Orthography, Reading, and Dyslexia

edited by

James F. Kavanagh, Ph.D.
Associate Director
Center for Research for Mothers and Children
National Institute of Child Health and Human Development
National Institutes of Health

and

Richard L. Venezky, Ph.D.
Unidel Professor of Educational Studies
University of Delaware

University Park Press
Baltimore
1980

Chapter 13

HOW DOES ORTHOGRAPHIC STRUCTURE FACILITATE READING?

Dominic Massaro

It is well-known that the psychological processing of letter strings is facilitated to the extent the strings approximate the natural orthography, the rules of spelling of the written language. This result implies that readers utilize in some form the structural constraints given by rules during one or more processing stages in reading. The letters in a string conforming to the orthography are partially predictable or redundant when the visual information is incomplete. The goal of this chapter is to evaluate how knowledge about orthographic structure is represented and how it supplements visual letter information. This assessment is carried out within a general model of the processing stages in reading.

PROCESSING STAGES IN READING

A schematic representation of processing stages in reading is shown in Figure 1. At each stage of processing, memory and process components are represented. The memory component corresponds to the information available to the system at a particular stage of processing. The process component corresponds to the procedural operations on the information held by the memory component. In addition to the temporary memory structures, long-term memory contributes an additional source of information at some of the processing stages. The model will be elaborated simultaneously with the evaluation of how orthographic structure facilitates reading processes.

FEATURE DETECTION

A printed pattern is first transduced by the visual receptor system and the feature detection process detects features which are then stored in preperceptual visual storage. One of the simplest views of the contribution of orthographic structure is that experience with written language modifies the

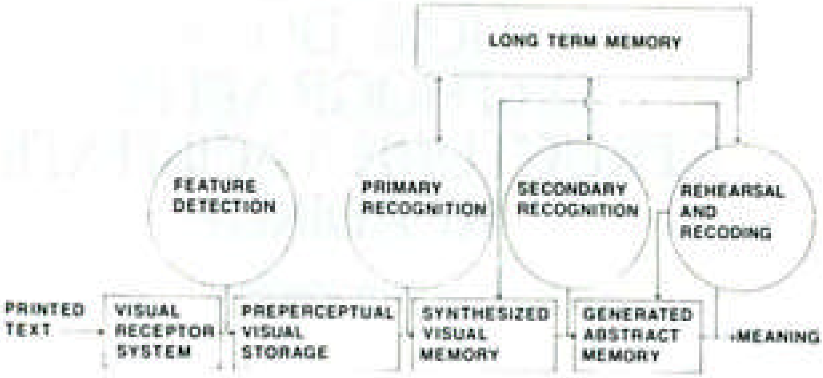

Figure 1. Schematic representation of information processing model for reading.

feature detection process. Given this view, it is possible that orthographic structure enhances the visual feature analysis of the letters in a string. Familiar or frequent orthographic contexts might facilitate feature analysis of the component letters. Orthography may also be exploited to guide feature analysis; recognition of some letters may guide the feature analysis of other letters in the pattern. Accordingly, readers would resolve a greater number of letter features or would obtain better resolution of the features to the extent that the string contains common letter sequences. Another possibility is that experience with certain letter patterns allows the establishment of visual features that are defined across adjacent letters. These supraletter features would make available additional visual information to the extent the letter string is redundant. The distinguishing characteristic of these mechanisms is that orthographic structure directly modifies the featural processing of visual information that the reader has available in reading.

Enhanced Feature Detection

Although it may seem reasonable, there is no evidence that orthographic redundancy enhances the detection of visual features of letters. The most direct test of the idea was carried out about a decade ago. Earhard (1968) asked subjects to indicate whether or not one of the letters in a string was thinner than the other letters. The letter strings were either words or random letter strings. The strings were presented at four different durations sufficient to cover the range from chance to perfect performance. There was absolutely no difference in tachistoscopic recognition of a thin letter between the word and random letter strings at any of the four exposure durations. These results argue that the visual system is *not* more finely tuned to visual features in common rather than in uncommon, letter patterns.

Earhard's results antedated two independent studies carried out by Krueger and Shapiro (1978) and Massaro (in press). In the Krueger and Shapiro study, subjects monitored lists of items presented one item at a time at varying rates of presentation. One-half of the lists of items were composed of six-letter words; the other one-half were lists of six-letter nonwords. In one experiment, subjects indicated whether a mutilated *A* or mutilated *E* appeared in a list. Performance was more accurate for word than for nonword lists. This result might be interpreted to mean that the word context enhanced the feature detection of the mutilated letter; however, enhanced performance with words may have resulted from the redundancy of the word context contributing to the interpretation of whether a mutilated letter was an *A* or an *E*. (The role of redundancy and interpretation is developed more fully in the section "Primary Recognition.")

To provide a more direct test of whether orthographic redundancy enhances feature detection, Krueger and Shapiro replicated their study but now eliminated any possible contribution of an interpretation advantage for words. In this experiment, subjects had to detect simply whether or not a mutilated *A* occurred. This task, therefore, directly taps how well the subject detects a mutilation and performance cannot be influenced by what the mutilated letter is interpreted to be. No performance differences were found as a function of word or nonword lists. This result indicates that detection of a mutilation was not modified by orthographic context, in agreement with the idea that orthographic context does not modify feature analysis.

Massaro's experiment involved the independent variation of the visual information about a letter and its orthographic context in a letter perception task. Consider the lowercase letters *c* and *e*. As can be seen by reading down a column of six items in Figure 2, it is possible to gradually transform the *c* into an *e* by extending the horizontal bar. To the extent the bar is long, the letter resembles *e* and not *c*. If the letter is now presented as the first letter in the context *-oin*, the context would support *c* but not *e*. Only *c* is orthographically legal in this context since three consecutive vowels would usually violate English orthography. This condition is defined as *e* illegal and *c* legal ($\bar{e}Ac$). Only *e* is valid in the context *-dit* since the cluster *cd* is an invalid initial English pattern. In this case, the context *-dit* favors *e* ($eA\bar{c}$). The contexts *-tso* and *-ast* can be considered to favor neither *e* nor *c*. The first remains as illegal context whether *e* or *c* is present ($\bar{e}A\bar{c}$), and the second is orthographically legal for both *e* and *c* (eAc).

The experiment factorially combined six levels of bar length with these four levels of orthographic context, giving a total of 24 experimental conditions. The test letter was also presented at each of the four letter positions in each of the four contexts giving a total of 96 items. The test string was presented for a short duration followed after some short interval by a masking stimulus composed of random letter features. In all cases, the subject indicated whether an *e* or *c* was presented in the test display.

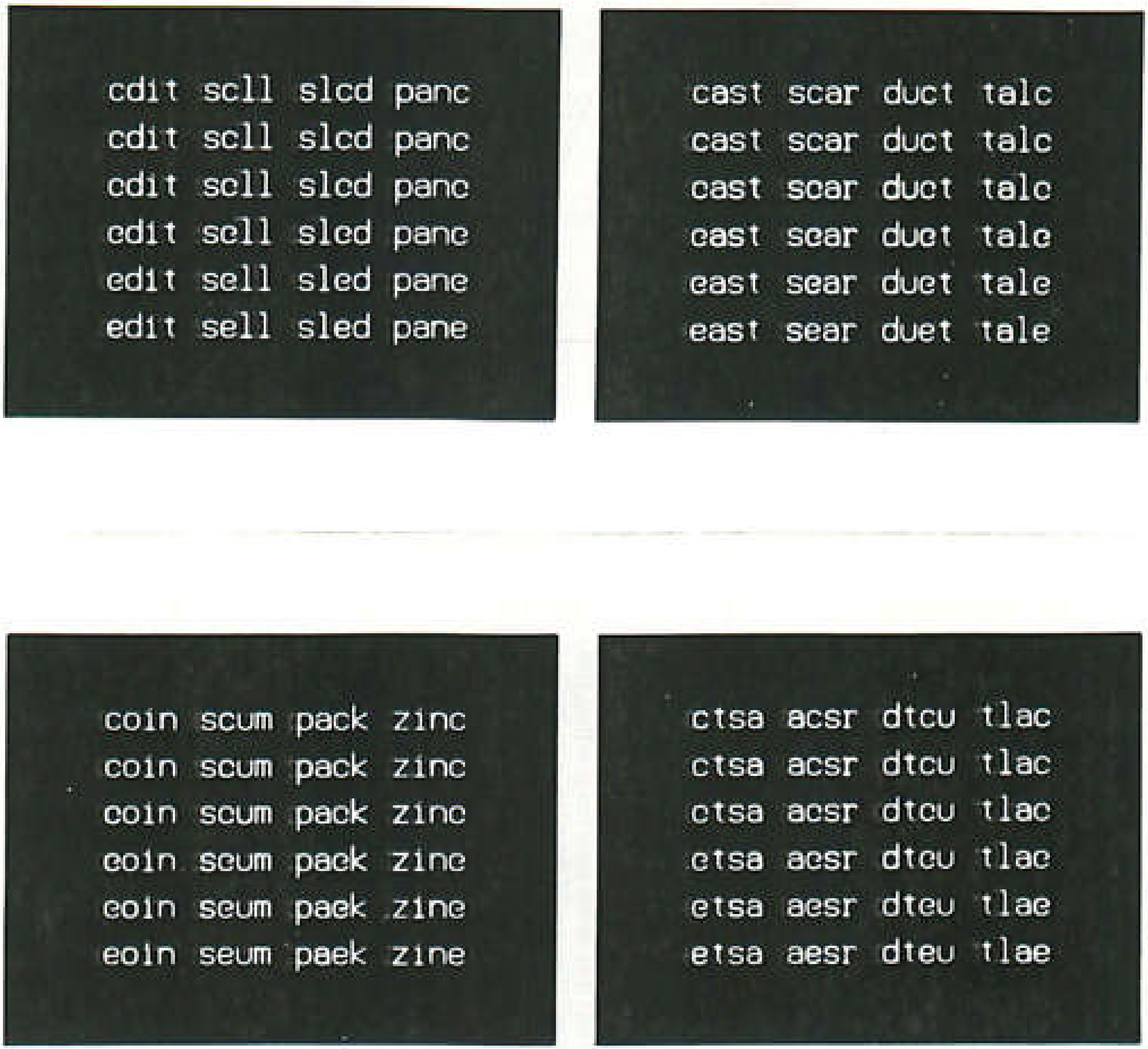

Figure 2. The 96 test items generated by the factorial combination of six bar lengths of the test letter, four serial positions of the test letter, and four orthographic contexts.

The experiment provides a direct test of whether orthographic context influences the feature detection process. One direct measure of feature resolution in the present experiment is the degree to which the reader can discriminate the bar length of the test letter. This discrimination can be indexed in the present experiment by the degree of differential responding to the successive levels of the bar length of the test letter. Better resolution of the test letter is assumed to occur to the extent the subject responds *e* to one length and *c* to another. In the (eʌc) context, both letters spell words whereas neither letter spells a word in the (ēʌc̄) context. If the word context influences feature detection, then the discrimination of bar length should differ in the word and nonword contexts.

The visual resolution of the test letter should be critically dependent on the orthographic context if the latter modifies feature detection. An index of the discriminability of the bar length of the test letter is given by d' and can be derived from the identification responses (Braida & Durlach, 1972). The probabilities of responding *e* to each of the six levels of bar length are transformed to z-scores. The d' value between two adjacent levels of the bar length is simply the difference between the respective z-scores. Cumulating these successive d' distances across the levels of bar length gives a cumulative d' discrimination function. The subject shows good discrimination of bar length to the extent that cumulative d' values are large.

As can be seen in Figure 3 there is no consistent effect of orthographic context on the cumulative d' values. These results indicate that the discrimination of bar length of the test letter did not change with context. The observed equivalence between the both legal and both illegal contexts is direct evidence against the idea that context modifies feature detection. Subjects should have performed differently in the both legal than in the both illegal contexts since these models assume that the context modifies lower-level feature analyses, and therefore, the discriminability of bar length.

One might question whether the cumulative d' values are sensitive measures of the visual resolution of the target letter. Evidence on this question can be derived from the effect of some other variable on the cumulative d' values. It is well known that visual resolution improves with the processing time available for a test stimulus. Processing time was varied in this experiment by varying the blank interstimulus interval between the short test display presentation and the masking stimulus. The cumulative d' values in Figure 4 show a very consistent and large effect of processing time. Discriminability as measured by cumulative d' values increased with increases in the available processing time. The fact that processing time enhanced resolution of bar length shows that the failure to find an effect of orthographic context on discriminability cannot be due to an insensitive test. This conclusion is also supported by the good description of the data by a model based on the assumption that discriminability of the bar length is independent of orthographic context (see the section "Primary Recognition").

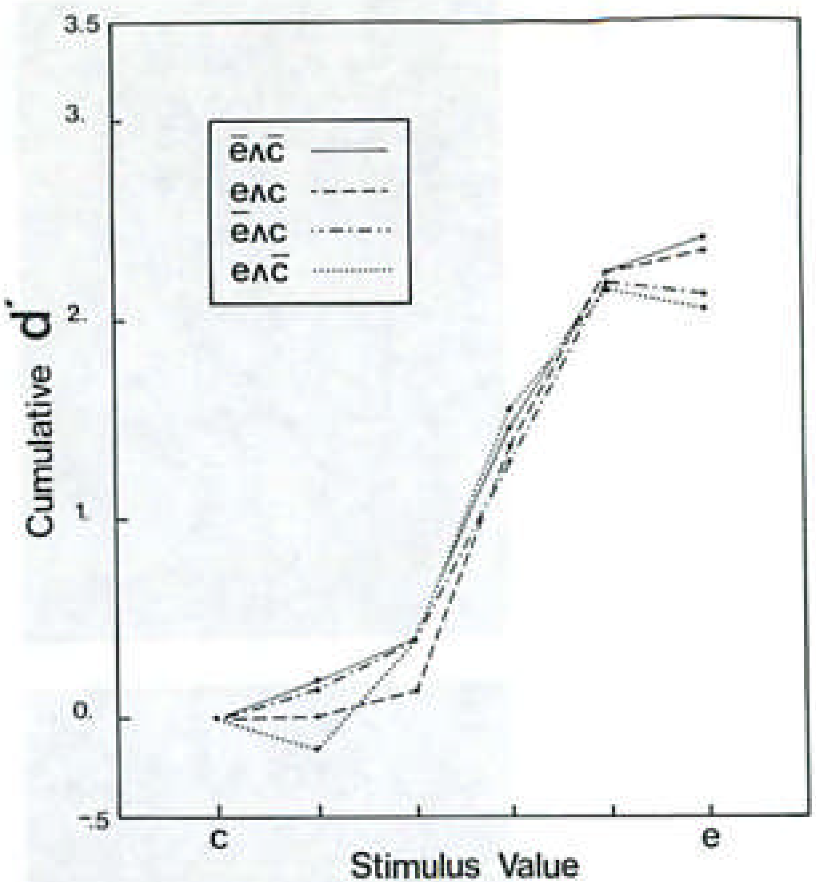

Figure 3. Cumulative d' values as a function of the bar length (stimulus value) of the test letter and the orthographic content.

Supraletter Features

There appears to be sufficient empirical data to reject the idea that orthographic structure enhances feature analysis of letter features. Another more popular alternative is that orthographic structure allows the establishment of *supraletter* features that represent the overall shape and configuration of common letter patterns or words. However, there is no evidence for this idea (Anderson & Dearborn, 1952; Gibson & Levin, 1975; Huey, 1908/1968). One of the strongest arguments against the idea of supraletter features is the small potential contribution of supraletter features to reading. Overall word shape, for example, does not sufficiently differentiate among

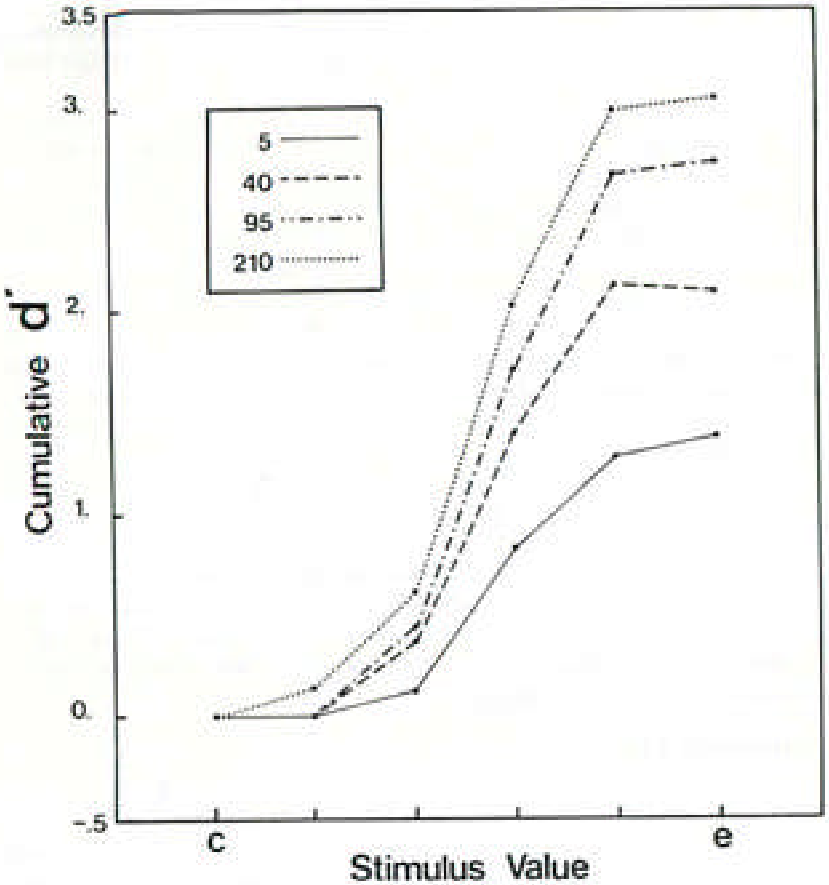

Figure 4. Cumulative d' values as a function of the bar length (stimulus value) of the test letter and the interstimulus in msec.

the words of a language. Groff (1975) examined the shapes of high-frequency words taken from school book sources. The shape was defined by drawing a contour around the letters so that elephant would be elephant. Only 20% of the 283 words were represented by a unique shape. Groff rightly concludes that the small number of words that can be represented by a unique shape precludes the utilization of this cue for accurate word recognition. It should be noted that although words may be recognized from just partial information about the letters, this does not necessarily imply that supraletter features were responsible. Nonvisual information about orthography might supplement the partial information from the letters.

There is also experimental evidence against the idea of word recognition based on supraletter features. Marchbanks and Levin (1965) and Williams, Blumberg, and Williams (1970) presented subjects a test pseudoword followed after a short delay by several other pseudowords. Subjects were required to choose the item that most resembled the test word. First-grade readers tended to choose the item that matched in first or final letter; they seldom matched on the basis of overall shape. Prereaders showed no consistent choices. These results show that beginning readers remember and recognize words on the basis of letter identity rather than overall word shape. Thompson and Massaro (1973) and Massaro (1973) found for adult readers that visual confusability between letters was equivalent in single letter and word presentations. Two letters likely to be confused for one another were just as confusing in single letter and word presentations. If recognition of words involved the utilization of different features than those contained in the component letters, different degrees of letter confusability in single letter and word presentations would have been expected.

The role of supraletter features has been evaluated in a number of studies by determining whether mixing the type fonts of letters eliminates the advantage of word over nonword letter strings. It should be pointed out, however, that attenuating the contribution of orthography by mixing type fonts does not necessarily implicate supraletter features as the cause of a word advantage. As will be discussed in the section "Primary Recognition," utilizing orthographic redundancy to narrow down the number of possible alternative interpretations of partial featural information may produce a word advantage. It is not unreasonable to assume that mixing type fonts may also disrupt the contribution of this process since the partial featural information is much more ambiguous when a variety of type fonts are represented in a letter string.

Although reducing the perceptual advantage of letter strings high in orthographic structure by mixing type fonts does not necessarily implicate supraletter features, an equivalent advantage with mixed type fonts would demonstrate that supraletter features were not responsible. Adams (1979) studied the tachistoscopic recognition of words, pseudowords very high in orthographic structure, and nonwords very low in structure. The items were presented in a single type font or the items were constructed from a variety of fonts. Figure 5 presents examples of the words, pseudowords, and nonwords in single and mixed type fonts. Performance was more accurate for words than pseudowords and poorest for nonwords. Most importantly, the size of the differences among the three types of items did not change when the letters of the items were presented in a variety of type fonts. If supraletter features or whole-word cues contribute to the perceptual advantage of well-structured strings, the advantage of the word and pseudoword strings should have been drastically attenuated in the mixed case presentation.

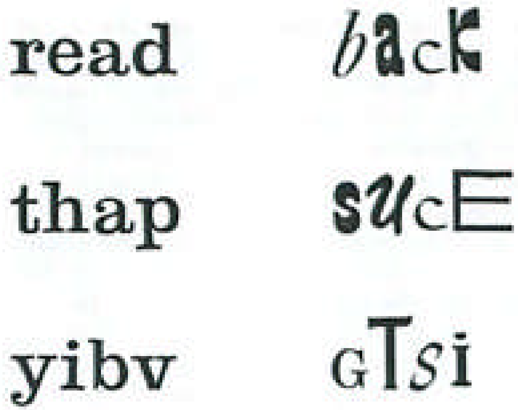

Figure 5. The word, pseudoword, and nonword items presented in a single type font or mixed type fonts (after Adams, 1979).

PRIMARY RECOGNITION

Faced with the set of features in preperceptual visual storage, the primary recognition process evaluates and integrates these features, producing a synthesized visual percept in synthesized visual memory. The primary recognition process operates on a number of letters simultaneously (in parallel). The visual features evaluated at each spatial location narrow down the set of possible letters for that position. The recognition process must choose from this candidate set the most likely letter alternative for each position. Knowledge in long-term memory can contribute to the primary recognition process. The accomplished reader knows not only what visual features define each letter but also something about the orthographic structure of the language. The primary recognition process, therefore, utilizes both the visual information in preperceptual storage and knowledge about the structure of legal letter strings. In the present model, it is assumed that the two sources of information make independent contributions to the recognition process (Massaro, 1973, 1975, in press; Thompson & Massaro, 1973).

This view of primary recognition was developed on the basis of experiments carried out using variants of the Reicher paradigm (Reicher, 1969; Wheeler, 1970). Subjects were presented with either a word or a single letter for a short duration followed immediately by a masking stimulus and two response alternatives. The response alternatives would both spell words in the word condition; for example, given the test word *word*, the response alternatives ---*d* and ---*k* would be presented. The corresponding letter condition would be the test letter *d* followed by the response alternatives *d* and *k*. Performance was about 10% better in the word than in the single-letter condition.

Given the two-alternative forced-choice control, it was argued that the reader was able to utilize orthographic context to eliminate possible alternatives during the perception of the test display before the onset of the masking stimulus (Thompson & Massaro, 1973). As an example, given recognition of the context *wor* and a curvilinear segment of the final letter, the reader could narrow down the alternatives for the final letter to d, o, and q. Given that o and q are orthographically illegal in the context *wor*, d represents an unambiguous choice. The reader will therefore perceive the word *word* given just partial information about the final letter. If the reader has recognized the same curvilinear segment in the corresponding letter condition, however, any of the three letters (*d*, *o*, and *q*) are still possible and the perceptual synthesis will result in *d* only one out of three times. What is critical in this analysis is that a word advantage is obtained even though the visual featural information available to the primary recognition process is equivalent in the word and letter conditions. The orthographic context of the word simply provides an additional but independent source of information. The featural information available to the recognition process does not change with changes in orthographic context. In this view, although orthographic context facilitates word perception, it does not modify the feature analysis of the printed pattern.

Quantitative Test

An experiment was carried out to test a quantification of this view of primary recognition (Massaro, in press). The logic of the experiment centered on the question of whether orthographic context and featural processing make independent contributions to letter perception. The experiment was described in the section "Feature Detection." Subjects viewed four-letter strings that differed in terms of the visual information about the critical letter and the orthographic context of surrounding letters. In the *e*–*c* identification task, the horizontal bar of the critical letter was presented at six different lengths covering the range of a good *c* to a good *e*. Each test letter was also presented at each of the four letter positions in each of four orthographic contexts (see Figure 2).

Figure 6 presents the probability of an *e* response at each of the experimental conditions. The results showed large effects of stimulus information and orthographic context on the identification of the test letter. The significant interaction of these two variables revealed that the magnitude of the context effect was large at the more ambiguous levels of visual information. Although the context effect decreased some with experience in the task, it was still highly significant on the fourth day of the experiment.

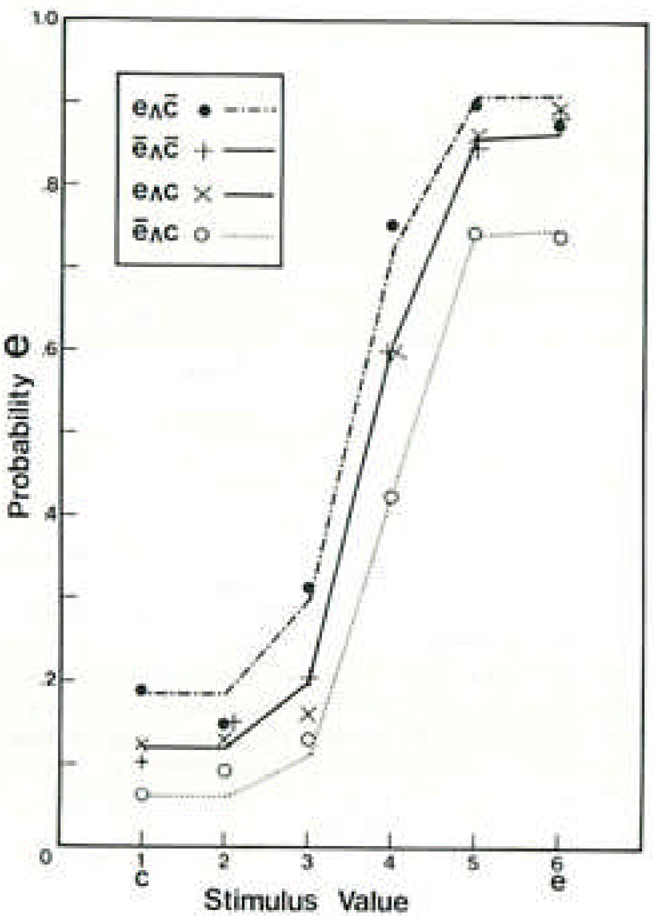

Figure 6. Observed (points) and predicted (lines) probabilities of an *e* response as a function of the bar length (stimulus value) of the test letter and orthographic context.

The model provides a straightforward interpretation of the experimental situation. Two independent sources of information are available: the visual information from the critical letter and the orthographic context. The first source of information can be represented by V_i where the subscript i indicates that V_i changes only with bar length. For the *e*-*c* identification, V_i specifies how much *e*-ness is given by the critical letter. This value lies between zero and one and is expected to increase as the length of the bar is increased. With these two letter alternatives that differ only in bar length, it is

reasonable to assume the visual information supporting *c* is simply one minus the amount of *e*-ness given by the same source. Therefore, if V_i specifies the amount of *e*-ness given by the test letter, then $(1 - V_i)$ specifies the amount of *c*-ness given by that same test letter.

The orthographic context provides independent evidence for *e* and *c*. The value C_j represents how much the context supports the letter *e*. The subscript j indicates that C_j changes only with changes in orthographic context. The value of C_j lies between zero and one and should be large when *e* is legal and small when *e* is illegal. The degree to which the orthographic context supports the letter *c* is indexed by D_j and is independent of the value of C_j. The value of D_j also lies between zero and one and should be large when *c* is legal and small when *c* is illegal.

Faced with two independent sources of information, the reader evaluates the amount of *e*-ness and *c*-ness from these two sources. In order to arrive at a combined value of *e*-ness and of *c*-ness, it is necessary to integrate the two sources. The amount of *e*-ness and *c*-ness for a given test display can therefore be represented by the conjunction of the two independent sources of information:

$$e\text{-ness} = (V_i \wedge C_j) \quad (1)$$
$$c\text{-ness} = ((1 - V_i) \wedge (D_j)) \quad (2)$$

The *e*-ness and *c*-ness values given by both sources can be determined once conjunction is defined. Research in other domains has shown that a multiplicative combination provides a much better description than an additive combination (Massaro & Cohen, 1976; Oden, 1977; Oden & Massaro, 1978). Applying the multiplicative combination, Equations (1) and (2) are represented as:

$$e\text{-ness} = V_i \times C_j \quad (3)$$
$$c\text{-ness} = (1 - V_i) \times (D_j) \quad (4)$$

A response is based on the *e*-ness and *c*-ness values: a choice of *e* is assumed to be made by evaluating the degree of *e*-ness relative to the sum of the *e*-ness and the *c*-ness values. This choice rule is a direct application of Luce's (1959) choice axiom. The probability of an *e* response, P(e), is expressed as

$$P(e) = \frac{V_i C_j}{V_i C_j + (1 - V_i)(D_j)} \quad (5)$$

To derive P(e) for the four orthographic contexts, a simplifying assumption about context is that a given alternative is supported to the degree *x* by a legal context and to the degree *y* by an illegal context, where $1 \geq x \geq y \geq 0$. The values *x* and *y* do not have subscripts since they depend only on the legality of the context. Therefore, C_j is equal to *x* when *e* is legal in a particular context and equal to y when e is illegal. Analogously, D_i is equal to x when c is legal in a particular context and equal to y when c is illegal.

Given the context with e legal and c illegal

$$(eA\bar{c}):\ P(e) = \frac{V_i x}{V_i x + (1 - V_i)y} \tag{6}$$

since the e-ness is given by $V_i x$ and the c-ness by $(1 - V_i)y$. Analogous expressions for the other three contexts are

$$(eAc):\ P(e) = \frac{V_i x}{V_i x + (1 - V_i)x} = V_i \tag{7}$$

$$(\bar{e}A\bar{c}):\ P(e) = \frac{V_i y}{V_i y + (1 - V_i)y} = V_i \tag{8}$$

$$(\bar{e}Ac):\ P(e) = \frac{V_i y}{V_i y + (1 - V_i)x} \tag{9}$$

Equations 6 and 9 predict an effect of context to the extent a legal context gives more evidence for a particular test letter than does an illegal context, i.e., to the extent $x > y$. A second feature of this model is that P(e) is entirely determined by the visual information when the context supports either both or neither of the test alternatives; Equations 7 and 8 both predict $P(e) = V_i$.

This form of the independence model was tested against the observed response probabilities. Figure 6 gives the observed and predicted values. In order to fit the model to the data, it was necessary to estimate 6 values of V_i for each level of bar length of the critical letter and an x value for a legal context and a y value for an illegal context. The parameter values were estimated using the iterative routine STEPIT by minimizing the squared deviations between predicted and observed probability values (Chandler, 1969).

The model provided a good description of the results. In addition, the parameter estimates are psychologically meaningful. The value of V_i increased with increases in the length of the bar of the critical letter. The values were 0.11, 0.11, 0.19, 0.60, 0.85, and 0.86 for the six respective levels. The value of x was 0.76 for legal context and the value of y was 0.40 for the illegal context. The root mean squared deviation between the predicted and observed response probabilities was 0.02. It is clear that an additive combination of the two sources would fail since the curves would have to be parallel. Supporting this, the description was 2.5 times poorer for the same model based on the additive combination of the independent sources.

Johnston (1978) offers what he considers a critical test of a general form of this interpretation of the facilitating effects of orthographic redundancy. His goal was to provide a direct test of the assumption that a word

context facilitates letter identification because its identity is constrained by the identity of the surrounding letters. Two classes of words were chosen and presented for tachistoscopic identification. Words were chosen to have either a high or low constraining context on a particular letter position. Consider the word pairs *date-gate* and *drip-grip* when the initial letter position is tested. Nine four-letter words end in ate whereas only three end in rip. If the primary recognition process utilizes information from the orthographic context to supplement featural information, Johnston reasoned that performance should be better when tested on the first letter of either of the first pair of words than the second, when the alternatives are limited to *d* and *g* in the Reicher-Wheeler task. Although the standard word advantage was found, Johnston found no performance difference between these two classes of words.

Johnston's experiment is limited in a number of ways and, therefore, cannot be taken as evidence against the idea that orthographic redundancy provides an additional source of information at primary recognition. With respect to his structure manipulation, Johnston varied lexical constraints at a given position, but there is no evidence that the psychological representation of orthography is limited to lexical constraints (Massaro, 1975). For example, although three times as many words are possible given the context *-ate* than given the context *-rip*, almost as many letters are legal in the second context as in the first. Our more recent studies have shown that the contribution of orthographic redundancy can be captured by descriptions that are independent of whether the letter string is lexically represented (Massaro, Taylor, Venezky, Jastrzembski, & Lucas, in preparation). Johnston equated the position-sensitive bigram frequency of the high and low constraint items and since we have shown that this variable captures nearly all of the contribution of orthographic redundancy, it is not surprising that there were no differences between the two classes of words.

SECONDARY RECOGNITION

The secondary recognition process transforms the synthesized visual percept into a higher-order code in generated abstract memory. Synthesized visual memory holds a sequence of letters that are analyzed by the secondary recognition process with the goal of closing off the letter string into a phonological, lexical, and/or meaning code. The secondary recognition process makes this transformation by finding the best match between the letter string and codes in long-term memory. Orthographic redundancy might facilitate the transformation from a visual to a higher-order code.

Mewhort (1974) presented first-order or fourth-order pseudowords one letter at a time horizontally across a visual display. Each letter was presented for 3 msec and the interletter interval was varied from 0 to 100 msec.

Although increasing the interletter interval from 0 to 50 msec did not change performance on the first-order strings, it decreased performance on the fourth-order strings. Because increasing the rate of presentation did not lower performance on the first-order strings, Mewhort argues that the disruption of performance on the fourth-order strings cannot be due to simple character identification. These results are taken by Mewhort to support the idea that orthographic structure aids transfer from a visual store to a verbal short-term memory.

Although feature analysis may not be influenced by increasing the interletter interval from 0 to 50 msec, the utilization of orthographic structure during perceptual recognition probably requires having featural information from adjacent letters simultaneously available. Therefore, the utilization of orthographic structure at primary recognition can also account for Mewhort's finding that increasing the interletter interval decreases performance on fourth-order but not first-order letter strings. Although Mewhort's results are ambiguous, it is still reasonable to expect that orthographic structure can facilitate the transfer of a visual percept into a higher-order code since a well-structured string will require fewer codes than a poorly-structured string.

REHEARSAL-RECODING

Recoding and rehearsal processes build and maintain semantic and syntactic structures at the level of generated abstract memory. Generated abstract memory corresponds to the short-term or working memory of most information processing models. In our model, this memory is common to both speech perception and reading. It is also possible to go from meaning to a visual or auditory percept in our model. The recoding operation can transform the meaning of a concept into its surface structure and auditory or visual form.

Baddeley (1964) questioned the perceptual contribution to the original psychological study of orthography carried out by Miller, Bruner, and Postman (1954). The authors had subjects reproduce letter sequences, eight letters in length, corresponding to different approximations to English based on Shannon's (1948) algorithms. The displays were exposed for durations of 10 to 500 msec and the number of letters reported increased with display duration. Also, performance was a systematic function of the order of approximation to English. By correcting for redundancy of the strings, the amount of information transmitted was shown to be equivalent for the four different approximations.

Baddeley (1964) observed that performance in the task was unlikely to be a direct index of how well the letter sequences were perceived. Observing that performance improved at a negatively accelerated function of log exposure duration, it was possible to extrapolate from the curves and observe that an exposure duration of one or two hours would be required for correct report of all eight letters. Baddeley argued that the results of Miller et al. (1954) may have reflected differences in memory for the sequences rather than visual perception. To test this idea, he presented the eight-letter sequences of Miller et al. at a duration that was sufficient for that subject to name each of the eight letters. Presentation times ranged between one and two seconds. The contribution of orthographic redundancy to performance was essentially identical to that reported by Miller et al. (1954). Baddeley concluded that both interletter redundancy and exposure time allow a more effective coding and, therefore, better memory and recall of the letter sequence.

Krueger (1971) evaluated the influence of orthographic redundancy on short-term recognition memory. In one study, a string of six letters was presented one letter at a time at a rate of four letters per second. The letters were presented either left to right or right to left across the display screen. Words produced only a 3% performance advantage over corresponding nonwords whereas the left-to-right presentation gave about a 7% advantage relative to the right-to-left presentation for both types of strings. Slowing the presentation rate to 2.5 letters/sec eliminated the left-to-right advantage but not the word advantage. These results replicate Baddeley's findings that orthographic structure can improve short-term memory and extend them to a recognition memory situation that eliminates output interference present in free recall studies.

SUMMARY

There is now good evidence that orthographic structure does not modify featural processing of the visual information in reading, although it contributes a significant source of information to perceptual recognition. Knowledge of the orthography allows the reader to constrain the set of admissable letters in a letter string when only partial featural information is available. Orthographic structure not only facilitates perceptual recognition, the phenomenal experience of a visual pattern, it enhances accessing higher-order codes for the pattern, and maintenance and elaboration of the codes in memory.