The word superiority effect: Dependence on short-term memory factors

IRA B. APPELMAN

Stanford University, Stanford, California 94305

Two experiments were conducted to test whether the word superiority effect, that letters in words are perceived more accurately than letters in nonwords, could be attributed to short-term memory (STM) factors. One hypothesis attributed the word superiority effect to superior maintenance of words in STM. Another hypothesis was that letters in STM have considerable positional uncertainty which is overcome by the orthographic characteristics of the words. Both experiments utilized a simultaneous same-different task, where subjects compared two four-letter strings, one on top of the other, which were presented tachistoscopically. In Experiment I, the two presented strings were either both words or both nonwords and a word superiority effect was obtained. This result was interpreted as disconfirming the STM maintenance hypothesis. In Experiment II, letters were removed from one of the two letter strings, making the serial position of the comparison unambiguous. The word superiority effect disappeared. This result was interpreted as supporting the positional uncertainty hypothesis.

Several investigators have found that under tachistoscopic presentation, letters in words are perceived more accurately than single letters or letters in nonwords (Aderman & Smith, 1971; Baron & Thurston, 1973; Estes, 1975a; Reicher, 1969; Smith & Haviland, 1972; Wheeler, 1970; Spoehr & Smith, Note 1). A typical task is the brief presentation of a word, nonword, or single-letter display such as "PLOW," "PWLO," or "P" followed by the forced-choice alternative "P/S——," that is, "Was there a 'P' or 'S' in the first position of the display?" Note that for the word display either alternative creates a word and for the nonword display neither alternative creates a word. In most experiments with this postcue procedure, subjects choose correctly more often for the word than for the nonword and single-letter displays, hence the result that letters in words are apparently perceived more accurately than letters in nonwords or single letters.

The reason the word superiority effect is surprising is that (it is claimed) the effect is "perceptual." The advantage of the postcue forced-choice task introduced by Reicher (1969) is that if the subject could identify the letters "—LOW" of the above sample display, it would not be possible to guess "PLOW" at more than chance level since "SLOW" is also a word. It is also well known that subjects can remember letters in words better than letters in arbitrary nonword strings (Baddeley, 1964; Miller, 1956). Thus, it is important to show that the

advantage for words does not simply result from the subject's guessing strategy or ability to maintain words better over time.

A modification of the forced-choice postcue task is presentation of the alternatives before presentation of the word, nonword, or single-letter display. In this detection task, the advantage of words over nonwords disappears and the advantage of words over single letters reverses (Bjork & Estes, 1973; Estes, 1975a; Estes, Bjork, & Skaar, 1974; Massaro, 1973; Thompson & Massaro, 1973). Both the appearance of a word superiority effect in the postcue task and the disappearance of the word superiority effect in the precue detection task have been replicated many times by different investigators and have become the major results that require theoretical explanation.

EXPERIMENT I

Some experimenters (Aderman & Smith, 1971; Bjork & Estes, 1973; Mezrich, 1973; Reicher, 1969) have recognized that the postcue forced-choice task does involve a short-term memory component. Subjects must maintain the letter strings in memory long enough to compare them with the delayed presentation of the forced-choice alternatives. Thus, the word superiority effect may arise from the short-term memory requirements of the task. The disappearance in the detection task of the word superiority effect would be expected since the subject knows the possible alternatives prior to presentation of the word or nonword and need not maintain the display in STM waiting for the choice alternatives. Reicher originally discarded the STM maintenance hypothesis because he was the only one to get a word superiority effect in the precue detection task.

In line with a STM explanation of the word

This research was supported by a National Science Foundation grant (NSFGJ-443X3) while the author was a National Science Foundation Graduate Fellow. Richard C. Atkinson and Edward E. Smith provided useful comments on the experiments and the manuscript. Until September 1977, reprint requests should be sent to Ira B. Appelman, Department of Psychology, Stanford University, Stanford, California 94305.

superiority effect, Mezrich (1973) proposed that the representation of words and letters in STM was different, the former being verbal and the latter being visual. Since STM is largely acoustical rather than visual (Conrad, 1964), the verbally coded words would have an advantage over the visually coded letters. He presented words and single letters to the subjects in a forced-choice postcue task. There was a 1.5-sec delay between stimulus offset and onset of the choice alternatives. In that interval, subjects were required to pronounce aloud what they believed had been tachistoscopically presented. Mezrich believed that this procedure would force the subjects to maintain the single letters and words in the same verbal format. Under these conditions, the word superiority effect disappeared (and, in fact, reversed).

Another result suggestive of the STM explanation is that the advantage of pronounceable nonwords over unpronounceable nonwords is as great as the advantage of words over unpronounceable nonwords. There is no difference between words and pronounceable nonwords (Baron, 1974; Baron & Thurston, 1973; Smith & Spoehr, 1974; Spoehr & Smith, Note 1). Thus, the word superiority effect appears to be pronounceability effect. [However, Manelis (1974) has recently found differences between words and pronounceable nonwords.] Since pronounceability is the ease with which a verbal code can be generated, superior maintenance of pronounceable items in STM is implicated.

The purpose of Experiment I was to test in a more direct way whether the necessity to maintain the display in STM, such as in the postcue condition, was responsible for the word superiority effect. Subjects were presented with two word or two nonword letter strings simultaneously, one on top of the other, and asked to decide if they were the same or different. The same-different task eliminates the STM component of the postcue task without providing information about the comparison prior to presentation as in the detection task. If the word superiority effect is a STM maintenance effect, then it should not appear in the simultaneous same-different task.

Method

Subjects. Twenty-four Stanford undergraduates served as paid subjects.

Materials. The displays consisted of two four-letter strings, one on top of the other, which were either both words or both nonwords. For "same" trials, the top and bottom strings were identical. For "different" trials, the top and bottom strings differed by one letter. The "same" string that was changed for the "different" trials was placed on the bottom for half the sets and on the top for the other half of the sets. Twenty-four words were arbitrarily chosen. Each was paired with a word differing in one letter such that one-fourth of the word pairs differed in each of the four letter positions. Nonwords corresponding to the word pairs were arranged by scrambling the three letters of the word pairs which never differed so that the nonwords were not pronounceable. A sample set of word and nonword, same and different displays is presented in the left half of Table 1. The displays were typed on index cards in

Table 1
Stimulus Displays

	Experiment I		Experiment II	
	Word	Nonword	Word	Nonword
Same	WORD	RWOD	D	D
	WORD	RWOD	WORD	RWOD
Different	WORK	RWOK	K	K
	WORD	RWOD	WORD	RWOD

uppercase Orator typeface on an IBM Selectric typewriter. The letter strings were presented in an Iconix tachistoscope with a space between each letter. The entire display subtended slightly less than 1 deg horizontal angle and 1 deg vertical angle.

A 2 by 2 factorial design was employed with two levels of string type (word and nonword) and two levels of response type (same and different). The four factorial conditions and 24 different stimulus sets combined to create a complete replication of 96 trials. The trials were arranged into six blocks of 16 trials each with an equal number of word-nonword, same-different trials in each block.

Procedure. Each subject served in one session which consisted of 12 practice trials followed by 192 experimental trials, two replications of the six blocks in random order where the same block never ended the first replication and began the second. At the beginning of the session, the subjects received instructions about the same-different judgment and the trial sequence. At the beginning of each trial, the subject looked into the tachistoscope at the fixation field which consisted of the corners of a rectangle. On the experimenter's signal, the subject pressed a button initiating the trial. One second later, the fixation field was replaced by a brief presentation of the display at the center of what was the fixation field, followed by return of the fixation field. The corners of the rectangle were far enough apart to indicate the position of the display yet not overlap or mask it, and thus were preferred over a fixation point. The duration of the display was adjusted after each block so that the subjects made approximately 25% errors. The subject decided whether the top and bottom strings were the same or were different at the end of the trial. Subjects were not told whether their decisions were correct or incorrect. The session lasted approximately 45 min.

Results

The main results, percentage correct for the word-nonword, same-different conditions are presented in the left half of Table 2.

The reported percentages are means for 24 subjects, and each is based on a total of 1,152 observations. The relevant finding is the 7.1% advantage for words over nonwords. However, the word superiority effect is highly dependent on response condition (12.9% in the same condition, while only 2.2% in the different condition) as evidenced by the significant interaction between

Table 2
Percentage of Correct Choices for Word-Nonword
Same-Different Conditions

	Experiment I			Experiment II		
	Word	Non-word		Word	Non-word	
Same	77.2	64.3	70.8	67.4	65.4	66.4
Different	77.0	74.8	75.9	78.6	80.6	79.6
	77.1	70.0		73.0	73.0	

Table 3
Percentage of Correct Choices as a Function of Serial
Position for the "Different" Conditions

	Experiment I				Experiment II			
	Serial Position				Serial Position			
	1	2	3	4	1	2	3	4
Word	91.3	70.5	59.0	87.1	86.5	80.9	70.5	76.7
Nonword	91.3	69.1	54.9	84.0	86.1	79.2	77.1	79.9

stimulus and response type. A two-way analysis of variance indicated a highly significant effect of stimulus type (word vs. nonword) [$F(1,23) = 24.8$, $p < .001$, $MS_e = 54.5$], a nonsignificant effect of correct response type (same vs. different) ($p > .10$), and a significant interaction [$F(1,23) = 10.7$, $p < .005$, $MS_e = 63.9$].

As each word and nonword was four letters in length, it is possible to examine errors in the "different" condition as a function of the serial position of the difference. The left half of Table 3 lists the percentage correct for the word and nonword different conditions.

It is important to realize that effects of serial position per se are not meaningful, since no effort was made to control for difficulty of comparison across serial position. That is, the words "SELL" and "FELL" differ in the letters "S" and "F," which are relatively distinct, while the words "PACE" and "PAGE" differ in the letters "C" and "G," which are relatively indistinct. Therefore, a serial position effect (or lack of serial position effect) could result from unequal difficulty of comparison across serial position. The important comparison is between the word and nonword serial position curves. A two-way analysis of variance indicated a significant effect of serial position [$F(3,69) = 37.6$, $p < .001$, $MS_e = 309.9$] and a nonsignificant effect of stimulus type (word vs. nonword) ($p > .10$), with a nonsignificant interaction ($p > .25$). The mean display time for the 24 subject means was 22.0 msec, with a range of 9.4 to 38.2 msec.

Discussion

The hypothesis that the word superiority effect arises from a greater capacity to maintain words vs. nonwords in STM over time is clearly not supported by the data. The advantage for words appeared in the simultaneous same-different task, where necessity to maintain stimuli over time is minimized. Unlike the postcue procedure, the subject need not maintain the display in STM until the choices appear.

One issue relevant to the same-different task is: At what level of abstraction are the two letter strings being compared (Posner, 1969)? Two possible matching strategies are letter (character) matching and word (name) matching. Eichelman (1970) conducted a same-different task similar to

Experiment I, except he recorded reaction time, the time it took for the subject to make the same-different judgment. Eichelman varied the number of letters that differed in the two letter strings and also the type font, upper- vs. lowercase, of the two letter strings. Reaction time in the "different" condition was a decreasing function of the number of different letters. Reaction time for letter strings in different type font was greater than reaction time for strings of the same type font. These two results suggest that a name match was not being conducted, since the physical features of serial position and type font affected the comparison times.

As pointed out previously, the exact character of the serial position effects in Table 3 cannot be taken seriously because there was no attempt to control the similarity of letter comparisons across serial position. However, the fact that there are serial position effects is not consistent with a name match. That is, if the strings were being compared by string name, there should be no serial position effects. The results here and those of Eichelman (1970) indicate that a more physical type of match is taking place. Another result of interest is that, although there are serial position effects, they are basically of the same character for both word and nonword displays; there is no interaction. This result suggests that whatever comparison process is going on, it is the same for both word and nonword displays.

Most of the studies of the word superiority effect require the subject to make a binary decision, postcue forced-choice or precue detection, and time of presentation is brief enough so that the subject makes a large number of errors. Smith and Spoehr (1974) have called these studies "accuracy" studies and distinguished them from "reaction time" studies in which time of presentation is not brief and the time it takes the subject to make the binary decision is measured. There is a methodological issue of whether reaction time and accuracy studies will yield the same results; does the choice of measure affect the character of the results? The simultaneous same-different task has been used by Baron (1974) and Eichelman (1970) in a reaction time task. The results were essentially the same as those of Experiment I. Reaction time for words was less than for nonwords. Smith and Spoehr (1974) have pointed out that in most cases reaction time and accuracy measures have yielded similar results as in the simultaneous same-different task they now have.

EXPERIMENT II

The results from Experiment I suggest that the word superiority effect does not arise from an advantage over time of words vs. nonwords in STM. Estes (1975a, b) has proposed that the word superiority effect results from positional uncertainty

of letters in STM. The orthographic characteristics of words reduce positional uncertainty and result in superior performance for words over nonwords. Estes (1975a) presented words and nonwords tachistoscopically and required the subjects to recall the letter in a postcued position. He separated errors into three types: (1) omissions—failure to recall a letter, (2) intrusions—producing a letter that was not in the display, and (3) transpositions—recalling a letter in the display but from a position other than the one cued. The word superiority effect resulted entirely from less transposition errors for words than for nonwords. Thus, Estes attributed the superiority effect to positional uncertainty in STM overcome by the orthographic characteristics of the words.

In a related experiment, Johnston and McClelland (1974) presented words and nonwords to subjects in a postcue forced-choice task. On half the trials, the subjects were instructed to look at the whole item which was a four-letter word or nonword. On the other half of the trials, the subjects were informed prior to each trial which position in the display would contain the critical letter. On nonword trials, as expected, the position precue increased performance accuracy. However, paradoxically, on word trials, knowing the position of the critical letter decreased performance accuracy. The authors point out the scarcity of studies where knowing where in the display an item would occur decreased performance. These paradoxical results are in line with a positional uncertainty explanation of the word superiority effect. In the nonword condition, there is a lot of positional uncertainty and precuing the subject for position reduces that uncertainty and thus increases accuracy. In the word condition, positional uncertainty is not that great. Providing position precues reduces uncertainty by allowing the subject to attend to a particular display position, but increases uncertainty by nullifying the value of the orthographic characteristics of the words. Even if the paradoxical decrease in word accuracy does not hold up and there is simply no increase in word accuracy with position precues, position cues will still reduce the difference between word and nonword performance which is the word superiority effect.

It was pointed out above that the word advantage is really a pronounceability advantage. Pronounceability was suggestive of a verbal code which could be better maintained in STM. Actually, most investigators have not distinguished between pronounceability and orthographic characteristics. Adelman and Smith (1971) controlled specifically for orthographic characteristics. They constructed letter strings which contained either three spelling patterns (Gibson, 1965, p. 1071), e.g., "SWILG," or unrelated letters, e.g., "LGISW." For each item, they obtained ratings of pronounceability. They obtained a main effect of orthographic vs. unrelated letter strings, but the

correlation between accuracy and rated pronounceability was extremely small (.007). It is not, however, advisable to claim too much for a distinction between two factors (orthographic characteristics and pronounceability) that are so highly confounded. Because the positional uncertainty hypothesis relies on the orthographic characteristics of words to explain the word superiority effect and does not attach importance to pronounceability, Adelman and Smith's results are very compatible with the positional uncertainty explanation.

If positional uncertainty gives rise to the word superiority effect, then the results of Experiment I could be explained in the following way. In the simultaneous same-different task, as the letters are entered in STM before comparison, there is considerable positional uncertainty which would affect which letters are compared with which during the comparison stage. The orthographic characteristics of the words decrease positional uncertainty and thus improve the comparison process resulting in the word superiority effect.

In the same way that Johnston and McClelland (1974) reduced word performance and increased nonword performance by providing positional precues, the positional uncertainty hypothesis implies a reduction in the word superiority effect for the simultaneous same-different task if positional cues are provided. This was accomplished in Experiment II for the simultaneous same-different task by providing unambiguous cues in the display for which position contained the critical comparison. If the positional uncertainty hypothesis is correct, position cues should reduce or eliminate the word superiority effect.

Method

Subjects. Twenty-four Stanford undergraduates who had not participated in Experiment I served as paid subjects.

Materials and Procedure. The stimuli and procedure were exactly the same as in Experiment I, with the following exception. For each of the 24 word-nonword, same-different sets, there was a critical letter that was the same in both strings for the "same" trials and which changed in one string for the "different" trials. In Experiment II, the display consisted of only one word or nonword string, with the critical comparison letter either above or below the corresponding letter in the string. The right half of Table 1 presents a sample set of the modified display. In effect, three of the four letters of the second string were removed and the position of the critical comparison became more obvious.

Results

The main results, percentage correct for the word-nonword, same-different conditions, are presented in the right half of Table 2. The reported percentages are means for 24 subjects, and each is based on a total of 1,152 observations. A two-way analysis of variance indicated a nonsignificant effect of stimulus type (word vs. nonword) ($p > .25$), a highly significant effect of correct response type (same

vs. different) [$F(1,23) = 18.6$ $p < .001$, $MS_e = 225.0$], and a nonsignificant interaction ($p > .10$).

Errors for the different condition as a function of serial position are listed in the right half of Table 3. There was a significant effect of serial position [$F(3,69) = 10.5$, $p < .001$, $MS_e = 122.7$], a nonsignificant effect of display type (word vs. nonword) ($p > .25$), and a nonsignificant interaction ($p > .10$). The mean display time for the 24 subject means was 14.9 msec, with a range of 8.8 to 26.8 msec.

Discussion

Implicit in most discussions of the word superiority effect is that its cause is the same no matter where it occurs. Such an assumption should be treated with caution. However, the fact that the word advantage can be eliminated in the simultaneous same-different task by providing position cues is in line with positional effects in the postcue forced-choice (Johnston & McClelland, 1974) and the postcue recall (Estes, 1975a) tasks. These results implicate positional effects as important in the explanation of the word advantage.

The positional uncertainty explanation of the word superiority effect was proposed by Estes (1975a, b) as a part of a general model of letter perception. As conceptualized here, the major results are explained with the following assumptions: (1) letters entering STM have considerable positional uncertainty, and (2) orthographic characteristics provide usable information that reduces positional uncertainty. With these assumptions, it is possible to predict the word superiority effect by assessing the positional uncertainty in displays before and after an experimental manipulation.

Most theoretical discussions of the word superiority effect have been phrased in terms of two competing hypotheses: inference and unitization. The inference hypothesis is that "wordness" supplies redundant information which can be used by the perceptual system. The unitization hypothesis is that "wordness" allows for processing of higher order, and thus fewer, units (e.g., syllables), thus requiring less work by the perceptual system. Most problems in word perception have centered around controlling for inference, since, if the subject can guess better in the word condition, then an obtained word superiority effect could not be considered "perceptual." Reicher and others using the postcue task have claimed that the forced-choice procedure where both alternatives create a word does not allow an advantage for words due to inference. Thompson and Massaro (1973) using the precue task claimed that whether there could be inference or not was dependent on processing assumptions. For example, if the word and nonword displays were "WORD" and "RWOD," respectively, if the subject sees "WOR" and a curve in Position 4, the subject

could either know that the fourth letter was "D" or at least eliminate alternative letters that also have curves in them but do not make a word. However, if the subject sees "RWO" and a curve in the fourth position, it is not possible to eliminate the alternatives. Thompson and Massaro have therefore claimed that the postcue is not an adequate control for redundancy and that the detection task is.

Smith and Haviland (1972) tried to meet the above objections by making the nonwords more redundant. They constructed nonwords according to rules well known by their subjects and used the postcue forced-choice procedure. If the word advantage resulted from inference allowed by redundancy in word displays and if the subjects could use the redundancy in the nonword displays which was well explained to them, then the word advantage should disappear. The word superiority effect did not disappear, and, in order to show that the subjects simply were not well acquainted with the redundancy rules for the nonwords, they tested the subjects over several sessions and showed that the word advantage did not diminish even though the subjects were undoubtedly becoming more familiar with the rules for generating the nonword displays.

The appearance of the word superiority effect when redundancy is controlled in the postcue forced-choice task is in conflict with the Thompson and Massaro (1973) version of the inference hypothesis. The conflict is resolved (Massaro, 1973) by assuming that even though Smith and Haviland's subjects participated in the experiment for 4 days, the word and nonword trials were not arranged in sufficiently long blocks for the subjects to become proficient in using the nonword redundancy rules.

The positional uncertainty hypothesis can easily accommodate the apparent conflicts in the results. In the detection task, the subject knows the letter he is looking for prior to seeing the display. Therefore, the fact that there is more positional uncertainty in nonwords than in words does not matter because position is not important in the task. In the postcue forced-choice task, the position of the tested letter is critical and thus there is an advantage for the word condition. Varying redundancy in the nonword displays with artificial rules known to the subjects does not affect the word advantage because it does not affect the positional uncertainty in the display.

REFERENCE NOTE

1. Spoehr, K. T., & Smith, E. E. *Perceptual units and parsing processes in word perception*. Paper presented at the annual meeting of the Psychonomic Society, St. Louis, Missouri, November 1972.

REFERENCES

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

- BADDELEY, A. D. Immediate memory and the "perception" of letter sequences. *Quarterly Journal of Experimental Psychology*, 1964, **16**, 364-367.
- BARON, J. Successive stages in word recognition. In S. Dornic & P. M. A. Rabbitt (Eds.), *Attention and performance* (Vol. 5). New York: Academic Press, 1974.
- BARON, J., & THURSTON, I. An analysis of the word-superiority effect. *Cognitive Psychology*, 1973, **4**, 207-228.
- BJORK, E. L., & ESTES, W. K. Letter identification in relation to linguistic context and masking conditions. *Memory and Cognition*, 1973, **1**, 217-223.
- CONRAD, R. Acoustic confusion in immediate memory. *British Journal of Psychology*, 1964, **55**, 75-84.
- EICHELMAN, W. H. Familiarity effects in the simultaneous matching task. *Journal of Experimental Psychology*, 1970, **86**, 275-282.
- ESTES, W. K. The locus of inferential and perceptual processes in letter identification. *Journal of Experimental Psychology: General*, 1975, **104**, 122-145. (a)
- ESTES, W. K. Memory, perception, and decision in letter identification. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium*. Potomac, Md: Erlbaum Associates, 1975. (b)
- ESTES, W. K., BJORK, E. L., & SKAAR, E. Detection of single letters and letters in words with changing vs. unchanging mask characters. *Bulletin of the Psychonomic Society*, 1974, **3**, 201-203.
- GIBSON, E. J. Learning to read. *Science*, 1965, **148**, 1066-1072.
- JOHNSTON, J. C., & MCCLELLAND, J. L. Perception of letters in words: Seek not and ye shall find. *Science*, 1974, **184**, 1192-1194.
- MANELIS, L. The effect of meaningfulness in tachistoscopic word perception. *Perception & Psychophysics*, 1974, **16**, 182-192.
- MASSARO, D. W. Perception of letters, words, and nonwords. *Journal of Experimental Psychology*, 1973, **100**, 349-353.
- MEZRICH, J. J. The word superiority effect in brief visual displays: Elimination by vocalization. *Perception & Psychophysics*, 1973, **13**, 45-48.
- MILLER, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 1956, **63**, 81-97.
- POSNER, M. I. Abstraction and the process of recognition. In J. T. Spence & G. H. Bower (Eds.), *Advances in learning and motivation* (Vol. 3). New York: Academic Press, 1969.
- REICHER, G. M. Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, 1969, **81**, 275-280.
- SMITH, E. E., & HAVILAND, S. E. Why words are perceived more accurately than nonwords: Inference versus unitization. *Journal of Experimental Psychology*, 1972, **92**, 59-64.
- SMITH, E. E., & SPOEHR, K. T. The perception of printed English: A theoretical perspective. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition*. Potomac, Md: Erlbaum Press, 1974.
- THOMPSON, M. C., & MASSARO, D. W. Visual information and redundancy in reading. *Journal of Experimental Psychology*, 1973, **98**, 49-54.
- WHEELER, D. D. Processes in word recognition. *Cognitive Psychology*, 1970, **1**, 59-85.

(Received for publication January 27, 1975;
revision received June 9, 1975.)