Perception & Psychophysics
1974, Vol. 16, No. 1, 182-192

The effect of meaningfulness in tachistoscopic word perception*

LEON MANELIS
Stanford University, Stanford, California 94305

Studies of tachistoscopic word perception were reviewed under two theoretical headings: the structural approach, in which the variables of interest are linguistic relations among letters, and the lexical approach, where the interest is in the availability of words in lexical memory. The results of a recent tachistoscopic recognition study question the importance of lexical availability by finding no difference in performance between meaningful words and well-structured, pronounceable nonwords. In the present study, further comparisons between words and pronounceable nonwords were performed, and a meaningfulness effect was demonstrated. The generality of this finding was discussed, and alternative models accounting for the effect were considered. Two of these were capable of explaining structural effects as well as the meaningfulness effect: a translation model and a lexical discrimination net.

This paper deals with the perception of isolated words that are presented tachistoscopically. Many studies of tachistoscopic word perception can be classified according to two theoretical perspectives, which will be known in this paper as the structural approach and the lexical approach. The two perspectives are distinguished by the types of variables that are considered as determinants of perceptual accuracy. In the structural approach, the variables of interst are linguistic relations among component letters. The letters in a string of random consonants, for example, are relatively unrelated, whereas the letters in a pronounceable nonword are related according to regularities of spelling and pronunciation. These regularities seem to be the same as those that characterize meaningful words. However, structural models make little distinction between meaningful words and well-structured nonwords that contain the same regularities. Perception of a word is said to be facilitated by relations among its component letters, and its presence in the lexicon is considered relatively unimportant.

There are two lines of research that follow the structural approach. The first is the work of Gibson and her associates on spelling patterns (Gibson, 1965; Gibson, Pick, Osser, & Hammond, 1962; Gibson, Shurcliff, & Yonas, 1970). The spelling pattern was originally defined as a cluster of letters having an invariant pronunciation in a given environment within words. The validity of the spelling pattern as a structural unit that facilitates word perception has been supported by evidence showing that items composed entirely of spelling patterns (e.g., GLURCK) are easier to identify than items which violate the regularities of spelling patterns and are less pronounceable (e.g., CKURGL) (Gibson et al, 1962). More recently, the spelling pattern has been redefined in terms of orthographic regularity rather than correspondence to pronunciation (Gibson et al, 1970).

*The author is grateful to Jonathan Baron, Paul G. Matthews, and Edward E. Smith for helpful discussion and for commenting on preliminary versions of the manuscript.

A second line of research following the structural approach is that of Spoehr and Smith (Spoehr, 1973; Spoehr & Smith, 1972, 1973). In their model, strings of letters are said to be parsed into syllable-like units, each composed of a vocalic nucleus and one or more consonants. Identifying the visually presented string involves translating these "vocalic center groups" into a phonological code. In accord with the model, words composed of two vocalic center groups have been shown to be harder to identify than words composed of only one (Spoehr & Smith, 1973). Other effects based on the parsing and translation of phonologically defined units within strings of letters have also been demonstrated (Spoehr, 1973; Spoehr & Smith, 1972).

The other theoretical perspective on tachistoscopic word perception is the lexical approach. Unlike the structural approach, the concern here is not explicitly with relations among the component letters of an item but, rather, with the availability of the whole item in lexical memory. Variables of interest in tachistoscopic recognition include the meaningfulness and familiarity of the items tested and the probability that an item (necessarily a word) is suggested by the context of other words. A common variable explored in studies following the lexical approach is word frequency. Under some circumstances, words that are more frequent in print are identified more readily than infrequent words (e.g., Newbigging, 1961). A model that very clearly exemplifies the lexical approach is Morton's logogen model (Morton, 1969). A logogen can be regarded as a lexical entry; there is a separate logogen for each word in lexical memory. It has a value which is increased by both sensory and contextual cues specific to the word it represents. When the value exceeds a threshold, the word becomes available as a response. High-frequency words are said to have lower thresholds and thus are more available than low-frequency words. Semantic information derived from the context in which a word appears is said to increase the value of its logogen, again making the word more available.

The two theoretical approaches are not contradictory.

It is possible that a complete model of tachistoscopic word perception would require consideration of both structural and lexical factors. However, most recent studies of visual word perception have focused on structural variables, and some evidence has been presented that casts doubt on the validity of the lexical approach. In an extensive set of experiments on tachistoscopic recognition of letter strings, Baron and Thurston (1973) tested structural variables and a crucial lexical variable. Their results showed that pronounceable items following the structure of English spelling are better perceived than unpronounceable items, but they found no difference in performance between words and pronounceable nonwords. The pronounceable nonwords were closely matched to the words; for example, both types of items were composed of acceptable spelling patterns, and both types contained one syllable. Thus, according to structural considerations alone, no difference in performance would be expected. On the other hand, if the presence of an item in lexical memory facilitates tachistoscopic perception, then performance on words should be better than on pronounceable nonwords. Baron and Thurston's finding of no difference between words and pronounceable nonwords implies that in the tachistoscopic task, the presence of an item in the lexicon—its meaningfulness—is unimportant. The main purpose of the present study is to perform further comparisons of words with pronounceable nonwords and thus to provide further tests of the role of meaningfulness.[1]

As Smith and Spoehr (1974) point out in their review of word perception, this kind of test of meaningfulness is especially significant for a lexical theory which they call "feature redundancy theory." Their discussion is based on the work of F. Smith (1971) and Rumelhart and Siple (1974). The central assumption of the theory, as Smith and Spoehr describe it, is that whole words are represented in memory by lists of visual features. The features are the same as those that characterize letters, with the addition of a marking for position. Identification of a word proceeds by matching the set of features extracted from the physical stimulus against the feature lists stored in memory. The result of the process is identification of the stimulus as one of the verbal categories corresponding to the stored feature lists. (Rumelhart and Siple describe a specific decision rule for selecting one of the categories.) It is not necessary to match every feature in a word in order to identify it because the features of a word are redundant. Only a relatively small number of features out of all those present in a word may be needed to distinguish it from other words.

Because of their vast experience with reading, it is reasonable for literate adults to have feature lists in memory for many words. But they are unlikely to have feature lists for nonwords, which rarely, if ever, are seen in print. Feature matches for words would be far more likely than for nonwords. To account for identification of nonwords at all, processes other than feature matching are invoked. F. Smith (1971, p. 180), for example, suggests that nonwords are identified by retrieving phonological features for word fragments contained in the nonwords. Because an immediate feature match will fail for nonwords, and other processing will be necessary, nonwords should be more difficult to identify. It is important to note that this prediction holds even for the most well-structured nonwords, for no matter how closely they conform to orthographic and phonological regularities, nonwords are not represented in lexical memory.

A comparison of words with well-structured, pronounceable nonwords is a critical test for feature redundancy theory. And because this is a test of meaningfulness with degree of structure controlled, it is significant for the validity of the lexical approach in general. In addition to the Baron and Thurston (1973) experiments, three older studies have included manipulations of meaningfulness with degree of structure controlled. The older studies should be interpreted with caution, however, because they all rely on some form of free-report task, which allows S's responses to be influenced by possible report biases. As part of a larger experiment, Postman and Rosenzweig (1956) found no difference in tachistoscopic recognition thresholds for words and pronounceable nonwords. However, most of the words tested were of low frequency and some of the nonwords were meaningful affixes. On the other hand, Gibson, Osser, and Pick (1963) found that tachistoscopically presented words were reported more accurately than pronounceable nonwords by three out of four groups of children. A subsequent study by Gibson and others found that meaningful acronyms (e.g., IBM) had lower recognition thresholds than similarly unpronounceable but meaningless trigrams (e.g., MBI) (Gibson, Bishop, Schiff, & Smith, 1964). However, it is unclear whether words are processed in the same way as acronyms. (The same study also found that meaningless, pronounceable trigrams had lower thresholds than the acronyms, and that a small number of meaningful words had still lower thresholds than the pronounceable trigrams.)

The study by Baron and Thurston (1973) used a forced-choice test and thus was not subject to report biases. The fact that their experiments found no meaningfulness effect but did demonstrate an effect of spelling regularity is of interest in itself, for it suggests that even if a meaningfulness effect does exist, it may not be produced by the same mechanism that causes the spelling regularity effect. Baron and Thurston raise this possibility and at one point conclude that their procedures at least separated the effects (p. 220). They also state that in tasks other than their own, such as reaction-time tasks, additional processes may be involved that allow Ss to use information beyond their knowledge of spelling regularities (p. 226). [Further discussion of this last issue is given in Baron (in press).] In the present study, although some changes in the design features of the Baron and Thurston experiments were made, the task was essentially the same: forced-choice testing of items that are presented tachistoscopically and free of context. The changes that were made were aimed at providing a more powerful test of a possible meaningfulness effect in this task, for some aspects of the Baron and Thurston design could have made the critical difference between words and pronounceable nonwords difficult to detect. Taken individually or in combination, one or more of these design features may have reduced or eliminated the effect.

One aspect of the design was the nature of the stimuli. A small number of words were tested, and they were unusual in two ways. They were mostly uncommon; the median frequency of occurrence in print for the 12 words of Baron and Thurston's Experiment I was 5 per million; for the 16 words of Experiment II, the median frequency was 11 per million (Kucera & Francis, 1967). In a reaction-time task requiring Ss to categorize items as words or nonwords, it takes longer to respond to uncommon words than to common words (Rubenstein, Garfield, & Milliken, 1970; Rubenstein, Lewis, & Rubenstein, 1971a, b). With the brief exposure of a tachistoscopic presentation, lexical access thus may be less likely to succeed for uncommon words, and the possible superiority of words over nonwords would be less apparent. Another unusual feature of the words tested was that most of them (8 in each experiment) were inflected. It may be that inflected words are less available in lexical memory than base forms (Kintsch, 1972a, b).

A second aspect of the Baron and Thurston design was the proportion of words relative to the other stimuli. In each experiment, only one-fourth of the stimuli were words, one-fourth were pronounceable nonwords, and one-half were irregularly spelled, unpronounceable nonwords. The item types were randomly intermixed. If lexical access is a component of word perception which can be induced to occur or not to occur, then the small proportion of real words could have made lexical access unlikely.

A third aspect of the design was the large number of presentations of each stimulus to each S. Assuming that each member of the stimulus set was presented equally often, there would have been 32 tachistoscopic presentations in Experiment I and 16 in Experiment II. Each stimulus was also shown in an undegraded form as feedback following S's response. In Experiment I, the number of presentations was multiplied several times because each stimulus was also a distractor among the test alternatives that were shown following the tachistoscopic display, and because one-third of the distinct words and nonwords were represented twice in the stimulus set. The large number of presentations would have reduced the difference in familiarity between words and pronounceable nonwords that existed at the outset of the experiments. It is possible that Ss developed perceptual strategies specific to the stimuli that were tested. In particular, the stimuli could have been identified by accessing memory for the stimulus set rather than by accessing lexical memory. A related phenomenon occurred in a study of word frequency (Pierce, 1963). When Ss were able to memorize the stimuli before testing, the usual word frequency effect on recognition thresholds was obliterated. Baron and Thurston (1973, p. 220) acknowledge the problem of multiple presentations. They also point out that repetitions did not remove the effect of pronounceability. However, it may be that the influence of repetitions is more destructive to a possible meaningfulness effect than it is to the pronounceability effect.

The present paper reports experiments that incorporated modifications of the design features discussed above.[2] The basic comparison between words and pronounceable nonwords was performed several times because, in light of the Baron and Thurston data, there is an empirical question of the existence of a meaningfulness effect and because the effect is theoretically significant, not only for feature redundancy theory, but for the lexical approach in general.

EXPERIMENT I

The aim of this experiment was to optimize the possibility of finding a meaningfulness effect in the forced-choice tachistoscopic recognition task. Only common words were tested, and a large number were included. Words and pronounceable nonwords were presented on separate blocks of trials in order to maximize the tendency to access lexical memory for words. And each S saw each stimulus only once in order to preclude familiarity with the stimuli before presentation.

Method

Stimuli. The stimuli consisted of 96 four-letter, one-syllable words and 96 four-letter, one-syllable pronounceable nonwords. They are listed in Table 1. The words have frequencies of 50 or more per million (Kucera & Francis, 1967). They were selected to form pairs in which each member differs from the other by only one letter (e.g., BAND, LAND). The two words of each pair served as the test alternatives when either one was presented tachistoscopically. The test was, in effect, a probe of one letter position. For each pair of words, a matched pair of nonwords was formed. The contrasting letters in the nonword pair were the same as in the word pair and were located at the same position. The nonwords differed from the words by one letter (e.g., BANT, LANT). The two members of the nonword pair also served as test alternatives when either nonword was presented. In the nonwords, the letter that differed from the corresponding words was at least two positions away from the probe position. This restriction was followed on the assumption that letters forming structural units are likely to be contiguous. Thus, a unit containing the probed letter position in a word pair was likely to be preserved in the nonword pair. Altogether, there were 48 quartets of matched words and nonwords. Each quartet was probed at a single letter

Table 1
Stimulus Sets for Experiments I and II

Words		Nonwords		Words		Nonwords		Words		Nonwords	
BAND	LAND	BANT	LANT	COLD	HOLD	COND	HOND	NAME	GAME	NARE	GARE
LONG	SONG	LONK	SONK	HAIR	PAIR	HAIM	PAIM	WEST	BEST	WEFT	BEFT
MARK	DARK	MARD	DARD	NEED	FEED	NEEK	FEEK	GAIN	RAIN	GAIR	RAIR
FALL	CALL	FALD	CALD	LACK	BACK	LACT	BACT	TOWN	DOWN	TOWD	DOWD
SPOT	SHOT	SPON	SHON	BALL	BILL	BALP	BILP	BOAT	BEAT	BOAP	BEAP
FARM	FORM	FARN	FORN	WALL	WELL	WALF	WELF	SHOW	SLOW	SHOG	SLOG
HALL	HILL	HALK	HILK	ROAD	READ	ROAK	REAK	WENT	WANT	WENK	WANK
MUST	MOST	MUSP	MOSP	LOST	LIST	LOSH	LISH	FUND	FIND	FUNT	FINT
SHIP	SHOP	THIP	THOP	FILM	FIRM	TILM	TIRM	LOSE	LOVE	FOSE	FOVE
GREW	GROW	FREW	FROW	ROLE	ROSE	WOLE	WOSE	LIFE	LINE	HIFE	HINE
HAND	HARD	TAND	TARD	NICE	NINE	BICE	BINE	FAST	FACT	SAST	SACT
HOLE	HOPE	FOLE	FOPE	THEN	THAN	SHEN	SHAN	MADE	MAKE	HADE	HAKE
WORD	WORK	MORD	MORK	ROOF	ROOM	SOOF	SOOM	SEND	SENT	SIND	SINT
PART	PARK	GART	GARK	HELP	HELD	HALP	HALD	KIND	KING	TIND	TING
THIS	THIN	WHIS	WHIN	PAIN	PAID	BAIN	BAID	FOOD	FOOT	NOOD	NOOT
SEEN	SEEM	FEEN	FEEM	DEAD	DEAR	JEAD	JEAR	HEAD	HEAT	VEAD	VEAT

position, and a different set of 12 quartets was probed at each of the four positions.

Design. The stimuli were divided into two forms, A and B, each presented to a different group of Ss. A form contained only one member of each word pair and one member of each nonword pair from a stimulus quartet. Within a form, only these two stimuli from the quartet were presented tachistoscopically, although both members of each pair were presented as test alternatives. Under this arrangement, for any S, the tachistoscopic presentation of a stimulus was the first time that the stimulus was shown.

The 96 stimuli in each form were divided into two blocks of 24 words and two blocks of 24 nonwords. The stimuli were randomized within each block. For half the Ss presented each form, the blocks were given in the order of words-nonwords-nonwords-words; for the other half of the Ss, the order was nonwords-words-words-nonwords.

Materials and Equipment. The stimuli were constructed by applying dry transfer letters (Letraset Univers 59, 60-pt. uppercase) onto white cards. The test alternatives were typed side by side on white cards with an IBM Executive Registry typewriter having a similar type font. Stimuli and test alternatives were presented in a three-field Iconix tachistoscope. The stimuli subtended a visual angle of 1.0 deg vertically and 2.3 deg horizontally; the test alternatives subtended .2 deg vertically and .9 deg horizontally and were spaced .8 deg apart.

Procedure. In the instructions, Ss were informed of the two types of items that would be presented. They were told that only one type of item would occur during a block of trials, and before each block, the item type was announced. During the first 48 trials of the experimental session, the exposure duration of the stimuli was adjusted to achieve 75% accuracy for each S. The exposure durations ranged from 13 to 38 msec, with a mean of 20.2. The adjustment trials comprised one block of 24 words and another block of 24 nonwords that were different from those presented on the subsequent four blocks of test trials. At the outset of each trial, S viewed a fixation field containing two horizontal gray lines bounding the area where the stimuli would appear. After E's ready signal, S pressed a button to initiate the trial. Half a second after the button was pressed, the fixation field went off and the stimulus appeared for the adjusted exposure duration. The stimulus was immediately followed by a mask (a cross-hatched pattern) in the same area and the test alternatives above it. S chose between the alternatives by pressing one of two response buttons. A response was required on all trials, even if it was a guess, but there was no time constraint. When either response button was pressed, the field containing mask and alternatives went offf and the fixation field reappeared.

Subjects. Twenty Ss were solicited from the Stanford University S pool; all were native speakers of English. Ten Ss were tested with one form and 10 with the other. Each S served for one session of about 45 min.

Results and Discussion

The proportion of correct responses for words was .786; for nonwords, .741. The statistical significance of the difference was assessed by a technique that yields a slight underestimation of a quasi-F ratio (min F') by treating both Ss and items as random effects variables (Clark, 1973; Winer, 1971, p. 375). This technique allows generalization to both a population of Ss and a population of items. (The levels of significance reported for min F's in this paper are more conservative than those associated with the more traditional F tests, which treat either Ss or items, but not both, as random effects.) The only significant effect was meaningfulness; min $F'(1,65) = 3.54$, $p < .07$.[3] There was no effect of forms and no interaction of forms with meaningfulness; min F's < 1. The superior performance on words relative to pronounceable nonwords is contrary to the results of Baron and Thurston (1973).

A noteworthy aspect of the data was revealed by forming differences between the number of correct responses to words and the number of correct responses to nonwords. These differences indicate the magnitude of the meaningfulness effect. They were computed for each S (summing across stimuli) and for each stimulus quartet (summing across Ss). The distribution of differences among the stimuli was unimodal and very nearly symmetrical. The distribution of differences among Ss, however, was bimodal. The differences ranged from −5 (negative meaningfulness effect) to +6 (strong meaningfulness effect). (Forty-eight words and 48 nonwords were presented to each S.) No S had a difference of +2. The mean of the differences among the 12 Ss above +2 was 4.50; the mean for the 8 Ss below +2 was −.88. The bimodality suggests the possibility of dichotomizing Ss into those who do interpret pronounceable letter strings by accessing lexical memory and those who do not. However, it may be that the bimodality resulted from the two forms administered to separate groups of Ss. Of the 12 Ss with differences greater than +2, 8 had Form B, but among the 8 Ss below +2, 6 had Form A. This seems to suggest that the bimodality could be attributed to the items of each form rather than to the Ss; the items of Form B would seem to produce a larger meaningfulness effect. The same forms were used in Experiment II, however, and there Form A, not Form B, produced the larger effect. Nevertheless, the bimodality appeared again in a very similar way. This result will be discussed further.

EXPERIMENT II

It is conceivable that Ss access lexical memory only when words can be expected consistently from trial to trial, as with the blocked presentation schedule in Experiment I. If so, there should be no meaningfulness effect when words and nonwords are randomly intermixed and Ss cannot anticipate which type of stimulus will appear on each trial. The present experiment tests this possibility.

Method

The stimuli were the same as in Experiment I, and they were divided into two forms in the same way. Procedural details were also the same, except that words and nonwords were randomly intermixed. Twenty Ss who had not served in Experiment I were solicited from the Stanford University S pool; all were native speakers of English. Ten Ss were tested with Form A and 10 with Form B.

Results and Discussion

The proportion of correct responses for words was .727; for nonwords, .698. The difference of .029 was smaller than the difference of .045 found in Experiment I. Statistical analysis for the present experiment indicated no significant effects; for the meaningfulness variable, min $F'(1,64) = 1.42$, $p > .10$; for the effect of form and interaction of form with meaningfulness, min $F's < 1$. Because the same stimuli and the same procedure were used in Experiments I and II, a statistical analysis was performed on the combined data. The only significant effect was meaningfulness; min $F'(1,77) = 4.50$, $p < .05$.[4]

Differences between the number of correct word responses and the number of correct nonword responses were computed for each S, as in Experiment I. The distribution of differences was again bimodal. The differences ranged from −3 to +7, and again no S had a difference of +2. The mean of the differences above +2 and the mean of the differences below +2 were very close to the corresponding means found in Experiment I. The mean for the eight Ss above +2 was 4.38 (4.50 in Experiment I); the mean for the 12 Ss below +2 was −.58 (−.88 in Experiment I). Of the 8 Ss above +2, 5 had Form A; of the 12 below +2, 7 had Form B. As noted earlier, this arrangement of forms was opposite to that found in Experiment I, and it suggests that item differences between the forms do not account for the bimodality.

Both Experiments I and II, then, exhibit a similar bimodality, which suggests that some Ss accessed lexical memory in interpreting the stimuli and some Ss did not. These experiments were not designed to investigate individual differences, however, and dichotomizing Ss in general on the basis of the data collected here would be hasty. Nevertheless, the bimodality should be considered in comparing the results of Experiment I with the results of Experiment II. Although Experiment I did show a meaningfulness effect with words and nonwords separately blocked, the effect was smaller and not statistically significant in Experiment II, with words and nonwords randomly intermixed. The discrepancy may indicate a strategy difference. With blocked presentation, it is likely that S will access lexical memory on word trials but not on nonword trials. With randomized presentation, all trials will be approached similarly. Because lexical retrieval would fail on half the trials, S's tendency to access lexical memory would be reduced and words would often be processed as nonwords. In light of the similar bimodality observed in Experiments I and II, however, another interpretation is possible. If some Ss have a tendency to access lexical memory and others do not, it may be that by chance more Ss of the first type were tested in Experiment I than in Experiment II. To rule out this possibility, it is necessary to test the same Ss under both blocked and randomized schedules. This design was incorporated in Experiment III.

EXPERIMENT III

In addition to resolving the question of individual differences, Experiment III was designed to improve the level of statistical significance from the prior experiments by testing each S and each item more thoroughly. More thorough testing was accomplished by presenting all the stimuli in a quartet to each S. In order to prevent familiarity with the stimuli as a result of showing them as test alternatives, only single letters were used as alternatives.

Pronounceability ratings for the items tested were also obtained. It has been shown that pronounceability is a good predictor of tachistoscopic performance for a wide variety of item types (Smith & Spoehr, 1974; Spoehr, 1973). The rating data collected here were used to investigate this relationship.

Method

Stimuli. The 96 words and 96 nonwords were mostly the same as in the previous experiments, but several of the stimuli were replaced in order to satisfy two criteria. First, either letter at the probe position should not occur at any other position. Because the test alternatives were single letters rather than complete

Table 2
Summary of Results for Experiments I Through III

Experiment	I	II	III	
	B	R	B	R
Meaningful Words	.786	.727	.774	.799
Pronounceable Nonwords	.741	.698	.689	.734
Difference	.045	.029	.085	.065

Note–The letters under the experiment numbers denote blocked or randomized scheduling of words and nonwords. The proportions in the body of the table are uncorrected for guessing.

strings, occurrence of one of the alternatives at another position could have biased S to choose that alternative. Second, two items were replaced in order to insure that a letter occurring twice at adjacent positions in a word (e.g., ROOF, ROOM) would always occur in the same way in the corresponding nonword (e.g., SOOF, SOOM). It is possible that these geminate clusters function as units and thus should be preserved in the nonwords.

Design. Each S was tested with 96 blocked stimuli on one day and 96 randomized stimuli on another day. The blocked stimuli were divided into two blocks of 24 words and two blocks of 24 nonwords. For half the Ss, the order of presentation was words-nonwords-nonwords-words; for the other half of the Ss, the order was nonwords-words-words-nonwords. Half the Ss had the session of blocked trials first; the other half had the session of randomized trials first. On each day, there were also 48 warm-up trials, which were blocked or randomized, according to that day's schedule for the test trials. On the first day, the warm-up trials were used to adjust the exposure duration for a level of 75% correct responses for each S. On the second day, the same exposure duration was used. The range of exposure duration was 12.3 to 26.0 msec, with a mean of 15.5.

Materials and Procedure. The materials and the procedure were the same as in the previous experiments, with one exception. The test cards each contained four dashes in a row, corresponding to the four letter positions of the stimuli. Above and below the dash representing the probe position were the two alternative letters. This array subtended a visual angle of .6 deg vertically and 1.4 deg horizontally.

Subjects. Twenty-four Stanford University students served as Ss for two sessions of about 45 min each; all Ss were native speakers of English.

Pronounceability Ratings. The 192 test stimuli presented tachistoscopically were also rated for pronounceability by a separate group of 24 Stanford University students. The stimuli were randomized and presented on two sheets of paper. The instructions asked Ss to pronounce each item to themselves and decide how easy or difficult it was to do that, relative to the other items, on a 5-point rating scale. It was emphasized that the ratings should reflect only the relative ease of pronouncing the items and not their meaningfulness.

Results

Tachistoscopic Test. The main results are shown in Table 2. Words were perceived more accurately than nonwords, and the size of the effect seems to be larger under blocked presentation than under randomized presentation. Unexpectedly, blocking produced poorer performance than randomization. Statistical analysis indicated that the effect of meaningfulness was highly significant; min $F'(1,58) = 13.22$, $p < .001$. Contrasts testing the difference between words and nonwords were also significant within the randomized condition and within the blocked condition; respectively, min $F'(1,156) = 6.55$, $p < .025$, and min $F'(1,156) = 10.90$, $p < .005$.

Pronounceability Ratings. Words were rated more pronounceable than nonwords. On the 5-point scale, the mean rating for words was 3.36; the mean rating for nonwords was 2.91. The difference was highly significant; min $F'(1,34) = 18.03$, $p < .001$. The correlation between tachistoscopic accuracy and mean pronoucneability ratings for each word and nonword was also significant; $r(190) = .16$, $p < .05$.

Discussion

The superiority of words over pronounceable nonwords was clearly demonstrated. It is important to note that in the present experiment words were perceived significantly better than nonwords under the randomized condition. This result indicates that the meaningfulness effect is present when S cannot anticipate which type of stimulus will be presented and thus cannot use different strategies for processing words and nonwords; the same system that processes words should account for poorer performance with nonwords.

As in the previous experiments, the magnitude of the meaningfulness effect was larger under blocked presentation than under randomized presentation. The difference in the magnitude of the effect was obtained within Ss, suggesting that the difference between Experiments I and II was not due solely to variation among Ss and that it can be attributed to the influence of the presentation schedule. The effect of presentation schedule would be a reduction of the tendency to access lexical memory in the randomized condition. This interpretation should be considered tentative, however, because the interaction of presentation schedule and meaningfulness was not statistically significant.

The bimodality observed in Experiments I and II may still indicate variation among Ss in the tendency to access lexical memory. However, the variation would not be independent of task conditions; more Ss may be induced to access lexical memory under the blocked schedule than under the randomized schedule. In the present experiment and in other experiments conducted by the author, distributions of the differences between accuracy scores for words and nonwords were examined, but the distributions were not consistently bimodal, and most Ss did show the meaningfulness effect. Thus, although Experiments I and II raise the possibility of S variability in modes of processing pronounceable letter strings, when considered together with the rest of the data, the evidence is only suggestive.

As shown in Table 2, the size of the meaningfulness effect in the present experiment was considerably larger for both blocked and randomized conditions, as compared with the corresponding figures for Experiments I and II. The only major difference from the prior experiments was the use of single letters rather than complete strings as test alternatives. It is not clear why this change should have amplified the effect; with either form of test alternative, only a single letter position was probed. The form of the test should be unimportant when a stimulus is identified completely, for an accurate representation of the stimulus would be available in short-term memory and could be probed in a variety of ways, all of which would yield correct information. Thus, either form of test should be valid as a measure of complete identification of the stimuli. When only a fragmentary representation of a stimulus is available, however, perhaps it is matched more accurately with test alternatives that are complete strings rather than single letters.[5] Fragmentary representations are presumably more common for nonwords than for words. Thus, the form of the test alternatives would affect performance on nonwords more than performance on words, and the size of the meaningfulness effect would be larger when the alternatives are single letters.

The data from the pronounceability rating task indicate that visually presented words are judged more pronounceable than nonwords and that the ratings correlate with accuracy scores. Spoehr (1973) has also found that words were rated more pronounceable than pronounceable nonwords. A related result has been obtained by Baron (in press), who found that in a task requiring same-different judgments based on the sound of visually presented items, words produced faster reaction times than pronounceable nonwords. Similarly, Forster and Chambers (1973) found that the latency for pronouncing a word is less than that for a nonword. These results indicate that in some sense it is easier to generate a phonological code for visually presented words than for nonwords. This idea will be discussed below in greater detail as a possible explanation for the meaningfulness effect.

GENERAL DISCUSSION

The purpose of this study was to compare performance accuracy in perceiving tachistoscopically presented words and pronounceable nonwords. The earlier experiments of Baron and Thurston (1973) did not detect a difference between the two kinds of stimuli, but it was suggested that the conditions of those experiments were unfavorable to lexical access: the words were unusual, most of the items presented were nonwords, and the items were repeated several times. In the present experiments, these conditions were changed: the words were common, no more than half of the items were nonwords, and they were not repeated. Under these conditions, superior performance was observed for words in several experiments. (The effect has also been replicated by the author in several experiments not reported here.) The fact that words, which are represented in lexical memory, are perceived more accurately than structurally matched nonwords, which are absent from lexical memory, agrees with feature redundancy theory and supports the validity of the lexical approach in models of tachistoscopic word perception.

Table 2 summarizes the results that were obtained. It is notable that the magnitude of the difference between words and nonwords was larger when the stimuli were separately blocked than when they were randomly intermixed. Although the interaction between meaningfulness and presentation schedule was not statistically significant, the direction of the results was consistent and can be noted in Experiments I and II and in Experiment III. (Experiments I and II tested the same stimuli, and both experiments incorporated complete words and nonwords as test alternatives. In Experiment III, the same stimuli and the same Ss were tested under both blocked and randomized schedules, and the test alternatives were single letters.) This pattern of results is in keeping with the effectiveness of lexical access; under the randomized schedule, Ss would be less likely to process stimuli as words. Because the interaction between meaningfulness and presentation schedule was not statistically significant, however, further tests are needed in order to draw a firm conclusion. The appropriate experiments would involve manipulating S's expectancy for real words, as in the paradigm of Aderman and Smith (1971) or by varying the proportion of words intermixed with nonwords.

Generality of the Results

The meaningfulness effect would be less likely to hold under conditions unfavorable to lexical access, such as those discussed with reference to the Baron and Thurston (1973) study. Two other studies have included similar conditions and have also found no difference in performance on words and nonwords. In a study by Bjork and Estes (1973), words and nonwords (anagrams of the words) were presented tachistoscopically and tested by a two-alternative forced-choice probe of one letter position. The alternatives, however, were the same two letters for every stimulus presented. Thus, the task was detection of one of the two possible single letters. In order to perform the task, Ss need only have tested for the presence of those letters rather than identify the display as a whole. In a similar study by Massaro (1973, Experiment II), words and nonwords (consonant strings) were presented tachistoscopically and tested by a four-alternative forced-choice probe of one letter position. The position and the four alternatives were the same on every trial. Again, there was no difference in performance for words and nonwords, and the same interpretation can be made: the task involved detection of one of a small set of known alternatives, and there was no need to identify the whole display.[6]

In both of these studies, the purpose of the testing procedure was to control for S's knowledge of the constraints which characterize letter sequences in words and thereby isolate perceptual effects. From the viewpoint of the present study, on the other hand, the constraints in the stimuli are of great interest. The question addressed here was what type of constraint does S utilize in processing strings of letters. Previous results have demonstrated that the organization of a string of letters into spelling patterns and syllables facilitates processing (Gibson et al, 1962; Spoehr & Smith, 1973). The present results suggest that the whole word may be an effective unit as well. The findings of Bjork and Estes (1973) and Massaro (1973) suggest that when the task involves detection of one letter from a small set of alternative targets used consistently from trial to trial, the linguistic structure of the string in which the target is embedded has no effect. But when S is not so familiar and well practiced with the targets, as in the present study, linguistic structure does facilitate performance. It is significant that the task in the present study did not specifically require or encourage interpretation of the stimuli as linguistic units, for the test following stimulus presentation probed only a single letter position. The meaningfulness effect implies that the stimuli were linguistically interpreted nevertheless. Whether the effect should be considered perceptual is a moot point, because the term "perceptual" is not well defined. It may be that the locus of the effect is subsequent to what is usually considered a basic perceptual process: the extraction of visual features from the physical display. The results of Bjork and Estes (1973) and Massaro (1973) suggest that linguistic variables do not affect feature extraction, for if they did, it seems likely that even the detection task would be influenced by them. Accordingly, it may be more accurate to consider the meaningfulness effect and other linguistic effects as the result of encoding processes operating on the feature input. The feature input itself would be degraded very rapidly. Thus, when a string of letters is only briefly available, encoding processes would be utilized to preserve stimulus information in a form that cannot be so readily degraded. The encoding processes would operate in a way that captures more of the stimulus information necessary to choose between probe alternatives when the stimuli are words than when they are nonwords. Possible forms for these processes will be considered in the following section.

Alternative Models Accounting for the Meaningfuless Effect

A model that accounts for the meaningfulness effect should also account for the structural effects that have been obtained by others. Previous theoretical discussions have focused on either the structural approach (Gibson, 1965; Gibson et al, 1962, 1970; Spoehr & Smith, 1973) or the lexical approach (Morton, 1969; F. Smith, 1971) without integrating the two. In the following discussion, four types of models that could account for the meaningfulness effect are presented, and their ability to account for previously obtained structural effects is considered as well.

Critical Substring Units. Although the words and nonwords in this study were closely matched, they did differ at one letter position, thus making the letter clusters of the words different from those of the nonwords. The nonwords were all well-structured, pronounceable strings, but conceivably, the letter replaced in the words to form the nonwords could have disrupted certain substring units that facilitate word perception. This idea is very similar to the spelling pattern hypothesis. It is unsatisfactory as an explanation of the meaningfulness effect, however. All of the nonwords in this study seem to contain acceptable spelling patterns; it is thus not clear what the critical substring units are or how they could be enumerated. As an approximate measure of the extent to which the stimuli contained possible critical units, summed frequencies of the bigrams and trigrams in each word and nonword were computed. The norms that were consulted give frequencies for bigrams and trigrams at specific positions within four-letter words (Mayzner & Tresselt, 1965a, b). The bigram and trigram frequencies were greater for words than for nonwords. However, the frequencies were unrelated to accuracy scores from the tachistoscopic task. For each of the words and nonwords in Experiments I and III, Pearson product-moment correlations between frequencies and accuracy scores were computed, and none was significantly greater than zero. For words, the correlations with bigram frequency were .02 in Experiment I and .10 in Experiment III; the correlations with trigram frequency were .03 and .02. For nonwords, the correlations with bigram frequency were −.14 in Experiment I and −.16 in Experiment III; the correlations with trigram frequency were −.43 and −.18. For words and nonwords together, the correlations with bigram frequency were .03 and .02; with trigram frequency, .00 and .01. Thus, the critical substring hypothesis was not supported.

Feature Redundancy Theory. According to this theory, as discussed above, words are identified by matching visual features extracted from the stimulus against feature representations of words in lexical memory (Smith & Spoehr, 1974). For nonwords, an immediate feature match will fail, and other processing will be necessary. Feature redundancy theory thus predicts that words will be more accurately perceived than nonwords, and the theory is confirmed by the results of this study. As Smith and Spoehr point out, however, the prediction is only that stimuli represented in memory will be perceived better than those that are not. The basic theory does not consider structural differences among nonwords (or among words). Thus, the theory at least needs to be amplified.

Phonological Encoding. Many of the structural effects that have been obtained can be summarized as a general effect of pronounceability. Pronounceability has been measured by having Ss rate items that are presented visually, as in Experiment III of the present study (also Gibson et al, 1962, 1970; Spoehr, 1973). Strings that are rated easier to pronounce are also easier to perceive when presented tachistoscopically (Spoehr, 1973). For example, pronounceable nonwords are perceived more accurately than nonwords consisting of a vowel surrounded by unacceptable consonant clusters, and these items in turn are perceived more accurately than strings of consonants. The rating data from Experiment III indicate that words and pronounceable nonwords fit into the overall pattern; words are more pronounceable as well as more perceptible than pronounceable nonwords. The correlation between pronounceability and perceptibility suggests the possibility that, in the tachistoscopic task, S attempts to encode the stimuli phonologically. The model of Spoehr and Smith is based on this principle; a visually presented letter string is said to be processed by mechanisms that ultimately translate the stimulus into a phonological code (Spoehr, 1973; Spoehr & Smith, 1973). The model has been supported by several results. For example, two-syllable words were found to be perceived less accurately than one-syllable words (Spoehr & Smith, 1973), and words containing more letter-sound correspondences were found to be perceived less accurately than words of very similar structure (e.g., STARK vs SHARK) (Spoehr, 1973).

The meaningfulness effect can be accommodated within a translation model. The fact that the syllable effect and the effect of number of letter-sound correspondences have been demonstrated with real words suggests that words are at least partially translated. However, a complete translation, based on the application of letter-sound correspondences to all the letters in a word, would be unnecessary. Instead, a phonological code could be retrieved from lexical memory, thus shortening the translation process. A partial phonological code generated by letter-sound correspondences could serve as a cue for retrieving the complete code from lexical memory. This type of translation process for words could be more efficient than the translation of nonwords because of the reduction of the number of letter-sound correspondences applied to words. A prediction arising from this explanation is that phonological variables—such as number of syllables or number of letter-sound correspondences represented in an item—should have less effect on words than on nonwords because of the abbreviated translation.

Lexical Discrimination Net.[7] Another way of summarizing many of the results that have been obtained is with the generalization that strings of letters are perceived better to the extent that they approximate the structure of real words. For example, words contain vowels and certain consonant clusters. When a well-structured, pronounceable nonword is altered either by removing the vowel or by substituting unacceptable consonant clusters, its perceptibility is reduced, and when both alterations are made, perceptibility is reduced further (Gibson et al, 1962; Spoehr & Smith, 1972). The pattern of results like these indicates that the effects could be accounted for by a mechanism that processes words most efficiently and fails to the extent that strings of letters differ from words. A mechanism that would satisfy this criterion is a discrimination net for identifying words on the basis of feature tests for specific letter positions.

In a discrimination net, a series of tests are performed on an input stimulus in such a way that the outcome of a given test determines which tests are made subsequently (Feigenbaum, 1963; Hintzman, 1968). As a word-perception device, the discrimination net could accept as input a set of visual features, with the features marked for position within the letter string. The net would be organized to take advantage of the feature redundancies of real words. At any point within the net, tests would be made for the presence of those features that distinguish the input from other words. After a series of such tests, a word would be uniquely identified. For example, if features for the letters –OST in a four-letter word had already been identified, subsequent tests would be directed toward distinguishing only among C, H, L, M, or P at the first position, for these are the only possibilities that would form a word.

Nonwords could not be processed as efficiently in the discrimination net because they do not exhibit the same constraints as words. To account for the identification of nonwords, two assumptions seem necessary. First, the tests should be able to indicate that a feature whose presence was being tested was in fact not there. This would stop the sorting process and prevent nonwords from being identified as words. Second, the results of completed tests should be retained so that when a nonword is stopped within the net, the information accumulated to that point would be available. Further processing would not be able to exploit the redundancies of real words. The perception of nonwords, then, would be facilitated to the extent that they can be sorted through the net. Nonwords that approximate the structure of real words would be sorted further through the net and would be identified more accurately. Words themselves would be sorted completely through the net and so would be identified more accurately than nonwords.

The basic characteristic of a discrimination net is that the outcomes of certain tests determine which tests are made subsequently. This basic idea has been suggested by Wheeler (1970) to account for his and Reicher's (1969) finding that words are more perceptible than single letters, and it has been suggested by Aderman and Smith (1971) as a mechanism to explain the spelling pattern effect. Smith and Spoehr (1974) have raised an objection to this idea, however, in light of a recent finding. Shiffrin and Gardner (1972) demonstrated that simultaneously presented visual characters are detected as accurately as sequentially presented characters, suggesting that feature extraction is not subject to attentional control. Smith and Spoehr point out that this argues against the directed testing that is characteristic of a discrimination net. In the present discussion, however, the discrimination net is assumed to operate on the set of features already extracted from the display by prior mechanisms. It would function to identify a word when given a feature input, and thus, rather than being a mechanism for feature extraction, it would function as a retrieval device for lexical memory. By embodying the redundancy among the features of real words, the discrimination net could process nonwords to the extent that they exhibit the same constraints. Thus, the same mechanism that would serve to retrieve words from lexical memory may also be able to account for a variety of structural effects. In order to support this mechanism as a plausible model, it would be necessary to construct a discrimination net based on the structure of real words and to simulate structural effects for nonwords.

Conclusion. Of the four models tentatively considered as explanations for the meaningfulness effect, two remain as viable alternatives: the phonological encoding model and the lexical discrimination net. At present, there are no results on which to base a decision between the two models. There are some drawbacks associated with each, however, and the problems seem somewhat less serious for phonological encoding than for the discrimination net. In the phonological model, S generates a phonological code for the visually presented stimulus. It has been shown, however, that deaf Ss report pronounceable nonwords more accurately than stimuli that violate acceptable spelling patterns (Gibson et al, 1970). If this result is not simply an artifact of report biases, a special model that is not based on phonological processes would be needed to explain its occurrence among deaf Ss. In the discrimination net model, as it has been sketched here, there is no apparent way to account for facilitating effects of semantic context. [Such effects on perceptual accuracy have been demonstrated only with a free-report task, however (e.g., Tulving & Gold, 1963).] It is also not clear whether the structure embodied in a discrimination net could account for the apparently phonological effects of syllables (Spoehr & Smith, 1973) or number of letter-sound correspondences (Spoehr, 1973). Finally, the discrimination net does not provide an explanation of how a pronounceability effect can be obtained in the absence of a meaningfulness effect (Baron & Thurston, 1973), whereas in the phonological model, translation may be able to proceed without lexical access.

REFERENCES

Aderman, D., & Smith, E. E. Expectancy as a determinant of functional units in perceptual recognition. Cognitive Psychology, 1971, 2, 117-129.

Baron, J. Successive stages in word recognition. In S. Dornic and P. M. A. Rabbitt (Eds.), *Attention and performance V*. New York: Academic Press, in press.

Baron, J., & Thurston, I. An analysis of the word-superiority effect. Cognitive Psychology, 1973, 4, 207-228.

Bjork, E. L., & Estes, W. K. Letter identification in relation to linguistic context and masking conditions. Memory & Cognition, 1973, 1, 217-223.

Clark, H. H. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. Journal of Verbal Learning & Verbal Behavior, 1973, 12, 335-359.

Feigenbaum, E. A. The simulation of verbal learning behavior. In E. A. Feigenbaum and J. Feldman (Eds.), *Computers and thought*. New York: McGraw-Hill, 1963.

Forster, K. I., & Chambers, S. M. Lexical access and naming time. Journal of Verbal Learning & Verbal Behavior, 1973, 12, 627-635.

Gibson, E. J. Learning to read. Science, 1965, 148, 1066-1072.

Gibson, E. J., Bishop, C. H., Schiff, W., & Smith, J. Comparison of meaningfulness and pronunciability as grouping principles in the perception and retention of verbal material. Journal of Experimental Psychology, 1964, 67, 173-182.

Gibson, E. J., Osser, H., & Pick, A. D. A study of the development of grapheme-phoneme correspondences. Journal of Verbal Learning & Verbal Behavior, 1963, 2, 142-146.

Gibson, E. J., Pick, A., Osser, H., & Hammond, M. The role of grapheme-phoneme correspondence in the perception of words. American Journal of Psychology, 1962, 75, 554-570.

Gibson, E. J., Shurcliff, A., & Yonas, A. Utilization of spelling patterns by deaf and hearing subjects. In H. Levin and J. P. Williams (Eds.), *Basic studies on reading*. New York: Harper & Row, 1970.

Hintzman, D. L. Explorations with a discrimination net model for paired-associate learning. Journal of Mathematical Psychology, 1968, 5, 123-162.

Kintsch, W. Abstract nouns: Imagery versus lexical complexity. Journal of Verbal Learning & Verbal Behavior, 1972a, 11, 59-65.

Kintsch, W. Notes on the structure of semantic memory. In E. Tulving and W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972b.

Kucera, H., & Francis, W. N. *Computational analysis of present-day American English*. Providence, R.I: Brown University Press, 1967.

Massaro, D. W. Perception of letters, words, and nonwords. Journal of Experimental Psychology, 1973, 100, 349-353.

Mayzner, M. S., & Tresselt, M. E. Tables of single-letter and digram frequency counts for various word-length and letter-position combinations. Psychonomic Monograph Supplements, 1965a, 1, 13-32.

Mayzner, M. S., & Tresselt, M. E. Tables of trigram frequency counts for various word-length and letter-position combinations. Psychonomic Monograph Supplements, 1965b, 1, 33-78.

Morton, J. Interaction of information in word recognition. Psychological Review, 1969, 76, 165-178.

Newbigging, P. L. The perceptual redintegration of frequent and infrequent words. Canadian Journal of Psychology, 1961, 15, 123-132.

Paivio, A., & O'Neill, B. J. Visual recognition thresholds and dimensions of word meaning. Perception & Psychophysics, 1970, 8, 273-275.

Pierce, J. Some sources of artifact in studies of the tachistoscopic perception of words. Journal of Experimental Psychology, 1963, 66, 363-370.

Postman, L., & Rosenzweig, M. R. Practice and transfer in the visual and auditory recognition of verbal stimuli. American Journal of Psychology, 1956, 69, 209-226.

Reicher, G. M. Perceptual recognition as a function of meaningfulness of stimulus material. Journal of Experimental Psychology, 1969, 81, 274-280.

Rubenstein, H., Garfield, L., & Milliken, J. A. Homographic entries in the internal lexicon. Journal of Verbal Learning & Verbal Behavior, 1970, 9, 487-494.

Rubenstein, H., Lewis, S. S., & Rubenstein, M. Homographic entries in the internal lexicon: Effects of systematicity and relative frequency of meanings. Journal of Verbal Learning & Verbal Behavior, 1971a, 10, 57-62.

Rubenstein, H., Lewis, S. S., & Rubenstein, M. Evidence for phonemic recoding in visual word recognition. Journal of Verbal Learning & Verbal Behavior, 1971b, 10, 645-657.

Rumelhart, D. E., & Siple, P. Process of recognizing tachistoscopically presented words. Psychological Review, 1974, 81, 99-118.

Shiffrin, R. M., & Gardner, G. T. Visual processing capacity and attentional control. Journal of Experimental Psychology, 1972, 93, 72-82.

Siegel, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.

Smith, E. E., & Spoehr, K. T. The perception of printed English: A theoretical perspective. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition*. Potomac, Md: Erlbaum Press, 1974.

Smith, F. *Understanding reading*. New York: Holt, Rinehart & Winston, 1971.

Spoehr, K. T. Linguistic processes in the perception of letter strings. Unpublished doctoral dissertation, Stanford University, 1973.

Spoehr, K. T., & Smith, E. E. Perceptual units and parsing processes in word perception. Paper presented at the 13th Annual Convention of the Psychonomic Society, St. Louis, 1972.

Spoehr, K. T., & Smith, E. E. The role of syllables in perceptual processing. Cognitive Psychology, 1973, 5, 71-89.

Tulbing, E., & Gold, C. Stimulus information and contextual information as determinants of tachistoscopic recognition of words. Journal of Experimental Psychology, 1963, 66, 319-327.
Wheeler, D. D. Processes in word recognition. Cognitive Psychology, 1970, 1, 59-85.
Winer, B. J. *Statistical principles in experimental design.* New York: McGraw-Hill, 1971.

NOTES

1. The term "meaningfulness" is used here only to refer to the presence of an item in the lexicon; the sense of the term when it is applied to production of associations (e.g., Paivio & O'Neill, 1970) is not intended.

2. Spoehr and Smith (1972) have also compared words with pronounceable nonwords. In one experiment, words were perceived more accurately, but in a second experiment, there was no difference. As in the Baron and Thurston (1973) experiments, however, no more than one-fourth of the stimuli were meaningful words.

3. The F' itself was slightly larger; $F'(1,65) = 4.06$, $p < .05$. Min F' is simpler to compute than F', which requires breaking down the data according to treatments, items, and Ss. In the present experiments, this type of breakdown also produces the problem of having a single dichotomous observation—a correct or incorrect response—in each cell. Accordingly, min F's were computed throughout the rest of this paper. For the few cases in the subsequent experiments where a nonsignificant min F' was at all close to a significant value, a maximum value of F' was also computed (Clark, 1973). In none of those cases did max F' reach significance. Similarly, in each of the case studies presented by Clark, min F' and max F' were very close and their significance levels did not differ.

4. Because of the bimodality among Ss, nonparametric tests were performed on the data from Experiment I and the data from Experiments I and II combined, treating Ss as the units of observation. Wilcoxon matched-pairs signed-ranks tests (Siegel, 1956, p. 75) indicated that the meaningfulness effect was significant. The probability levels ($ps < .01$, two-tailed tests) were smaller than those for the corresponding analyses of variance.

5. If the representation were an incomplete list of features, for example, a single letter test alternative at the probe position might be incorrectly matched with features extracted from other positions in the original stimulus display. This problem could be reduced with complete strings if letters at the other positions in the test alternative were matched with the feature representation and if these matches prevented false matches for the probe letter.

6. Smith and Spoehr (1973) make a similar argument with reference to other studies.

7. The author is indebted to Paul G. Matthews for suggesting this type of model.

(Received for publication February 18, 1974;
revision received April 8, 1974.)