Journal of Educational Psychology
1980, Vol. 72, No. 6, 730–742

Reading Ability and Utilization of Orthographic Structure in Reading

Dominic W. Massaro and Glen A. Taylor
University of Wisconsin

A perceptual-recognition task was used to assess whether utilization of orthographic structure in letter recognition varies with reading ability. Anagrams of words were made to create strings that orthogonally combined frequency and regularity measures of orthographic structure. These strings and the original words were used as test stimuli in a letter-recognition task. Good and poor college readers showed equally large effects of orthographic structure on task accuracy, whereas poor sixth-grade readers did not utilize orthographic structure to the same degree as very good sixth-grade readers. To facilitate the teaching of orthographic structure, some of the important constraints in written English and various games for teaching these constraints are presented.

One of the primary concerns of educators is the wide range of reading ability across both child and adult populations. Good and poor readers have been shown to differ in performance on a variety of general tasks, but until recently little attempt has been made to explore how children of different reading abilities might differ in basic word-recognition processes. If poor readers reveal fundamental deficits in certain recognition processes, then instruction for the poor readers could be directed at these processes. The rationale for the present research is that the study of basic word-recognition processes can aid in assessing word-recognition abilities of good and poor readers at both the elementary school and adult levels.

The goal of the present research was to determine how the reader's knowledge and utilization of orthographic structure varied with reading ability. Orthographic structure refers to the constraints describing how letters are sequenced in the written language. Previous research (Massaro, Venezky, & Taylor, 1979) established the psychological reality of orthographic structure, and the present objective was to evaluate its contribution to letter processing as a function of reading ability. The framework for the present study was an information-processing model that has been developed and tested over the past few years (Massaro, 1975, 1978, 1979b).

Figure 1 presents a schematic representation of the stages of processing in reading. Given the information-processing model, an important issue is determining the stage or stages of processing responsible for reading difficulty. In the present model, deficits could exist in either the temporary processing stages, the long-term knowledge base, or in both. Feature detection, primary recognition, secondary recognition, and rehearsal–recoding stages might each be responsible for various reading difficulties. In the present study, we concentrated on those processes involved in letter and word recognition.

During an eye-fixation period, the light pattern of the letters is transduced by the visual receptors. A feature-detection process detects and transmits visual features to preperceptual visual storage. The primary

This research was supported by funds from the National Institute of Education to the Wisconsin Research and Development Center for Individualized Schooling.

David Klitzke programmed the experiments, assisted in testing the subjects, and contributed to the data analyses. Peter Lucas and Jim Jastrzembski also made significant contributions to the data analyses.

Glen A. Taylor is now at Bell Laboratories, Homdel, New Jersey.

Requests for reprints should be sent to Dominic W. Massaro, who is now with the Program in Experimental Psychology, University of California, Santa Cruz, California 95064.

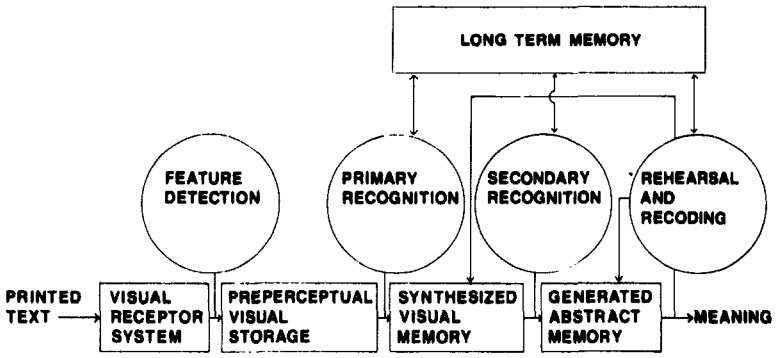

Figure 1. Information-processing model of reading printed text.

recognition process attempts to synthesize the isolated features in preperceptual visual storage into a sequence of letters and spaces in synthesized visual memory. To do this, the primary recognition process can utilize information held in long-term memory, which for the accomplished reader includes a list of features for each letter of the alphabet along with information about the orthographic structure of the language. The primary recognition process evaluates the features in preperceptual visual storage and compares or matches these features with perceptual prototypes in long-term memory. A perceptual prototype represents each of the letters as it would ideally be represented in preperceptual visual storage. The primary recognition process seeks for each letter position the perceptual prototype that provides the best match to the featural information in preperceptual visual storage. This process operates on a number of letters simultaneously (in parallel). Featural information is resolved by feature detection at different rates, and there is some evidence that gross features are available before more detailed features (Breitmeyer & Ganz, 1976; Massaro & Schmuller, 1975).

The primary recognition process is therefore faced with a succession of partial information states. These may be regarded as yielding candidate sets of letters that are successively more consistent with the available featural information for each letter position in the string being synthesized. The reduction of these candidate sets to a single best matching letter (or to no match) relies not only on the incoming visual features but also on a knowledge of orthographic structure. If, for example, an initial *th* has been resolved in a letter string, and the features available for the next letter are consistent with either *c* or *e*, the process might synthesize *e* without waiting for further visual information, since initial *thc* is not acceptable, whereas initial *the* is. Conversely, the actual presence of an irregular string or substring such as *thq* might demand more exhaustive featural information for acceptance of each of the three letters because of the illegality of the sequence. It is assumed that the two sources of information, visual features and orthographic structure knowledge, make independent contributions to the recognition process (Massaro, 1973, 1975, 1979a; Thompson & Massaro, 1973). In this view, orthographic context facilitates word perception, but it does not modify the feature analysis of the printed pattern (Massaro, 1979a).

The primary recognition process transmits a sequence of recognized letters to synthesized visual memory. Figure 1 shows how the secondary recognition process transforms this synthesized visual percept into meaningful form in generated abstract memory. We assume secondary recognition attempts to close off the letter string into a word. The secondary recognition process makes this transformation by finding the best match between the letter string and a word in the lexicon in long-term memory. Knowledge of orthographic structure can also contribute to secondary recognition; word recognition can occur without complete recognition of all of the component letters.

Given the letters *bea* and the viable alternatives *l* and *t* in final position, only *t* makes a word, and therefore word identification (lexical access) can be achieved. Each word in the lexicon contains both perceptual and conceptual codes. The word that is recognized is the one whose perceptual code gives the best match and whose conceptual code is most appropriate in that particular context.

Although Anderson and Dearborn (1952) acknowledged the effects of familiarity and set on word perception in their comprehensive book, *The Psychology of Teaching Reading,* the utilization of orthographic structure is not addressed. Nowhere do the authors explicitly recognize that the perception of some letters can provide information about other letters or that some letters are more likely to occur in some positions and contexts than others. Familiarity and set effects were assumed to occur at the word level or even higher at the phrase or sentence level, but not at the sublexical level. One of the first studies of the utilization of orthographic constraints in reading was carried out by Miller, Bruner, and Postman (1954). These authors had subjects reproduce letter sequences flashed in a tachistoscope. The eight-letter strings represented different approximations to English based on Shannon's (1948, 1951) algorithms. The authors found that performance improved with increases in the degree to which the letter strings approximated English. By correcting for the statistical redundancy in the strings, the amount of information transmitted was shown to be equal for the various approximations. This result is consistent with more recent empirical and theoretical work demonstrating that orthographic structure provides an independent source of information to the reader (Massaro, 1979a). Massaro (1980) provides a review of how orthographic structure facilitates reading and Massaro, Taylor, Venezky, Jastrzembski, and Lucas (1980) discuss various descriptions of orthographic structure in English.

To assess how good and poor readers utilize knowledge about the structure of written language, it is necessary to state various descriptions of this structure and then determine how well these descriptions capture reading performance. Venezky and Massaro (1979) and Massaro et al. (1979) have distinguished between two broad categories of orthographic structure: *statistical redundancy* and *rule-governed regularity.* The first category includes all descriptions derived solely from the frequency of letters and letter sequences in written texts. The second category includes all descriptions derived from the phonological constraints in English, scribal conventions for sequencing letters in words, or both. Massaro et al. (1979) contrasted a specific statistical-redundancy description with a specific rule-governed description in a series of experiments utilizing a target-search task. The statistical-redundancy description was the summed single-letter positional frequencies of the letters in the test string. This summed frequency measure was derived from counts that give the number of occurrences of each letter by word position and word length in a sample of written text (Mayzner & Tresselt, 1965). For test strings of six letters, the frequency measures would be the sum of the frequency of the string's first letter as the first letter in six-letter words plus the second letter's frequency as the second letter in six-letter words, and so on. Although the rule-governed description was not as systematically developed, a set of rules was formulated to classify strings as orthographically regular or irregular. The selection rules are given in Massaro and Taylor (Note 1).

To create the test stimuli, Massaro et al. (1979) chose six-letter words and made anagrams of each of them to find four strings that orthogonally combined high and low single-letter positional frequency and regular and irregular orthographic structure. This 2×2 factorial design provides a direct assessment of the contribution of frequency and regularity to letter perception. Mason (1975) had previously evaluated the contribution of summed single-letter positional frequency in a letter search task. Good and poor readers indicated whether or not a six-letter string contained a predetermined target letter. In one experiment (Experiment 2; Mason, 1975), good readers averaged 63 msec faster for strings high in positional

Words	winter charge turned (1293)	Positional Frequency	
		High	Low
Orthographic Regularity	Regular	triwen chager drunet (1438)	trewin greach tredun (540)
	Irregular	wrntei ahcger rdnuet (1421)	rntewi hreagc edtrnu (526)

Figure 2. The five types of items, examples of each type, and the average single-letter positional frequency for each type.

frequency than for those low in positional frequency. Mason found that 100 practice trials in a follow-up experiment reduced the performance advantage for high positional frequency strings to an average of about 22 msec.

In the target search task of Massaro et al. (1979), good sixth-grade readers averaged about 12 msec faster for regular than for irregular strings and about 12 msec faster for high than for low positional frequency; however, neither result was statistically significant. Our interest in this issue led us to conduct an extensive sequence of target search experiments contrasting a reaction-time task with an accuracy task. In the accuracy task, the test-letter string is presented for a short duration and followed by a masking stimulus. The subject indicates whether or not the target letter was present in the test string. The results of this research are reported in Massaro et al. (1979) and Massaro et al. (1980). The most salient finding for the present experiments was that accuracy tasks revealed stronger effects of orthographic structure than did reaction-time tasks. Since the previous experiments did not incorporate the variable of reading ability, the present research used the powerful accuracy tasks to probe whether individuals of varying reading ability use orthographic structure differently.

The two experiments reported here were direct replications of the Massaro et al. (1979) accuracy task with a new set of items and the inclusion of corresponding word stimuli. Figure 2 illustrates the five types of items used in the experiments. College sophomores having good reading scores were contrasted with peers having poor reading scores in the first experiment. Sixth graders at two reading levels and adults comprised the three groups in the second experiment.

Experiment 1

Method

Subjects. A class of 248 college students in introductory psychology was given the Nelson-Denny Reading Test (Forms C & D, 1973) under the "cut-time administration" limits for adults. These limits permit 7.5 minutes for the 100 vocabulary items and 15 minutes for the 72 comprehension questions. The resulting distribution of student scores closely paralleled the standard norms for this administration procedure with a mean score for our sample of 80.9 and a standard deviation of 21.8. Students scoring in the top 15% (raw scores of 106–140) and the bottom 15% (raw scores 60–17) of our sample were selected as our "good reader" and "poor reader" populations, respectively. Of these two groups, 19 good readers and 15 poor readers agreed to participate in the experiment and were paid $2.00 for their participation. Of these 34 subjects, 16 good readers and 11 poor readers achieved experimental accuracy scores between 65%–85%. Five of these good readers were randomly excluded to maintain equal group sizes. Thus, 11 good readers and 11 poor readers formed the final subject groups. The mean raw reading score for the final good reader group was 113.5, and the mean for the poor reader group was 50.3.

Stimulus list. Listings of the 100 highest and 100 lowest positional frequency anagrams were obtained for several hundred common six-letter words selected from the Kučera and Francis (1967) word list. The anagram listings were produced by a computer program that generated all 720 possible permutations for each six-letter word and computed the positional frequency for each permutation from the Mayzner and Tresselt (1965) single-letter counts for six-letter words by summing the position-dependent frequency for each letter. Words with repeated letters were not used. From the 200 anagrams for each six-letter word, two were chosen with high positional frequencies and two were chosen with low positional frequencies. Within each frequency pair, one member was chosen to be orthographically irregular and one to be orthographically regular. The rules for judging regularity are given in Massaro et al. (1980) and Massaro and Taylor (Note 1). Those sources also give the 200 stimulus items, that is, the 40 words and the four anagrams of each word. From 24 additional words, 120 practice stimuli were generated. These words were not as evenly matched for regularity and positional frequency as the experimental words, and a few contained repeated letters.

Two occurrences of each of the 120 practice stimuli were presented in random order for each subject's practice session, and two occurrences of each of the 200 experimental stimuli were presented in random order for each of two experimental sessions. In the two occurrences of each stimulus, one occurrence was tested as a target trial and one as a catch trial. On target trials, the target letter was selected randomly from among the six letters in the test string. For catch trials, a target was selected randomly from a set of 21 letters weighted by their probability of occurrence in the stimulus set. If the selected letter was present in the display string, additional drawings with replacement were made until an appropriate target letter was selected. The letters *j, k, q, x,* and *z* did not occur in the experimental stimulus strings and, therefore, were never tested.

Apparatus. The visual displays were generated by a Digital Equipment Corporation LSI-11 computer under software control and presented on Tektronix Monitor 604 oscilloscopes (Taylor, Klitzke, & Massaro, 1978a, 1978b). These monitors employ a P31 phosphor with a decay to .1% of stimulated luminance within 32 msec. The alphabet consisted of lowercase sans-serif letters resembling the type font Univers 55. For an observer seated comfortably at an experimental station, the six-letter displays subtended about 1.9 degrees of visual angle horizontally, and the distance from the top of an ascender to the bottom of a descender was about .4°. This computerized laboratory facility permitted testing up to four subjects in parallel. Each subject received the same visual display on the display monitor, and the computer collected individual subjects' responses.

Procedure. A trial began with the presentation of a single point in the center of the screen that served as a fixation point. After 250 msec the fixation point was replaced by the test string followed after a variable blank interval by the masking stimulus. The masking stimulus was made up of a montage of random letter features. The mask changed from trial to trial and covered the exact area of the test stimulus on that trial. The fixation point returned, followed by the target letter. The onset of the masking stimulus always occurred 70 msec after the onset of the test string, and the onset of the target letter always occurred 180 msec after the onset of the masking stimulus. The target letter remained present until all subjects responded or for a maximum of 4 sec. The intertrial interval between the offset of the target letter and the onset of the fixation point was 500 msec.

The durations of the test strings and masking stimuli were adjusted to keep average performance across all conditions and the subjects being tested together as close as possible to 75% correct. The target duration could vary between 10–39 msec, and the mask duration could vary between 1–30 msec. The summed duration of the target and mask was always 40 msec, and the durations were traded off using an adaptive algorithm. The durations were modified after every block of 20 trials, during which time each trial type (target or catch) by display type (word, *R–H, R–L, I–H, I–L*) occurred exactly twice. Therefore, all conditions were tested an equal number of times under the same stimulus conditions.

Each subject was tested on a single day for about 90 minutes. The testing session consisted of 240 practice trials and two 400-trial experimental sessions. Subjects were given rest and feedback after the practice trials and a rest period between the two experimental sessions.

Results

Figure 3 presents the average percentage of correct responses as a function of the five display types; reader ability is the curve parameter. There was an overall 26% difference across display types, $F(4, 80) = 169$, $p < .001$, but this difference was identical for good and poor readers, $F(4, 80) = .37$. Responses were 13% more accurate for words than for regular-high items, 9% more accurate for regular than for irregular items, and 4% more accurate for high-frequency than for low-frequency items. These results replicated other studies with these items (Massaro et al., 1980) and, in addition, showed absolutely no effect of reading ability on the role of orthographic structure.

Although the role of orthographic structure did not change with reading ability, good and poor readers were influenced dif-

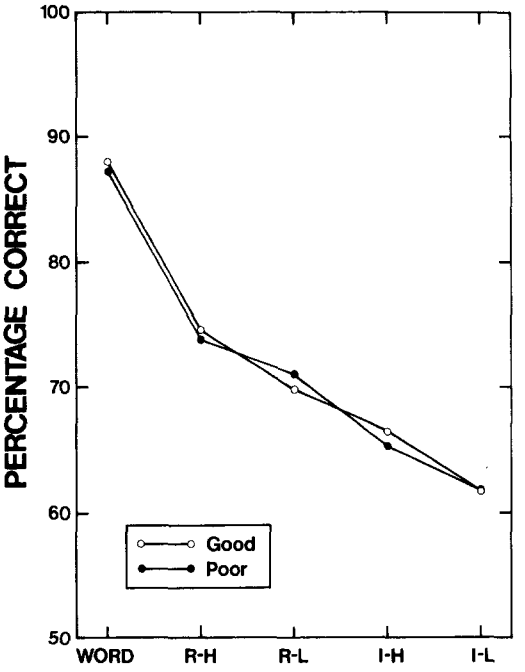

Figure 3. Percentage of correct responses for good and poor college readers as a function of display type. (R-H = regular high; R-L = regular low; I-H = irregular high; I-L = irregular low.)

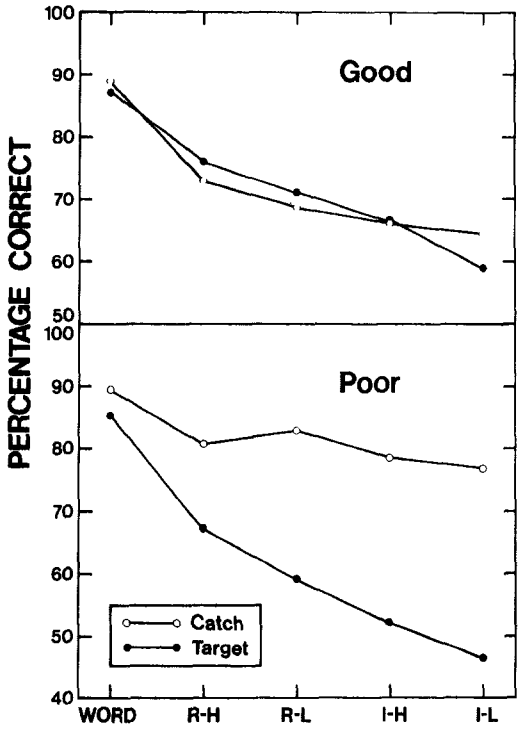

Figure 4. Percentage of correct responses on target and catch trials for good and poor college readers as a function of display type. (R-H = regular high; R-L = regular low; I-H = irregular high; I-L = irregular low.)

ferently by the type of trial. Figure 4 plots performance accuracy of good and poor readers for target and catch trials as a function of display type. Good readers were about equally accurate on target and catch trials, whereas poor readers averaged 20% better on catch than on target trials, $F(1, 20) = 5.87, p < .025$. This result may reflect a conservative bias for poor readers. Given incomplete visual information, poor readers may be more reluctant to say that they saw a target letter even though their evidence is just as good as that for good readers.

Experiment 2

Method

Subjects. The subjects were 47 sixth-graders selected from the Madison Public Schools and 14 college students in introductory psychology. All subjects, sixth-graders and adults, volunteered to participate and were paid $3.00 for participation. The sixth graders ranged in chronological age from 11.6 to 13.3 years. Five of the sixth-grade readers had to be eliminated, since they failed to perform between 65%–85% correct in the task or because of equipment malfunctions. The sixth graders were administered the comprehension subtest of the Gates-MacGinite Reading Test (Survey D, Form 1, 1965). Comprehension grade-level scores varied from Grade 3.5 (raw score of 19) to 11.9+ (perfect score of 52), with an average of 7.2. In addition, scores on the Sequential Tests of Educational Progress (Step; Addison-Wesley, 1977) were available for all but four of the sixth graders. We did not have many very poor readers in our sample of sixth graders. Two groups of sixth-grade readers were created by including readers with the most extreme reading scores. The "poor" reader group included all children with a comprehension score at or below Grade 6.2 on the Gates test and at or below the 63rd percentile on the STEP test. The "good" reader group included all students at or above Grade 10.8 on the Gates test and at or above the 98th percentile on the STEP test. Given these criteria, 10 readers were included in each of the two groups. The average Gates grade levels were 5.48 and 11.8 for the poor and good reader groups, respectively. The corresponding average percentiles on the STEP test were 32.7 and 98.8. Finally, the 10 adults performing closest to 75% correct in the target search task were chosen for the adult reading group. All participants, sixth graders and adults, were paid $3.00 for participation.

Stimuli and apparatus. The test displays and apparatus were identical to those in Experiment 1.

Procedure. The procedure was identical in all respects to that in Experiment 1 except for two differences. First, the target letter was given before rather than after the test display. Each trial began with a fixation point for 250 msec, followed by a target letter for 500 msec followed by the test display and masking stimulus. The fixation point then replaced the masking stimulus and remained on until all subjects responded or for a maximum of 4 sec. The next trial began after an intertrial interval of 500 msec. Second, the durations of the test and masking stimuli were adjusted for each of the subjects individually rather than across all of the subjects being tested together. This procedure is much more sensitive and allows a direct assessment of the durations needed for each subject.

The sixth graders were given the 20-minute reading test prior to the experimental test session; the adults received only the test session. The testing session consisted of 120 practice trials and two 200-trial experimental sessions. Subjects were given rest and feedback after the practice trials and a rest period between the two experimental sessions. The complete experiment took about 90 minutes for the sixth graders and 60 minutes for the adults.

Results

Figure 5 plots the percentage of correct responses as a function of orthographic structure for each of the three groups of readers. Performance improved an average of 11% with increasing orthographic structure, $F(4, 108) = 30.4, p < .001$, but the

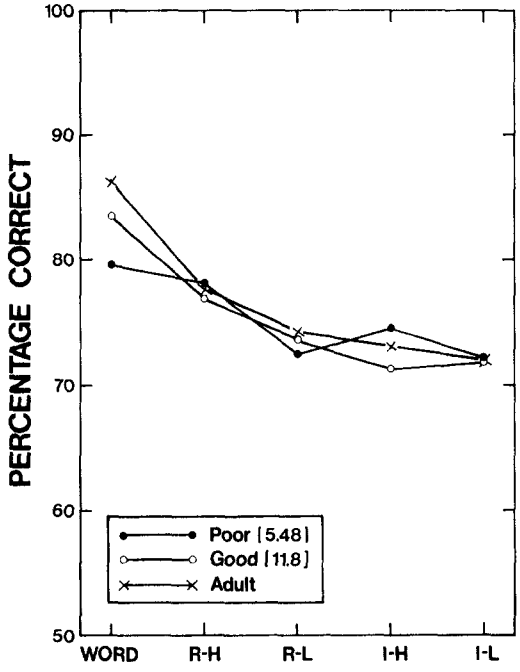

Figure 5. Percentage of correct responses for adult readers, good sixth-grade readers, and poor sixth-grade readers as a function of display type. (R-H = regular high; R-L = regular low; I-H = irregular high; I-L = irregular low.)

overall interaction of orthographic structure and reading level missed statistical significance, $F(8, 108) = 1.54, p > .2$. The range of structure effects was 14% for adults, 12% for the good sixth-graders, and only 7% for the poor sixth-grade readers. Specific comparisons between pairs of reader groups were carried out. The comparison between the adults and poor sixth-grade readers produced a significant interaction of reader group and display type, $F(4, 72) = 2.89, p < .05$. These specific comparisons showed that adults and good sixth-grade readers were significantly better at utilizing orthographic structure than were poor sixth-grade readers.

Overall, subjects were about 9% more accurate on target than on catch trials, $F(1, 27) = 15.7, p < .001$, but this result did not interact with reader ability, or with orthographic structure.

To provide a more detailed evaluation of orthographic structure, two additional analyses were carried out. First, an analysis was performed on the 4 classes of anagrams, treating regularity and frequency as factors. Regular anagrams were recognized about 3% better than irregular anagrams, $F(1, 27) = 13.85, p < .001$. Performance was about 3% more accurate on high- than on low-frequency anagrams, $F(1, 27) = 14.46, p < .001$. The significant interaction of regularity and frequency, $F(1, 27) = 4.73, p < .001$, reflected the greater advantage of the regular-high items relative to the other three classes of anagrams. None of these effects interacted with reading ability.

The second analysis was carried out to contrast the words with the regular-high anagrams. Words were recognized about 5.5% better than the regular-high anagrams, $F(1, 27) = 22.56, p < .001$, and this effect interacted with reading ability, $F(1, 27) = 3.33, p = .05$. Poor sixth-grade readers showed a 1.5% word advantage, their good reading colleagues showed a 5.5% advantage, and adults showed an 8.5% advantage. This result showed that utilization of the lexical status and/or the orthographic structure of words was a direct function of reading level.

To assess whether the stimulus durations varied across the three groups of readers, an analysis was carried out with the letter-string durations as the dependent measure. For each subject, the duration could be changed after every block of 20 trials, giving a total of 20 duration values across the 400 trials. Blocks of trials and reader group were the factors in the analysis of variance. The durations decreased from about 25 to 24 msec across the 20 blocks of the experiment, $F(19, 513) = 5.16, p < .001$. The duration also varied systematically with reading level, $F(2, 27) = 23.37, p < .001$. The average durations were 27, 25, and 18 msec for the poor sixth-grade, good sixth-grade, and adult readers, respectively. Even the 2 msec difference between the poor and good sixth-grade readers was statistically significant, $F(1, 18) = 5.83, p < .05$. Although the task appeared to be more difficult for the poorer readers, it should be emphasized that the duration differences as a function of reading level do *not* necessarily implicate differences at a visual level. If, for any reason, poor readers make more errors in the task than good readers, then duration differences

would be observed. As an example, it could have been the case that poor readers forgot the target letter or hit the wrong button by mistake more often than good readers. In general, performance differences in any complex task do not implicate the reason(s) for the differences.

The orthographic structure effects were significantly smaller in the precue task of Experiment 2 than in the postcue task of Experiment 1. The 26% difference across item types for the college sophomores was reduced to a 14% difference. This results is consistent with the idea that the target letter precue may encourage the reader to rely less on knowledge about orthographic structure. Rather than attempting to resolve the letter string into a word or wordlike string, the subject may simply look for critical features of the target letter (Massaro et al., 1980). This 12% difference between the two tasks is about twice as large as the difference observed by Massaro et al. (1980). The interaction of orthographic structure with reading ability may have been even more apparent in Experiment 2 if the overall effect of orthographic structure had been of the same magnitude as it was in the postcue task.

Correlation Analyses

It is reasonable to expect that post hoc correlations of various measures of orthographic structure with performance on each item would provide a more sensitive evaluation of the contribution of orthographic structure. Rather than looking for differences between classes of items, the post hoc correlations allowed an assessment of differences among all of the items in the task. Massaro et al. (1980) found that some post hoc descriptions of orthographic structure provided a better description of performance in the target search task than did either single-letter positional frequency or the binary classification of regularity. Accordingly, it is possible that these post hoc measures of orthographic structure would provide a more sensitive assessment of its utilization as a function of reading ability.

A number of frequency measures and one measure of regularity were correlated with performance on each of the 200 items used in the present experiments. The source of all of the frequency measures comes from the Massaro et al. (1980) analysis of a word corpus sampled by Kučera and Francis (1967). Counts were determined for single-letter, bigram, and trigram units. Tables were prepared by counting the occurrence of each unit by position in words of a given length. Massaro et al. (1980) provide the tables along with the details of their preparation. The single-letter positional frequency of a given letter in the test string is measured by the frequency with which the letter occurs in the same serial position in words of six letters. The summed single-letter frequency of the complete letter string is the sum of the positional frequencies of each of the individual letters in the string. Summed bigram and trigram frequencies are analogously defined in terms of bigram and trigram letter combinations. Position-insensitive counts give the cumulative frequencies of the units without regard to position in words of six letters.

All unit frequencies were based on word token rather than type counts. A type count counts each particular word only once regardless of how often it occurs, whereas a token count counts the number of occurrences of each word. The two counts are highly correlated in the English language. Each position-sensitive frequency of a given unit was transformed to a logarithmic value before the summed counts for the letter strings were computed. The log of 0 was defined as 0. Massaro et al. (1980) and the present study found that these log counts correlated more highly with the performance measures than did the linear frequency counts.

The word-frequency counts were taken directly from the Kučera and Francis (1967) count. Log word frequencies were used in the correlations, since these correlated higher than linear word frequency. The value 0 was assigned to all nonwords.

Finally, as an attempt to quantify regularity, Massaro et al. (1980) computed a simple count of orthographic irregularities for the 200 test items. An irregularity was counted for each impermissible vowel cluster as well as for each phonologically illegal consonant cluster when considered as part

Table 1
Correlations of a Number of Predictor Variables With Overall Accuracy Performance as a Function of Reading Level in Experiments 1 and 2

	Experiment				
	1		2		
Predictor	Good	Poor	Adults	Good	Poor
Position-sensitive frequency					
Single letter	.513	.492	.294	.232	.249
Bigram	.671	.678	.423	.413	.256
Trigram	.704	.720	.434	.474	.246
Position-insensitive frequency					
Single letter	.203	.229	.157	.130	.080
Bigram	.520	.543	.339	.342	.155
Trigram	.644	.645	.374	.424	.193
Word frequency	.601	.634	.346	.358	.167
Regularity	.514	.508	.280	.286	.206

Note. With $df = 199$, correlations greater than .18 are significant at $p < .01$.

of a monosyllable. The counts were made negative so that the expected correlations would be positive. The exact rules for the irregularity count was given in Massaro et al. (1980) and Massaro and Taylor (Note 1).

The dependent variable for the correlations was the average performance on each of the test items. For each of the 200 stimulus items, an average percentage correct score was computed by averaging the scores for each item across the subjects in each group of readers. Table 1 gives the correlations of the predictor variables with overall accuracy performance as a function of reading level in Experiments 1 and 2. Position-sensitive summed trigram counts provided the overall best predictor of performance. All of the predictor variables were correlated with performance to some degree and what was true for one measure was generally true for the others. The predictor variables were significantly correlated with each other. For example, the correlation of word frequency with performance could be completely accounted for by trigram frequency. Word frequency and position-sensitive trigram frequency correlated .78 in the present list of letter strings.

For the postcue task in Experiment 1, position-sensitive summed trigram frequency accounted for about 50% of the variance in overall accuracy on each of the 200 test items. Consistent with the main effects in the factorial design, the correlation did not differ for good and poor adult readers. The large correlations found in the postcue task of Experiment 1 were significantly attenuated in the precue task of Experiment 2. For adult readers, the best predictor accounted for only 19% of the variance. However, the correlations with orthographic structure now varied systematically with reading level. The correlations were very similar for the adults and good sixth-grade readers and significantly smaller for the poor sixth-grade readers. For example, position-sensitive summed trigram frequency predicted about three times as much variance for the adults and good sixth-grade readers as for the poor sixth-grade readers ($p < .001$). For every measure of orthographic structure except single-letter frequency, the correlations for adults and good sixth-grade readers were significantly larger than those for poor sixth-grade readers ($p < .001$). Therefore, the reading level differences found in the factorial analyses in Experiment 2 were also apparent in the correlational analyses.

Discussion

The present results revealed significant differences in the utilization of orthographic structure as a function of reading level. Poor sixth-grade readers do not appear to

utilize structure to the same degree as do their good-reader peers or adult readers. In addition, good sixth grade readers appear to have already reached an asymptotic level of the utilization of orthographic structure. Finally, good and poor college readers utilize orthographic structure equally well in the target search task.

Mason and Katz (Experiment 2; 1976) found similar results with sixth-grade readers. The good readers were reading at about Grade 9 level, whereas the poor readers were reading at about Grade 3. Performance on a nonredundancy condition was compared with a redundancy condition in a target search task, using Greek and other symbols as test items. On each trial, a target symbol was presented, followed by a test string of 6 symbols; subjects indicated whether or not the target symbol was present in the test string. In the no-redundancy condition, any of the 12 symbols was equally likely to occur in any of the six positions in the test string. In the redundancy condition, the symbols were constrained to occur only in some positions. Given the no-redundancy condition, reaction times did not differ for the good and poor readers. Good and poor readers given the redundancy condition did differ, however, in that good readers benefited from redundancy, whereas poor readers did not.

The results of Experiment 2 along with those of Mason and Katz (1976) encourage the belief that the utilization of orthographic structure is related to reading level of young readers. The failure to find differences between good and poor college readers in Experiment 1 does not necessarily weaken this conclusion, since they represent a different population of readers. Although good and poor college readers may utilize orthographic structure to similar degrees in the target search task, these poor readers may have spent more time and effort in mastering the use of this structure in learning to read. If this were the case, other skills necessary for good reading may not have developed to a sufficient degree.

Classroom Practice

Given the establishment of orthographic structure as a psychological construct and its relationship to reading ability, it may not be premature to discuss some implications of the present research for learning to read. If the utilization of orthographic structure in word recognition is an important component in reading, then it may be profitable to facilitate the child's understanding of this structure. First, we present some of the important constraints in written English that could be profitably taught. Although these constraints are based on letter occurrences in written text for adults (Kučera & Francis, 1967), an analysis of written texts for Grades 3–9 (Carroll, Davies, & Richman, 1971) indicated that orthographic constraints remain relatively constant for texts across reading levels.

One of the most noticeable constraints in English orthography is the difference between where consonants and vowels occur in words. Vowel sounds are relatively infrequent in initial or final position in English words. Therefore, the reader can expect words to begin and end with consonants, except for final *e*, which is not pronounced. For example, when the letters *a, e, i, o, u,* and *y* occur in five-letter words, they occur in first position only 9% of the time. The letters *a, i, o,* and *u* are found in final position in five-letter words on only 1% of their occurrences. If the orthography were unstructured, then the expected occurrence at one of the five positions would be 20%. Therefore, since most words contain at least one vowel, vowel letters can be expected more often in the medial positions than in the initial and final positions (except for *e* and *y*, which are relatively frequent in final position; when these two letters occur in five-letter words, they occur 35% of the time in final position). The digraph vowels, *ea, ee, ie, oa,* and *oo* also occur most often in medial positions. Similarly, in short words that are usually monosyllabic, consonants are more likely to be found in initial and final positions.

Consonant clusters are another key feature of English orthography. There are great constraints on the frequency and location of these clusters in English words. Table 2 presents the most frequently occurring bigram consonant clusters in initial

Table 2
Most Frequently Occurring Consonant Clusters in Word-Initial and Word-Final Positions in English Words of 4–7 Letters

Word initial	Word final
th	*ng*
wh	*th*
st	*st*
fr	*ch*
sh	*ld*
pr	*ll*
ch	*nd*
gr	*nt*
tr	*rs*
pl	*ts*
cl	*ht*
sp	*ds*
br	*ss*
kn	*ck*
dr	*rt*
sc	*rd*
cr	*ns*
bl	*gh*
sm	*ct*
fl	*ls*

Note. Consonant clusters are listed in order of frequency of occurrence.

and final position in English words of four through seven letters. The clusters *th, st,* and *ch* are frequent in both initial and final positions. The cluster *sh* occurs about four times more often in initial than in final positions. However, all other clusters frequent in one position are relatively infrequent or actually illegal and nonoccurring in the other. Of the 20 frequent consonant clusters in initial position, 13 of them do not occur at all in final position. Three others (*sp, sc,* and *sm*) seldom occur in final position. Of the 20 frequent consonant clusters in final position, 14 of them never occur in initial position. Three others (*ll, ts,* and *gh*) only occur a few times and might be considered illegal in initial position. In this way, there is extremely high predictability with regard to where most consonant clusters can be found in English words. It should be noted that the final cluster *ht* derives solely from *ght* and probably should be taught on this basis.

The implications of the fact that many consonant clusters in initial and final positions are illegal if the letters are reversed in order are interesting. All of the initial consonant clusters in Table 2 are illegal if the letters are reversed. The clusters *th* and *wh* are frequent in initial position, but *ht* and *hw* cannot occur in initial position. Similarly, many of the clusters in final position are illegal and do not occur when the letter order is reversed. This constraint could nicely compensate for a potential visual difficulty in reading. There is some evidence that readers may correctly recognize the letters in a string but mistake their relative positions. These transposition errors have been observed in experiments by Estes, Allmeyer, and Reder (1976) and others. Transpositions of letters have even been proposed as one of the primary causes of dyslexia—the child reading the word *was* as *saw* (Orton, 1925). Utilization of the constraints on letter clusters could conceivably compensate for difficulty in resolving the relative spatial position of letters. A reader having resolved the letters *c* and *h* would know they must occur in the order *ch* and not in the order *hc*. Similarly, the vowel clusters *ea* and *oa* can be clarified on recognition of the letters, since *ae* and *ao* seldom occur in English.

Children should be able to learn some of the constraints in English orthography at about the same time that they usually learn phonics. Instead of drill and practice, a game format could be employed to teach various aspects of orthographic structure. We present only a few possible games, since teachers will probably prefer to develop their own variations. The goal of the games is to teach children common letter patterns and where in words these patterns are most likely found.

One possible game is very similar to the experiments presented here. Students could be asked to search for letters and letter clusters in common English words. It is not necessary (and, in fact, not desirable) that all of the words be in the child's written vocabulary. Students could be given a vertical list of words and asked to search the list from the top down, checking each word that contains a particular letter or cluster of letters. A stopwatch could be used to measure the search time, and some score could be given in terms of the number of letters found and the search time. Following the game, the children can discuss where they have found

certain letters and clusters and what general rules of thumb might be helpful. The benefits of learning the rules should be readily apparent to the children. Knowing that the consonant cluster *wh* must occur at the beginning of a word would greatly facilitate searching a list of words for this cluster. The word lists could be varied to teach many of the constraints we have discussed. The differences between vowels and consonants, the common consonant and vowel clusters, the positions in which letters and letter clusters are usually found, and the unique ordering of the letters in most clusters can be illuminated in the target search game.

Letter strings can also be created by arranging and rearranging a small number of letters and letter clusters on a magnetic board. Here the game could be made more interesting by having the children create new words to describe certain events. The goal would be to create the word visually rather than orally, although the child could also attempt to pronounce the new word during or after it is spelled. Students also could be asked to categorize letter strings as possible or impossible sequences. If the string is called impossible, the student could further indicate what is wrong with it and how it might be corrected.

A well-known game currently available on many hobby computers and on the Texas Instrument's Speak and Spell computer is "Hangman." The child (or adult) is given a mystery word of a certain length. The letters that make up the word are guessed one at a time. A correct guess is rewarded with the letter in the position(s) in which it appears in the mystery word. To win, the participant must guess all of the letters before making a certain number of incorrect guesses. Obviously, guesses should not be random but can be optimized on the basis of the structure of English spelling. Discussion of good and bad guesses might provide valuable insights into the structure of written language.

Obviously, this is not a comprehensive list of games and we are confident that teachers will want to develop others. Initial reactions of reading teachers have been very enthusiastic. It is generally accepted that children should be made increasingly aware of language, and although orthographic structure is just one small attribute of written language, it will be to a student's advantage to understand and utilize this structure in reading and writing.

Reference Note

1. Massaro, D. W., & Taylor, G. A. *Reading ability and the utilization of orthographic structure in reading.* (Tech. Rep. No. 515). Madison: Wisconsin Research and Development Center for Individualized Schooling, University of Wisconsin, July 1979.

References

Anderson, I. H., & Dearborn, W. F. The psychology of teaching reading. New York: Ronald Press, 1952.

Breitmeyer, B. G., & Ganz, L. Implications of sustained and transient channels for theories of visual pattern masking saccadic suppression, and information processing. *Psychological Review,* 1976, *83,* 1–36.

Carroll, J. B., Davies, P., & Richman, B. *The American Heritage word frequency book.* New York: American Heritage, 1971.

Estes, W. K., Allmeyer, D. H., & Reder, S. M. Serial position functions for letter identification at brief and extended exposure durations. *Perception & Psychophysics,* 1976, *19,* 1–15.

Gates-MacGinitie reading test. New York: Teachers College Press, 1965.

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

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.

Mason, M., & Katz, L. Visual processing of nonlinguistic strings: Redundancy effects and reading ability. *Journal of Experimental Psychology: General,* 1976, *105,* 338–348.

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

Massaro, D. W. (Ed.). *Understanding language: An information-processing analysis of speech perception, reading, and psycholinguistics.* New York: Academic Press, 1975.

Massaro, D. W. A stage model of reading and listening. *Visible Language,* 1978, *12,* 3–26.

Massaro, D. W. Letter information and orthographic context in word perception. *Journal of Experimental Psychology: Human Perception and Performance,* 1979, *5,* 595–609. (a)

Massaro, D. W. Reading and listening (Tutorial Paper). In P. A. Kolers, M. Wrolstad, & H. Bouma (Eds.), *Processing of visible Language* (Vol. 1). New York: Plenum Press, 1979. (b)

Massaro, D. W. How does orthographic structure facilitate reading? In J. F. Kavanagh & R. L. Venezky (Eds.). *Orthography, reading, and dyslexia.* Bal-

timore, Md.: University Park Press, 1980.

Massaro, D. W., & Schmuller, J. Visual features, preperceptual storage, and processing time in reading. In D. W. Massaro (Ed.), *Understanding language: An information-processing analysis of speech perception, reading, and psycholinguistics.* New York: Academic Press, 1975.

Massaro, D. W., Taylor, G. A., Venezky, R. L., Jastrzembski, J. E., & Lucas, P. A. *Letter and word perception: Orthographic structure and visual processing in reading.* Amsterdam, The Netherlands: North-Holland, 1980.

Massaro, D. W., Venezky, R. L., & Taylor, G. A. Orthographic regularity, positional frequency, and visual processing of letter strings. *Journal of Experimental Psychology: General,* 1979, *108,* 107–124.

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,* 1965, *1,* 13–32.

Miller, G. A., Bruner, J. S., & Postman, L. Familiarity of letter sequences and tachistoscopic identification. *Journal of General Psychology,* 1954, *50,* 129–139.

Nelson-Denny reading test. Boston: Houghton-Mifflin, 1973.

Orton, S. "Word-blindness" in school children. *Archives of Neurology and Psychiatry,* 1925, *14,* 581–615.

Shannon, C. E. A mathematical theory of communication. *Bell System Technical Journal,* 1948, *27,* 379–423; 623–656.

Shannon, C. E. Prediction and entropy of printed English. *Bell System Technical Journal,* 1951, *30,* 50–64.

Taylor, G. A., Klitzke, D., & Massaro, D. W. Considerations in software character generation. *Behavior Research Methods and Instrumentation,* 1978, *10,* 787–788. (a)

Taylor, G. A., Klitzke, D., & Massaro, D. W. A visual display system for reading and visual perception research. *Behavior Research Methods and Instrumentation,* 1978, *10,* 148–153. (b)

Sequential Tests of Educational Progress. Reading, Mass.: Addison-Wesley, 1977.

Thompson, M. C., & Massaro, D. W. The role of visual information and redundancy in reading. *Journal of Experimental Psychology,* 1973, *98,* 49–54.

Venezky, R. L., & Massaro, D. W. The role of orthographic regularity in word recognition. In L. Resnick & P. Weaver (Eds.), *Theory and practice of early reading.* Hillsdale, N.J.: Erlbaum, 1979.

Received May 29, 1979 ■