The Role of Phonetic and Orthographic Similarity in Picture-Word Interference*

Stephen J. Lupker
University of Western Ontario

ABSTRACT Picture-word interference refers to the fact that when a picture (i.e., line drawing) is presented with a word superimposed, picture-naming latency is longer than when the same picture is presented alone. This phenomenon, like the Stroop phenomenon, seems to be strongly influenced by the nature of the superimposed word. In the present paper the effects of phonetic and orthographic similarity between the word and the picture's name were investigated in order to get a clearer idea of the role these factors play in the picture-naming process. Both orthographic and phonetic similarity were found to facilitate picture naming in comparison to an unrelated word condition although not with respect to a picture alone condition. Further, the data suggest that this facilitation is not an output phenomenon and, as such, can be separated from the response competition processes that lead to the basic interference effect. Instead, the locus of these effects appears to be the name-retrieval process with orthographic and phonetic information from the word aiding in the search for the picture's name.

RÉSUMÉ L'interférence mot-image se rapporte au fait que lorsqu'une image (i.e., un dessin) est présentée avec un mot en surimpression, la latence d'identification de l'image est plus longue que lorsque la même image est présentée seule. Ce phénomène, comme celui de Stroop, semble être fortement influencé par la nature du mot en surimpression. Les effets dus à la similitude phonétique et orthographique existant entre le mot et le nom de l'image ont été étudiés dans le but de clarifier le rôle joué par ces facteurs dans le processus d'identification de l'image. Les similitudes orthographiques et phonétiques facilitent toutes deux l'identification de l'image comparativement à la condition où un mot non relié est utilisé mais non par rapport à la situation où l'image seule doit être identifiée. De plus, les données suggèrent que cette facilitation n'est pas liée à la sortie de l'information et peut, en ce sens, être dissociée des processus de compétition de la réponse qui conduisent à l'effet d'interférence. Les effets trouvés résideraient plutôt dans le processus de recouvrement du mot dont la similitude orthographique et phonétique favorise la recherche du nom de l'image.

What are the mechanisms involved in the simultaneous processing of two inputs when the observer must respond to only one of those inputs? This question, in many shapes and forms, has been with us at least since the

*This research was supported by Grant A6333 from the Natural Sciences and Engineering Research Council of Canada. The author would like to thank Anne Kennedy, Eric Cartman and Mike Callahan for their help in the data collection and analyses and Albert Katz for his many helpful suggestions on earlier drafts of this manuscript. Address reprint requests to Stephen J. Lupker, Department of Psychology, University of Western Ontario, London, Ontario, Canada, N6A 5C2.

time of Stroop (1935) and, today, lies at the heart of psychology's current interest in interference paradigms.

The classic interference paradigm involves the Stroop colour-word phenomenon. The stimuli are words which name colours printed in incongruent ink colours. The subject's task is to name the ink colour, ignoring the word. Typically, colour-naming latencies will be much longer in this situation than in a control condition where the stimuli are colour patches whose colours are to be named.

The present research focuses on what appears to be a very similar interference paradigm. The stimuli are pictures, represented as line drawings, with words superimposed. The subject's task is to ignore the word and name the picture. Typically, picture-naming latency is longer in this condition than in a control condition in which the same pictures are presented alone. This effect has been called the picture-word interference phenomenon.

The standard explanation of both the Stroop and picture-word interference phenomena have centred on response competition processes (Dyer, 1973; Klein, 1964; Lupker, 1979; Posner & Snyder, 1975). While the subject is actively processing the relevant input (the colour or the picture), he or she is also passively (or automatically) processing the word. Because reading a word is such an overlearned behaviour, the word's name will become available before the name of the relevant input. This creates a situation where the word's name will occupy a preeminent position in the single motor-output channel. In order to produce a response the subject must clear this channel by suppressing the tendency to say the word's name, a process which takes time. Thus, it is this suppression process which presumably accounts for the difference between the Stroop and picture-word interference conditions and their respective control conditions.

The processes involved in responding to these two component stimuli, however, are a bit more complex than this simple explanation implies. In particular, the amount of interference created by a given word is a function of a number of factors. Using the Stroop task, Klein (1964) has reported that words which name colours (the standard condition) are more interfering than colour-related words (e.g., lemon, grass, fire, sky) which, in turn, are more interfering than common words (e.g., put, take, heart, friend). Similarly, using the picture-word interference task, Rosinski (1977) reported that words in the same semantic category as the pictures they appear on cause more interference than words crossing category boundaries. Conversely, Seymour (1977) and Stirling (1979) have both demonstrated a response facilitation in the Stroop task when the word is semantically related to its own ink colour. Apparently, the words in these tasks automatically provide not only their phonetic codes but also a certain amount of semantic information as well. The nature of this information as well as its relationship to the relevant input (the colour or the picture) both seem to be determinants of the time needed to produce the appropriate response.

Clearly, the output process is a very important process in tasks of this sort. However, because of results such as those mentioned above a number of investigators have recently suggested that this process can not be the sole determinant of the time needed to respond to a two component stimulus (Hock & Egeth, 1970; Lupker & Katz, 1981; Seymour, 1977; Stirling, 1979; Williams, 1977). Focussing on the picture-word interference task, Lupker and Katz examined two pre-output stages in picture processing in order to determine their potential contributions to the observed interference. Their conclusions were as follows. First, an input process characterized by the resolution of figure and ground accounts for little, if any, interference. Typically, the size of this effect tended to be approximately 10 msec and was probably due to lateral masking. Second, a decision process in which the subject must evaluate the acquired pictorial information in order to make the decision the task demands (here, retrieving the picture's name) can be affected by semantic factors in certain situations. In particular, when the semantic information automatically available from the word is congruent with the pictorial information being evaluated, subjects appear to have difficulty determining that the word's name is inappropriate for the picture, prolonging decision time. When the word and the picture are unrelated, no such problems appear to arise. Finally, the lion's share of the interference does appear to be due to an output process. This process, as described above, involves the suppression of the tendency to say the word and it does appear to be sensitive to factors which affect the strength of this tendency.

The purpose of the present paper is to investigate another set of factors which appear to affect picture-word interference in order to understand their influence on picture processing. Recently Rayner and Posnansky (1978; Posnansky & Rayner, 1977, 1978) have reported results which suggest that the amount of interference observed in the picture-word interference task is a function of figural, orthographic, and phonetic factors. In particular, Rayner and Posnansky observed a facilitation in picture naming when (1) the superimposed letter string maintained the overall shape of the picture's name (e.g., "loaf" or "loef" on the picture of a leaf), (2) the superimposed letter string had the same first letter as the picture's name, and (3) a superimposed nonword was pronounced the same as the picture's name (e.g., "burd" on the picture of a bird). Thus, it appears that these three forms of similarity may somehow act to speed up the picture-naming process in the picture-word interference context.

Unfortunately a number of methodological problems in Rayner and Posnansky's studies make these effects a bit difficult to interpret. The most serious of these is their practice of blocking their experimental conditions. In their task, subjects undoubtedly became aware very quickly of the relationship between the pictures and the letter strings appearing on all subsequent trials in a block. In the phonetic similarity conditions (i.e., "burd" on the picture of a bird), knowledge of this relationship probably led the subjects to simply pronounce the nonword and ignore the picture entirely. In the same first letter conditions, this knowledge may have induced the subjects to focus on the first letter and to use the phonetic information from it to initiate a response, hence, tripping the voice key and stopping the timer before picture processing had been completed. Thus, in light of these potential problems it appears that the influence of phonetic and orthographic similarity on the normal processing of pictures remains unclear.

Even if subjects were not using the strategies outlined above, there would still be some problems in interpreting many of Rayner and Posnansky's results. The first problem pertains to the figural similarity effect. Unless a substantial amount of orthographic and phonetic similarity is involved as well (e.g., loaf-leaf), this effect tends to be restricted to situations in which stimulus on-time is quite brief. Yet, in these situations letter confusions should be quite frequent suggesting that this effect may be partly due to subjects mistakenly perceiving some of the actual letters in the picture's name. Thus, it becomes somewhat difficult to separate this effect from any effects of orthography and, potentially, even phonetics. The second problem, which the authors acknowledge, is that it is difficult to manipulate orthographic and phonetic similarity independently. Rayner and Posnansky's phonetic conditions generally also involved a substantial amount of orthographic similarity (burd-bird). Likewise in those conditions which can be used to argue for an orthographic effect (i.e., their same first letter conditions), a substantial amount of phonetic similarity is also involved. Thus, it appears that, at present, it is not possible to separate the effects of orthography and phonetics from one another as well as from the effects of subject strategies.

The present paper will focus on the effects of orthographic and phonetic similarity on picture processing in an attempt to explain those effects within the framework of a general processing model. If it can be shown that phonetic similarity actually does facilitate picture naming, there would be two potential loci for this effect. The first would be the process of determining the picture's name. Numerous investigators have argued for the role of semantic processing in retrieving a picture's name (e.g., Nelson, Reed, & McEvoy, 1977; Smith & Magee, 1980) suggesting that semantic information from the word could influence this process, as demonstrated by Lupker and Katz (1981). However, very few models have suggested a possible role for phonetic factors in this search process. One exception would be the semantic/lexical network model of Collins and Loftus (1975). In this model there is a semantic network organized along the lines of semantic similarity and a corresponding lexical network organized along the lines of phonetic similarity. Concept names are stored in the lexical network. Initially, pictures may only allow access to the semantic network; however, it would be necessary to access the lexical network as well in order to retrieve the picture's name. Words, on the other hand, would allow rapid and direct access to the lexical network. This access would be followed by a spread of excitation among phonetically similar concept names, serving to increase their accessibility. Thus, it would be easier for a phonetically similar picture's name to be located, resulting in a shorter picture-naming latency.

The other process which could be facilitated by phonetic similarity would be the output process. As described earlier, this process involves the suppression of the tendency to say the word and it does seem to be sensitive to factors which influence the strength of this tendency. For example, pronounceable nonwords which are orthographically regular and phonetically viable would not be expected to produce the same tendency toward pronunciation as common, unrelated words, due to their unfamiliarity. In fact, Lupker (1979) has reported that unrelated words typically do produce 15-20 msec more interference than pronounceable nonwords. Further, Lupker and Katz (1981) have reported that when the influence of the output process is controlled the word-nonword difference disappears, implicating this process as the locus of the effect. Phonetic similarity may influence this process in a somewhat similar manner. That is, suppression of the tendency to say the word's name may be simpler when fewer components of its articulatory code must be changed in order to produce the correct response. Thus, this process, as well, may be the locus of any phonetic facilitation.

If it can be demonstrated that orthographic similarity facilitates picture naming, the locus of the effect would be much clearer. Various aspects of a word's phonetic code may determine how easily its name can be suppressed. However, the identities of the letters giving rise to those codes should be irrelevant to this process. Thus, any effects of orthographic similarity should be attributable to the name-retrieval process. In fact, Collins and Loftus (1975) have suggested, at least parenthetically, that the structure of their lexical network may be based to some degree on orthographic, as well as phonetic, similarity. Thus, a priming explanation similar to that suggested above may explain any effects of orthographic similarity on the name-retrieval process.

EXPERIMENT I

The main focus of Experiment I will be the orthographic similarity effect, whereas the influence of phonetic similarity will be investigated in Experiment II. In both of these studies two methodological considerations will always be in force. First, conditions will not be blocked, so that subjects will not be aware of when an upcoming trial involves a word carrying useful information. Second, none of the various manipulations will involve the word's first letter. Thus, the initial phoneme which the subjects produce to stop the timer will always be different from the first phoneme in the word.

In an attempt to evaluate the contributions of orthographic similarity independent of the contributions of phonetic similarity, the orthographic effect was looked at in two different ways. The first involved the creation of an orthographic similarity condition using real words. For this condition pairs of words were selected such that their spellings were identical, except for the first letter, but their single vowel sounds were different in their standard pronunciations. Examples would be bear-year, hand-wand, etc. One member of each pair became the word and the other was the picture in Condition 1. In the other conditions, the same pictures were used. In Condition 2, the unrelated word condition, the same words appeared on different pictures (e.g., "YEAR" was superimposed on the picture of a hand). In Condition 3, the pictures appeared without a letter string superimposed. These two conditions will provide baselines against which to evaluate the orthographic similarity condition. Condition 2 will provide a measure of how interfering this particular set of words is, whereas Condition 3 will provide a measure of how rapidly this particular set of pictures can be named.

Ideally, in order to investigate the effects of orthographic similarity, the word pairs selected for this experiment should have been totally dissimilar phonetically. However, due to the nature of the English language it was not possible to find more than a few orthographically similar word pairs which met so strict a criterion. The criteria which were used, while not quite as strict, did guarantee that any phonetic similarity which did arise would be minimal and would be confined to the very last phoneme or phonemes. Thus, any contribution of phonetic factors to the "orthographic" effect in Experiment I should be quite minimal. Nonetheless, in order to evaluate the effects of orthographic similarity more fully, a second orthographic manipulation was introduced. For this manipulation a new letter string was generated from the name of each picture. All letters in the picture's name were replaced by Xs except the letters beyond the final vocalized vowel (e.g., "XXR" was generated from oar, "XXRT" was generated from dart). Thus, a context was created in which relevant orthographic information was available but the generation of a phonological representation was essentially impossible (Condition 4). This condition can be compared to a second condition (Condition 5) in which these same letter strings appeared on different pictures. A difference between conditions would be evidence for orthographic facilitation in the absence of phonetic information. The lack of a difference would not only indicate that orthographic similarity does not facilitate picture naming, but it would also point toward the phonetic information from these final letters as the cause of much of the facilitation in the other orthographic condition.

A second question being examined in Experiment I is whether figural similarity alone plays a role in any orthographic similarity effect. As noted earlier, Rayner and Posnansky (1978) have argued that, when using lower case letters, word shape information can facilitate picture naming. Although this effect may be due solely to subjects mistakenly perceiving certain of the letters, the possibility exists that figural similarity may have an effect above and beyond that of orthographic similarity. In the present studies the availability of word shape information was quite limited since only upper case letter strings were used. However, it may be that figural similarity at the individual letter positions can also facilitate responding. If so, any facilitation observed in the orthographic similarity conditions in Experiment I may be a function of both the letters' forms as well as their identities.

In order to examine the question of whether figural similarity between the superimposed letters and the picture's name facilitates picture naming, two additional conditions were created. In Condition 6, each letter of the picture's name was replaced by the letter it was most frequently confused with in the data of Townsend (1971). In Condition 7, these same letter strings were superimposed on other pictures. These letter strings consisted almost entirely of consonants and were essentially unpronounceable. Thus, hereafter, these conditions will be referred to as the consonant string conditions. In both of these conditions only a minimal tendency toward pronunciation should be evoked; thus, little, if any, output competition should arise. More importantly, if simple figural similarity between the letters in the letter string and those in the picture's name facilitates responding, the first of these conditions should lead to shorter picture-naming latencies than the second.

Method

Subjects: Twenty University of Western Ontario undergraduate volunteers (6 males and 14 females) received course credit for participating in this experiment twice in one 45 minute session. All were native English speakers.

Materials and Equipment: Twelve word pairs were selected which met the following criteria: (1) their spellings are identical except for the first letter; and (2) their single vowel sounds are different in their standard pronunciations. Line drawings ("pictures") were obtained for one member of each pair from a children's colouring book. Seven sets of these twelve pictures were produced, and each picture was glued on a 23 x 25.6 cm card. These seven sets of pictures corresponded to the seven experimental conditions.

To create Condition 1, the names of the other members of the pairs were superimposed on the appropriate pictures in the picture set. In Condition 2, the picture set contained the same set of words superimposed on different pictures. In Condition 3, the picture alone condition, no letters appeared on the pictures. For the fourth and fifth conditions, strings of Xs were created from each picture's name. This was done by replacing each letter of the name by an X except those letters beyond the final vocalized vowel. In Condition 4, each string of Xs was superimposed on the picture used to generate it. In Condition 5, each of these strings of Xs was superimposed on one of the other pictures. To create the sixth and seventh conditions, letter strings were created from each picture's name by replacing each letter of the name by the letter it was most frequently confused with in the confusion matrix data reported by Townsend (1971). In Condition 6, each letter string was superimposed on the picture used to generate it. In Condition 7, these same letter strings were superimposed on one of the other pictures. The names of the pictures, the words and the other letter strings used are reported in the Appendix.

A Ralph-Gerbrands Co. (Model 1-3B-1C) three-field tachistoscope was used to present the stimuli. Viewing distance was 77 cm and viewing was binocular. The letters typically subtended a visual angle of .24° horizontally and .36° vertically. The pictures subtended visual angles between 1.90° and 5.74° horizontally and 3.84° and 5.74° vertically. A Hunter Klockounter (Model 120) timer was used to time the subject's vocal, picture-naming responses. An ElectroVoice Inc. (Model 621) microphone was positioned approximately 7 cm away from the subject's mouth. The microphone was connected to a Lafayette Instruments Co. (Model 18010) voice-activated relay which stopped the timer at the initiation of the subject's vocal response.

Procedure: Each subject was tested individually. The subjects were informed they would be seeing a series of pictures, some of which would have words superimposed, and their job would be to name the pictures as rapidly as possible without making any mistakes. They were also informed beforehand of the names of the pictures they would be seeing although they were not shown an example picture-word stimulus. The subjects then responded to each of the 84 stimuli in a random order. Each stimulus remained in view for 750 msec, regardless of the subject's reaction time. The response-stimulus interval was then used by the experimenter to record the picture-naming latency and reset the equipment for the next trial. Thus, this time was not held constant but was generally around five sec. Errors were recorded and those pictures were randomly placed back into the set of to-be-presented stimuli. Following a brief rest, subjects again responded to the same stimuli in a different random order. The entire procedure took about 45 min.

Results

As is typically the case in vocal reaction time tasks errors were virtually nonexistent (less than 1% in all conditions). Thus, the few error trials were not analyzed.

TABLE I

Reaction times (msec) as a function of experimental conditions in Experiments I and II

	Experiment I (Example picture: foot)						
	Ur	Or	PA	X/Or	X/Ur	CS/F	CS/NF
$\overline{RT}$ (msec)	754	698	676	665	690	692	697
(Example letter strings:	BAR	BOOT		XXXT	XXR	TQQL	QRP)
	Experiment II (Example picture: plane)						
	Word Conditions						
	Ur	Ph	Or/Ph	$\bar{X}$	PA		
$\overline{RT}$ (msec)	701	678	646	675	632		
(Example words:	POWER	BRAIN	CANE)				
	Pronounceable Nonword Conditions						
	Ur	Ph	Or/Ph	$\bar{X}$			
$\overline{RT}$ (msec)	680	665	627	657			
(Example nonwords:	VOOSE	TAIN	NANE)				
$\bar{X}$	691	672	636				

Or = orthographically similar letter string; Ph = phonetically similar letter string; Ur = unrelated letter string; PA = picture alone; CS = consonant string; F = figural similarity; NF = no figural similarity; X = string of Xs.

The mean correct reaction times for the seven conditions are presented in the upper panel of Table I. Each of these data points is based on 480 observations. As is obvious, the main effect of conditions was highly significant ($F(6, 114) = 18.24, p < .001$).[1] This effect was examined more closely through the use of five pairwise planned comparisons. The first three involved the orthographic similarity, unrelated word, and picture alone conditions. The mean in the unrelated word condition was significantly larger than the mean in either of the other two conditions ($p < .01$). The 22 msec difference between these latter two conditions was also significant ($p < .05$). The other two comparisons involved the four nonword conditions. The mean in the Xs with orthography condition was significantly less than that in the Xs without orthography condition ($p < .01$). However, the 5 msec difference between the two consonant string conditions did not approach significance.

The main effect of trial block was also significant ($F(1, 19) = 52.24, p < .001$) indicating that subjects improved with practice. However, this factor did not interact with conditions ($F < 1.0$).

Discussion

There were two major results in Experiment I. The first was the highly significant difference between the unrelated word condition and the orthographic similarity condition. Words orthographically similar to the picture's name produced picture-naming latencies 56 msec faster than those in the unrelated word condition. The second was the significant difference between the two Xs conditions. The Xs with orthography condition produced picture-naming latencies 25 msec faster than the Xs without orthography condition. The size of this difference should, of course, be smaller than that between the orthographic similarity and unrelated word conditions, since less orthographic information is available to facilitate the picture-naming process. In any case, these two results demonstrate that even when phonetic similarity is controlled orthographic similarity is an important factor in picture-word interference.

These findings support Rayner and Posnansky's conclusion that orthographic similarity facilitates the picture-naming process in the picture-word interference context. The facilitation observed, especially in the orthographic similarity condition, is not facilitation in the normal sense of the word, however, since the latency in this condition is not less than that in the picture alone condition. In line with the earlier discussion, it appears that orthographic information from a letter string facilitates retrieval of the picture's name. Thus, the name becomes available sooner in these conditions than in any other condition. However, in the orthographic similarity condition, as

[1]Due to the arguments presented by Wike and Church (1976) and others, stimulus materials was not treated as a random factor as suggested by Clark (1973) in this or any subsequent analysis.

in the unrelated word condition, the tendency to produce the word's name must still be suppressed before the correct response can be emitted. In the picture alone condition no such suppression is necessary. Therefore, because of the time needed to suppress a competing response, the orthographic similarity condition loses its advantage over the picture alone condition. In the Xs conditions little, if any, interference was attributable to phonetic factors since generating a phonetic code was so difficult. However, a small amount of interference probably arose due to lateral masking, which also was not a problem encountered in the picture alone condition. Thus, although the Xs with orthography condition was significantly faster than the Xs without orthography condition, it showed little facilitation with respect to the picture alone condition.

The second aim of Experiment I was to determine whether simple figural similarity between the superimposed letters and the letters in the picture's name contributes to the orthographic similarity effect. Two consonant string conditions were created, one using letters high in figural similarity to those in the picture's name and one using letters figurally quite dissimilar to those in the picture's name. The observed difference between these two conditions was a nonsignificant 5 msec, indicating that this type of figural similarity is irrelevant to the picture-naming process.

Coincidentally, this lack of a difference also suggests that letter string length is irrelevant to the picture-naming process. Due to the use of only upper case letters, length is the only word shape cue available in the present studies. However, it is a 100 per cent valid cue in the figural similarity condition (as well as the Xs with orthography condition) while being an invalid cue ten of twelve times in the figurally dissimilar condition (as well as the Xs without orthography condition). Thus, if it were a useful cue it should have affected the difference between the two consonant string conditions. Clearly it did not. Thus, it seems unlikely that it could have had any effect in the two Xs conditions either. As such, it seems safe to conclude that the facilitation observed in the Xs with orthography condition is based on the identities of the letters themselves and not on their figural properties or those of their letter strings.

EXPERIMENT II

The purpose of the second experiment was to examine the effects of phonetic similarity in the picture-word interference task in order to provide an explanation of these effects within the framework of a general processing model. First, it was necessary to determine whether phonetic similarity actually does facilitate the naming of picture-word stimuli when orthographic cues are not available. To this end, nine pairs of words were selected such that they rhymed, in the sense that other than the initial phoneme they would be identically pronounced; but, they had different first letters and they had identical letters in no more than one other position. Thus, the words in each pair were orthographically the same in, at most, one letter position. Examples would be pairs like plane-brain, flower-hour, etc. One member of each pair became the word and the other member the picture in the first condition. A comparison of this condition to an unrelated word condition should indicate whether phonetic similarity, in what is essentially the absence of orthographic similarity, does facilitate picture-naming in the picture-word interference context.

If it can be demonstrated that phonetic similarity is an important factor in picture-word interference, the next step would be to determine the locus of this effect. One possibility would be that, like orthographic similarity, phonetic similarity provides a cue to aid in the retrieval of the picture's name. Alternatively, the similarity between the motor aspects of the word's name and the picture's name may reduce the amount of competition between the two inputs, and hence, speed up the picture-naming process. In an effort to distinguish between these alternate interpretations, a second manipulation was introduced. As noted earlier, pronounceable nonwords typically yield picture naming latencies 15-20 msec shorter than unrelated words (Lupker, 1979). Based on the fact that this difference disappears when the contributions of the output process are controlled (Lupker & Katz, 1981) it seems quite likely that the output process is the locus of this effect. Thus, additive factors logic (Sternberg, 1969) can be used to localize any effects of phonetic similarity.

Based on the nine pictures selected earlier, nonword conditions were created to parallel the phonetic similarity and unrelated word conditions in the present experiment. In the first condition (the phonetic similarity condition), the nonwords rhymed with the picture's name, but had few, if any, letters in common. In the other condition, the nonwords had no orthographic or phonetic relationship to the pictures on which they appeared. The predictions for this manipulation are straightforward. If phonetic similarity aids in the name-retrieval process, its effects should be independent of any word-nonword differences. However, if phonetic similarity aids in the suppression process, it would be expected that these two factors would not produce additive effects.

In an effort to get a firmer grip on the locus of both phonetic and orthographic facilitation another phonetic condition was added. Words and nonwords in this condition not only rhymed with the picture's name but also, other than the first letter, or consonant bigram, their spellings were identical to the spellings of the picture's name. Thus, these letter strings were both phonetically similar (like those in the other phonetic condition) and orthographically similar (like the words in Experiment I) to the names of the pictures on which they appeared.

The purpose of this condition was to examine orthographic similarity in conjunction with phonetic similarity. These letter strings are as similar to the pictures' names as possible without involving the initial letter position and, thus, the initial phoneme. Therefore, a comparison of these conditions to their respective unrelated letter string conditions should indicate the maximum amount of facilitation which these two sources can provide in the present context. It would be expected that this condition would provide, at the very least, the same amount of facilitation found in the orthographic similarity condition in Experiment I. In addition, the present condition also allows the previously made claims about the locus of orthographic facilitation and the locus of the word-nonword difference to be examined. If, as argued, these two effects can be localized at different stages, adding orthographic similarity to phonetic similarity should not change the nature of the phonetic similarity by type of irrelevant input (word or nonword) interaction. That is, the word-nonword difference in the phonetic similarity condition should be equivalent to the word-nonword difference in the orthographic plus phonetic similarity condition. If this result is obtained, it will be taken as fairly strong evidence that orthographic similarity does facilitate name retrieval whereas the word-nonword difference is an output phenomenon.

Method

Subjects: Twenty University of Western Ontario undergraduate volunteers (2 males and 18 females) received course credit for appearing in this experiment and another, unrelated experiment in the same one hour session. All were native English speakers.

Materials and Equipment: Nine word pairs were selected which met the following criteria: (1) the two words rhyme in the sense that other than their initial phoneme they are identically pronounced according to Webster's New Collegiate Dictionary (1977); (2) the words have different first letters; and (3) the words have the same letter in no more than one other position. Line drawings ("pictures") were obtained for one member of each pair from a children's colouring book. Seven sets of these pictures were produced and glued on 23 x 25.6 cm cards. These seven sets of pictures corresponded to the seven experimental conditions.

To create the first condition, the names of the other members of the pairs were superimposed on the appropriate pictures in the picture set. To create the second condition, a word was selected for each picture based on the following criteria: (1) the word's name and the picture's name rhyme according to Webster's New Collegiate Dictionary (1977); and (2) other than the first letter or consonant bigram, the word's name and the picture's name are spelled identically. The third condition, the unrelated word condition, was created by using words from the first two conditions (five from the first condition, four from the second) and superimposing them on different pictures.

Three pronounceable nonword conditions were then created to parallel the three word conditions. In the first pronounceable nonword condition, nonwords were selected which rhymed with the picture's name but had few, if any, letters in common. In the second pronounceable nonword condition the nonwords selected also rhymed with the pictures' name and, other than the first letter or consonant bigram, were also spelled the same as the picture's name. (The "pronunciation" of the nonwords was determined by an informally selected group of eight colleagues, secretaries, and graduate students. Complete unanimity of opinion was required before any nonword was used.) The third nonword condition was created by using nonwords from the first two nonword conditions (five from the first condition, four from the second) and superimposing them on different pictures. Finally, a picture alone condition was created with no letters appearing on the pictures.

Because, unlike Experiment I, the words in the three word conditions were not identical, a further consideration in their selection was that the three sets of words be essentially equivalent on the dimensions of imageability and printed familiarity according to Paivio's (Note 1) norms. The mean imageability ratings for the words in Conditions 1, 2, and 3 were 5.44, 5.49, and 5.43, respectively. The mean familiarity ratings for the three sets of words were 5.67, 5.63, and 5.67, respectively. Additionally, the mean word length in all three word conditions and the length of the nonwords used in the three nonword conditions were approximately the same. The names of the pictures, the words and the pronounceable nonwords used are reported in the Appendix.

The tachistoscope, timer, microphone, and voice-activated relay were the same as used in Experiment I. Also, the letters and pictures were the same size as in Experiment I.

Procedure: The procedure was very similar to that of Experiment I. Subjects were told to name the pictures as rapidly and accurately as possible and then were shown an example picture-word stimulus. The subjects responded to each of the 63 stimuli in Experiment II in a random order. Following a brief rest they next responded to each of the stimuli in another, unrelated experiment. Following another brief rest they again responded to the stimuli in Experiment II in a different random order. Again, errors were recorded and those pictures were randomly placed back into the set of to-be-presented stimuli. The entire procedure took about one hour.

Results

As in Experiment I errors were virtually nonexistent, less than 1% in all conditions, and error trials were not analyzed. The mean correct reaction times for the seven conditions are presented in the lower panel of Table I. Each of these data points is based on 360 observations.

The basic design of this experiment was a 3 (conditions) by 2 (word or nonword) by 2 (trial blocks) design with an additional control condition, the picture alone condition. Ignoring the control condition, the correct reaction times were submitted to a 3 X 2 X 2 ANOVA. All three main effects, conditions, ($F(2, 38) = 23.22, p < .001$), word vs. nonword, ($F(1, 19) = 17.57, p < .001$), and trial blocks, ($F(1, 19) = 23.98, p < .001$), were highly significant. As expected, pronounceable nonwords were less interfering than words, and subjects were faster in the second trial block. A subsequent Newman-Keuls analysis indicated that all conditions were significantly different from one another. The phonetic similarity condition facilitated reaction time by approximately 20 msec in comparison to the unrelated letter string condition ($p < .05$), while the orthographic plus phonetic similarity condition led to an additional facilitation of 35 msec ($p < .01$). Finally, none of the two-way interactions nor the triple interaction even approached significance (each $p > .15$). In particular, there were no indications that the relationship between the three conditions varied as a function of whether the letter string was a word or a nonword ($F < 1.0$). Thus, the phonetic and orthographic factors important in picture-word interference seem to be independent of the factors producing the word-nonword difference.

Discussion

There are two points to be made about the results from Experiment II. First, when orthographic similarity was kept to a minimum the effects of phonetic similarity were somewhat small (approximately 20 msec) but reliable. In particular, the difference between the phonetic similarity and unrelated word conditions in the present experiment was considerably smaller than the 300+ msec difference observed by Rayner and Posnansky (1978) in their most comparable conditions. Part of this disparity is certainly attributable to different degrees of phonetic similarity being employed in the different experiments. In Rayner and Posnansky's task the pronunciation of the phonetically similar letter string and the pronunciation of the picture's name were identical, whereas in the present study these two pronunciations always differed in their first phoneme. Nonetheless, it is also quite likely that concomitant orthographic manipulations and subject strategies, arising from their use of blocked conditions, contributed substantially to the "phonetic" effects Rayner and Posnansky observed. As such, while Rayner and Posnansky were correct in concluding that phonetic similarity does facilitate picture-naming in the picture-word interference context, its effects seem to be somewhat limited.

The second point to be made is that, while there was a small but reliable difference in the amount of interference produced by words and nonwords, this difference did not vary as a function of the orthographic and phonetic manipulations. This lack of an interaction implies that the word-nonword difference can be localized at a stage different from that responsible for the effects of orthographic and phonetic similarity. In the case of orthographic similarity this result was as expected, since it had been argued previously that orthographic similarity facilitates the name-retrieval process whereas the word-nonword difference is an output effect. What was not necessarily anticipated was that the effects of phonetic similarity would be independent of the word-nonword difference. However, the lack of an interaction between these two factors does seem to indicate that the existence of a phonetic relationship between the two competing responses does not facilitate the suppression process.

The logical locus of the phonetic facilitation would be the name-retrieval process. Ideally, in order to test this hypothesis one would create an experiment in which the factors of phonetic similarity and orthographic similarity were varied factorially. Additive factors logic would again be applied and, hopefully, an interaction between these two factors would be obtained. Unfortunately, due to the constraints in the English language, an experiment of this sort is not feasible. However, the results from these experiments taken together do provide at least some evidence of an interaction of this sort.

The orthographic plus phonetic similarity condition in Experiment II facilitated picture naming by 55 msec with respect to the unrelated word condition. This facilitation is almost exactly the same as the facilitation observed in the orthographic similarity condition in Experiment I, 56 msec. While it cannot be assumed that this orthographic similarity condition was absolutely free of the influence of phonetic factors, it certainly involved much less phonetic similarity than the orthographic plus phonetic similarity condition in Experiment II. Yet, there is no evidence of any additional facilitation in the orthographic plus phonetic similarity condition even though phonetic similarity clearly does facilitate picture naming. Instead, it appears that when both phonetic and orthographic similarity are present, the effects of phonetic similarity are severely attenuated. Thus, although there clearly can be problems in making comparisons across experiments, it appears that orthographic and phonetic similarity do not produce independent effects. Therefore, it seems likely that the locus of the phonetic effect, like the locus of the orthographic effect, is the name-retrieval process.

GENERAL DISCUSSION

The purpose of the present paper was to evaluate the effects of orthography and phonetics in picture-word interference in order to understand how these factors might affect picture naming. Two conclusions can be offered. First, both factors facilitate picture naming in comparison to an unrelated word condition. Second, the locus of this facilitation appears to be the name-retrieval process. Apparently, orthographic and phonetic information from words, pronounceable nonwords, and even unpronounceable letter strings can provide cues to aid in the search for a picture's name, thus facilitating the naming process.

Based solely on the results in the present paper, these conclusions would, of course, have to be restricted to the final n-1 letter positions of a word. With respect to the first letter position, Posnansky and Rayner's (1978) results suggest that phonetic similarity, as well as orthographic similarity, does facilitate picture naming. If so, as argued earlier, the locus of the orthographic effect would very likely be name retrieval. However, the locus of this phonetic effect would be quite difficult to pin down. At least part of the effect could probably be localized at name retrieval. However, the remainder of this facilitation could easily have occurred, not because it is easier to suppress a word beginning with the same phoneme, but becasue when the two competing responses do begin with the same phoneme, the subject can start responding, hence tripping the voice key and stopping the timer before the suppression process is complete. Therefore, the effect of phonetic similarity in the first letter position was not investigated in the present paper, leaving this particular aspect of the facilitation as yet unresolved.

The semantic/lexical network model of Collins and Loftus (1975) would be one means of attempting to account for the present results. According to this model there is a lexical network organized along the lines of phonetic similarity. Concept names are stored in this network. Accessing a location in the network produces a spread of excitation to nearby locations, allowing those locations to be accessed more rapidly in the course of picture processing. However, while this model may be able to account for the present results in a general way, it also runs into some difficulties. To begin with it must account for the orthographic facilitation. It could only do so by adding the assumption that much of the network is structured along the lines of orthographic similarity. In fact, based on the rather small size of the phonetic effect in Experiment II, it would apparently be necessary to suggest that the network is primarily an orthographic network. Second, the network is supposedly a set of unitary concept nodes, much like a dictionary. Nonwords should not have representations in the network and, thus, should allow neither access to the network nor any subsequent spread of excitation to orthographically similar concept names. Accounting for the facilitation with pronounceable and unpronounceable nonwords would involve either extending the hypothesized network to include nodes for all possible letter strings or suggesting that word nodes may be accessed by nonwords as well as words. While the latter suggestion may have some merit when considering pronounceable nonwords, it seems quite unlikely to be true of the X strings used in Experiment I.

A better model of these processes would be one in which the retrieval of a concept's name is not necessarily viewed as a unitary operation. In fact, it seems evident from research on the tip-of-the-tongue phenomenon (e.g., Brown & McNeill, 1966) that a concept's name need not be retrieved in a unitary manner. That is, individuals in a tip-of-the-tongue state can often retrieve phonetic and orthographic information about a concept's name, (e.g., its first phoneme or first letter) without being able to retrieve its full name. How a name actually is retrieved probably depends on how memory is accessed. When accessed by a word stimulus, the name-retrieval process will probably appear to be unitary for a number of reasons: overlearning of retrieval routes, the close relationship between orthography and phonetics, etc. However, when it is accessed by a nonverbal stimulus such as a picture, retrieval of the name need not be a unitary operation, but may be a piece by piece retrieval of whatever relevant information can be found. If so, it may be the case that any relevant orthographic or phonetic information supplied by the irrelevant word may obviate the retrieval of the same pieces of information about the picture, thus facilitating the entire retrieval process.

The model being suggested is a simple extension of the reading model of Theios and Muise (1977). Memory is viewed as a set of unconnected locations, each corresponding to a particular concept. Each location is regarded as a file containing relevant information about that particular concept (e.g., its name, its semantic category, etc.). Both words and pictures may access these files and allow retrieval of whatever information is relevant to the task being engaged in.

Within the framework of this model of memory the description of the processes involved in picture-word interference would be as follows. When a picture-word stimulus is presented to a subject, both stimulus components allow access to the relevant locations in memory. Due to years of reading experience, accessing memory by means of a word automatically leads to the rapid retrieval of the concept's name. On the other hand, accessing memory by means of a picture necessitates a search through information about the concept before the name can be discovered. Thus, it takes longer to name a picture than to name a word (Fraisse, 1968). With respect to picture naming, what, apparently, are being searched for are any pieces of information relevant to the picture's name. The process would be much like that of putting together a puzzle. Relevant orthographic and phonetic information from the word can provide pieces to the puzzle, thus speeding up the search process. Presumably, as Rayner and Posnansky argue, a word's orthographic information is available and can therefore be used sooner than its phonetic information. As such, the search process should be facilitated more by an orthographic cue than a phonetic cue. A comparison of the relevant conditions in Experiments I and II suggests that this, indeed, was the case. Further, when both cues are available, as in the orthographic plus phonetic similarity condition in Experiment II, the orthographic cue may allow the picture's name to be retrieved before the phonetic cue can affect the retrieval process. As such, orthographic and phonetic cues together may be no more beneficial than orthographic cues alone.

At that point when the picture's name is finally available the word's name will already be available, and will be blocking the output channel. Pronounceable nonwords, although processed slightly differently will also produce a name code which will lead to a blockage of the output channel. In order to clear this channel, the tendency to produce the output code currently occupying it must be suppressed. Pronounceable nonwords, being novel graphic representations, should not evoke as strong a tendency toward pronunciation as common unrelated words making them slightly easier to suppress. Thus, the locus of the word-nonword difference would be the suppression process, implying that orthographic and phonetic factors should affect words and nonwords in the same way, as was found in Experiment II. Finally, as the subject completes the suppression process, the output channel is cleared and the production of the picture's name is possible.

The purpose of the present paper was to gain a better understanding of how the automatic processing of an irrelevant stimulus component, in particular, a word, can influence the processing of a relevant component, here, a picture. The results have provided some interesting information about the processes involved in naming pictures. However, in addition, they have served to point out once again the difficulty in answering the question posed in this paper's first sentence. Clearly, the mechanisms involved in the simultaneous processing of two inputs work quite interactively. Information from the irrelevant input can either drive the processing of the relevant input along or hold it back depending on what that information is and, presumably, when it becomes available. Hopefully, the present discussion has helped to provide a beginning for understanding just how those mechanisms might work.

REFERENCE NOTE

1. Paivio, A. *Imagery and familiarity ratings for 2448 words: Unpublished norms.* Department of Psychology, University of Western Ontario.

REFERENCES

Brown, R., & McNeill, D. The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior,* 1966, **5,** 325-337.

Clark, H.H. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior,* 1973, **12,** 335-359.

Collins, A.M., & Loftus, E.F. A spreading-activation theory of semantic processing. *Psychological Review,* 1975, **82,** 407-428.

Dyer, F.N. The Stroop phenomenon and its use in the study of perceptual, cognitive and response processes. *Memory & Cognition,* 1973, **1,** 106-120.

Fraisse, P. Motor and verbal reaction times to words and drawings. *Psychonomic Science,* 1968, **12,** 235-236.

Hock, H.S., & Egeth, H.E. Verbal interference with encoding in a perceptual classification task. *Journal of Experimental Psychology,* 1970, **83,** 299-303

Klein, G.S. Semantic power measured through the interference of words with color-naming. *American Journal of Psychology,* 1964, **77,** 576-588.

Lupker, S.J. The semantic nature of response competition in the picture-word interference task. *Memory & Cognition,* 1979, **7,** 485-495.

Lupker, S.J., & Katz, A.N. Input, decision and response factors in picture-word interference. *Journal of Experimental Psychology: Human Learning and Memory,* 1981, **7,** 269-282.

Nelson, D.L., Reed, V.S., & McEvoy, C.L. Learning to order pictures and words: A model of sensory and semantic encoding. *Journal of Experimental Psychology: Human Learning and Memory,* 1977, **3,** 485-497.

Posnansky, C.J., & Rayner, K. Visual-feature and response components in a picture-word interference task with beginning and skilled readers. *Journal of Experimental Child Psychology,* 1977, **24,** 440-460.

Posnansky, C.J., & Rayner, K. Visual vs. phonetic contributions to the importance of the initial letter in word identification. *Bulletin of the Psychonomic Society,* 1978, **11,** 188-190

Posner, M.I., & Snyder, C.R.R. Attention and cognitive control. In R.L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium.* Hillsdale, N.J.: Lawrence Erlbaum Associates, 1975.

Rayner, K., & Posnansky, C. Stages of processing in word identification. *Journal of Experimental Psychology: General,* 1978, **107,** 64-80

Rosinski, R.R. Picture-word interference is semantically based. *Child Development,* 1977, **48,** 643-647.

Seymour, P.H.K. Conceptual encoding and locus of the Stroop effect. *Quarterly Journal of Experimental Psychology,* 1977, **29,** 245-265

Smith, M.C., & Magee, L.E. Tracing the time course of picture-word processing. *Journal of Experimental Psychology: General,* 1980, **104,** 373-392.

Sternberg, S. The discovery of processing stages: Extension of Donders' method. *Acta Psychologica,* 1969, **30,** 276-315.

Stirling, N. Stroop interference: An input and output phenomenon. *Quarterly Journal of Experimental Psychology,* 1979, **31,** 121-132.

Stroop, J.R. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology,* 1935, **18,** 643-662.
Theios, J., & Muise, J.G. The word identification process in reading. In N.J. Castellan, Jr. and D. Pisoni (Eds.), *Cognitive theory,* Volume II. Potomac, Maryland: Erlbaum Associates, 1977.
Townsend, J.T. Theoretical analysis of an alphabetical confusion matrix. *Perception and Psychophysics,* 1971, **9,** 40-50.
Webster's New Collegiate Dictionary. Springfield, Mass.: G. & C. Merriam Co., 1977
Wike, E.L., & Church, J.D. Comments on Clark's "The language-as-fixed effect fallacy." *Journal of Verbal Learning and Verbal Behavior,* 1976, **15,** 249-255.
Williams, E. The effect of amount of information in the Stroop color word test. *Perception and Psychophysics,* 1977, **22;** 463-470.

Received 9 February 1981
Revision received 19 May 1981
Accepted 1 June 1981

APPENDIX

Stimuli in Experiment I

Picture	Or	Ur	X/Or	X/Ur	CS/F	CS/NF
BEAR	YEAR	WORK	XXXR	XXRK	DFRP	TQPX
BOW	COW	WAND	XXW	XXND	DQV	URHO
BOWL	FOWL	SHOE	XXWL	XXE	DQVI	UQF
DART	WART	GLOVE	XXRT	XXXVE	ORPL	HLQYF
FOOT	BOOT	BAR	XXXT	XXR	TQQL	QRP
FORK	WORK	YEAR	XXRK	XXXR	TQPX	DFRP
HAND	WAND	COW	XXND	XXW	URHO	DQV
HOE	SHOE	WART	XXE	XXRT	UQF	ORPL
KEY	THEY	FOWL	XXY	XXWL	XFV	DQVI
OAR	BAR	BOOT	XXR	XXXT	QRP	TQQL
STOVE	GLOVE	WORD	XXXVE	XXRD	HLQYF	HVQPO
SWORD	WORD	THEY	XXRD	XXY	HVQPO	XFV

Stimuli in Experiment II

	WORD CONDITIONS			NON WORD CONDITIONS		
Picture	Ph	Or/Ph	Ur	Ph	Or/Ph	Ur
BROOM	TOMB	ROOM	TRUCE	SUME	FOOM	VEHR
CHAIR	BEAR	AIR	WHITE	VEHR	ZAIR	TAIN
FIRE	CHOIR	SPIRE	WHALE	GUYER	NIRE	HOWER
FLOWER	HOUR	POWER	CHOIR	GAUER	HOWER	JITE
KITE	LIGHT	WHITE	BEAR	DIGHT	JITE	NIRE
MOOSE	TRUCE	NOOSE	LIGHT	THUCE	VOOSE	SUME
NAIL	WHALE	JAIL	ROOM	KALE	LAIL	ZAIN
PLANE	BRAIN	CANE	POWER	TAIN	NANE	VOOSE
TRAIN	CANE	BRAIN	NOOSE	HANE	ZAIN	KALE

Ph = phonetically similar letter string; Or = orthographically similar letter string; Ur = unrelated letter string; X = string of Xs; CS = consonant string; F = figural similarity; NF = no figural similarity.