The role of familiar units in perception of words and nonwords

JAMES L. McCLELLAND
University of California, San Diego, La Jolla, California 92093

and

JAMES C. JOHNSTON
Bell Telephone Laboratories, Murray Hill, New Jersey 07974

This paper investigates the effects of familiarity with whole-word units and letter-cluster units in perceptual encoding of letter strings. Subjects viewed brief, masked presentations of words and pronounceable pseudowords differing in letter cluster frequency. Identification of both display types was compared to control single letters. Perceptual accuracy was indexed by probe forced-choice responses and full verbal reports of the displays. Evidence that familiarity of whole-word units facilitated encoding was mixed but, on balance, favorable. Evidence that familiarity of letter-cluster units facilitated encoding was completely absent. This negative finding is surprising in view of the fact that we did obtain a large advantage of letters in pseudo-words as well as words over single letters. The discussion section considers an alternative to the view that perceivers use detectors for familiar letter-cluster units in the process of forming representations of pronounceable, orthographically regular letter strings.

Since the pioneering studies of Cattell (1886), research into the role of familiarity in perception has focused on the perception of printed words. In early studies, Cattell and others found that subjects were able to report the contents of a brief tachistoscopic display more accurately if words were shown than if the display consisted of unrelated-letter (UL) strings (Huey, 1908; Neisser, 1967). More recently, Reicher (1969) and other have shown that forced-choice identification of stimulus letters is more accurate for letters presented in words than for letters in unrelated-letter (UL) strings or even single letters presented alone (Kruger, 1975; Smith & Spoehr, 1974). Since postperceptual guessing and memory overload interpretations are ruled out by Reicher's testing procedure, it appears that familiarity affects the formation of a reportable representation of the presented stimulus and not merely the overt verbal response (Baron, in press). We will refer to the process of

forming such representations as the perceptual encoding process, and we will consider what perceivers know about words that facilitates perceptual encoding.

Familiarity with whole-word units is not the only kind of knowledge that can facilitate perceptual encoding. A large number of studies of perception of nonwords suggest that accuracy of perceptual encoding varies with the degree to which a letter string embodies the structural regularities of English. Different researchers have stressed different structural properties. It has been shown that nonwords are encoded (or at least reported) more accurately if they are high in sequential redundancy (Miller, Bruner, & Postman, 1954; Rumelhart & Siple, 1974; Smith, 1969), if they are orthographically regular (Baron & Thurston, 1973), or if they are made up of various types of subunits such as "spelling pattern units" (Aderman & Smith, 1971; Gibson, Pick, Osser, & Hammond, 1962; Gibson, Schurcliff, & Yonas, 1970) or "vocalic center groups" (Spoehr & Smith, 1975). These various properties tend to be correlated, and all tend to correlate with pronounceability. In addition, the frequency of occurrence of letter clusters is almost certainly correlated with these indices of structural regularity as well.

One hypothesis which might account for these findings states that the perceptual encoding of letter strings is influenced by familiarity with letter groups of various sizes, including whole words, component letter clusters in words, and of course single letters. Indeed, several investigators have suggested that

perceptual encoding of words and nonwords might be accomplished by a system of detectors for familiar word, letter-cluster, and single-letter units (Estes, 1975; LaBerge & Samuels, 1974; Smith, 1971). We call this hypothesis the *familiar units hypothesis.*

A system of familiar unit detectors could operate in a variety of ways that would facilitate perceptual encoding of letter strings. (1) Detection of only a few features might be sufficient for unambiguous activation of the detector for a higher order unit because of the redundancy of English words and letter clusters (Rumelhart & Siple, 1974; Smith, 1969, 1971). (2) Detectors for units of different sizes might operate at least partially independently of each other so that detectors for larger units would provide additional opportunities to pick up stimulus information not encoded by single-letter detectors. (3) Higher-order unit detectors might receive input from features not used for individual letters (cf. Wheeler, 1970). (4) Higher-order unit detectors might be able to maintain their activation in the presence of masking longer than single-letter detectors, permitting sufficient time for recoding of information into a form suitable for overt responding (Johnston, in press; McClelland, 1976).

Our approach is to look for evidence supporting the familiar-units hypothesis which is independent of any specific proposals. Instead, we rely only on the assumption that probability of activating the detector for a given unit is a monotonic increasing function of the frequency of exposure of the unit (Morton, 1969; Smith, 1971). From this assumption, two predictions follow: (1) Strings containing letter clusters of high familiarity should have an advantage over strings containing clusters of low familiarity, and (2) holding cluster familiarity constant, words should have an advantage over pronounceable and orthographically regular pseudowords.

Evidence Concerning Letter Cluster Detectors

Gibson et al. (1962) and Gibson et al. (1970) have shown that letter strings made up only of familiar letter clusters are reported more accurately if the clusters occur in their normal positions in a word. For example, GL at the beginning of GLURCK is reported more accurately than CK at the beginning of CKURGL (see also Aderman & Smith, 1971). To accommodate these findings, it appears necessary to modify the familiar units hypothesis to state that letter cluster detectors are position specific, receiving activation only from displays containing the given cluster in its typical position. Once this modification of the familiar units hypothesis is accepted, the fact that a number of studies (Gibson et al., 1970; Spoehr & Smith, 1975) have failed to find a correlation of accuracy of report and overall letter cluster frequency ceases to provide a source of embarrassment for the familiar units hypothesis.

All three of the studies mentioned which manip-

ulated position-specific letter cluster frequency (Aderman & Smith, 1971; Gibson et al., 1962; Gibson et al., 1970) confounded this factor with the overall orthographic regularity/pronounceability of the letter string. Thus, these studies do not provide direct support for position-specific cluster detectors over and above any other possible explanation of perception of nonword strings.

One way to try to separate position-specific cluster frequency from other potential factors is to use regression analyses. Gibson et al. (1970) tried this approach. They found that string length was a powerful predictor of report accuracy, accounting for 72% of the variance in item scores. An index of position- and word-length-specific bigram frequency in conjunction with string length accounted for 83% of the variance, but rated pronounceability in conjunction with length accounted for 87% of the variance, and inclusion of the bigram frequency index with both string length and pronounceability only added an additional, insignificant 1%. An index of position- and word-length-specific trigram frequency was a poorer predictor than the bigram frequency measure, and did not account for any of the variance after bigram frequency was taken into account. Gibson et al. concluded that pronounceability accounted for all of the variance that letter-cluster frequency accounted for, and more besides.

While suggestive, this pattern of results is hardly conclusive. Several difficulties are associated with Gibson et al.'s (1970) use of the Mayzner and Tresselt (1965) count of the frequency of occurrence of bigrams in running text: (1) The count was based on a sample of only 20,000 words, a small data base for establishing reliable estimates for hundreds of different bigrams in each of scores of position and word-length combinations. (2) If letter-cluster frequency is represented psychologically in a similar manner as word frequency, raw frequencies should have been transformed either logarithmically or exponentially (Catlin, 1971; Oldfield & Wingfield, 1965) to give more weight to differences between moderately frequent clusters and infrequent clusters. (3) Sensitivity of detectors for letter clusters may vary, not with the total number of occurrences of the cluster (token count), but with the number of different words in which it appears (type count; Furth, 1972).

Another way to disentangle cluster familiarity and regularity/pronounceability is to examine effects of bigram frequency within sets of word stimuli or sets of pseudoword stimuli which are pronounceable and regular. In one such study, Biederman (1966) reported a slight report advantage for words containing frequent letter sequences. However, he used a very small number of words, and it is quite possible that his effect was simply due to idiosyncracies of his population of words, since significance levels were not computed over stimuli.

Evidence for Word Detectors

Since Howes and Solomon (1951), a host of studies have shown that accuracy of verbal report of words varies with word frequency (Catlin, 1971; Neisser, 1967). However, this effect has often been attributed to postperceptual guessing (e.g., Pierce, 1963) and the effect may be partially contaminated by structural differences between high- and low-frequency words (Landauer & Streeter, 1973). Evidence for a word advantage over pronounceable, orthographically regular pseudowords has now been obtained by three investigations using Reicher's (1969) forced-choice test to control for guessing biases (Juola, Leavitt, & Choe, 1974; Manelis, 1974; McClelland, 1976), but other investigations which have used the same procedure (Baron & Thurston, 1973; Spoehr & Smith, 1975) have not found consistent word-pseudoword differences. All of these studies attempted to equate for pronounceability and regularity, but several did not equate for cluster frequency. Failure to control for this factor may partially account for the inconsistent pattern of results obtained.

EXPERIMENT 1

Our first experiment attempted to provide improved tests of the two predictions we derived from the familiar-units hypothesis: (1) higher performance on items composed of more familiar letter clusters, and (2) higher performance on words than on regular/pronounceable pseudowords. Our strategy was to manipulate the word-pseudoword factor and high vs. low letter-cluster familiarity orthogonally. Cluster familiarity was indexed by a position- and word-length-specific type count of bigrams in English words, and this measure was cross-validated by a type count and a token count based on the entire Kučera and Francis (1967) corpus of one million words. We did not investigate trigram frequency, but this factor is highly correlated with bigram frequency and did not contribute to the regression analyses of Gibson et al. (1970).

Our experiment also compared perceptual encoding of letters in words and pseudowords to single letters. We used a postdisplay patterned mask known to produce a word superiority over single letters (Johnston & McClelland, 1973). Since single letters and pseudowords have not been previously compared under such conditions, our experiment is the first to test directly whether the word advantage over single letters depends solely on whole-word familiarity.

As dependent measures, we used a verbal report of the letters presented, in addition to the probe forced-choice test of Reicher (1969). The verbal reports provide a more sensitive measure than the forced choice, because of the 50% performance floor for the latter measure. In addition, verbal reports permit an analysis of interdependencies in process-

ing letters in multiletter groups not possible with forced-choice data. It has been suggested that verbal reporting undermines the word advantage over single letters (Mezrich, 1973). However, we will show that forced-choice performance is not disturbed by making a verbal report before viewing the forced-choice alternatives.

Method

Materials

Data were collected from a total of 288 experimental stimuli: 96 words, 96 pronounceable pseudowords, and 96 single letters. The stimuli formed 24 matrices like the one in Table 1. In all matrices, the items consisted of two pairs of four-letter words, two pairs of four-letter pseudowords, and two single-letter pairs. Each pair of items differed by a single critical letter, and the critical letters were the same in all six pairs in the same matrix.

All word stimuli occurred at least once in the Thorndike-Lorge G count (Thorndike, 1944), though some may not have been familiar to all subjects. All the pseudowords were initially judged to conform to English orthography by both authors and an undergraduate assistant, and none occurred anywhere in the Thorndike-Lorge corpus.

One word pair and one pseudoword pair in each matrix were high in bigram frequency (high-BF), while the other pairs were low in bigram frequency (low-BF). Within each matrix, the mean bigram frequency of the high-BF members was at least double the mean frequency of the low-BF members of each matrix. Over all stimuli, the bigram-frequency distributions of words and pseudowords were closely matched. As many of the context letters were kept the same between the matched word and pseudoword pairs as possible, given the interlocking set of constraints on string construction.

The bigram-frequency count used in the selection of the stimuli (Rafferty type count) was derived from a crossword-puzzle dictionary (Rafferty, 1960). Words outside normal English usage (as judged by J.J.) were not counted. The bigram frequencies of the items selected were validated using two objective measures, a type and a token count derived from the four-letter words in the Kučera and Francis (1967) tabulation of one-million words of English text. The high-BF and low-BF stimuli differed by a 3:1 ratio on both type counts and more than a 5:1 ratio (.7 log unit) on the token count, but the mean frequencies for words and pseudowords were very similar at the same level of bigram frequency (Table 2).

Within the word stimuli, the high-BF and low-BF items were matched for number of words occurring more than 100 times per million ("AA" words), number between 50 and 100 times per million ("A" words), and for distribution of actual frequencies of occurrence for less frequent words (Thorndike-Lorge G count).

In order to maximize the difference in cluster frequency between the high-BF and low-BF items in each stimulus matrix, it was sometimes necessary to use a different critical letter position for the critical letter in the high-BF and low-BF pairs. Over the 24 matrices, however, the critical letter occurred in each position six times at each level of bigram frequency. To control for position-specific perceptibility differences against the background of the nonuniform masking pattern used in the experiment, the two

Table 1
One of the 24 Stimulus Matrices Used in Experiment 1

Bigram Frequency	Words		Pseudowords		Single Letters	
High	DILL	DELL	VILL	VELL	_I__	_E___
Low	ITCH	ETCH	ILCH	ELCH	I ___	E____

Note—Underlines were not present in the stimuli but were present in the choice alternatives to denote the critical letter position.

Table 2
Statistics for Bigram Frequency Measures

| | Critical Bigram Frequency[a] | | | | Mean Frequency of All Three Bigrams | | | |
| | Words | | Pseudowords | | Words | | Pseudowords | |
	M	SE	M	SE	M	SE	M	SE
				Rafferty Type Count[b]				
High BF	18.4	1.2	18.9	1.3	23.8	.9	22.7	.7
Low BF	6.9	.9	6.8	.9	8.5	.6	8.2	.6
				K-F Type Count[b]				
High BF	20.2	1.4	20.9	1.4	25.6	1.0	25.1	.8
Low BF	7.4	1.0	8.1	1.1	8.9	.7	9.1	.6
				K-F Token Count[c]				
High BF	2.97	.06	3.00	.06	3.09	.04	3.08	.03
Low BF	2.16	.12	2.18	.13	2.36	.08	2.18	.08

[a]For critical letters in the second and third positions, critical bigram frequency is the mean of the frequencies of the two bigrams containing the critical letter. [b]Number of four-letter-word types containing the bigram in the same position. [c]Log_{10} of the number of four-letter-word types containing the bigram in the same position.

single-letter pairs in each matrix occupied the same positions as the corresponding critical letters in the word and pseudoword stimuli.

Design

The two members of each item pair were each placed into one of two separate lists, each of which was presented to 12 subjects. Within each list, the words, pseudowords, and single letters were kept segregated, and subjects knew which kind of item was being presented. However, the items were arranged in random order with respect to critical letter position and bigram frequency. Each list was then divided into two cycles of three 24-trial blocks, and each cycle contained one block of items of each material type. Order of blocks within cycles was counterbalanced over subjects.

The dependent measures obtained in the experiment were based on verbal reports of the letter(s) presented and forced choices between the stimulus and its pairmate (see procedure). To assess the impact of the verbal report on forced-choice performance, reports preceded forced choice on half the trials in each block (first or second half, counterbalanced over subjects).

Subjects

The subjects were 24 paid University of Pennsylvania students with normal or corrected-to-normal vision.

Visual Conditions

Stimuli were typed on white cards using an IBM Model 12 electric typewriter with carbon ribbon and presented in a two-field Polymetric tachistoscope. At the viewing distance of 39 cm, stimulus letters were approximately .42 deg high and up to .33 deg wide. A four-letter stimulus subtended approximately 1.40 deg. The pre- and postexposure field consisted of a white card with black, pen-drawn contours in a rectangular array approximately 3.4 deg wide and 1.8 deg high. Both curved and jagged contours were present in an irregular pattern with approximately the same grain size as a capital letter (for a replica of the type of mask used, see Johnston & McClelland, 1973). The luminance of the stimulus field was approximately 1.6 log fL and that of the masking field 1.5 log fL, measured by an SEI photometer.

Procedure

At the beginning of each session, the experimenter explained the forced-choice task and defined the types of stimulus material that would appear. Pseudowords were described as strings of letters that were like words without being familiar or meaningful. Each subject was instructed to fixate the middle of the display area and to try to see all the letters presented, giving equal im-

portance to all positions. Trials began with a signal from the experimenter. When ready, the subject pressed a foot switch which started the stimulus presentation after a 300-msec delay. Five seconds after stimulus offset, a click cued the subject to look outside of the tachistoscope and choose between the two alternatives displayed just above the top edge of the tachistoscope at the same viewing distance as the stimuli. Critical letter position was indicated as in Table 1.

On trials in the report condition, the subject gave his report during the interval between the stimulus presentation and the forced-choice test. Each subject was told to report the letter or letters which best represented what he had seen, in the appropriate order. The letters reported did not have to correspond to the type of material presented but did have to contain the correct number of letters.

Each subject viewed 12 blocks of trials grouped into four cycles of three blocks. The first two cycles (20 trials per block) contained a separate list of stimuli and were used for practice and to find an exposure duration at which the subject made close to 75% correct forced choices. Cycles 3 and 4 contained the six 24-trial blocks of experimental items. Each block was preceded by two trials used to absorb transition effects. Exposure duration remained fixed within cycles in Cycles 3 and 4 but was adjusted between these cycles if performance on Cycle 3 was much above or below 75% correct forced choice averaged over all three material types.

Results and Discussion
Replication and Extension of the Word Superiority over Letters

Our experiment replicated the word superiority over single letters obtained previously in the forced-choice test by Johnston and McClelland (1973, pattern mask condition), Reicher (1969), Wheeler (1970), and others. Further, it extended this effect to pseudowords. Subjects were correct on 80% of the trials in the word condition, 78% in the pseudoword condition, and 66% in the single-letter condition. The main effect of material type was reliable both over subjects, $F_s(2,46) = 25.0$, $p < .001$, and items, $F_i(2,94) = 43.1$, $p < .001$. Both words and pseudowords had a reliable advantage over single letters over both subjects and items, $p < .01$ by Tukey's test, but the 2% word-pseudoword difference was not significant ($p > .1$).

Verbal Report Condition

Before turning to a consideration of the familiar-units hypothesis and other substantive issues, we consider three issues concerning the verbal reports.

Influence on forced choice. The 5-sec delay between the presentation of the stimulus and the forced-choice test did not noticeably alter the size of the word-letter difference obtained here (14%) compared to the effect obtained by Johnston and McClelland (1973) under similar verbal conditions but with a 1- to 2-sec delay (15%). Further, there was no effect of the report during the delay interval on the subsequent forced choice. Neither the main effect of report condition on forced-choice accuracy nor the interaction of report condition and material type were statistically reliable either over subjects or over stimuli, all Fs < 1.

Letter rearrangement in free reports. For letters in words and pseudowords, it is possible that subjects sometimes rearrange letters in their reports due to position uncertainty (Estes, 1975). A measure of letter rearrangement is provided by the probability of reporting the critical letter in some position other than the one in which it was presented. This probability was .10 for words, .09 for pseudowords. These probabilities are overestimates of rearrangement, however, because some of the letters scored as rearrangements by this procedure are simply incorrect reports of some other letter. While the exact rate of overestimation is difficult to measure, an estimate is provided by the probability of reporting the incorrect forced-choice alternative (not presented), instead of the critical letter, in some position other than the critical letter position. This probability was .06 for both words and pseudowords. Subtraction of the alternative reports from the critical letter reports gives an estimate of the "true" probability of reporting the critical letter in an incorrect position of only .04 for words and .03 for pseudowords.

Role of guessing on the basis of context. A third issue regarding verbal reports is the potential influence of guessing processes which can operate to fill in reports of letters which are not perceived, using information extracted from other letters and knowledge of contextual constraints (cf. Pierce, 1963). If there is much sophisticated guessing of this sort in verbal reports, accuracy should be higher for letters which are more easily predictable from context than for less predictable letters (Rumelhart & Siple, 1974). However, Johnston (in press) has reported no effect of predictability as measured by several different indices on accuracy of reports of letters in words, under conditions close to those used here. We have repeated these and several other analyses on the report results of the present experiment. Space prevents us from describing these

analyses in detail; suffice it to say that no effects of predictability emerged. Nor did any relationship between report accuracy and any of the predictability measures emerge in a series of regression analyses.[1] It may seem somewhat surprising that we obtained no predictability effects. However, three features of our experiment may have helped reduce the reliance on contextual information to guess unperceived letters. First, the use of postdisplay masking in the present study may limit the time available for formulating a good guess on the basis of fragmentary information (Estes, 1975). Second, subjects were explicitly told not to force their reports to conform to the material type presented. Third, they made their reports letter by letter and did not pronounce the whole stimulus as a unit. Perhaps as a result of these features of the procedure, many of the reports on word trials were not, in fact, words, and many of the reports on both word and pseudoword trials were not pronounceable.

Analysis of the Familiar-Units Hypothesis

The familiar-units hypothesis predicts an advantage for letters in words over letters in pseudowords and an advantage for letters in items containing high-frequency bigrams over letters in items containing bigrams of lower frequency. We now turn to a consideration of these predictions.

Word-pseudoword difference. There was a large word-pseudoword difference in probability of completely correct report (Table 3). This effect was reliable over subjects, $F(1,23) = 42.14$, $p < .001$, and over stimuli, $F(1,47) = 40.28$, $p < .001$. The difference in probability of correct critical-letter report was also reliable over subjects, $F(1,23) = 5.64$, $p < .05$, and over stimuli, $F(1,47) = 8.44$, $p < .01$. The forced-choice difference was in the same

Table 3
Dependent Measures from Experiment 1
Broken Down by Bigram Frequency

	Words	Pseudo-words	Single Letters
Forced Choice			
High Bigram Frequency	.81	.79	.67
Low Bigram Frequency	.78	.77	.64
Mean	.80	.78	.66
Critical Letter Report[a]			
High Bigram Frequency	.62	.50	.26
Low Bigram Frequency	.57	.51	.23
Mean	.60	.50	.24
Completely Correct Report[b]			
High Bigram Frequency	.38	.20	
Low Bigram Frequency	.41	.18	
Mean	.39	.19	

[a]Scored correct only if reported in the correct position.   [b]All four letters reported correctly in the correct position.

direction but was not reliable either over subjects, $F(1,23) = 1.36$, p > .10, or over stimuli, $F(1,47) = 1.58$, p > .10.

A further analysis indicates a relation of word frequency and probability of completely correct report (Table 4). There was a 19% difference between words in the highest frequency quartile (≥ 50 per million) and those in the lowest quartile (≤ 2 per million), $F(1,48) = 9.99$, p < .01. Also, the correlation of probability of completely correct report and word frequency was statistically reliable (r = .26, p < .05).

Bigram frequency effects. Bigram frequency did not appear to affect perceptual encoding of letters in either words or pseudowords (Table 3). The forced-choice advantage of high-BF words over control single letters was .14; for low-BF words, the difference was identical. Similarly, the advantage for high-BF pseudowords over single letters was .12, and for low-BF pseudowords the difference was .13. These differences reveal neither a main effect of bigram frequency nor an interaction with material type, all Fs < 1 over subjects and stimuli. Similar negligible trends are apparent in critical-letter reports. The word advantage over single letters was .36 for high-BF items, .34 for low-BF items; the pseudoword advantage was .24 for high-BF items and .26 for low-BF items. As before, neither the main effect of bigram frequency nor the interaction with material type were reliable; again, all Fs < 1. Single-letter controls do not apply to probability of completely correct report, but here the raw scores speak directly: the 2% advantage of high-BF pseudowords is offset by the 3% disadvantage of high-BF words, and once again neither the main effect of bigram frequency nor its interaction with material type are reliable, $F(1,23) = 1.23$ over subjects for the interaction; all other Fs < 1.

To provide the strongest possible opportunity for an effect of bigram frequency to emerge, we performed an item analysis on the probability of completely correct report of whole bigrams at the extreme ends of the frequency range. Probability of correct bigram report should presumably reflect any real effect of bigram frequency more directly than any other measure since the units in question are being directly assessed. Bigrams falling into the highest and lowest frequency quartiles (Kučera and Francis token count) were selected separately for initial and final letter-pairs in words and pseudowords. This selection procedure produced four groups of bigrams representing two very distinct levels of bigram frequency crossed with material type (word vs. pseudoword). The mean log frequency corresponded to 2,820 occurrences per million words for bigrams in the highest quartile and 43 occurrences per million for bigrams in the lowest quartile. Yet no

Table 4
Upper-Lower Quartile Analyses of Frequency Effects on Completely Correct Report of the Units in Question

	Word	Pseudo-word	Mean
Word Frequency[a]			
Upper Quartile	.49		
Lower Quartile	.30		
Bigram Frequency[b]			
Upper Quartile	.47	.32	.40
Lower Quartile	.49	.33	.41
Individual Letter Frequency[c]			
Upper Quartile	.62	.54	.58
Lower Quartile	.51	.39	.45

[a]*Mean probability of reporting all four letters in the correct position.*
[b]*Mean probability of reporting both letters in the bigram in the correct position.*
[c]*Mean probability of reporting the individual letters in the correct position.*

difference was found between levels of bigram frequency (Table 4). The word-pseudoword effect was highly reliable, $F(1,92) = 21.2$, p < .001, while the effects of bigram frequency and the interaction of material type with bigram frequency were both unreliable, $F(1,92) < 1$ in both cases.

Position-specific single-letter frequency. Although the experiment was not explicitly designed to look for effects of single-letter frequency, various indications in the literature suggested that this variable might be correlated with report and forced-choice performance (Mason, 1975; McClelland, 1976). To determine whether single-letter frequency was correlated with accuracy of verbal report, we performed an additional highest-lowest quartile analysis. The index of single-letter frequency was position- and word-length specific; we simply determined the total number of four-letter word tokens in the Kučera and Francis count containing the given letter in the given position. The single letters in each position falling in the highest and lowest quartile for their material type were included in the analysis. Mean log frequency for the upper quartile corresponded to 20,350 occurrences per million vs. 1,550 occurrences per million for the lowest quartile.

The analysis (Table 4) revealed a striking 13% advantage for letters in the highest frequency quartile over those in the lowest quartile, $F(1,92) = 32.5$, p < .001. A word-pseudoword difference was again obtained, $F(1,92) = 17.9$, p < .001. The interaction of letter frequency with the word-pseudoword factor was nonsignificant, $F(1,92) < 1$.

Interdependencies in Reports

Since we obtained reports of all four letters in our word and pseudoword stimuli, it was possible to look for interdependencies in reports. First, we asked

whether subjects reported 0, 1, 2, 3, or all 4 letters correctly in the proportions we would expect if each letter was perceived and reported independently of all the other letters. Probabilities expected, given independent report of each letter, were calculated from the binomial using the obtained probability of correctly reporting the letter presented in each position. Interdependencies in *arrangement* of letters are forced: if letter i is reported in letter j's position, letter j cannot be reported in the correct position. For this reason, we counted letters correct regardless of position.

The predicted and obtained results for this analysis are shown in Figure 1. For pseudowords, and to an even greater extent for words, there were considerably more reports of all four letters than would be expected assuming independence. In both cases, the deviation of the obtained curve from the expected curve was highly reliable ($\chi_4^2 = 353.68$ for words, 97.51 for pseudowords, both ps < .001).

We next determined the correlation (phi) of report accuracy for all pairs of letters in two different positions (Table 5). For both words and pseudowords, all of these pairwise correlations were significantly greater than zero (smallest value of $\chi_1^2 = 7.35$, p < .01). Correlations were stronger for each pair of positions in words than for the corresponding pair of positions in pseudowords (the smallest value of χ_1^2 obtained was 4.08, p < .05). Statistical comparisons within a material type were difficult because the different correlations were not all independent, but it is apparent that the strength of the correlation between two letters varied with the degree of their separation. In both words and pseudowords, all the correlations between letters in adjacent positions were stronger than all correlations between letters separated by one intervening letter, and these in turn were all stronger than the correlations between letters separated by two intervening letters.

Discussion

Experiment 1 provides some evidence that whole-word familiarity facilitates perceptual encoding. More importantly, the experiment demonstrates directly, for the first time, that the word advantage over single letters does not require familiarity with the stimulus as a word. Since the experiment provides no indication that letter-cluster familiarity facilitates perceptual encoding, some other structural factor appears to be responsible for the "pseudoword-letter" effect as well as the pseudoword advantage over unpronounceable unrelated letter (UL) strings obtained in previous studies (e.g., Aderman & Smith, 1971; Baron & Thurston, 1973; McClelland, 1976).

Experiment 1 also demonstrated strong interdependencies in reports of letters in words and pseudowords. These interdependencies may be important evidence for the encoding of multiletter
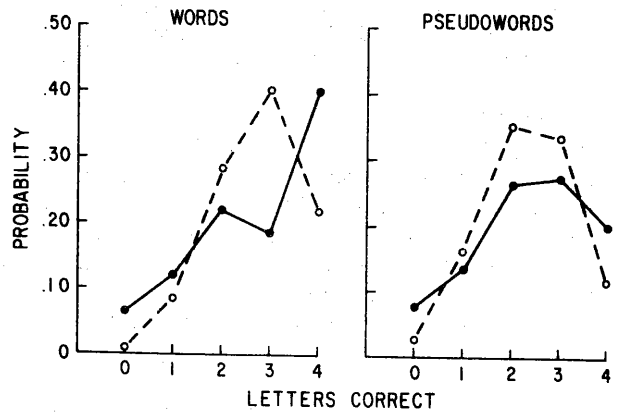

Figure 1. Probability of correctly reporting 0, 1, 2, 3, or all 4 letters correctly (solid line) and expected probabilities given independent letter identification (dotted line) for words and pseudowords, Experiment 1.

groups in orthographically regular stimuli. Alternatively, they may reflect extraneous sources of covariation in reports, such as trial-to-trial variability in attentiveness, subject differences, item differences, and the like.

EXPERIMENT 2

To control for these general sources of covariation in verbal reports, and to examine the importance of orthographic structure in perception of nonwords, Experiment 2 tested performance on unrelated-letter (UL) strings. Word and single-letter stimuli were included to facilitate comparison of performance on the pseudowords used in Experiment 1 with the UL strings used in Experiment 2.

Method

Experiment 2 was identical to Experiment 1 in all respects except that UL strings were used instead of pseudowords and only 12 subjects were tested. The UL strings were constructed using the set of letters occurring in the word stimuli, retaining the critical letters in the same positions. The noncritical letters were selected at random without replacement from the pool of

Table 5
Correlations (Phi) of Report Accuracy for
Pairs of Letters i and j in Experiment 1

Adjacent Letters		Letters Separated by			
		One Other Letter		Two Other Letters	
Words					
1 and 2	.299	1 and 3	.262	1 and 4	.254
2 and 3	.357	2 and 4	.277		
3 and 4	.373				
Mean	.343		.270		.254
Pseudowords					
1 and 2	.153	1 and 3	.148	1 and 4	.113
2 and 3	.177	2 and 4	.130		
3 and 4	.252				
Mean	.194		.139		.113

context letters used in the words, with the restriction that no item could be pronounced without adding sounds not indicated by the actual letters, as judged independently by each author.

Results and Discussion

Giving a verbal report before the forced choice had no detectable effect on forced-choice accuracy, just as in Experiment 1. The main effect of report condition (report vs. no report before the forced choice) was nonsignificant, $F < 1$ over subjects and over stimuli. The interaction of report condition with material type was likewise nonsignificant, $F(2,22) < 1$ over subjects, $F(2,285) = 1.21$, $p > .1$ over stimuli.

Results on the forced-choice, critical-letter report and completely correct report measures are presented in Table 6. For all three measures, the main effects of material type were highly reliable both over subjects and stimuli (smallest $F = 10.9$, $p < .01$).

The word advantage over single letters and UL strings reported first by Reicher (1969) was obtained once again. In the forced choice, our 14% advantage for letters in words over single letters is reliable by Tukey's test over subjects and stimuli, and it is exactly the same size as the difference obtained in Experiment 1. The 10% word advantage over UL strings is also highly reliable using either error term, and it is larger than the effects reported by Baron and Thurston (1973), Reicher (1969), and many others. The word advantages over the other material types showed up in critical-letter reports in considerably amplified form, and the size of the difference in probability of completely correct report between words and UL strings is striking.

The method we used to set individual exposure durations does not permit comparison of absolute performance levels between experiments, but using as benchmarks the performance on the word and single-letter stimuli common to both experiments, we can estimate that the pseudoword-UL string difference would have been about 8% had these two material types been compared directly. A direct comparison by McClelland (1976) under similar conditions produced just such a difference.

The relation between performance on UL strings and single letters is not so clear. The 4% UL advantage in the forced choice is not reliable ($p > .1$ in l.s.d. tests over subjects and stimuli), and another experiment by Johnston (in press) obtained a 2% forced-choice difference in the other direction, also nonsignificant. To add further to the confusion, the probability of reporting the critical letter in the correct position was 8% greater for single letters than for UL strings. This difference was significant by Tukey's test over stimuli, but not subjects.

The best interpretation of these results is that subjects had difficulty arranging letters in UL strings

Table 6
Probability Correct for Three Dependent Measures, Experiment 2

	Words	UL Strings	Single Letters
Forced Choice	.86	.76	.72
Critical Letter Report[a]	.72	.30	.38
Completely Correct Report[b]	.55	.04	

[a] Probability of reporting the critical letter in the correct position.
[b] Probability of reporting all four letters in the correct position.

in their reports but were able to use the identity of letters reported in incorrect positions in selecting the correct forced-choice response. Evidence of difficulty in report arrangement comes from the fact that the probability of reporting the critical letter in an incorrect position was .21, compared to a probability of only .11 of reporting the incorrect forced-choice alternative (not shown) in an incorrect position. Using the difference as an index of true probability of reporting identified letters in incorrect positions (see verbal report section in Experiment 1), the probability of reporting the critical letter in a UL string becomes .40, comparable to the value of .38 obtained for single letters.

Interdependencies in Reports

For UL strings, probability of reporting 0, 1, 2, 3, or all 4 letters was nearly exactly what would be expected given independent report of each letter, $\chi_4^2 = 1.16$, $p > .80$ (Figure 2). Correlational analyses revealed slight marginally significant negative correlations in accuracy of report for adjacent letters and near-zero correlations between nonadjacent letters (Table 7).

In contrast, the word reports revealed the same strong pattern of interdependencies as in Experi-
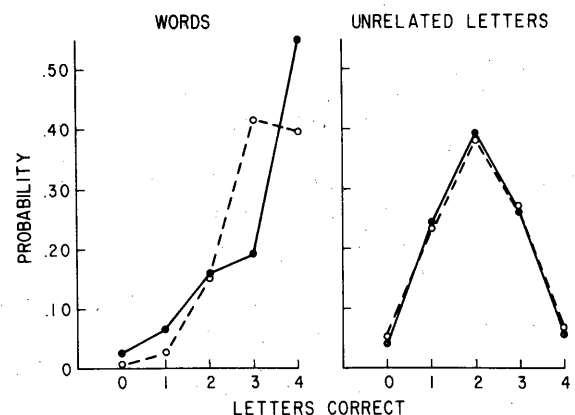

Figure 2. Probability of correctly reporting 0, 1, 2, 3, or all 4 letters correctly (solid line) and expected probabilities given independent letter identification (dotted line) for words and UL strings, Experiment 2.

Table 7
Correlations (Phi) of Report Accuracy for Pairs of Letters i and j
in Words and Unrelated Letter Strings, Experiment 2

| | Adjacent Letters | | Letters Separated by | | | |
			One Other Letter		Two Other Letters	
			Words			
1 and 2	.354	1 and 3	.257	1 and 4	.259
2 and 3	.249	2 and 4	.362		
3 and 4	.398				
Mean	.334		.310		.259
			Unrelated Letter Strings			
1 and 2	−.045	1 and 3	.065	1 and 4	.025
2 and 3	−.112	2 and 4	−.021		
3 and 4	−.056				
Mean	−.071		.022		.025

ment 1. Subjects reported all four letters correctly more often than would be expected assuming independence, and the deviation of the obtained frequencies of reporting 0, 1, 2, 3 and 4 letters correct from the expected frequencies given independence was highly reliable, $\chi^2_3 = 93.3$, p < .001. Further, there were substantial reliable (p < .01) positive correlations in accuracy of report between the two letters in each of the six possible pairs. As before, the correlations tended to be strongest for adjacent letters and weakest for those at the extremes.

Comparison of the pseudoword results from Experiment 1 with the UL results reveals that the correlations for adjacent letter pairs were reliably more positive for letters in pseudowords than for letters in UL strings in all three cases (p < .05). The correlation between the second and fourth letters was reliably more positive in pseudowords than in UL strings as well. The two remaining correlations, between the first and third letters and between the first and fourth letters, both showed a tendency to be more positive in pseudowords, but neither trend was reliable at the .05 level. Apparently, then, interdependencies in accuracy of reports of adjacent letters depend on the orthographic regularity of the environment in which those letters occur. It appears likely, in addition, that interdependencies between nonadjacent letters also depend on the orthographic regularity of the letter string, but the statistical evidence for this is not conclusive.

GENERAL DISCUSSION

Evidence Concerning the Familiar-Units Hypothesis

Our experiments were primarily designed to test two predictions of the familiar units hypothesis: (1) that accuracy of perceptual encoding would be greater for words than for pseudowords, and (2) that accuracy of perceptual encoding would vary with letter-cluster frequency. We also found evidence that bears on another prediction that a familiar units model might make, (3) that single-letter frequency

would be correlated with accuracy. We will first consider the two of these predictions which our results support and then turn to the one they failed to support.

The role of word familiarity. Previous experiments reporting whole-word familiarity effects have suffered from two drawbacks. Either they have used a whole report measure of perception without checking for sophisticated postperceptual guessing processes or, if they have used a forced-choice test, they have tended to find rather small word-pseudoword differences which have not always been statistically reliable, thus enhancing the suspicion that whole report effects are indeed due only to sophisticated guessing. In addition, many studies have not controlled effectively for structural properties. Our results make two relevant points here. (1) We found no indication that position-specific bigram frequency influenced accuracy of reports, so we may set aside this structural factor as a possible cause of word-pseudoword differences when they are obtained. (2) We found large word-pseudoword differences and a word-frequency effect on probability of completely correct reports, a smaller word-pseudoword difference in probability of correctly reporting a single critical letter and an even smaller, nonsignificant, difference in probability of making the correct forced-choice decision.

At first glance, the forced-choice result might seem to support a guessing interpretation of the word-frequency effect and of the word advantage over pseudowords obtained in the verbal reports. However, neither the present investigation nor Johnston (in press) has obtained any indication of a relationship between accuracy of verbal reports and letter predictability, as we would expect from a sophisticated guessing hypothesis (see Johnston, in press, for a full discussion of this point). Similarly, Estes (1975) has found no evidence of a role for guessing in a related paradigm. As he suggests, the use of postdisplay masking may limit the amount of time available for formulating a guess on the basis of fragmentary information.

In view of the fact that we did not obtain any evidence of a role for guessing in our verbal report results, it is worth noting that there are other explanations of the discrepancy between the verbal report and forced-choice results besides those which rely on guessing in the verbal report. Indeed, even if whole-word familiarity has a dramatic effect on the probability of encoding a letter string completely correctly in the absence of any guessing effects, very ordinary considerations would lead us to expect that word familiarity might have a very weak effect on accuracy of the forced-choice response. Familiarity of a stimulus as a whole word would be expected to affect encoding of the whole string correctly but would not be expected to affect encoding of a frag-

ment given that perception of the whole fails. Our results indicate that subjects usually encode fragments smaller than the complete letter string when they fail to encode the entire string correctly. Because of this fact, any measure of perceptual encoding of a single letter in a letter string should reflect influences on encoding the string as a whole much more weakly than a measure of the probability of encoding all of the letters. A correct forced-choice response can potentially be based on encoding of only the critical letter, or even just a fragment of the critical letter, and some correct forced choices undoubtedly result from pure guesses when no information about the critical letter is extracted from the stimulus at all. Thus, the fact that the word advantage in the forced choice was small is not particularly devastating to the view that word familiarity does influence perceptual encoding of our stimuli as wholes. So many investigators have now reported a small, but statistically reliable, forced-choice advantage for words over pseudowords (Juola, Leavitt, & Choe, 1974; Manelis, 1974; McClelland, 1976; Spoehr & Smith, 1975, Experiment 1) that it seems likely that failures to obtain this result (as in the present investigation; Baron & Thurston, 1973; Spoehr & Smith, 1975, Experiment 2) ought to be attributed to the relative insensitivity of the dependent measure.

The role of single-letter familiarity. Both the upper-lower quartile analysis and the regression analysis revealed a significant relationship between single-letter frequency (position and word-length specific) and accuracy of report of letters in both words and pseudowords. Such a relationship has also been found for letters in UL strings by McClelland (1976). Mason (1975), using quite a different task, had found effects of position-specific single-letter frequency as well. These effects support the view that perception of all types of letter strings involves a position-specific letter-analysis process sensitive to the frequency of occurrence of letters in different positions. There is an alternative possibility, however. Because the highest frequency initial and final letters tend to be consonants, while the highest frequency medial letters tend to be vowels, strings containing higher position-specific letter frequencies will tend to be strings containing medial vowels surrounded by consonants. According to the view that subjects encode as single units strings with vocalic nuclei (Spoehr & Smith, 1975), we would expect strings with medial vowels to be more easily encoded than those without. Thus, the single-letter frequency effect might possibly reflect a structural difference between strings rather than a direct effect of position-specific letter familiarity.

The role of cluster familiarity. There are now two studies in the literature which report no position-specific bigram frequency effects within sets of word,

pseudoword, and/or irregular nonword stimuli. Manelis (1974) found no correlation of position- and word-length-specific bigram frequency with probability of correct forced choice. Similarly, in an analysis of reaction-time data obtained in a same-different classification task, Chambers and Forster (1975) found no effect of either bigram or trigram frequency. However, both of these studies relied only on post hoc correlational analyses to reach these conclusions, and it is not clear whether position-specific bigram frequency varied in a wide enough range to show a significant effect. In addition, the dependent measures used may not be the most sensitive: the forced-choice measure of Manelis only shows word-pseudoword effects in some studies, as we have outlined above, and the same-different reaction-time measure has likewise proven fairly insensitive in at least one study (Baron, 1975).

Two previous studies which have used the more sensitive verbal report procedure have encouraged their subjects to make guesses from partial information and have found clear indications of a role for guessing in their results (Rumelhart & Siple, 1974; Smith, 1969). These studies both report correlations between a measure related to bigram frequency ("sequential letter predictability") and accuracy of report. However, the effects could be produced entirely by guessing biases, and they may not even reflect bigram frequency effects per se, since neither experiment attempted to control for orthographic regularity separately from bigram frequency.

Finally, there are a few verbal report studies which report an advantage for low-bigram-frequency words over high-bigram-frequency words when all the stimuli being compared are low in word frequency (e.g., Broadbent & Gregory, 1968; see also Rice & Robinson, 1975). These findings may be interpreted as a reflection of the fact that it is easier to guess relatively unique words from partial cues than it is to guess words which are not unique: Low-bigram-frequency words are more unique, on the average, than high-bigram-frequency words (Rumelhart & Siple, 1974).

In our experiments we have used a verbal report measure as well as the forced choice. Analyses of our verbal report data revealed little role for guessing. As we have already noted, the verbal report provides a very sensitive measure, and the absence of any indication of contamination by guessing biases should make it a very powerful tool for revealing any bigram frequency effects which might be present. Yet our experiments provide no evidence for a role of bigram familiarity. We found no difference between high- and low-bigram-frequency words or pseudowords on any of our dependent measures. It is hard to argue that we simply failed to vary bigram frequency over a sufficiently large range. The ratio

of mean bigram frequency for the high-BF stimuli to the mean for the low-BF items was 3 to 1 on both type-count indices, and there was a difference of more than .7 log units (5 to 1) on the token-count index. Since two of these indices were based on the occurrences of bigrams in a diverse sample of contemporary texts covering more than one-million words, it is unlikely that the counts were unreliable; and it is hard to argue that neither the type count nor the token count is a valid index of the true familiarity of letter clusters. The failure to find an effect of bigram frequency in the upper-lower quartile analysis of completely correct reports of whole bigrams, where the difference in mean bigram frequency between groups was about 1.8 log units (a factor of 60) strengthens the view that there simply is not a bigram frequency effect to be found.

The absence of an effect of bigram frequency poses problems for the hypothesis that the perceptual advantage for letters in pseudowords over single letters depends upon the activation of detectors for familiar subword letter-cluster units. To maintain a letter-cluster detector model of pseudoword perception in the absence of a letter cluster familiarity effect, one might argue that the detectors for all the letter clusters which we tested function asymptotically well because of the extensive reading experience of our subject population. The difficulty with this hypothesis is its failure to deal with the word and position-specific single-letter frequency effects which we did find. Single letters are considerably more frequent, on the average, than bigrams, and if extensive experience produces asymptotically effective detectors, it should have been even harder to find single-letter frequency effects than bigram-frequency effects. A letter-cluster detector model might be saved by pointing to the possibility that apparent single-letter frequency effects are due to some correlated structural factor, but this move would bring us round full-circle. The letter-cluster detector hypothesis was offered to explain the role of English word structure in perception. Now it must appeal to some other aspect of English word structure to account for its own failures.

While further attempts might be made to maintain a familiar letter-cluster detector model in the face of our negative evidence, we will not pursue them here. Instead, we will consider the possibility that subjects might use some mechanism other than detectors for familiar units in the process of perception.

Toward a Model of Word and Pseudoword Perception

In this section, we will outline a model to account for the word and pseudoword advantages over single letters. First we will describe a model which accounts for the word advantage over single letters. Since we found effects of word frequency and a word-nonword difference in accuracy of verbal report, this model will incorporate detectors for familiar words. Since we found an effect of position-specific single-letter frequency as well, the model will also include position-specific detectors for single letters. Then we will extend the model to account for the perceptual advantage of letters in pseudowords over single letters. Since we found no effects of position-specific letter-cluster frequency, our model will present a way to account for pseudoword perception without using detectors for letter-cluster units.

Our model relies on a hierarchical system of detectors for visual feature, letter, and whole-word units (Estes, 1975; McClelland, 1976). In this system, the output of the feature level feeds into the letter level and the output of the letter level feeds into the word level. Detectors at the feature and letter levels are assumed to be position-specific, although there may be cross-talk between positions. We assume that verbal report and forced-choice responses may be based on the output of either the letter level or the word level. Mere activation of a detector, however, is not sufficient for correct responding; some period of detector activation is needed for translation of a detector's output into a response buffer.

Johnston (in press) and McClelland (1976) have both already noted how such a system can account for the advantage of words over single letters. The account depends on the idea that the letter level will be more susceptible than the word level to disruption by the postdisplay feature mask typically used in word-perception studies. Letter detectors will receive direction disruption, since they receive the output of feature detectors activated by the mask. However, the letter detectors are interposed between the feature detectors and the word detectors and so may serve to protect any activation at the word level from the mask's effects. Thus, subjects will have more time to translate words into the response buffer than they will have to translate individual letters. This model permits us to explain the effects of word and single-letter frequency we have observed without resorting to the usual sophisticated guessing accounts of these effects (Rumelhart & Siple, 1974). We need only assume that the rate of activation of a detector, and hence the time available for encoding its output, is an increasing function of its frequency of activation.

This account of the word advantage over single letters relies on detectors for words to provide a source of information above the letter level. To extend the account to pseudowords, it would be natural to postulate detectors for higher order units contained in these stimuli as well. However, pseudowords are not themselves familiar units, and we have no evidence favoring the existence of detectors for

familiar letter cluster units, so we seem to be out of candidates. To extend the model to pseudowords without postulating detectors for higher order units, we may postulate that perceivers do form higher level representations of pseudowords, but they are novel representations constructed from the output of the letter level rather than familiar ones activated by the output of the letter level. Such representations would be relatively protected from the effects of masking compared to representations at the letter level, and would thus allow extra time for the translation of information about pseudowords into a response buffer, compared to the amount of time available for translation of representations of individual letters.

To make this proposal more concrete, we might hypothesize (in the spirit of Spoehr & Smith, 1975) that the representations constructed in this way are sequences of abstract phonological codes which could serve as the input to the process of articulation (Chomsky & Halle, 1968). The process of constructing these codes might be guided by rules specifying what sequences of such abstract phonological codes would actually be pronounceable.

What might such rules actually look like? One possibility is that they might simply list the specific combinations of letters that are pronounceable in each of several different kinds of orthographic contexts (Gibson et al., 1962). Alternatively, the rules might be given in terms of what classes of letters can occur together, where the classes are defined in terms of the features of the abstract phonological segments for which letters stand (Chomsky & Halle, 1968). For example, in word-initial position, one of the acceptable classes of letter clusters consists of a stop consonant (B, P, D, T, G, C, or K) followed by a liquid (R or L). A rule which specified that a letter string could begin with a stop followed by a liquid would therefore account for the orthographic regularity of the initial clusters BR, PR, DR, TR, GR, CR, KR, BL, PL, GL, CL, and KL.[2] It can be shown that rules of this sort can be incorporated into a device which constructs phonological-segment representations of most pronounceable pseudowords, and fails to construct representations of strings which are not pronounceable. The rules in such a system are generally like that stop-liquid rule described above in that they encompass letter-clusters varying widely in frequency of occurrence. Therefore, we would not expect to find much in the way of a correlation between accuracy of perceptual encoding and bigram frequency in a system which made use of such rules.

The hypothesis that subjects construct novel representations of pseudowords permits us to account for interdependencies in reports of letters in pseudowords as well as words. We would expect the formation of a higher level representation of an entire word or pseudoword stimulus to result in correct report of all four letters, while failure to form such a representation should lead to a lower probability of report of each letter. In light of this account, the fact that correlations of letter reports were stronger for adjacent letters than for nonadjacent letters may indicate that subjects often formed representations of small groups of adjacent letters. The fact that there was a reliable correlation in letter reports between the first and last letters in pseudowords as well as words suggests that representations of whole pseudowords were constructed some of the time.

The relationship between the construction of a phonological code for a word and the activation of a detector for the word requires some consideration. It is possible that these processes take place in parallel (Baron, 1974; Frederiksen & Kroll, 1976), so that activation of the word detector and construction of a phonological code would provide two possible sources for encoding whole words, while only one source would be available for pseudowords. This suggestion can be used to explain the fact that subjects report words completely correctly more often than pseudowords, since two sources should be expected to yield a greater chance of success than one.

To further this rule-guided constructive approach, we need a more precise specification of the nature of the constructed representations and of the process of construction itself. Whatever the details of the eventual model which may emerge from such a specification, our finding of a large pseudoword advantage over single letters, in the absence of any letter-cluster frequency effects, suggests that the process of perceiving stimuli which conform to regular structural patterns may be considerably more complex than the mere activation of detectors for familiar units.

REFERENCES

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

BARON, J. Phonemic stage not necessary for reading. *Quarterly Journal of Experimental Psychology*, 1973, **25**, 241-246.

BARON, J. Successive stages in word recognition. In S. Dornic & P. M. A. Rabbitt (Eds.), *Attention and performance V.* New York: Academic Press, 1975.

BARON, J. The word-superiority effect: Perceptual learning from reading. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 5). New York: Halsted Press, in press.

BARON, J., & THURSTON, I. An analysis of the word-superiority effect. *Cognitive Psychology*, 1973, **4**, 207-228.

BIEDERMAN, G. B. The recognition of tachistoscopically presented five-letter words as a function of digram frequency. *Journal of Verbal Learning and Verbal Behavior*, 1966, **5**, 208-209.

BROADBENT, D. E., & GREGORY, M. Visual perception of words differing in letter digram frequency. *Journal of Verbal Learning and Verbal Behavior*, 1968, **7**, 569-571.

CATLIN, J. *The interaction of stimulus information and bias in auditory word recognition.* Unpublished doctoral dissertation, University of Pennsylvania, 1971.

CATTELL, J. M. The time taken up by cerebral operations. *Mind*, 1886, **11**, 220-242.

CHAMBERS, S. M., & FORSTER, K. I. Evidence for lexical access in a simultaneous matching task. *Memory & Cognition*, 1975, **3**, 509-559.

CHOMSKY, N., & HALLE, M. The sound pattern of English. New York: Harper & Row, 1968.

ESTES, W. K. The locus of inferential and perceptual processes in letter identification. *Journal of Experimental Psychology: General*, 1975, **104**, 122-145.

FREDERIKSEN, J. R., & KROLL, J. F. Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, **2**, 361-379.

FURTH, I. *Components of reading disability*. Unpublished doctoral dissertation, University of New South Wales, Kensington, N.S.W., Australia, 1972.

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 & J. P. Williams (Eds.), *Basic studies in reading*. New York: Basic Books, 1970.

HANSEN, D., & ROGERS, T. S. An exploration of psycho-linguistic units in initial reading. In *Proceedings of the Symposium of the Psycholinguistic Nature of the Reading Process*. Detroit: Wayne State University, 1965.

HOWES, D. W., & SOLOMON, R. L. Visual duration threshold as a function of word probability. *Journal of Experimental Psychology*, 1951, **41**, 401-410.

HUEY, E. B. *The psychology and pedagogy of reading*. New York: Macmillan, 1908.

JOHNSTON, J. C. A test of the sophisticated guessing theory of word perception. *Cognitive Psychology*, in press.

JOHNSTON, J. C., & MCCLELLAND, J. L. Visual factors in word perception. *Perception & Psychophysics*, 1973, **14**, 365-370.

JUOLA, J. F., LEAVITT, D. D., & CHOE, C. S. Letter identification in word, nonword and single letter displays. *Bulletin of the Psychonomic Society*, 1974, **4**, 278-280.

KRUEGER, L. E. Familiarity effects in visual information processing. *Psychological Bulletin*, 1975, **82**, 949-974.

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

LABERGE, D., & SAMUELS, J. Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 1974, **6**, 293-321.

LANDAUER, T. K., & STREETER, L. Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 119-131.

MANELIS, L. The effect of meaningfulness in tachistoscopic word perception. *Perception & Psychophysics*, 1974, **16**, 182-192.

MASON, M. Reading ability and letter search time: Effects of orthographic structure defined by single letter positional frequency. *Journal of Experimental Psychology: General*, 1975, **104**, 146-166.

MAYZNER, M. S., & TRESSELT, M. E. Tables of single-letter and digram frequency count for various word-length and letter position combinations. *Psychonomic Monograph Supplement*, 1965, **1**, 13-22.

MCCLELLAND, J. L. Preliminary letter identification in the perception of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, **2**, 80-91.

MEZRICH, J. J. The word superiority effect in brief visual displays: Elimination by vocalization. *Perception & Psychophysics*, 1973, **13**, 45-58.

MILLER, G., BRUNER, J. S., & POSTMAN, L. Familiarity of letter sequences and tachistoscopic identification. *Journal of General Psychology*, 1954, **50**, 129-139.

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

NEISSER, U. *Cognitive psychology*. New York: Appleton-Century-Crofts, 1967.

OLDFIELD, R. C., & WINGFIELD, A. Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, 1965, **17**, 273-281.

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

RAFFERTY, K. *Crossword puzzle dictionary*. New York: Dell, 1960.

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

RICE, G. A., & ROBINSON, D. O. The role of bigram frequency in perception of words and nonwords. *Memory & Cognition*, 1975, **3**, 513-518.

RUMELHART, D. E., & SIPLE, P. Process of recognizing tachistoscopically presented words. *Psychological Review*, 1974, **81**, 99-118.

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*. New York: Halsted Press, 1974.

SMITH, F. The use of featural dependencies across letters in the visual identification of words. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 215-218.

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

SPOEHR, K. T., & SMITH, E. E. The role of orthographic and phonotactic rules in perceiving letter patterns. *Journal of Experimental Psychology: Human Perception and Performance*, 1975, **104**, 21-34.

THORNDIKE, E. L. *The teacher's word book of 30,000 words*. New York: Teacher's College, 1944.

WHEELER, D. D. Processes in word recognition. *Cognitive Psychology*, 1970, **1**, 59-85.

NOTES

1. Other facts worth brief mention which emerged from these analyses include: (1) The presence of curvature in the critical letter was rather strongly correlated with probability of correct report. This held up both for letters in words and pseudowords and for single letters alone. Johnston (in press) found a similar effect in his results. (2) Even when several other factors were controlled, no effects of position-specific bigram frequency emerged, confirming analyses to be presented in the text. (3) Word frequency and position-specific single-letter frequency effects were found, also confirming results presented in the text.

2. The attentive reader will note that the rule given in the text includes TL and DL, which are not orthographically regular in initial position in English. This overinclusiveness of the stop + liquid rule can be blocked by a simple selectional restriction, namely that a stop and a liquid occurring together in word-initial position cannot have the same place of articulation (D, T, and L all share the dental place of articulation). Similar selectional restrictions are needed to block some other clusters which conform to the general rules but which, like TL and DL, are not actually regular.