not more than about three-quarters of a second. This estimate is made up as follows. The response to the high probability targets comes on average about 430 msec after the word onset. Of this we could estimate a conservative 100 msec for the initiation and completion of the response. The remaining 330 msec has to be added to the time taken for the material between the crucial context words and the target. This could be as little as 400 msec in some sentences, meaning that the context has had at most 730 msec in which to have its effect. In further experiments we intend to explore this factor more thoroughly

Note that the result cannot be accounted for by any simple guessing explanation, for guessing would be equally applicable to the High and Low sentences so far as the target was concerned. In Low sentences the subject would guess wrongly, respond equally quickly and only realize later that the response was correct even though the target word was not that expected.

While we have demonstrated that the latency of phoneme identification is affected by the probability of the target word, we would not want to claim that all the previous results could be accounted for on the basis of this variable alone. However it is clear that some of the earlier results are susceptible. To start with we can consider the result of the early Foss and Lynch (1969) experiment. They showed that a target was responded to with longer latency in a self-embedded sentence than in a right branching sentence. The target was the phoneme /b/, in the word *broke* in the two sentences: *The store sold the whisky that intoxicated the rioter that broke the window* (right branching) and *The rioter that the whisky that the store sold intoxicated broke the window* (self-embedded).

Subjectively, real-time prediction is almost impossible with the self-embedded sentence. It is difficult enough to tell whether a sentence such as: *The cow that the architect that the builder commissioned milked designed the house* is sensible or not even when it is written. Whether or not the differences in response time found by Foss and Lynch can be accounted for entirely in terms of the transitional probability of the target word is not clear, but it must be a potent factor. Similarly, the advantage in *RT* enjoyed by targets occurring late rather than early in a sentence shown by Foss (1969), assuming it is not an artefact, could also be explained in terms of their greater predictability. In other experiments it is unlikely to play any role. Foss and Jenkins (1973) found an effect of lexical ambiguity on the detection of a target in the word following. Thus the response to the /b/ in *beside* was longer in the sentence *The child put his straw beside the machine* than in the sentence *The child put his fork beside the machine*. The difference is attributed to the extra processing time required for the ambiguous *straw* compared to the nonambiguous *fork*. It seems unlikely that *beside* would be differentially predictable in the two cases, if measured by the usual procedures of sentence completion. It is, however, possible that the real-time predictability differs in the two cases. This would happen if the mechanisms responsible for the predictions were also those which analyze the semantics of the preceding words. If the predictability differs, then so will the word identification time. This might be enough to account for differences in the phoneme identification times without involving the limited-capacity system in the phoneme identification itself. Equally with a model which is passive at the word level, such as the logogen model (Morton, 1969), the word identification itself would not implicate the limited capacity system except inasmuch as it uses the results of the operation of this system. These relationships are made more explicit in Fig. 2.

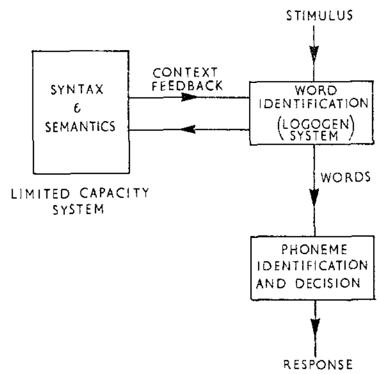


FIG. 2. A model of sentence comprehension in which the range of the limited capacity system is restricted to higher processes leaving word identification independent of capacity requirements. The effect of the contextual feedback provided by the limited capacity system is to speed the recognition of contextually more probable words.

Let us restate the two positions. Foss implicates a limited-capacity system in all the operations involved in sentence comprehension—phoneme identification, word identification, syntactic and semantic processing, and so on. We are suggesting that it is sufficient to restrict the range of operation of the limited capacity system to the higher processing, leaving both phoneme identification and word identification independent so far as capacity requirements are concerned. One function of the higher processes is to provide contextual cues which speed the identification of words. However, the amount of work involved is a function of context (that is of the material being processed) and is not affected by the word actually being recognized by the separate word identification system.

A decision between these alternatives will depend on future, more sensitive experiments. Meanwhile, our basic result, that variations in transitional probability of a target word produce differences in response times of 70 msec must cast doubt on the interpretation of results used to support the notion of limited capacity at all levels of analysis, derived from experiments in which the manipulation of the experimental variables affects the probability of the word containing a target phoneme.

## Appendix

*Sentences Containing Real Targets. The Low Transitional Probability Target Word Appears after the High*

*Plosive Targets*

He sat reading a *b*ook/*b*ill until it was time to go home for his tea.

A sparrow sat on the *b*ranch/*b*ed whistling a few shrill notes to welcome the dawn.

He had a drink of *b*eer/*b*randy in the hope that it would cheer him up.

Passing overhead was a *b*ird/*b*all which stood out starkly against the redness of the sky.

Slowly he opened the *d*oor/*d*ance for it was an occasion which required a little formality.

Quickly he ate his *d*inner/*d*ue; once finished however, he realized he was still very hungry.

On the path was a large *d*og/*d*rink which caused everyone some surprise at the time.

It was a most exciting *d*ay/*d*ress and she felt that her wish had come true.

On the wall was a *p*icture/*p*lant which was unusual for this region of the country.

At the sink she washed a *p*late/*p*an thinking of the time when she was younger.

He was ready to go to the *p*ictures/*p*eople, but was stopped by a guilty feeling.

On the train he looked at his *p*aper/*p*lan trying to remember his friend's old address.

In the sky was a white *c*loud/*c*olour which made the brothers think of winter snow.

Walking in, he took off his *c*oat/*c*lothes, glad to be home out of the rain.

Through the window he saw the *c*ar/*c*rowd suddenly drenched in an unexpectedly heavy spring shower.

Accidentally she broke the *c*up/*c*otton, mainly because she was too tired to do her work.

From a distance she heard a *t*rain/*t*eacher which brought back all her early childhood memories.

She played her favourite *t*une/*t*ragedy again, something which gave her a great deal of pleasure.

One cup was placed on a *t*able/*t*op not far from a group of playing children.

We had a cup of *t*ea/*t*ongue which was all we could offer one remaining survivor.

*Nonplosive targets*

She bought a new *h*at/*h*ome because the old one was now quite past its prime.

They went to see the new *h*ouse/*h*orse which was a birthday present from their father.

She ran quickly down the *h*ill/*h*all to give the bad news to the vicar's wife.

In the bathroom they washed their *h*ands/*h*eads clean from the dirt acquired under the car.

She reached up to dust the *l*edge/*l*ance, but unfortunately slipped and sprained her right ankle.

He bought a book at the *l*ibrary/*l*imit of the village fully intending to read it.

He tripped over a *l*og/*l*oad, cursed and decided never to go round that way again.

Together they had their *l*unch/*l*uxury and talked about the reasons which had brought them together.

Happily they listened to the *m*usic/*m*achine until it was time for them to go out.

Quickly he ate his *m*eal/*m*essage which all foreign spies have to do at some time.

She looked into her *m*irror/*m*ind and decided she could not go through with the deal.

The dog sat on the *m*at/*m*ound in the garden licking its injured paw and whining.

Thoughtfully he chewed a *s*weet/*s*tick but took it out of his mouth again to talk.

Angrily he threw a *s*tone/*s*ail into the pond wetting the young girl next to him.

He looked up at the *s*ky/*s*ign and wondered what he ought to do next.

It was a horrifying *s*ight/*s*cene and one which he dearly hoped not to experience again.

The glass is full of *w*ater/*w*ood put in by the young child during playtime.

She gazed thoughtfully through the *w*indow/*w*ash thinking of the time she felt so much younger.

He jumped over the *w*all/*w*oman, slipped and fell on his back causing himself severe pain.

Happily they listened to the *w*ireless/*w*edding until it became time for them to pack up.

## REFERENCES


CAIRNS, H. S., & FOSS, D. J. Falsification of the hypothesis that word frequency is a unified variable in sentence processing. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 41–43.

CLARK, H. H. The language-as-fixed-effect fallacy: a critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 335–359.

CORCORAN, D. W. J. An acoustic factor in letter cancellation. *Nature*, 1966, **210**, 658.

FOSS, D. J. Decision processes during sentence comprehension: Effects of lexical item difficulty and position upon decision times. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 457–462.

FOSS, D. J., & LYNCH, R. H. Decision processes during sentence comprehension: Effects of surface structure on decision times. *Perception and Psychophysics*, 1969, **5**, 145–148.

FOSS, D. J., & JENKINS, C. M. Some effects of context on the comprehension of ambiguous sentences. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 577–589.

FOSS, D. J., & SWINNEY, D. A. On the psychological reality of the phoneme: Perception, identification and consciousness. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 246–257.

MCNEILL, D., & LINDIG, K. The perceptual reality of phonemes, syllables, words and sentences. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 419–430.

MORTON, J. The effects of context on the visual duration threshold for words. *British Journal of Psychology*, 1964, **55**, 165–180.

MORTON, J. Interaction of information in word recognition. *Psychological Review*, 1969, **76**, 165–178.

RUBENSTEIN, H. & POLLACK, I. Word predictability and intelligibility. *Journal of Verbal Learning and Verbal Behavior*, 1963, **2**, 147-158.

SAVIN, H. B., & BEVER, T. G. The non-perceptual reality of the phoneme. *Journal of Verbal Learning and Verbal Behaviour*, 1970, **9**, 295–302.

THORNDIKE, E. L., & LORGE, I. *The Teacher's word book of 30,000 words.* New York: Bureau of Publications, Teachers College, Columbia University, 1944.

TULVING, E., MANDLER, G., & BAUMAL, R. Interaction of two sources of information in tachistoscopic word recognition. *Canadian Journal of Psychology*, 1964, **18**, 62–71.


(Received July 1, 1975)