Acta Psychologica 43 (1979) 413–426

THE ROLE OF LATERAL MASKING AND ORTHOGRAPHIC STRUCTURE IN LETTER AND WORD RECOGNITION*

Dominic W. MASSARO and David KLITZKE
Dept. of Psychology, University of Wisconsin, U.S.A.

Received September 1978

The present experiments evaluated the contribution of orthographic structure and lateral masking in the perception of letter, word, and nonword test displays. Performance was tested in a backward recognition masking experiment in which a masking stimulus followed the test display after a variable blank interstimulus interval. In agreement with previous findings across different experiments, words were recognized better than single letters at short interstimulus intervals, but the opposite was the case at long intervals. Performance on the nonwords resembled performance on letters at short masking intervals and performance on words at long masking intervals. The quantitative results were described by a processing model that incorporates the effects of lateral masking and orthographic structure in the dynamic processing of letter strings. Lateral masking tends to lower the potential perceptibility of letters whereas orthographic structure can reduce the uncertainty of the candidate letters in the letter sequence. The present model predicts that the quantitative contribution of each of these processes to performance is critically dependent upon the processing time available before the onset of the masking stimulus.

A persistent concern in reading-related research has been to what extent word recognition can be described in terms of processing the component letters of a word. One popular thesis has been that words are not processed in terms of their component letters but rather are recognized on the basis of supraletter features such as whole word length and shape (see Purcell *et al.* 1978, for the most recent statement of this thesis). Early in the history of reading research, word recognition was shown to have an advantage over single letter recognition.

---

* This research was supported in part by funds from the National Institute of Education, Department of Health, Education, and Welfare to the Wisconsin Research and Development Center for Cognitive Learning. The helpful comments of Glen A. Taylor are gratefully acknowledged.
Requests for reprints should be sent to Dominic W. Massaro, Department of Psychology, University of Wisconsin, Madison, Wisconsin 53706.

Erdmann and Dodge (1898, cited in Huey 1908) demonstrated that words could be read at distances too great to permit identification of their component letters when presented alone. Cattell (1886) found that a whole word could be named as rapidly as a single letter. When necessary these results were utilized to justify the sight-word method of teaching reading. Children were taught to recognize and name words as wholes without reference to the spelling patterns that make them up or concern for how particular spelling patterns are pronounced. The sight-word method did not remain popular for long, although it appears to return to the forefront periodically, not necessarily because of, but usually contemporaneous with another demonstration that words are special in the manner in which they are processed.

In the last decade, experimental psychology has seen a resurgence of reading-related research primarily in the form of letter and word recognition experiments. An influential experiment was carried out by Reicher (1969). Subjects were presented with either a single letter, a four-letter word, or a four-letter nonword flashed in a tachistoscope and had to report what they saw. Reicher's contribution to this century-old task was to constrain the subject's choice by presenting two letter alternatives after each trial. Both alternatives would spell words in the word condition so that performance would not benefit from some simple guessing strategy. Even with these constraints, Reicher found a 10% advantage for recognition of a letter in a word over recognition of a letter in a nonword or a letter presented alone.

This paper reports two experiments designed to illuminate two component processes that are functional in the recognition of letter strings. These are lateral masking between adjacent letters and the orthographic structure of letter sequences that spell words. Lateral masking occurs when the presence of one contour lowers the visibility of a neighboring contour. The perceptibility of adjacent letters in a letter string can be decreased because of lateral masking. The orthographic structure of letter strings that spell words refers to the fact that only a limited number of letter sequences are possible in words. In this case, the presence of one letter can reduce the uncertainty of a neighboring letter if the letter sequence spells a word. Accordingly, letters in words have the disadvantage of lateral masking and the advantage of orthographic structure and the interplay of these two processes is critically dependent on the processing time available in the task.

It is possible to describe the temporal course of letter perception by the equation

$$d' = \alpha\,(1 - e^{-\theta t}) \qquad (1)$$

where d' represents the degree to which the test letter is perceived and is expressed in z-score units analogous to the measure d' of signal detection theory (Massaro 1975). Letter perceptibility as measured by d' can be derived from the observed performance in a letter recognition task. In Equation 1, the parameter α represents maximal perceptibility with unlimited processing time, t the processing time measured from the onset of the test stimulus to the onset of a masking stimulus, θ the rate of processing, and e the natural logarithm. Perceptibility is a negatively accelerating growth function of the processing time available. The value α is dependent on the properties of the visual display and the acuity of the visual system. The value θ is dependent on the rate of processing this information. The value of θ can be expected to be dependent upon process variables such as selective attention and the degree to which the reader utilizes information about the orthographic structure of letter sequences.

In terms of this model, lateral masking and orthographic structure are accounted for by different parameters in the equation describing perceptibility as a function of processing time. Adjacent letters that degrade the perceptibility of a neighboring letter should decrease the α value for that letter. Adjacent letters that reduce the uncertainty of a given letter because of orthographic structure should increase θ, the rate of processing the letter. Consider the case in which the single letter *c* is presented. Perceptual resolution of the letter is assumed to be a temporally extended and continuous process. As an example, when t = 100 msec, the letter may be resolved sufficiently to reduce the alternatives to *c*, *e*, and *o*, whereas the letter *c* is not resolved completely until t = 200 msec. In the analogous word case, the letter *c* may be presented in the context *coin*. In this case, because of lateral masking *c* may not be resolved completely but may be seen at only 90% clarity even with unlimited processing time. Hence the α value for *c* would be lower in the word than in the single letter condition. Although the word context lowers the asymptotic perceptibility of the letter *c*, it should also enhance the rate of processing of the letter *c*. If the context *oin* is completely resolved and the alternatives for the first letter are limited to *c*, *e*, and *o*, no further visual processing is necessary given this visual information and information about orthographic structure. The strings *eoin* and *ooin* are illegal in English and, therefore, *c* is the only valid alternative for the first letter. In this hypothetical example, *c* can be recognized completely in the word context on the basis of the same visual information for that letter as in the case when it is presented alone and three alternatives are still viable.

Although lateral masking lowers the asymptotic perceptibility of a letter in a word, information about orthographic structure allows the reader to arrive at a correct decision more quickly with a word than a single letter. The operation of these two processes over time should produce performance differences between the letter and word conditions that are critically dependent on the processing time available. When processing time is maximal, rate of processing is unimportant and the letter condition should show an advantage because the perceptibility of a letter in a word is reduced by lateral masking. With intermediate processing times, the presence of orthographic structure in the word condition may enhance the rate of processing of the test letter offsetting the deficit of lateral masking. It follows that a performance advantage may be found for words at short processing intervals.

In the present experiments, processing time was controlled in a backward recognition masking task in which the test display was followed by a masking stimulus after a variable blank interval. Single letters, words, nonwords, and letters embedded in dollar signs were utilized as test items in the Reicher (1969) task. The two test alternatives, always presented 250 msec after the test item, were the complete test item and its corresponding foil that differed by one letter. Performance on nonwords should be equivalent to that on letters in dollar signs at all masking intervals. Both have the disadvantage of lateral masking and both possess no orthographic structure. The disadvantage of lateral masking on word trials should be overcome by the advantage of orthographic structure at short but not at long processing intervals.

Experiment 1

Method

On each trial, *Ss* saw a fixation point for 500 msec followed by one of four possible test items: a letter, a four-letter word, a four-letter nonword, and a letter with three \$ characters. The test item was followed after a variable blank interval by a masking stimulus and on some trials no mask was presented. The two response alternatives were presented above and below the position of the test item 250 msec after the onset of the test item. The alternatives consisted of the exact stimulus that was presented and a corresponding item that differed by just the critical letter. If the *S*s were given the test item *tart*, the alternatives would be *tart* and *wart*. The test item *t* would be followed by the alternatives *t* and *w*, and so on. The observers made their choice of the top or bottom alternative by pressing one of two push-buttons. The alternatives remained present throughout the response interval. The response interval lasted until each of the three *S*s made a response or terminated after four sec. The visual displays were generated under computer control and were presented on Tektronix Monitor 604 oscilloscopes. A P-31 phosphor was used; this phosphor declines to 0.1% of its intensity within 32 msec after it is turned off. Three *S*s were tested at a time, each seated in a separate darkened room.

The alphabet

The alphabet consisted of lower case non-serified letters very closely resembling the type font Univers 55. The lines were composed of dots so closely spaced that individual dots were not apparent. The line segments comprising the letter appeared continuous. The ratio of the height of ascenders and descenders to x-height letters was 3 to 2 as was the ratio of the height of an x-height letter to its most usual width. Interletter spaces were about 0.38 of the width of an x-height letter. Four letters subtended about 1.5 degree of visual angle horizontally and the distance from the top of an ascender to the bottom of a descender was about 1°.

In order to equate the stimulus intensity and duration of a given letter in each of the four context conditions, it was necessary to plot that letter with the same number of points and in the same amount of time independent of the presence or absence of other letters in the display. In the single letter condition, for example, 'blank' context letters had to be plotted. Plotting blank letters involved executing instructions that did not affect the display screen but which required the same execution time and which were interspersed with actual point intensification instructions as if other visible points were being plotted as context letters.

Test stimuli

The stimuli were constructed from 32 four-letter word pairs, each pair differing in only one letter position, with the critical letter occurring equally often at each of the four letter positions. Table 1 presents the words and nonwords used in the experiments. Yoked to each word pair was a corresponding nonword pair, a single letter pair, and a letter presented in the context of three dollar signs pair. Given these stimuli, it was possible to test the same letters at all four serial positions in each of the four context conditions. Accordingly, any differences in the context conditions could not be due to testing different letters or different positions.

The test stimulus was presented for 33 msec and was followed by a masking stimulus after an interstimulus interval of 5, 20, 40, 65, 95, 130, or 170 msec, or on some trials no mask was presented. The duration of the masking stimulus was also 33 msec. The intensity of the dots in the letters and mask was a variable in the plot-

Table 1
The word and nonword displays used in the experiments.

Word	Nonword	Word	Nonword	Word	Nonword	Word	Nonword
yarn	ynra	ayes	syae	lays	lsya	wily	ilwy
darn	dnra	ades	sdae	lads	lsda	wild	ilwd
dent	dtne	idle	ldie	kids	kdsi	bind	ibnd
sent	stne	isle	lsie	kiss	kssi	bins	ibns
seat	stae	asps	sspa	case	aesc	sees	eess
meat	mtae	amps	smpa	came	aemc	seem	eesm
moon	mnoo	smug	gmsu	same	aems	loom	oolm
noon	nnoo	snug	gnsu	sane	alns	loon	ooln
nick	nkci	snip	pnsi	lane	aenl	barn	abrn
kick	kkci	skip	pksi	lake	aekl	bark	abrk
kill	klli	skew	wkse	bike	eikb	dusk	usdk
till	tlli	stew	wtse	bite	eitb	dust	usdt
tart	ttra	stab	btsa	cots	csto	slot	oslt
wart	wtra	swab	bwsa	cows	cswo	slow	oslw
wear	wrae	ewes	swee	paws	pswa	plow	olpw
year	yrae	eyes	syee	pays	psya	ploy	olpy

ting routine. Also, the plotting time per letter was independent of the size of the letter. As with the 'blank' context letters, all letters were padded with extra instructions, if necessary, to insure that all letters required the same execution time for a single plotting. A unique masking stimulus was presented on each trial. The masking stimulus was composed of response letters by selecting random feature strokes from the letters of the alphabet. A nonsense letter of the mask was roughly as dense as the letter *g*. The randomly selected strokes in the mask were horizontally displaced with random increments so that they never formed a valid letter. The mask covered *only* the appropriate letter position in the letter alone test condition and covered all four letter positions in the other three context conditions.

Stimulus selection

There were 2,048 unique conditions (4 letter positions X 8 masking conditions X 4 context conditions X 8 stimulus pairs X 2 alternatives) in the experiment. These were presented in 8 sessions of 256 trials each. Within each block of 32 trials, each of the masking conditions was presented exactly four times. Within each of these masking conditions, each of the four context conditions was presented exactly once. Additional constraints were that all four letter positions were tested exactly once within each of the eight masking conditions in every block of 32 trials. Both members of a given pair of items (*e.g.*, ttra-wtra) were presented in the same 256-trial session with the constraint that one member was presented in the first half and the other member in the second half of the session.

Finally, an on-line algorithm was used to keep overall performance averaged over the three subjects being tested together at 75% correct. After each block of 32 trials, the intensities of the dots and number of dots painted per unit of time were adjusted based on the performance in that block of trials. Given that each of the 32 masking conditions X context conditions occurred exactly once in each block of 32 trials, this adjustment procedure insured that each condition was tested at the same intensity an equal number of times.

Subjects

Six *S*s were tested for a total of five days. Students from introductory psychology classes volunteered to participate for bonus points in the course. Two sessions of 256 trials each were given per day. The first day was considered practice. This gave a total of 64 observations at each of the 32 context X masking conditions.

Results and discussion

Table 2 presents the average percentage of correct responses as a function of processing time and context condition. The processing time is given in terms of stimulus onset asynchrony (SOA), the time between the onset of the test item and the onset of the masking stimulus. Performance improved from 54% correct at the 38 msec SOA to 91% at an SOA of 203 msec which was only 2% poorer than the no mask condition, $F(7, 35) = 118$, $p < 0.001$. Letter recognition was critically dependent on the context condition, $F(3, 15) = 7.36$, $p < 0.005$. Considering just performance at the four shortest SOAs, performance averaged about 6% higher for words than for single letters, $F(1, 15) = 6.7$, $p < 0.025$. Therefore, the present experiment was successful in replicating the word advantage found in previous studies which used short masking intervals (Johnston and McClelland 1973; Reicher 1969; Thompson and Massaro 1973; Wheeler 1970).

Table 2
Percentage of correct recognitions as a function of processing time and context.

Processing time SOA* (msec)	Context			
	Letter	Word	Nonword	Letter in $
38	55	54	53	56
53	58	65	58	58
73	66	74	62	67
98	77	87	78	74
128	89	92	83	82
163	95	91	83	87
203	94	93	86	90
no mask	98	93	89	91

* SOA refers to stimulus onset asynchrony, the time between the onset of the test stimulus and the onset of the masking stimulus.

In contrast, performance for letters alone and words did not differ at the three longest SOAs and the no mask condition, $F(1, 15) < 1$. This result replicates previous findings of no word advantage when no mask is presented (Johnston and McClelland 1973; Juola *et al.* 1974).

There was no significant difference between the nonword condition and the letter embedded in dollar signs, $F(1, 15) \approx 1$. This result is encouraging and shows that very little forgetting occurred in the present experiments. If forgetting had occurred, we would have expected much better performance for the letter in dollar signs than in the nonword condition, since there would have been much more to remember in the nonword condition.

The result of critical interest to the present study is the significant interaction of context and processing time, $F(21, 105) = 2.19$, $p < 0.001$. In order to test the model d' values were computed for the letter, word, and nonword conditions as a function of stimulus onset asynchrony, the time between the onset of the test display and the onset of the masking stimulus. The d' values were computed from the average percentage correct values given in table 2. The d' values were computed from the averages rather than the individual subject scores in order to increase the reliability of the estimates. In this analysis, the hit rates correspond to the percentage correct values and one minus these values are the false alarm rates. The nonword and letter-in-$ conditions were also averaged before the d' values were computed, since there was no significant difference between these conditions.

Fig. 1 plots the d' values as a function of context and the stimulus onset asyn-

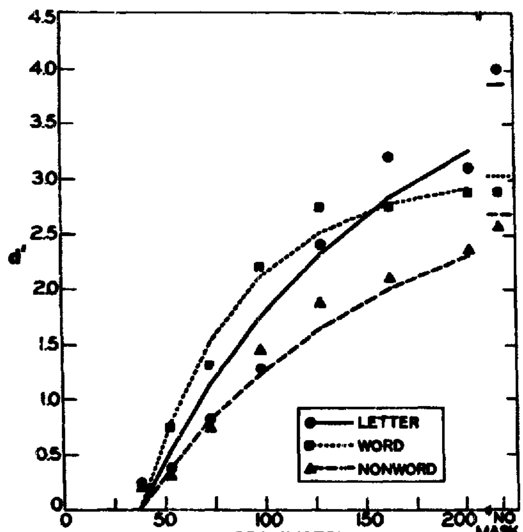

Fig. 1. Predicted (lines) and observed (points) performance in d' units as a function of letter context and the stimulus onset asynchrony (SOA) between the onset of the test display and the onset of the masking stimulus. Circles represent the single letter, squares the word, and triangles the average of the nonword and letter in $ displays.

chrony. A monotonic masking function was observed for each context condition but the interaction between context and the masking interval is the critical test of the model articulated in this paper. In terms of the model,

$$d' = \alpha (1 - e^{-\theta t}), \tag{1}$$

α should be larger for letters than words or nonwords because of the lateral masking of adjacent letters in the word and nonword conditions. The rate parameter θ, however, should be larger for words than for single letters or nonwords since orthographic structure allows the reader to arrive at a decision about which letters are present at a faster rate in the word than the single letter condition. The asymptotic clarity of the features for a letter in a word is less than for a letter presented alone but fewer of these features are necessary to arrive at a decision in the word relative to the single letter condition.

In fitting the model to the results, one value of α was estimated for the single letter condition and another value for the word and nonword conditions. The word and nonwords should have the same value of α since lateral masking should be equivalent in these two cases. With respect to the rate of processing θ, letters in words should be processed at a faster rate than letters in nonwords or a letter presented alone. Therefore, one value of θ was estimated for words and another for letters and nonwords. It was also necessary to estimate a dead time since the masking stimulus was more intense than the test stimulus. Therefore, the masking interval probably overestimated the true processing interval since the mask would have a faster arrival time at the visual processing center than would the test stimulus. Finally, it was necessary to estimate the duration that the display information was maintained in preperceptual visual storage, since it could not be available indefinitely. This duration, t_D, gives the available processing time when stimulus onset asynchrony is longer than t_D. If the masking interval exceeded t_D, then t_D was inserted in the equation.

The observed d' values were fit with the predictions of the model by estimating the six parameter values using the minimization subroutine STEPIT (Chandler 1969). The α value for the letter alone condition was 4.32, larger than the α value of 3.06 for words and nonwords. The θ value for words was 18.95, larger than the θ value of 8.45 for letters and nonwords. The dead time was estimated to be 37 msec and 266 msec was the estimated duration of preperceptual visual storage. The model provides a reasonably good description of the results, considering the fact that 24 independent data points were described with six parameter values. Although the model accurately describes the nonword condition and captures the significant crossover in the letter and word conditions, it overestimates performance in the letter condition at short SOAs. Future research will have to determine whether the deviations reflect primarily noisy data or basic inadequacies in the model. The average squared deviation between the predicted and observed values was 0.038. This fit was not improved much when unique values of α and θ were estimated for each of the three context conditions. With eight parameters, this description gave an averaged squared deviation of 0.034.

Experiment 2

One possible limitation with the present interpretation of experiment 1 is that the differences in the context conditions were observed at different levels of overall accuracy in the task (*cf.* Massaro 1975). The word advantage at short SOAs was observed when performance was poor whereas the letter advantage at long SOAs occurred with good overall performance. To eliminate the possibility that these results are unique to performance level, it is necessary to maintain average performance at a fixed level at each masking interval. If the interactions between context and processing time observed in experiment 1 are not due to the different levels of overall performance, the same results should be observed when average performance is constant at each masking condition. To this end, the target intensity was adjusted to maintain performance at 75% correct at each of the eight masking conditions.

Method

Eight *S*s from the same subject pool as those in experiment 1 were tested for five successive days. Four *S*s were tested in parallel. The on-line algorithm adjusted the stimulus parameters independently for each of the eight masking conditions maintaining overall performance averaged over the four *S*s being tested together at 75% correct at each masking condition. Therefore, plotting intensity was always the same for the four contexts at a given masking interval. Two *S*s were eliminated

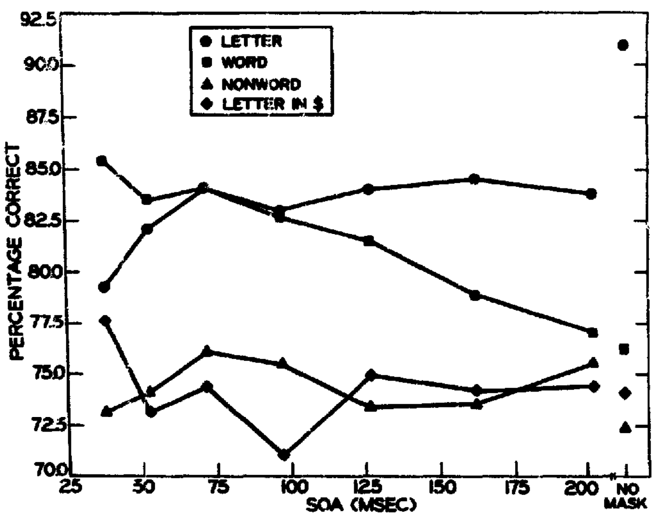

Fig. 2. Percentage of correct recognitions as a function of letter context and the stimulus onset asynchrony (SOA) between the onset of the test display and the onset of the masking stimulus. Circles represent the single letter, squares the word, triangles the nonwords, and diamonds the letter in $ displays.

from the data analysis because their overall percentage of correct responses averaged 89% and 52%, respectively, in the task. All other details were the same as in experiment 1.

Results and discussion

Fig. 2 presents the average percentage of correct identifications in the four context conditions as a function of processing time. There was a significant effect of letter context, $F(3, 15) = 21.3$, $p < 0.001$. Overall, letters were recognized about 3% more accurately when presented alone than when presented in words. Letters in words were in turn recognized about 7% more accurately than letters in nonwords or letters embedded in dollar signs. The significant interaction of the letter, word, and nonword conditions across the masking intervals replicates the results of experiment 1, $F(21, 105) = 1.81$, $p < 0.025$. The letter word difference was critically dependent upon the processing time available before the onset of the masking stimulus. Although a small word advantage was found at short masking intervals a letter advantage occurred at intervals longer than 95 msec. The word advantage decreased and the letter advantage increased with increasing processing time.

Given that overall performance was maintained at 75% correct at each masking interval, the degree of masking is indexed by the differences in the intensities of the displays as a function of processing time. In the present task, intensity is scaled between 0 and 1000, where 0 is a blank display and 1000 is maximum intensity. Table 3 gives the plotting values for each of the eight masking intervals. The results replicate the percentage correct measure in experiment 1. The intensity of the display had to be increased with decreases in the masking interval and the intensity at

Table 3
Average plotting values of the intensity of the test display as a function of processing time in experiment 2.

Processing time SOA* (msec)	
38	709
53	576
73	577
98	411
128	337
163	276
203	259
No mask	233

* SOA refers to stimulus onset asynchrony, the time between the onset of the test stimulus and the onset of the masking stimulus.

the longest masking interval was very close to that in the no mask condition. An analysis of variance on these values revealed a significant effect of masking interval, $F(7, 21) = 38.7, p < 0.001$.

General discussion

Although significant lateral masking effects were observed in the present experiments it should be pointed out that the standard experimental conditions of letter and word experiments do not always give significant lateral masking effects. Reicher (1969), for example, found no difference between single letter and nonword displays, indicating that no significant lateral masking occurred. We chose the prototypical experimental conditions in the present experiments in order to replicate previous work. However, the amount of lateral masking could be increased by decreasing the spacing between the letters and/or utilizing peripheral rather than foveal vision. Extending the present study to conditions giving larger effects of lateral masking would provide an even stronger test of the present model.

The results of the present experiment replicate the word advantage found over single letters when a masking stimulus immediately follows the test display (Hawkins *et al.* 1976; Reicher 1969; Thompson and Massaro 1973; Wheeler 1970) and a letter advantage when no masking stimulus was used (Juola *et al.* 1974). Johnston and McClelland (1973) also reported results that are consistent with the findings of the present experiments. Using a Reicher-Wheeler task, they found a large word advantage over letters when a pattern mask followed the test stimulus immediately, but a slight letter advantage in a no-mask condition. One hypothesis was that there is *less* lateral interference (masking) with a patterned mask than with no mask. Weisstein (1968) for example, locates lateral masking and backward masking at a level of the same neural mechanism. Saturating the neural mechanism with one type of masking could supposedly attenuate the amount of masking from the other type. This explanation would require an additional free parameter representing the amount of lateral masking to be estimated at each backward masking condition. The model tested in the present experiments predicts that less lateral masking will be observed at short processing intervals without requiring a direct (neural) tradeoff between lateral masking and backward masking. Lateral masking simply lowers the asymptotic perceptibility of the test display (α) and the absolute amount of lateral masking relative to a single letter control is a function of the rate of processing (θ) and the available time for processing the display before the onset of the masking stimulus. The good description of the results by the present model favors it over the interpretation offered by Johnston and McClelland (1973).

Estes (1975a,b) interprets the advantage of words over nonwords in terms of the cues to spatial position provided by familiar letter sequences and orthographic patterns. Positional information from context combined with feature input determines the perception of a letter in a particular display location. In terms of this model, the pattern mask can disrupt the cues to the location of a target letter. Estes interpreted Johnston and McClelland's results in this way; the large pattern mask in their study supposedly disrupted position cues to the target letter and the subjects may have responded on the basis of visual information from incorrect positions in the display. This explanation is a reasonable one for Johnston and McClelland's results since the area of their pattern mask was 44 times larger than the area of a single letter. In the present experiments, however, the pattern mask covered only the locations of the letters in the test display; therefore, on single letter trials, the mask occurred only at the exact location of the single test letter. It is difficult to see how the positional uncertainty hypothesis can account for the word advantage over letters in this situation since the mask is a reliable cue to the position of the test letter on single letter trials.

References

Cattell, J. M., 1886. The time it takes to see and name objects. Mind 11, 53–65.

Chandler, J. P., 1969. Subroutine STEPIT – finds local minima of a smooth function of several parameters. Behavioral Science 14, 81–82.

Estes, W. K., 1975a. Memory, perception, and decision in letter identification. In: R. L. Solso (ed.), Information processing and cognition: the Loyola Symposium. Potomac, Md.: Erlbaum Associates.

Estes, W. K., 1975b. The locus of inferential and perceptual processes in letter identification. Journal of Experimental Psychology: General 104, 122–145.

Hawkins, H. L., D. M. Reicher, M. Rogers and L. Peterson, 1976. Flexible coding in word recognition. Journal of Experimental Psychology: Human Perception and Performance 2, 380–385.

Huey, E. B., 1908. The psychology and pedagogy of reading. New York: Macmillan. [Republished by M. I. T. Press, 1968.]

Johnston, J. C. and J. P. McClelland, 1973. Visual factors in word perception. Perception and Psychophysics 14, 365–370.

Juola, J. F., D. D. Leavitt and C. S. Choe, 1974. Letter identification in word, nonword, and single-letter displays. Bulletin of the Psychonomic Society 4, 278–280.

Massaro, D. W., 1975. Experimental psychology and information processing. Chicago: Rand McNally.

Purcell, D. G., K. E. Stanovich and A. Spector, 1978. Visual angle and the word superiority effect. Memory and Cognition 6, 3–8.

Reicher, G. M., 1969. Perceptual recognition as a function of meaningfulness of stimulus material. Journal of Experimental Psychology 81 (2), 275–280.

Thompson, M. C. and D. W. Massaro, 1973. Visual information and redundancy in reading. Journal of Experimental Psychology 98, 49–54.

Weisstein, N., 1968. A Rashevsky-Landahl neural net: stimulation of metacontrast. Psychological Review 75, 494–521.

Wheeler, D. D., 1970. Processes in word recognition. Cognitive Psychology 1, 59–85.