Journal of Experimental Psychology:
Human Perception and Performance
1986, Vol. 12, No. 1, 18-35

0096-1523/86/$00.75

Perceptual Interactions in Two-Word Displays: Familiarity and Similarity Effects

James L. McClelland
Carnegie-Mellon University

Michael C. Mozer
University of California, San Diego

Previous studies have demonstrated the existence of perceptual interactions in the processing of two-word displays such as *SAND LANE.* When postcued to report one of the two words, subjects often make *migration errors,* in that the report of the specified word includes a letter of the other word (e.g., *LAND* or *SANE* instead of *SAND*). We find that migrations depend on the abstract, structural similarity of the strings, but not on the physical similarity; on whether the strings are words; and on whether the possible migration responses are words. We also rule out an interpretation of migration errors that attributes them to a guessing strategy. Our findings are interpreted in terms of models in which both strings simultaneously access high-level structural knowledge, that is, knowledge about what sequences of letters fit together to form familiar wholes.

The role of structure and familiarity in visual perception has usually been studied using displays consisting of a single stimulus object. It is generally observed that perception of the components of these objects is more accurate when the objects are coherent wholes than when they are random unstructured arrays. Furthermore, components of coherent objects are perceived better when they occur in these objects than when they are presented alone. For example, perception of a letter is more accurate when it occurs in a word or pseudoword than when it occurs alone or in an unrelated context (Baron & Thurston, 1973; Johnston & McClelland, 1973, 1974; Massaro & Klitzke, 1979; McClelland & Johnston, 1977; Reicher, 1969; Rumelhart & McClelland, 1982; Spoehr & Smith, 1975; Wheeler, 1970). Likewise, perception of a line segment is more accurate when it occurs as a component of a structured geometrical figure than when it occurs either alone or in a context with which it does not interact to form a coherent whole (McClelland, 1978; McClelland & Miller, 1979; Weisstein & Harris, 1974; Williams & Weisstein, 1978). Although aspects of the latter findings might be attributed to innate perceptual structures, in the case of words the inference is inescapable that acquired knowledge of stimulus structure influences perceptual processing of the constituents of these stimuli.

Many models of the process whereby familiar stimuli make contact with representations in memory have been proposed. Most of these models (e.g., Estes, 1975; Johnston & McClelland, 1980; LaBerge & Samuels, 1974; McClelland & Rumelhart, 1981; Paap, Newsome, McDonald, & Schvaneveldt, 1982; Rumelhart & McClelland, 1982) apply specifically to letter-string stimuli, although related models have been proposed by other investigators for other kinds of stimuli (McClelland, 1978; Palmer, 1975). In all of these models, perception involves a set of detectors for familiar units in a hierarchy of levels. For example, most of the word recognition models posit detectors for visual features, letters, and words. In some models (McClelland & Rumelhart, 1981; Palmer, 1975; Rumelhart, 1977), perception of elements at one level (e.g., the letter level) is facilitated by feedback from the next higher level (e.g., the word level). Similar hierarchies of detectors have been proposed by workers concerned with such issues as the development of automaticity of detection and related phenomena (Deutsch & Deutsch, 1963; Norman, 1968; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977).

These models deal primarily with the perception of single stimulus items. Although there is no conclusive evidence stating that more than one item can be processed in parallel, there is a reasonable amount of evidence suggesting that, at very least, irrelevant and unattended items are often processed (Allport, 1977; Bradshaw, 1974; Willows & MacKinnon, 1973). Further, the fact that information appearing in parafoveal vision can facilitate the processing of foveal information appearing shortly thereafter (Rayner, 1978; Rayner, McConkie, & Ehrlich, 1978; Rayner, McConkie, & Zola, 1980) seems to imply that in reading, information is extracted from several regions of the visual field simultaneously. Thus, it is relevant and potentially important to see whether models of the sort described above can be extended to deal with the parallel perception of multiple stimulus items, in particular, words.

The simplest possibility to consider is that each word is analyzed independently of the others. Independent processing could occur if the words are analyzed in separate, noninteracting "copies" of the same processing network.

There is, however, evidence of interactions among items in multiword displays. Such interactions were reported in Woodworth (1938) and have since been extended by Allport (1977), Shallice and McGill (1978), and Mozer (1983). The interactions take the following form: Reports of the words present in the display often contain what Mozer (1983) referred to as *migration errors;* that is, letters present in one word in the display often showed up in reports of the other word. For example, given the

---

This research was supported by National Science Foundation Grant BNS-79-24062 and by a National Institute of Mental Health Research Scientist Career Development Award MH00385 to the first author.

We would like to thank Thomas Carr, Kenneth Paap, Anne Treisman, and an anonymous reviewer for their helpful comments on an earlier draft of the manuscript; Autumn Chapman, Risa Lee Podolsky, and Sarah Archibald for assistance in running subjects; and Rebecca Duxbury for assistance in manuscript preparation.

Correspondence concerning this article should be addressed to J. L. McClelland, Department of Psychology, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213.

tachistoscopic display *SAND LANE,* subjects cued to report the item on the left often reported *LAND* or *SANE* instead of *SAND;* there was a similar tendency to make migration errors when the item on the right was cued as well. The probability of reporting *LAND* or *SANE* instead of *SAND* was considerably reduced when *SAND* was presented in the context of *BANK,* indicating that some of the migration errors were indeed a result of the presence of the letters *L* and *E* in the context word *LANE.* Interestingly, migration errors were significantly less frequent when the two display items did not share letters in common, for example, *SAND LOVE* versus *SAND LANE,* a result we term the *surround-similarity effect.*

McClelland (1985, in press) and Mozer (1984) have each proposed models of multiple-object perception that account for the basic migration effects observed in the letter-migration data. Both models were motivated by an attempt to see how connectionist processing mechanisms—that is, networks of simple, highly interconnected computing elements—might be designed to analyze several stimuli simultaneously and construct representations of tokens of the stimuli bound to particular locations in visual space. We will return to these models in the General Discussion, but for now, the important point is that although the two models differ in many details, they share a common prediction: Migrations of letters between two words result from interactions that occur in accessing representations of high-level structural knowledge, that is, knowledge about what sequences of letters fit together to form familiar wholes. In the McClelland model, there is a convergence of activation from the two stimulus items on central structures that contain the model's knowledge of words. In the Mozer model, the convergence occurs on structures that represent common letter clusters.

It seems somewhat counterintuitive that migrations of single letters could result from interactions at a level of processing where representations consist of multiletter sequences. An alternative, and perhaps more natural, interpretation of the letter-migration phenomenon has recently been offered by Treisman and Souther (1986) based on the feature-integration theory of attention (Treisman & Gelade, 1980; Treisman, Sykes, & Gelade, 1977). According to this theory, perceptual processing is divided into two stages, a preattentive stage at which each stimulus item is analyzed along a number of independent feature dimensions, and a postattentive stage at which unitary representations of stimuli are constructed by combining the set of features that belong together. The construction of unitary representations is guided both by focal attention, which limits the flow of information from the first stage to the second to features arising in a restricted spatial region, and by past experience, which specifies how features might be reasonably combined. If we assume that each letter position of a word defines a separate feature dimension, then the letter-migration phenomenon might be explained as follows: When *SAND LANE* is presented, the component letters of each word are registered independently. If attention is not focused on one word in particular, letters from both words will be available to the postattentive stage. Thus, the *S* and the *L* both activate detectors for *S* and *L* in the first position, the two *A*'s activate the detector for *A* in the second position, the two *N*'s activate the detector for *N* in the third position, and the *D* and the *E* activate detectors for these letters in the fourth position. The letter-level activations, in such a case, would lose track of which word contained the *S* and which the *L,* as well as which contained the *D* and which the *E.* It would thus be impossible for the feature-integration stage to tell which letters belonged to which of the two words, resulting sometimes in perceptions such as *LAND* or *SANE* instead of *SAND.*

To summarize, the three models discussed above all suggest that migrations result from an overlap in the processing of the two words. However, the McClelland and Mozer models propose that the overlap takes place at a fairly high level, when the contents of multiple display locations make some simultaneous contact with processing structures embodying learned information about words. In contrast, the Treisman and Souther model proposes that the earliest point of overlap is at a relatively low level, occurring when information about the location of individual letters is lost.

All three models are able to explain the basic phenomenon. Additionally, the McClelland and Mozer models have natural accounts for the surround-similarity effect in terms of the amount of overlap between stimulus items (see General Discussion); feature-integration theory has some problem explaining how the global similarity of the stimulus items affects the low-level encoding, though Treisman and Souther (1986) do offer several suggestions. The models diverge most significantly when we consider the role of higher order structure of the stimuli (hereafter, *familiarity*) in the formation of migrations. The McClelland and Mozer models predict a strong relation between familiarity and migrations, whereas a straightforward reading of feature-integration theory predicts that migrations, which result from low-level leakages in the selection process, should not depend on higher order stimulus structure. Further, the McClelland and Mozer models predict that, if migrations do occur with letters embedded in unfamiliar surrounds, the similarity of the surrounds should have no effect. The reasoning behind this is that similarity will cause increased interactions only when similar stimuli require similar processing structures. However, because the shared structures that result in migrations for familiar stimuli involve higher level knowledge, these structures will not come into play for unfamiliar stimuli, and similarity will have no effect on migrations.

The experiments presented here were designed to study the role of familiarity and similarity of the stimuli, in the hope that such data will help to distinguish or reconcile the two classes of models. In Experiment 1 we examined whether migration errors, and the effect of surround similarity on these errors, depend on the fact that the migrating letters fit together with the surround in which they occur to form familiar higher order units. In Experiment 2 we replicated the results of Experiment 1, using a slightly different paradigm, and in addition, ruled out an explanation of the phenomenon in terms of postperceptual guessing strategies. In Experiment 3 we examined the role of lexicality, independent of orthographic regularity, by comparing word stimuli to orthographically regular nonword stimuli (*pseudowords*). Finally, in Experiment 4 we examined the role of physical, as opposed to abstract, similarity of the stimuli.

Experiment 1

Words have several structural properties: They are, of course, familiar wholes; they contain familiar letter clusters; and they conform to the rules of English orthography. In Experiment 1 we examined whether these structural properties influence the production of letter-migration errors, and the effect of surround similarity on these errors.

The design of Experiment 1 involved comparing migration errors with word stimuli to migration errors with a set of control strings designed to eliminate all higher order structural properties. We spent some time considering possible control strings. We initially carried out pilot experiments using random four-consonant strings, for example, *FXVB*. However, two problems with these strings arose. First, subjects had a difficult time keeping track of which letter appeared in which position within each string (as previous studies have shown, cf., e.g., Estes, 1975; Ratcliff, 1981; Wolford, 1975). To obtain performance levels comparable to words, target-mask interstimulus intervals in the 500-ms to 1-s range were required. Second, given the difficulty that subjects had with both within-item and between-item position, the number of possible intrusion responses was greater for stimuli that had no letters in common than for those with several letters in common (e.g., *FXVB MRLT* vs. *FXVB MXVT*), making analysis difficult. To alleviate these problems, we decided to use strings composed of a target letter embedded in three random digits, for example, 2J13. The use of these letter-in-digits (LID) strings eliminated the above problems, because there was only a single letter in each string.

As in the earlier letter migration experiments, we wished to examine for letter-in-digits strings whether the surround in which a letter was embedded influenced migrations. We thus compared *same-surround* to *different-surround* stimulus pairs. Same-surround pairs were identical except for one letter, for example, 2J13 2K13; different-surround pairs differed in all four letters and digits, for example, 2J13 5K94.

However, after running several sets of subjects with these letter-in-digits strings, it occurred to us that subjects might not have processed the digit surrounds sufficiently to have appreciated surround similarity or lack thereof, and that surround-similarity effects might depend on the extent to which this was so. Subsequently, we found that several pilot subjects were unable to discriminate same from different surrounds significantly better than chance. Subjects felt as if it should have been an easy task and were somewhat surprised by how poorly they did. Of course, the ability to judge similarity consciously may not be a necessary condition for a surround-similarity effect, but the fact that subjects could not judge surround similarity makes it unclear whether the similarity of the digit stimuli registered at all.

This problem led us to redesign the stimuli so that each string was composed of one letter and three tokens of the same digit, for example, 2J22 2K22 for same-surround pairs, and 2J22 5K55 for different-surround pairs. The hope was that repeating a single digit would emphasize the similarity or dissimilarity of the stimuli. To verify that the similarity manipulation indeed had some effect on perception, Experiment 1 included trials in which subjects made similarity judgments.

The main point of Experiment 1 was to compare migrations of letters embedded in digit strings (LID) to migrations of letters in words (LIW). As for LID, Experiment 1 included both same- and different-surround LIW pairs. Same-surround pairs were identical except for one letter, for example, *LAMP LIMP;* different-surround pairs differed in all four letters, for example, *LAMP HINT.* With the surround type (LIW vs. LID) and surround similarity (same vs. different) conditions, there was a total of four stimulus types.

A modification of Mozer's (1983) procedure was used to facilitate comparison of performance on LIW and LID displays. On each trial, the subject was cued to report one letter of one item. For example, the subject might be cued to report the second letter in the left string. We call the cued letter the *target letter,* the item containing the cued letter the *target string* or *target word,* and the other item the *context.* As in Mozer's original experiment, the cue occurred after the stimulus presentation, so that subjects could not focus their attention until after the termination of the display.

If the target letter were *A* in the *LAMP–LIMP* or *LAMP–HINT* example, a *correct response* would be *A,* a *migration response* would be *I,* and any other response would be considered an *other* error. However, a migration response could occur for reasons other than the effect of the context. To demonstrate that at least some of the migration responses were, in fact, due to the presence of this letter in the context, it was necessary to estimate the number of times the migration response was reported when it was *not* in the context. We call such responses *pseudomigrations.* Pseudomigrations were measured by generating two context items to go along with each target item. For example, the target *LAMP* occurred with contexts *LIMP* and *LUMP.* If a subject was cued to report the second letter of the first string but reported *U* when shown *LAMP–LIMP* or *I* when shown *LAMP–LUMP,* the response was classified as a pseudomigration. These responses are ones that would have been classified as migrations had the other context been presented, but clearly the "migration" is not due to the presence of the letter in the context. The existence of "true" migrations would be indicated by a difference between the raw migration rate and the pseudomigration rate. We call this difference the *migration difference score.* An analogous procedure was used to measure migrations in the LID condition.

Method

Stimuli. The words used in this and subsequent experiments were selected from the set of four-letter words contained in the Kučera and Francis (1967) corpus. Foreign words, abbreviations, acronyms, and plurals were excluded from the sample; proper nouns were not.

LIW stimuli were generated with the aid of a computer program as follows. For each of the four target-letter positions, the program considered each word in the corpus as a potential target and attempted to find four matching context words, two of same surround (*Same 1* and *Same 2*), two of different surround (*Diff 1* and *Diff 2*). We call the set consisting of target and four context words a *target-context set.* Several conditions were required of the set: (a) Same 1 had to have the same letter in the target position as Diff 1; (b) Same 2 had to have the same letter in the target position as Diff 2; (c) the surround letters of Diff 1 and Diff 2 had to be identical and had to be distinct from the homologous letters in the target; (d) the letters in the target positions of the four contexts could combine with the target surround to form a new word (a migration); (e) no surround letter of Diff 1 or Diff 2 could combine with the target to form a new word; (f) no surround letter of the target could combine with Diff 1 or Diff 2 to form a new word. A valid set of contexts to go with *LAMP* would be *LIMP, LUMP, HINT,* and *HUNT.*

LID stimuli were generated as follows: A target item was generated by randomly selecting a letter from the set {B, D, F, G, J, K, P, Q, V, W, X, Z} and randomly selecting a digit from the set {1, 2, 3, 4, 5, 6, 7, 9} and combining the letter with three tokens of the digit. One position in the string was designated as the target position, and this is where the letter was placed. Four context items, two with same surround, two with different surround, were then generated meeting the following criteria: (a) Same 1 had to have the same letter in the target position as Diff 1; (b) Same 2 had to have the same letter in the target position as Diff 2; (c) the digit surrounds of Diff 1 and Diff 2 had to be identical; (d) the digit in the surround of Diff 1 and Diff 2 had to be different from the digit in the target surround. A valid set of contexts to go with 2J22 would be 2K22, 2Q22, 5K55, and 5Q55.

For LIW, 192 target-context sets were selected, and for LID, 224 target-context sets, with an equal number of sets for each target position. Each target-context set specified four *trials*, a trial consisting of the target paired with one of the contexts. These trials were divided up so that each subject saw each target presented with only one of the contexts. The trials were split into blocks of 32. Within a block, there were exactly two trials for each combination of target letter position (1, 2, 3, or 4), target item location (left or right), and surround type (same or different). The order of trials within a block was randomized.

For LIW, there were three further constraints: (a) A word could be presented to a subject no more than once as target. (b) A word could be presented to a subject no more than a total of two times as target, context, or possible migration response. Only 2.9% of same-surround and 4.4% of different-surround targets also appeared as contexts. (c) The mean frequency (Kučera & Francis, 1967) of the contexts in the two surround conditions was matched as closely as possible. Same-surround contexts had a mean word frequency of 30.7, and different-surround of 25.3.

For both LIW and LID stimuli, 24 practice target-context pairs were generated following the same general rules as for the experimental stimuli.

Design. The LID versus LIW manipulation was between subjects. Originally, we planned to run 8 subjects in each condition. However, a hint of a surround-similarity effect in the LID condition after 8 subjects had been run led us to run an additional 8 subjects in that condition for a total of 16 subjects.

Procedure. Subjects were tested individually. Each subject sat with the experimenter in a soundproof chamber in front of an AED 512 Color Graphics/Imaging Terminal (manufactured by Advanced Electronics Design, Inc.). The practice trials were presented first, followed by 6 blocks of the experimental trials. Each trial proceeded as follows. Subjects were asked to fixate on a centrally located fixation point, and to say "go" when they were ready. The experimenter then pressed a key that caused a pair of stimulus strings to be displayed for 66 ms. The pair was followed by a variable duration blank field, which in turn was followed by a random dot mask for 200 ms. Finally, a cue was presented beneath the location of the target string, and a second cue beneath the location of the target letter. Subjects were instructed to report whatever they had seen in the target letter position, and were told that it was permissible to guess if they had any clue as to the target letter's identity. A response was not required if the subject had no idea, though subjects reported some letter on more than 99% of the trials.

For LID subjects, performance in Trial Blocks 3 through 6 was measured both on the target-letter identification task and on judgments of whether the digit surrounds of the two stimulus strings were the same or different. During these trials, subjects were required to make a forced-choice surround-similarity judgment (same or different) following their target-letter reports. The LID subjects were shown a seventh block of trials, in which they were asked to omit the target-letter identification task and judge only surround similarity. During this block, the blank-field duration was held constant at the final duration of the preceding block of trials.

The fixation point, stimulus strings, and cue were contained within a green rectangle that was 7.4 cm high × 1.9 cm wide. The fixation point appeared in the center of the rectangle; one stimulus string appeared to the left of fixation, and the other to the right. Each letter or digit was 7 pixels wide and 9 pixels high; there was a 2-pixel space between characters, and a 14-pixel space between strings. This made the stimulus strings 3.0 cm wide, measured from the first character of the leftmost string to the last character of the rightmost string, with each four-letter string occupying 1.25 cm, and a space of 0.5 cm between the strings. The fixation point and stimuli were colored green, the cue red. At the viewing distance of 56 cm, each two-word display subtended a visual angle of 3.1°.

The duration of the blank field following stimulus presentation was adjusted to obtain a mean performance level of 70% correct, averaged across experimental conditions. The initial duration of the blank field was 250 ms. The duration was automatically lowered or raised after every 10 trials, practice trials included, to keep actual performance in line with the desired performance level.

Subjects. Twenty-four University of California, San Diego, undergraduates participated in this experiment for pay or to satisfy course requirements. There were 16 subjects in the LID condition and 8 in the LIW.

Results

The overall performance levels obtained were 70.2% correct target identification for LID and 67.8% for LIW. The average blank-field duration was 241 ms for LID subjects and 223 ms for LIW subjects.

Table 1 shows a summary of the results. The results in the LIW condition replicate previous findings. First, there is an effect of surround type on the overall percent correct: as in the Mozer (1983) study, subjects were less accurate in the same-surround condition than in the different-surround—65.9% versus 69.8%, $F(1, 7) = 19.9$, $p < .01$—though the difference is weaker here. Second, the experiment replicates the basic letter migration phenomenon: The number of migration responses exceeds the number of pseudomigration responses in both same- and different-surround conditions—same: 16.2% versus 1.7%, $F(1, 7) = 52.6$, $p < .001$; different: 9.8% versus 3.2%, $F(1, 7) = 7.4$, $p < .05$. Third, the experiment replicates the surround-similarity effect: The migration rate was reliably higher for the same-surround condition than for the different-surround condition, both for raw migrations—16.2% versus 9.8%, $F(1, 7) = 75.0$, $p < .001$—and in the migration difference scores, that is, the difference between the migration and the pseudomigration response percentages—14.5% versus 6.5%, $F(1, 7) = 52.6$, $p < .001$.

The effects generalize only weakly, if at all, to LID stimuli. Overall accuracy was at least as high in the same-surround condition as in the different-surround—71.1% versus 69.3%, $F(1, 15) = 1.19$, $p > .25$. Though overall accuracy was comparable in the LID and LIW conditions, far fewer migration errors occurred for LID—5.0% versus 13.0%, $F(1, 22) = 28.7$, $p < .001$, for this between-subjects comparison. Nonetheless, the percentage of migration responses was somewhat higher than the percentage of pseudomigration responses—same surround: 5.6% versus 2.5%, $F(1, 15) = 13.7$, $p < .01$; different surround: 5.0% versus 2.0%, $F(1, 15) = 10.2$, $p < .01$—indicating that the letter presented in the context string did sometimes influence reports of the target. There was a slight, unreliable tendency for raw migration errors to occur more frequently in the same-surround condition—5.6% versus 5.0%, $F(1, 15) < 1$—but this tendency was further attenuated in the migration difference score—3.1% versus 2.9%, $F(1, 15) < 1$.

Our confidence in these LID results is bolstered by data from two sets of pilot subjects. One set of 16 subjects was shown the "letter-in-distinct-digits" stimuli (e.g., 2J13). Although the sub-

Table 1
Percentage of Responses in Each Response Category by Experimental Condition: Experiment 1

	Responses						Migration difference score	
	Correct		Migration		Pseudomigration			
Condition	Same	Diff	Same	Diff	Same	Diff	Same	Diff
LIW	65.9	69.8	16.2	9.8	1.7	3.2	14.5	6.5
LID	71.1	69.3	5.6	5.0	2.5	2.0	3.1	2.9

Note. LIW = letter in word; LID = letter in digits. Diff = different.

jects had somewhat more trouble with these stimuli, as evidenced by an average blank-field duration of 306 ms required to achieve a performance level of 70% correct responses, the overall pattern of results was identical to that described above for the LID condition. For the second set of 16 pilot subjects, who also viewed letter-in-distinct-digits stimuli but at a performance level of 50%, the pattern of results was similar, though the raw migration rates were somewhat higher. In all cases, the migration difference score indicated that there were some true migration errors with LID, but the frequency of such errors did not appear to be strongly influenced by surround similarity.

Target presentation position. The results in Table 1 have been collapsed across target presentation position (target appearing on left or right side of display). No significant interactions involving target presentation position were found. However, for LIW, performance was slightly better for left targets—69.9% versus 65.8%, $F(1, 7) < 1$—while the trend was reversed for LID—67.5% versus 72.9%, $F(1, 15) = 1.6$, $p > .2$. For LIW, the migration difference score was slightly higher for left targets—11.5% versus 9.5%, $F(1, 7) < 1$—and this effect was magnified for LID—4.5% versus 1.5%, $F(1, 15) = 5.8$, $p < .05$.

Similarity judgments for LID. Similarity judgments obtained in the LID condition indicated that subjects were able to distinguish between same- and different-surround trials. During Trial Blocks 3 through 6, in which both target identification responses and *same–different* judgments were required, 77% of the judgments were correct; and during the final block, in which only the *same–different* judgment was required, 86% of the judgments were correct.

Discussion

In Experiment 1 we replicated the basic letter migration phenomenon and surround-similarity effect reported by Mozer (1983), using a variant of Mozer's task. However, the frequency of letter migration errors was much higher with words than with the unfamiliar LID strings, and the surround-similarity effect was found only with words. It would thus appear that although migration errors are not entirely dependent upon familiarity of the stimuli, similarity-induced migration errors do depend on the fact that letters in stimulus strings fit together with their neighbors to form superordinate wholes. Whether it is important for these superordinate wholes to be words or not is the subject of later experiments.

The McClelland (1985) and Mozer (1984) models, which attribute migration errors and the surround-similarity effect to the mechanisms involved in word recognition, predicted that migration errors and the surround-similarity effect should be eliminated for LID stimuli. Although the strong role of familiarity was upheld by this experiment, LID migrations did nonetheless occur. In the General Discussion, we consider possible explanations for these migrations.

Experiment 2

Although the method used to assess performance in Experiment 1 produced interesting data, it is susceptible to various biases. The pseudomigration measure is a sufficient control to demonstrate that subjects do intrude letters from the context in their report of the target string, but it does not rule out one plausible guessing strategy. Migration errors might simply result from a strategy of reporting letters from the context string when the subject fails to see the target letters. That is, subjects might simply choose to report the homologous letter from the context string, even if they had no uncertainty about its location, when they fail to identify the target letter correctly. Similar interpretations might be given for migration errors in other experiments as well (Allport, 1977; Mozer, 1983; Shallice & McGill, 1978; Treisman & Souther, 1986).

This guessing strategy can easily account for the surround-similarity effect with letters in words and its absence with letters in digits. For letters in words, subjects might well notice that the words sometimes share letters in common and sometimes do not. If so, they might plausibly decide to invoke the guessing strategy only when what they have seen of the two strings includes some letters in common between the two words. For letters in digits, however, matching digit surrounds seem less likely to induce subjects to think that the single letters embedded in each of the two strings of digits might be the same.

The above interpretation does not accord with subjects' phenomenological reports, nor does it fit well with the fact that migration errors are often made with high confidence (Mozer, 1983) or that they often occur rather early among subjects' responses when they are asked to report as many items as they can from a multi-item display (Treisman & Souther, 1986). Instead, it appears that migration errors reflect a real misperception, in which the letter is experienced as occurring in some location where it was not in fact presented. If this conclusion is correct, we would expect that conditions that produce frequent migration errors would result in poor accuracy of letter localization in a forced-choice test. That is, the perception of a letter as having occurred in an incorrect location should reduce the accuracy of letter localization.

In an attempt to confirm this conclusion, we ran a modification of Reicher's (1969) two-alternative forced-choice procedure. Our version allowed us to measure both the accuracy of letter identification and the accuracy of letter localization. As in Experiment 1, subjects viewed two-item displays, such as *LAMP–HINT*, and were tested on one letter in one of the two items, such as the *I* in *HINT*. This letter, the *target letter*, was never repeated in the other string. To test accuracy of letter identification, subjects were given two alternatives, the target letter and a *distractor letter*, say *U* in the above example. To test accuracy of letter localization, subjects were asked to choose in which of the two strings the target letter had occurred. The alternatives were presented as follows:

$$\underset{\mathrm{U}}{\overset{\mathrm{I}}{-}}-- \quad \underset{\mathrm{U}}{\overset{\mathrm{I}}{-}}--.$$

The dashes served to indicate the locations where letters or digits had appeared in the stimulus display, and thus alerted the subject to the position in the string that was being tested.

Subjects simply had to indicate one of the four alternatives as their choice. For displays containing words, the target and distractor letters could be inserted into either stimulus word to form a word, thereby preventing any bias toward reporting words from influencing responses. (Half the subjects actually saw *LAMP–HUNT* with the alternatives *I* and *U*, in order to balance out any possible preference for one of the two alternatives.)

In this experiment, subjects were presented with both LIW and LID stimuli in separate, alternating blocks of trials. The blank-field duration was adjusted independently for LIW and LID stimuli to obtain the same overall accuracy of *identity* responses in both conditions. This allowed a within-subject comparison of localization accuracy at the same overall level of letter identification accuracy.

Based on the results of Experiment 1, we expected that the target letter would occasionally be perceived but in the incorrect location, or perhaps in two locations at once. Consequently, we expected responses in which the letter identity was correct but the letter location incorrect. Such responses should be more frequent for LIW trials than for LID trials, and surround similarity should have an effect on these responses for LIW, but only slightly, if at all, for LID.

Method

Stimuli. For the LIW condition, 128 target-context sets were selected from the ones used in Experiment 1. For the LID condition, 128 target-context sets were generated in the same manner as in Experiment 1. Unlike Experiment 1, in which two pairs of contexts were matched to each target, we switched the roles of target and context. Hence, there were two pairs of targets (Same 1–Same 2 and Diff 1–Diff 2) associated with each context. This allowed us to use the same set of forced-choice alternatives for all four of the target-context pairings; one of the alternatives was correct for the Same 1 and Diff 1 targets; the other was correct for the Same 2 and Diff 2 targets. For example, the LIW context *LAMP* had associated with it the two same-surround targets *LIMP* and *LUMP*, and different-surround targets *HINT* and *HUNT*; the response alternatives were *I* and *U*. Same-surround LIW targets had a mean word frequency of 28.6 and different-surround of 22.7. The mean word frequency of the context was 25.2 in both conditions.

Four *trials* were formed from each target-context set, a trial consisting of the context paired with one of the targets. The trials were then divided up so that each subject saw each context presented with only one target. The LIW and LID trials were each split into 4 blocks of 32. Within a block, there were exactly two trials for each target letter position (1, 2, 3, or 4), target string location (left or right), and surround type (same or different). The order of trials within a block was randomized for each subject.

Practice stimuli consisted of 16 LIW and 16 LID pairs generated by following the same rules as for the experimental stimuli.

Procedure. The basic procedure was the same as that of Experiment 1, except for response collection and the exposure adjustment procedure.

On each trial, subjects were shown a stimulus pair, followed by a variable-duration blank field, followed by a mask, followed by the four-alternative choice display shown above. The target-letter response appeared above the dashes as often as below in each condition.

The experiment began with the practice LID trials followed by the practice LIW trials, and then continued with alternating blocks of LID and LIW trials. Following the end of a block, a dummy trial would take place in which no stimuli were shown; instead the display contained question marks. Subjects were then reminded what type of trial would come next.

The blank-field duration was adjusted independently for the LIW and LID conditions to obtain 85% accuracy of target-letter identification in each condition. Blank-field duration adjustments were made after every 10 LIW or LID trials. For most subjects, the blank-field duration was initially set to 167 ms for the LIW condition, and 333 ms for the LID condition. Several subjects received initial durations of 250 and 417 ms, respectively. We made one minor change from Experiment 1 in the stimulus exposure duration: We allowed the stimulus exposure to be reduced from 66 ms to 33 ms when subjects were performing too well even with a blank-field duration of 0 ms. By reducing the stimulus exposure, we were able to attain lower interstimulus intervals between target and mask.

Subjects. Subjects were drawn from the same pool as in Experiment 1. We planned to run 16 subjects using counterbalanced lists of materials. However, subjects were replaced if overall accuracy of letter identification fell below 80% or above 90% in either the LIW or LID condition or if the difference in accuracy between conditions was more than 8%. It was difficult to satisfy these stringent criteria; consequently, 27 subjects needed to be replaced. Thus a total of 43 were run in all. Of the replaced subjects, 3 performed too well, 21 too poorly, and 3 had too great a difference between LIW and LID conditions.

Results

On each trial, the four response alternatives specified four response categories. One category represented the target letter in the target location (*IcLc*—identity correct, location correct), one the target letter in the context location (*IcLi*—identity correct, location incorrect), one the distractor letter in the target location (*IiLc*—identity incorrect, location correct), and one the distractor letter in the context location (*IiLi*—identity incorrect, location incorrect). The IcLc category corresponds to a correct response, the IcLi category to a migration response. The distribution of responses over the four categories is shown in the upper half of Table 2 for the 16 subjects who met the stringent acceptance criteria. For these subjects, the average blank-field duration was 129 ms for LIW, and 540 ms for LID. Such a long interval in the LID condition presumably amounts to the equivalent of a no-mask condition for most of the subjects.

The accuracy of target identification, abbreviated *Ic* and shown in the fifth row of Table 2, was computed by summing the IcLc and IcLI response rates. This quantity was virtually identical for

Table 2
Percentage of Responses by Response Category and Condition: Experiment 2

Alternative type	LID		LIW	
	SS	DS	SS	DS
Accepted subjects				
IcLc	76.1	76.4	71.1	75.4
IcLi	7.4	6.4	12.6	7.9
IiLc	12.9	13.2	10.6	11.6
IiLi	3.6	4.1	5.7	5.1
Ic	83.5	82.8	83.7	83.3
Rejected subjects				
IcLc	65.5	63.8	63.7	67.9
IcLi	10.4	10.5	14.3	9.8
IiLc	15.8	18.2	15.5	16.4
IiLi	8.3	7.6	6.5	5.8
Ic	75.9	74.3	78.0	77.8

Note. LID = letter in digits; LIW = letter in word. SS = same surround; DS = different surround. IcLc = identity correct, location correct; IcLi = identity correct, location incorrect; IiLc = identity incorrect, location correct; IiLi = identity incorrect, location incorrect.

the LIW and LID conditions (83.5% vs. 83.2%); our attempt to match the two conditions on this quantity thus appears to have been successful. In addition, it appears that surround similarity had no effect on target-identification accuracy, either for LIW or LID.

There were, however, differences between conditions in the accuracy of letter localization. The overall percentage of IcLi responses was 10.3% for LIW and 6.9% for LID. This difference was reliable, $F(1, 15) = 5.99$, $p < .05$, and is in the same direction as the results of Experiment 1.

Surround similarity had an effect on LIW performance, but not on LID. With LIW, IcLc responses occurred reliably less often in the same-surround condition—71.1% versus 75.4%, $F(1, 15) = 5.17$, $p < .05$—and IcLi responses occurred reliably more often—12.6% versus 7.9%, $F(1, 15) = 24.0$, $p < .001$. With LID, there was virtually no difference between IcLc response rates in the same- and different-surround conditions—76.1% versus 76.4%, $F(1, 15) < 1$—nor between IcLi response rates—7.4% versus 6.4%, $F(1, 15) = 2.05$, $p > .15$.

The percentage of same-surround LIW IcLi responses was significantly higher than the percentages in the other three conditions—same-surround LIW versus same-surround LID: $F(1, 15) = 12.2$, $p < .005$; same-surround LIW versus different-surround LID: $F(1, 15) = 14.1$, $p < .005$—and there was no reliable difference among the other three conditions—different LIW versus same LID: $F(1, 15) < 1$; different LIW versus different LID: $F(1, 15) = 1.09$, $p > .25$.

The incorrect identity responses (IiLc and IiLi) were not evenly distributed between the correct and incorrect location. Indeed, for all conditions, subjects were more likely to place incorrectly identified letters in the "correct" location than in the "incorrect" location. The probable reason for this is simply that subjects often identified the context letter in its correct location and therefore were able to rule out that location, even when they were unable to identify the target letter.

Target presentation position. Although the results of Table 2 have been collapsed across target presentation position (target appearing on left or right of display), there was an interaction of Target Position × Surround Type for IcLi responses, $F(1, 15) = 7.96$, $p < .02$. IcLi responses were more frequent for LIW when the target appeared on the left (11.5% vs. 9.0%), but the opposite was true for LID (6.2% vs. 7.6%). In migration terms, there was a predominance of right-to-left migrations for LIW but not for LID. It appears that we have replicated the usual advantage for words presented to the right of fixation (Rayner, McConkie, & Ehrlich, 1978), and have found that the effect does not extend to unfamiliar strings.

Rejected subjects. Data from the rejected subjects are shown in the lower half of Table 2. For these subjects, the exposure adjustment procedure failed to achieve the desired accuracy level of 85%: Letter-identification accuracy was only 78.2% for LIW and 75.1% for LID, achieved with mean blank-field durations of 324 ms and 603 ms, respectively. The difference in letter-identification accuracy between LIW and LID conditions also reflects the failure of the adjustment procedure for these subjects. In other respects, the data from these subjects demonstrate the same effects as the data from the subjects who met the inclusion criteria. Overall, there were more IcLi errors with LIW than with LID, though the cell means indicate that this difference was due to the higher IcLi error rate for same-surround LIW. For LIW, the same-surround condition produced a lower IcLc response rate and a higher IcLi rate, whereas for LID, there was no difference between same- and different-surround conditions.

Estimating the true accuracy of letter identification. The response probabilities in the correct-identity cells (IcLc and IcLi) of Table 2 overestimate the true accuracy of letter identification, because on trials where subjects failed to perceive the target letter, they must often have guessed the identity correctly in the forced choice. A corrected estimate of letter-identification accuracy can be obtained, based on the following assumptions: (a) When subjects fail to detect the identity of the target letter, response choices are distributed evenly between the two alternative letters (the target and the distractor),[1] and (b) on trials when the target identity is guessed, localization accuracy is independent of identification accuracy. Note that we do not assume that location responses are equally distributed over the two alternatives when identity is guessed, because of the possibility that the subject might have correctly identified the context letter in its correct location, as noted above.

On the basis of these assumptions, an estimate of the true probability of correct letter identification for any condition can be obtained by subtracting the probability of choosing the incorrect alternative from the probability of choosing the correct alternative. This is, of course, the standard correction for guessing in a two-alternative forced choice. Here we simply apply this correction separately to the location-correct and location-incor-

---

[1] In a study using only a single word, pseudoword, or random letter string, McClelland and Johnston (1977) tested this assumption and found that it closely approximated their data. Subjects gave letter identity reports before making a forced choice, and when the identity reports were incorrect, the forced choice was correct 52% of the time.

rect responses, based on Assumption (b) above. The corrected results are shown in Table 3. Based on these estimates, there was some confusion of location information in all conditions. This confusion was greatest for same-surround LIW: Same-surround LIW produced reliably more IcLi responses than did different-surround LIW or different-surround LID, $F(1, 15) = 12.6, p < .01$; $F(1, 15) = 6.8, p < .05$, respectively, and tended to produce more such responses than same-surround LID, $F(1, 15) = 3.04$, $p = .10$. There were no reliable differences emerging between any of the other conditions, although the difference between same- and different-surround LID approached significance, $F(1, 15) = 3.33, p = .088$.

Discussion

Experiment 2 reinforces our confidence that the letter migration phenomenon and its sensitivity to surround similarity are not merely the result of postperceptual guessing strategies. Instead, the data support the view that these phenomena reflect, at least in part, some illusory perceptions in which presented letters are seen as having occurred in incorrect locations.

As in Experiment 1, the effect of surround similarity is reliable for LIW, but not for LID. There remains a hint of a surround-similarity effect in the latter stimuli, at least in the data from the accepted subjects, but the data from the rejected subjects fail to bear out this slight effect. (Note that these subjects were rejected only because they failed to perform approximately equally well on letters in words and letters in digits, and there is nothing wrong with their data from the point of view comparing same vs. different surround stimuli within the LID condition.) Although the evidence may hint at a slight surround-similarity effect with LID stimuli, it is clearly not as strong or reliable as the effect visible with words in both Experiments 1 and 2.

Beyond ruling out the guessing interpretation of migration errors and replicating previous results, Experiment 2 adds a within-subject comparison of the presentation conditions needed to achieve 85% correct forced-choice performance on letters-in-words and on letters-in-digits. This comparison indicates a reliable word-superiority effect, in that blank-field durations required for criterial performance were much shorter for LIW than for LID. Of course, a perceptual advantage for letters in words over letters in an unfamiliar context is not unexpected, based on earlier work with displays containing a single word or control string (cf. Rumelhart & McClelland, 1982, for a comparison of performance on letters in words and letters in digits). In fact, evidence that the effect holds up with two-word displays was established in the study in which the word-superiority effect was first reported (Reicher, 1969).

Experiment 3

Experiments 1 and 2 have shown that letter migrations are at least partially dependent on higher order properties of the stimuli, but they fail to identify which properties are critical. In particular, two properties are confounded: lexical status and orthographic regularity. Do migration error patterns depend on the lexical status of a letter string (either as target, context, or possible migration error response), over and above its orthographic regularity? In Experiment 3 we examined the role of lexicality, independent of orthographic regularity, by comparing word stimuli to orthographically regular nonword stimuli (pseudowords).

Table 3

Target Identity Responses Corrected for Guessing Over Correct and Incorrect Locations

Response type	LID SS	LID DS	LIW SS	LIW DS
	Accepted subjects			
IcLc	63.2	63.2	60.5	63.8
IcLi	3.8	2.3	6.9	2.8
Ic	67.0	65.5	67.4	66.6
	Rejected subjects			
IcLc	49.7	45.7	48.2	51.5
IcLi	2.1	2.9	7.8	4.7
Ic	51.8	48.6	56.0	56.2

Note. LID = letters in digits; LIW = letters in words. SS = same surround; DS = different surround. IcLc = identity correct, location correct; IcLi = identity correct, location incorrect.

It is well-known that the word-superiority effect extends to pseudowords (Baron & Thurston, 1973; McClelland & Johnston, 1977; Spoehr & Smith, 1975); all three of these articles suggest that word-superiority effects do not depend so much on whole-word familiarity of the stimuli as on conformity of the stimuli to orthographic or phonotactic rules. On the other hand, several investigators have reported a slight advantage for words over matched pseudowords (Manelis, 1974; McClelland, 1976; McClelland & Johnston, 1977), supporting some role of whole-word familiarity in the word-superiority effect, and McClelland and Rumelhart (1981; Rumelhart & McClelland, 1981, 1982) have argued that perceptual facilitation observed with pseudowords may depend on interactions with detectors for familiar words.

Because the perception of pseudowords appears to benefit from processing structures responsible for the perception of words, the McClelland (1985) and Mozer (1984) models of multiple-word perception predict that migrations, which result from interactions among these structures, will occur with pseudowords.

Experiment 3 consisted of trials in which the targets, contexts, and possible migration responses might be either words or pseudowords. There were, in fact, eight types of trials, each occurring equally often. The trials were classified according to whether (a) the target was a word or pseudoword, (b) the context was a word or pseudoword, and (c) the potential migration responses were words or pseudowords. A sample of the eight trial types is given in Table 4. There are two targets, one word and one pseudoword, and four contexts per target.

Method

Stimuli. The allowable words for this experiment were the monosyllables of the corpus used in Experiment 1. The allowable pseudowords were selected in the following manner. A grammar was developed to generate four-letter monosyllabic strings. The grammar was complete enough that it generated nearly all of the monosyllabic four-letter words in Kučera

Table 4
Example of the Eight Trial Types: Experiment 3

Trial type	Target	Context	Possible migration responses
WtWcWm	WAKE	WOVE	WAVE, WOKE
WtWcPm	WAKE	WISE	WASE, WIKE
WtPcWm	WAKE	WODE	WADE, WOKE
WtPcPm	WAKE	WICE	WACE, WIKE
PtWcWm	COSE	CAPE	COPE, CASE
PtWcPm	COSE	CUBE	COBE, CUSE
PtPcWm	COSE	CADE	CODE, CASE
PtPcPm	COSE	CUZE	COZE, CUSE

Note. Wt = word target. Pt = pseudoword target. Wc = word context. Pc = pseudoword context. Wm = word migrations. Pm = pseudoword migrations.

and Francis. The missed words were those having uncommon spellings, such as *disc,* which ends in *sc.* The list of strings generated by the grammar was examined, and the following strings were eliminated from the list: (a) real words, (b) homophones of real words, (c) strings that looked "strange" to either author, (d) strings strongly suggesting a word to either author, and (e) strings containing (position-specific) digrams or trigrams that did not occur among the set of four-letter words in Kučera and Francis. The strings that were not eliminated formed the set of allowable pseudowords. For every word and pseudoword, an approximation-to-English (*ATE*) rating score was computed using the formula of Rumelhart and McClelland (1982). The measure represents the sum of the conditional probabilities of each letter based on its left context and of the conditional probabilities of each letter based on its right context:

$$ATE = p(L1|S0) + p(L2|L1) + p(L3|L2C1) + p(L4|L3C1C2) + p(L4|S5) + p(L3|L4) + p(L2|L3C4) + p(L1|L2C3C4).$$

Here Li refers to the letter in position i ($i = 1, 2, 3,$ or 4), Ci refers to the category (consonant or vowel) of the letter in position i, and Si refers to the spaces preceding ($i = 0$) or following ($i = 5$) the string. See Rumelhart and McClelland (1982) for further details. The measure correlates strongly with judged orthographic regularity, as well as with perceptual facilitation in the Reicher forced-choice letter-identification task, and takes both position in word and pattern of consonant-vowel alternation into account.

We allowed migrations to occur in two letter positions, which we call the *target letter positions* or *TLPs*. With four-letter words, there are six TLP pairs: 1–2, 1–3, 1–4, 2–3, 2–4, and 3–4. For each of the TLP pairs, every word was considered as a possible target. We used a computer program to find four context strings to go along with the target, one of which was a word and formed word migrations in conjunction with the target (*WcWm*), one which was a word and formed pseudoword migrations (*WcPm*), one which was a pseudoword and formed word migrations (*PcWm*), and one which was a pseudoword and formed pseudoword migrations (*PcPm*). The letters in the non-TLPs (the *surround* letters) of each context were required to be the same as the homologous letter in the target, and neither letter in the TLPs of the context could be identical to the homologous letter in the target.

This same procedure was repeated for the pseudoword list. The final product was two lists, one containing word targets and the other pseudoword targets, along with their four contexts. Next, for each TLP pair, all targets from the word-target (*Wt*) set were matched with targets from the pseudoword-target (*Pt*) set so as to minimize the difference between their *ATE* ratings. Note that the matching word and pseudoword targets did not have identical surround letters; however, the matching targets always had the same TLPs.

Each pair of matched targets and their associated context strings formed a *target-context set.* Forty-four target-context sets were then selected according to the following criteria: (a) The target-context sets were chosen in approximately equal numbers from each TLP set. (b) A word or pseudoword could be presented to a subject no more than once as target and once as a context, or twice as a context, or four times as a potential migration response (see next paragraph for an explanation of how the target-context sets were divided up among subjects). In the final stimulus lists, about 5% of the targets also appeared as contexts; 10% of the contexts and 18% of the migration responses were repeats. (c) An attempt was made to match the *ATE* ratings of word (W) and pseudoword (PW) targets, W and PW contexts, and W and PW migrations. In the final stimulus lists, there were no significant differences between the target *ATE* ratings. The PW contexts had slightly lower *ATE* ratings than the W contexts (PW M = .48; W M = .59), and the PW migrations had slightly lower *ATE* ratings than the W migrations (PW M = .53; W M = .65). The differences are slight, relative to the range of values within each class of items (average SD = .32). All in all, then, words and pseudowords used in the experiment were comparable in their orthographic regularity. (d) An attempt was made to match the word frequency of word targets, contexts, and migrations across all conditions. For example, we wanted the word frequency of word contexts in the WtWcWm (word target, word context, word migration) condition to be the same as that in the PcWcWm condition. In the final stimulus lists, there were no significant differences in the word frequency or log(word frequency) of targets, contexts, or migrations across conditions.

Eight *trials* were formed from each target-context set, a trial being one of the targets paired with one of the contexts. The trials of the target-context sets were divided into two lists, called *stimulus lists,* in such a way that all the Wc trials stayed together and all the Pc trials stayed together. Thus, each target would appear twice in a stimulus list, and its contexts would be either both words or both pseudowords.

Each stimulus list consisted of 176 (44 target-context sets × 4) trials. The two stimulus lists were subdivided into two blocks, each block containing one occurrence of each target. Within a block, all eight trial types were equally represented. For each subject, trials within a block were ordered randomly. The order of presentation of the two blocks was counterbalanced across subjects.

Target presentation position was balanced so that there were the same number of left and right target presentations for each condition within each block. Further, one presentation of a given target always occurred on the left, and the other on the right. Target presentation position of individual trials was counterbalanced across subjects.

Twenty-four practice trials were also generated, three of each trial type, in the same manner as were the experimental trials.

Procedure. The procedure was the same as for Experiment 1, except that subjects were required to report the entire target word, not just a single target letter. Thus, the probe following each trial indicated which of the two words to report, but did not specify a letter. Subjects were instructed to report whatever they felt they had seen. Subjects were asked to spell out their responses to ensure that the experimenter did not misinterpret.

We aimed for an average performance level of 60% correct responses across all conditions. Subjects were replaced if their performance fell below 45% or above 75%. The initial blank field duration was 417 ms; the mean duration over all subjects and trials was 415 ms.

Subjects. Sixteen subjects from the same pool as Experiments 1 and 2 participated in this experiment. Six subjects had to be replaced because their performance level was out of range; three performed too well and 3 too poorly.

Results

Table 5 shows a summary of the correct and migration response percentages by condition. A response was classified as

Table 5
Percentage of Correct Responses and Migration Responses by Condition: Experiment 3

Target	Condition PcPm	WcPm	PcWm	WcWm
	Correct responses			
Pt	53.7	52.0	51.7	51.7
Wt	72.7	73.3	74.2	73.3
	Migration responses			
Pt	5.4	5.7	17.0	12.2
Wt	5.1	3.4	7.7	8.2

Note. Pt = pseudoword target. Wt = word target. Wc = word context. Pc = pseudoword context. Wm = word migration. Pm = pseudoword migration.

being correct if it matched the target exactly. A response was classified as being a migration if a migration occurred in *one* of the two TLPs and if all other letters matched the target. Responses in which migrations occurred in both TLPs (i.e., the subject reports the context) are not included because if they were, there would be valid word migrations in the WcPm conditions and valid pseudoword migrations in the PcWm conditions. However, double migrations occurred on fewer than 0.4% of the trials.

The correct response data tell a very simple story: Pseudoword targets are reported less accurately than word targets—52.3% versus 73.4%, $F(1, 15) = 132$, $p < .001$—but there are no effects of the lexical status of the context or of the potential migration responses; nor are there any significant interactions involving either of these factors.

The pattern of results for migration errors is quite different. More migrations occur when the target is a pseudoword than when it is a word—10.1% versus 6.1%, $F(1, 15) = 24.9$, $p < .001$—and more migrations occur when the potential migration responses are words—11.3% versus 4.9%, $F(1, 15) = 52.8$, $p < .001$. Further, there is a Target × Migration Response interaction, $F(1, 15) = 5.55$, $p < .05$: The word status of the target makes little difference if the potential migration responses are pseudowords, but it makes a big difference if the potential migration responses are words.

It is interesting to note that although migrations are far more likely in the Wm condition than in the Pm, overall performance in these two conditions is nearly identical (Pm = 62.9%; Wm = 62.7%).

An analysis of migrations and overall errors was also performed with the inclusion of target presentation position (left or right) as a factor. As in earlier experiments, performance was lower for targets presented on the left—52.0% versus 73.6%, $F(1, 15) = 40.2$, $p < .001$—and the migration rate was higher—10.7% versus 5.5%, $F(1, 15) = 18.8$, $p < .001$. There was a significant interaction involving target presentation position and migration response type on migration errors: The magnitude of the Pm–Wm contrast is simply larger for left targets—for left targets 6.4%–14.9%, and for right targets 3.4%–7.7%, $F(1, 15) = 4.75$, $p < .05$. There is a similar tendency for the interaction of target presentation position and target type.

Computing a migration difference score. As in the earlier experiments, we would like to be sure that the observed migration rates reflect the presence of letters in the context rather than a bias to report particular strings independent of the occurrence of such letters in the context. Consequently, a measure of the pseudomigration error rate is needed. The measure we used was computed as follows. Subjects saw each target presented twice during a session, once with a context that combines with the target to form word migrations (Wm), and once with a context that combines with the target to form pseudoword migrations (Pm). When presentation of the target with the Wm context yields responses that would count as migrations when the target was presented with the Pm context, these could be considered pseudomigration errors for the Pm context, and vice versa. For example, suppose the target *WAKE* were presented with Wm context *WOVE* and Pm context *WISE*. Responses of *WOKE* or *WAVE* to the Pm context should be counted as pseudomigration responses for the Wm context (as well as errors for the Pm context). Similarly, responses of *WIKE* or *WASE* to the Wm context should be counted as pseudomigration responses for the Pm context.

Table 6 shows the percentage of pseudomigrations and the migration difference score by condition. The migration difference score was computed by subtracting the percentage of pseudomigrations from the percentage of raw migrations. These results provide clear evidence that a large fraction of the raw migration responses were induced by the presence of the "migrating" letter in the context word. Further, the migration difference score reveals the same significance results as the raw migration data: More migrations occurred when the target was a pseudoword than when it was a word, $F(1, 15) = 14.3$, $p < .01$, and more occurred when the potential migration responses were words than when pseudowords, $F(1, 15) = 39.0$, $p < .001$; however, the interaction did not quite reach significance, $F(1, 15) = 3.40$, $p < .10$. The word status of the context was not a significant factor, nor were any interactions involving this factor.[2]

---

[2] There are two reasons why it might be misleading to present the usual migration difference score based on these pseudomigration rates. To begin with, if the overall error rates for Wm and Pm conditions were different, it would not be valid to use errors in the Wm condition as a baseline for the Pm, and vice versa. To see why this is so, suppose that one condition, say the Pm, had a much higher error rate. A higher error rate in the Pm condition would offer a disproportionate opportunity for Wm pseudomigrations, compared with the opportunity for migrations in the Wm condition. Fortunately, however, error rates are the same in Wm and Pm conditions.

Another potential problem with the pseudomigration rate is that it might depend on whether the potential migrations are words or not. Consider the presentation of target *WIKE* either with context *WOPE* or *WUBE*. The pseudomigration rate for the *WIKE–WUBE* (Pm) trial is based on the *WIKE–WOPE* (Wm) trial, and vice versa. If the migration rate turns out to be higher in the Wm condition, the *WIKE–WOPE* trial will produce relatively more migrations, and hence the opportunity for making pseudomigration responses in this condition might be reduced relative to the Pm condition. Thus, the pseudomigration rate might overestimate migrations in the Wm condition relative to the Pm. However, an examination of the pseudomigration rate (Table 6) shows that it is less than 3% in all conditions. As one would expect, it is higher for words than for pseudowords. The rate of word pseudomigrations surely repre-

Table 6
Pseudomigration Rates and Migration Difference Scores by Condition: Experiment 3

	Condition			
Target	PcPm	WcPm	PcWm	WcWm
	Pseudomigration rates			
Pt	1.4	0.6	3.1	2.0
Wt	0.0	0.8	2.2	1.1
	Migration difference scores			
Pt	4.0	5.1	13.9	10.2
Wt	5.1	2.6	6.5	7.1

Note. Pt = pseudoword target. Wt = word target. Wc = word context. Pc = pseudoword context. Wm = word migration. Pm = pseudoword migration.

Discussion

Experiment 3 demonstrates that whole-word familiarity is a factor in the production of migration error responses: Subjects made more migration errors when the target was a pseudoword than when it was a word and also when the potential migration responses were words than when they were pseudowords. Despite this evidence for the role of whole-word familiarity, the migration difference scores show that true migration errors were produced in every condition of the experiment. It therefore appears that the lexical status of the target and potential migration responses modulate but do not strictly determine the presence or absence of migration error responses.

Combining the results of Experiments 1 through 3, it appears that migrations of letters in orthographically regular strings are influenced by two factors: the lexical status of the target and potential migration error responses, and the similarity of the target and context strings. In the General Discussion, we consider mechanisms by which both types of modulatory influences arise from the interaction of perceptual information with detectors for familiar higher order units (which represent more than single letters).

Experiment 4

The previous experiments have helped to establish some of the higher-level properties that influence migration errors. In Experiment 4, we switch gears and examine whether lower-level properties are important. Given that higher-level properties of the stimuli are so influential in causing migrations, perhaps lower-level properties are not. In particular, it seems plausible that physical features of the stimuli are irrelevant.

There is some prior evidence related to this point, from an experiment reported briefly by Shallice and McGill (1978). They examined migration errors in displays consisting of words entirely in one type face, as in *RIDE–ROPE,* or alternating case within a word, as in *RiDe–rOpE,* where letters in homologous positions of the paired strings were always in contrasting cases. Though overall performance was somewhat worse with alternating case stimuli, subjects made approximately equal numbers of migration errors in both conditions. Although the data are suggestive, no statistical analyses were reported regarding this manipulation. Thus, we felt a more thorough examination of the issue was in order.

Experiment 4 seeks further evidence on the role of physical versus abstract similarity in letter migrations by examining migrations between words that match in case and between words that differ in case, as in *RIDE–ROPE* or *ride–rope* versus *RIDE–rope* or *ride–ROPE.* This manipulation preserves contrasting cases of letters from Shallice and McGill, while at the same time avoiding within-word case alternation, which is known to have a disruptive effect on letter identification performance (Adams, 1979; McClelland, 1976). If physical similarity of the stimuli is even partially responsible for the surround-similarity effect, different-case stimuli should produce fewer migrations than same-case stimuli. However, if migrations are influenced by abstract, structural similarity between letter strings, there should be as many migrations for different-case words as for same-case words.

Method

Stimuli. The stimuli were chosen from the set of four-letter words used in Experiment 1, excluding any word that contained repeated letters.

As in Experiment 3, we allowed letters to migrate from two positions in a word, the target-letter positions (TLPs). For each TLP pair, a computer program examined each word in the corpus as a potential target and attempted to find two matching contexts that met the following requirements: (a) The surround letters of the two contexts had to be identical to the homologous letters of the target; (b) the letters in the TLPs of the two contexts could migrate to the target to form new words; and (c) the two contexts had to have different letters in the target positions.

Of the target-context sets generated, 192 were selected for use as stimulus sets, 32 per TLP pair. Eight trials were formed from each target-context set, a trial being the target in either upper- or lowercase paired with one of the two contexts in either upper- or lowercase. In selecting target-context sets and distributing trials among subjects, there were several constraints: (a) A word could be presented to a subject no more than once as target; (b) a word could be presented no more than once as a context if it also appeared as a target; (c) a word could be presented no more than twice as a context. A word could appear any number of times as a potential migration, but in practice, words seldom appeared more than once.

The 192 trials for each subject were divided into two blocks of 96 trials, with equal numbers of items. Within a block, there were 16 trials for each TLP pair. Within this set of 16, there were exactly two items for a given target case (upper vs. lower), context case, and target presentation position (to the left of fixation or to the right). The order of trials within a block was randomized.

Twenty-four practice trials were generated in a manner similar to the experimental trials.

Procedure. The procedure for this experiment was identical to that of Experiment 3, except we aimed for an average performance level across conditions of 70% correct responses. The initial blank-field duration was 333 ms; the average duration over all subjects and trials was 310 ms.

Subjects. Sixteen subjects from the same pool as in the previous experiments participated in this experiment.

---

sents an upper bound on the rate of pseudoword pseudomigrations, and it is always lower than the raw migration rate in the corresponding condition. That is, the word pseudomigration rate is always smaller than the pseudoword migration rate in corresponding target and migration type conditions.

Table 7
Percentage of Responses by Response Catgeory and Condition: Experiment 4

Target case	Context case	Correct responses	Migrations responses	Pseudo-migrations	Migration difference score
Lower	Lower	72.5	6.9	2.9	4.0
Lower	Upper	66.5	8.7	2.5	6.2
Upper	Lower	70.7	10.3	1.3	9.0
Upper	Upper	68.8	9.2	2.1	7.2
Lower	Average	69.0	7.8	2.7	5.1
Upper	Average	69.8	9.8	1.7	8.1
Average	Upper	67.6	9.0	2.3	6.7
Average	Lower	71.6	8.6	2.1	6.5
Same		70.6	8.1	2.5	5.5
Different		68.6	9.5	1.9	7.6

Results

Table 7 shows a summary of the results. The overall performance level was 69.6%, though it varied slightly from one condition to the next. A response was classified as being *correct* if the target word was reported, as a *migration* if a migration occurred in one or both of the TLPs and all other surround letters matched the target, as a *pseudomigration* if the response would have been classified as a migration had the alternate context been presented, or as an *other* error. The first four lines of the table present the results for each pairing of target case and context case. These conditions are then combined in different ways to highlight any possible influence of the case of the context, the case of the target, and of the congruity between them.

The results of the experiment were largely insensitive to the case manipulation. If anything, there were slightly more migrations for different-case than for same-case stimuli, $F(1, 15) = 3.67, p < .10$, contrary to what we would expect if physical similarity were a determinant of the surround-similarity effect. In addition, there was a slight tendency for overall accuracy to be higher for same- than for different-case stimuli, $F(1, 15) = 2.90$, $p > .10$. Neither effect was particularly strong, though the former was reliable in the migration difference score, $F(1, 15) = 7.33$, $p < .05$.

There was a greater frequency of migrations to uppercase targets than to lowercase targets—migration difference score: 8.1% versus 5.1%, $F(1, 15) = 5.64, p < .05$. The only other reliable trend in the data was a higher overall accuracy for lowercase contexts, $F(1, 15) = 4.80, p < .05$.[3]

Discussion

Migration errors between words differing in case are as frequent as they are between words in the same case, if not more so, and overall letter identification accuracy is no better for different-case words than for same-case words. This is true despite the fact that when target and context are in different cases, they are less similar in appearance; for example, compare *RIDE–ROPE* to *RIDE–rope*. Thus, manipulating the visual similarity of the words does not have an effect comparable to that observed in previous experiments when surround similarity was manipulated. We can therefore conclude that the surround-similarity effect observed in two-word displays does not depend so much on the visual similarity of stimuli as on the fact that they are similar at a more abstract level of description—that is, they consist of similar sequences of letters. This aspect of the results seems to favor the notion espoused by McClelland (1976) and Adams (1979) that the constituents of subjects' representations of words are abstract units rather than particular visual forms, and that what is familiar about a word is the sequence of letter identities, rather than the global visual configuration it makes or even the simple conjunction of the shapes of the constituent letters.

However, one result appears to support some possible role for shape information; namely, there are fewer migrations into lowercase words than into uppercase words. One explanation for this is that subjects detect the outline shape of lowercase words, and use the constraints imposed by this shape information to filter out certain migration errors. For example, with *ride–ROPE*, if subjects noticed that the first string had an ascender in the third position and no descenders, *ripe* could be ruled out, because *ripe* has a descender rather than an ascender in the third position. If this explanation were correct, we would expect fewer migration errors that change the pattern of ascenders and descenders in a letter string. To examine this possibility, we performed a post hoc analysis of the migration-error rate as a function of whether the migration would preserve or distort the shape of the target. Obviously, this issue arises only with lowercase targets, because all uppercase letters have the same shape. Considering lowercase targets, then, the migration rate was 6.7% when the migration would not change the target's shape, but only 3.5% when the migration would change the shape. This difference is reliable, $F(1, 15) = 10.3, p < .01$. To ensure that the analysis was picking up effects of shape rather than some uncontrolled factor confounded with shape, we checked whether the same effect materialized with uppercase targets, where shape would not be a factor.

---

[3] Because the overall error rates were slightly different among conditions, we felt it worthwhile to examine the migration difference score conditional upon having made an error. This quantity is simply the migration difference score divided by the total number of errors in a particular condition for a particular subject. Results were similar to the earlier analysis, although the same- versus different-case contrast did not quite reach significance—19.6% versus 25.6%, respectively, $F(1, 15) = 4.1, p > .05$.

It did not; there were 6.2% migration errors for the stimuli corresponding to the *different* lowercase shapes, and 6.7% migration errors for the stimuli corresponding to the *same* lowercase shapes ($F < 1$).

It is worth reflecting on this state of affairs for a moment. It appears that the shape a letter *would* have, in the target case, influences its tendency to migrate into the target word. However, the shape of the letter as presented in the context and the congruity of that shape with the shape it would have in the target case do not appear to be important. The results are reminiscent of a recent finding by Virzi and Egeth (1984) in an experiment in which subjects had to report the identity and the color of adjectives displayed in different colors of ink. In this experiment, subjects reported illusory perceptions of ink colors that had been presented as names and of words that had been presented as ink colors, suggesting, as our findings do, that what is "migrating" is a very abstract representation indeed.

Our results add one further twist: The tendency for abstract properties to migrate is influenced by the "reasonableness" of the result of the migration, considered in physical (i.e., shape-preserving) terms. Although there are other possibilities, our results could be interpreted as indicating that misplaced abstract letter identities are influencing the construction of a representation of the visual form of a word. For the moment, however, we leave an examination of this interesting issue to further research.

General Discussion

Summary of Findings

The studies reported in this article have replicated the letter-migration phenomenon reported by Allport (1977), Mozer (1983), and Shallice and McGill (1978). These studies have gone further in showing that migration errors are not simply a reflection of a bias to report a letter known to have occurred in the context string in place of a target letter that has not been perceived; rather, migration errors reflect a true uncertainty concerning the location in which a presented letter occurred. This point was demonstrated in Experiment 2, in which we were able to assess the accuracy of letter localization independently of the accuracy of identification.

The present studies have also added a number of further findings about the conditions under which migration errors are obtained. Although migration errors occur on some fraction of trials with all types of stimuli we have used, letters are more likely to migrate (a) between orthographically similar words than between words sharing no letters in common (the surround-similarity effect, Experiments 1 and 2), (b) when embedded in words than when embedded in digit strings (Experiments 1 and 2), (c) when the target item is an orthographically regular pseudoword than when it is a word (Experiment 3), and (d) when the potential migration error responses are words than when they are pseudowords (Experiment 3).

The studies have also helped to characterize the nature of the similarity required for obtaining the surround-similarity effect. Because letters are no more likely to migrate between words of the same-case type than between words of different-case type (Experiment 4; Shallice & McGill, 1978), it is clear that similarity is defined not in terms of physical properties of the stimuli. Further, the fact that we found no surround-similarity effect for letters in digits (Experiments 1 and 2) suggests that the letters must fit together with their neighbors to form some type of higher order structure before a surround-similarity effect can be obtained.

Another set of findings concerns the effect of a second string in the display on the accuracy of identifying a target string. It is known that the presence of a context string reduces accuracy (Kahneman, Treisman, & Burkell, 1983), but the extent of the reduction appears to depend on the exact nature of the two strings. Experiments 1 and 2 have shown that when the target and context strings share several letters in common, accuracy is reduced relative to when they share no letters in common. As with the effect of surround-similarity on migrations, it is similarity at an abstract level that matters because accuracy of identifying the target word when flanked by a context word of the same-case type was no worse than when flanked by a context word of different-case type (Experiment 4). Also, as with the effect of surround-similarity on migrations, the effect of accuracy depends on the presence of higher order structure in the displays: Accuracy of identifying a target letter embedded in a digit string was unaffected by whether the context string was composed of an identical or a different set of digits (Experiments 1 and 2).

Lastly, we note that the present experiments have incidentally replicated several findings related to the word-superiority effect that have predominantly been explored in single-word displays. These findings are as follows: (a) an advantage of letters embedded in words over letters embedded in an unfamiliar surround, as examined in a forced choice test of letter-identification accuracy (Experiment 2), and (b) an advantage of words over pseudowords, as examined in a free-report task of whole-word stimuli (Experiment 4). These effects have been widely reported in the extensive literature on perception of single words or control strings (Baron & Thurston, 1973; McClelland & Johnston, 1977; Rumelhart & McClelland, 1982); in fact, their first extension to multiword displays occurred in Reicher's (1969) report introducing the word-superiority effect.

The present results differ from some recent findings of Treisman and Souther (1986). These authors used displays consisting of four three-letter strings—either words, pseudowords, or all-consonant strings. Two tasks were used; in one, subjects were required to indicate whether a particular target string had appeared in the display; in the other, they were simply asked to report as many of the strings as they could.

Certain aspects of our present results were replicated. In particular, subjects produced more "combination errors" (migrations) than "control errors" (pseudomigrations) in both tasks; that is, subjects falsely reported seeing a string more often when all its component letters were present in the display than when they were not.

The major difference in results occurred in comparing performance on trials in which the potential migration errors were words as opposed to pseudowords. Treisman and Souther did find a significant effect of the lexical status of the potential migration response, but it was considerably smaller than the effect we obtained in Experiment 3. Another result is somewhat discrepant from our conclusions, although it does not directly conflict with our data. This is the fact that, using the target-detection task, migration errors seemed to occur as often with unpronounceable consonant-consonant-consonant (CCC) stimuli as with pronounceable consonant-vowel-consonant (CVC) stimuli. This finding contrasts with our observation that migration errors are much less frequent when letters are embedded in an unfamiliar surround.

In general, it appears that the lexical and orthographic status of target and migration strings had a larger impact in our study than in that of Treisman and Souther. Their results lead them to stress the relative similarity of results across conditions, whereas ours lead us to stress the relative dissimilarity. However, the differences are really only differences of degree. Both sets of studies found migration errors in all conditions, and both sets found some effects of lexical status of target and potential migration errors. Our studies do bring out one point that Triesman and Souther did not investigate: the surround-similarity effect, and more important, the fact that the effect depends on abstract, structural properties of the stimuli.

We now turn to a consideration of possible interpretations of these and other aspects of our results.

Interpretations of Familiarity and Similarity Effects

Earlier, several models were introduced to account for the letter-migration data. Each suggests that migrations result from an overlap in the processing of the two words. The Treisman and Souther model proposes that the earliest point of overlap is at a relatively low level, occurring when information about the location of individual letters is lost. In contrast, the McClelland and Mozer models propose that the overlap takes place at a higher level, when the contents of multiple display locations make some simultaneous contact with processing structures embodying learned information about words.

Treisman and Souther's model accounts for the data they report in their experiments, and it can account for many aspects of our results, but it fails to account for two of our central findings. First, we find that migration errors occur more frequently in words than in unfamiliar letter-in-digits (LID) strings. If migration errors did indeed result from low-level leakages in the selection process, then such errors should not be affected by higher order stimulus structure. Second, we find that migrations depend on the *abstract* similarity of *familiar* stimuli. If surround-similarity effects were obtained both with words and LID strings, we might still attribute migration errors to preattentive processes, operating independently of higher level knowledge; similarity-dependent cross-talk between perceptual channels is clearly a possible basis for the surround-similarity effect. But the fact is that a surround-similarity effect was obtained with words but not with LID strings. This fact inescapably leads to the conclusion that migration errors are due, at least in part, to a familiarity-dependent process. This point is reinforced by the finding that migration errors are as frequent between words of different cases as between words of the same case, indicating that the similarity that determines the likelihood of migration is not simply visual but is defined in terms of the learned categories to which different visual forms are assigned. Again, the conclusion that some familiarity-dependent process influences migration errors seems to be required by the results.

In what follows, we offer two alternative approaches to account for the familiarity-dependent nature of the letter-migration phenomenon.

One suggestion (McClelland, 1985) begins with the observation that a connectionist network, for example, the interactive-activation model of word perception (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982), has information stored in it in the form of connections among the units in the network. Information actively before the mind is represented by a pattern of activation over the units, but the information in long-term store that allows familiar stimuli to produce specific effects is stored in the interconnections among the units. For example, the fact that *M* in the first position of a letter string, *A* in the second, *K* in the third, and *E* in the fourth activate the word *MAKE* is due to connections among the units for the letters and the unit (or units) for the word.

The essential idea in McClelland's model is that it is possible to process multiple display locations simultaneously if local networks of processing elements, tied to spatial position, each have the necessary connection information. One way of achieving this is to have independent copies of the same processing network, with the identical connection information coded in each network. However, this scheme requires much redundancy and is inflexible if connections need to be changed. Rather than hard wiring the connections, connections in McClelland's model are programmed by a central knowledge source.

This idea is illustrated in Figure 1, wherein a hard-wired connectionist network is shown on the left next to a programmable one on the right. The hard-wired net (on the left) is set up with connections that allow two-letter strings from the limited alphabet {I, N, O, S} to activate units for the five words that can be made from these letters. On the right is shown a network of the same size that can be programmed to do this job or many other jobs by way of a second input to each of the connections. The connections in this net are different from the ones in the left hand net, because they transmit the product of the two signals they receive to the unit they connect to. This means that an external input essentially sets the strength of the connection in the local net. If the external input is 0, it is as if there is no connection; if it has some value, say α, greater than 0, then it is as if α is the strength of the connection between the two units.

When some subset of the connections in the right-hand net have been turned on, we can see it as having been programmed to process input patterns in a particular way. For example, by turning on the connections that are hard-wired in the left-hand net, we could program the right-hand net to do the same work. By turning on different connections, we could program it to process the words made of the *INOS* letters in some other language. If we imagined that the level below was reprogrammed so that that letter-level units were activated by letters in some other alphabet, we could program the net to process words in Greek, Russian, or whatever simply by external inputs to appropriate connections.

The model assumes that a perceptual processing system consists of a large number of these local networks and that they can be programmed with connection information to process different things. What programs them is, of course, a central network. If a single pattern is shown, it activates input units to one of the local networks and simultaneously projects to the central network. In the central network it produces a pattern of activation that is not itself the basis of the perception of the pattern, because it is not tied to a position in space. Rather, it is the basis for activating connections in the local networks. When a single pattern is presented for processing, the connections activated will be those that are appropriate for processing the pattern shown.

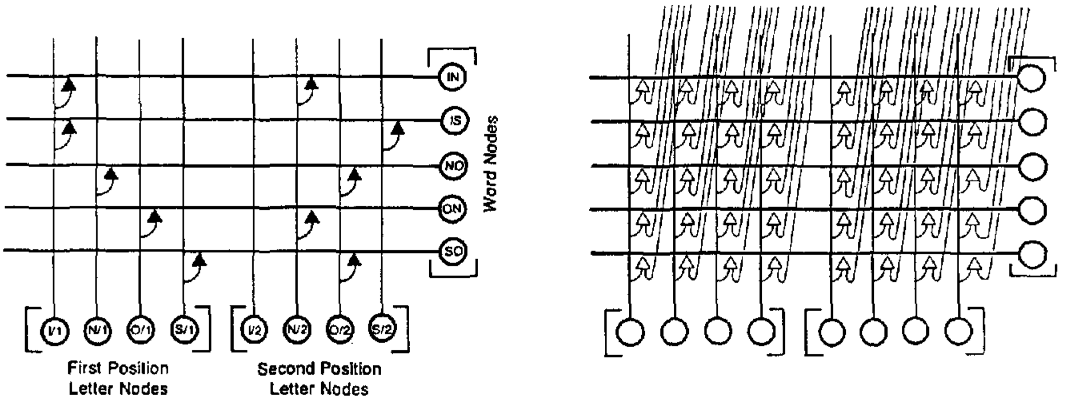

Figure 1. A hard-wired net for processing the words *in, is, no, on,* and *so* (on the left), and a net that can be programmed to do the same job. (See text for explanation.)

Connections for processing other, similar patterns will be activated but to a reduced extent.

When multiple patterns are shown, they project simultaneously to the central network. There the pattern of activation they produce is one that captures relative location within word, but fails to capture which letters occurred in which word, as in the case we considered in the introduction, in which two patterns projected simultaneously onto the same net. This superposition of patterns is, however, not the basis of perception, but only the basis for programming the local networks. The local networks will be programmed to allow inputs to activate detectors for all the words that lie in the superposition of the two input words (e.g., *LAND, SANE, SAND,* and *LANE* in the case of the input *LAND SANE*) but will be programmed less effectively to allow inputs to activate detectors for other words (e.g., *BAND* or *LINE*) that do not. The result is that each programmable network shows a tendency for words whose letters are all present in the superposition of two stimuli to become more active than words whose letters are not all present in the superposition. In the elaborated version of this model described in McClelland (in press), these active word units send feedback support to their corresponding letter units, thereby increasing the tendency for letters present in the context string to be produced as errors.[4] Attention to just one item in a multi-item display would reduce this tendency, and so would account for the low frequency of migration errors under conditions of focused attention (Treisman & Schmidt, 1982).

In this model, the surround-similarity effect arises because of the fact that duplicated letters in both strings cause stronger activation of central word nodes, which in turn causes stronger activation of connections in the local processing modules. The result of this is stronger top–down support from local word nodes to local letter nodes when the target and context strings share letters than when they do not. Such an effect would not be obtained with letters in digits, of course, because there are no central nodes for letter–digit combinations as there are for letter combinations that make words.

The letter-migration errors that occur for LID strings are accounted for, on this model, by assuming that the same type of mechanism that programs local word-nodes to be activated by local letter-nodes programs local letter-nodes to be activated by local feature-nodes. Such an assumption accounts for the fact that visual similarity appears to influence letter transposition errors obtained in displays consisting of a single row of consonants (Morrison, 1983).

The model shows a strong tendency to make more migration errors when the potential migration responses are words than when they are not. The reason is that migration errors are based almost entirely on partial activations of word units, even if response selection occurs by reading out activations from the letter level. Thus, if, for example, *LAND* is presented as target with *SANE* as context, detectors for *LANE* and *SAND* are strongly activated in the competition with the detector for *LAND,* and therefore *E* and *S* receive considerable feedback support from the word level. If, on the other hand, *HAND* is presented as target with *CANE* as context, because there are no detectors for the pseudowords *HANE* or *CAND,* there will be little word-level activation which supports the letter-level units that correspond to the migration error responses. This will be true, in general, whether or not the displayed items are themselves words. Thus, the model predicts that migration errors will be much more frequent when the error forms a word than when it forms a pronounceable nonword.

Because the letter units in the McClelland model are assumed to represent letter-identity information abstracted from specific visual details such as case, font, and so forth (cf. McClelland, 1976), it follows that the surround-similarity effect should not depend on the physical similarity of the surrounds used in each of the display items, but on the fact that they are similar at the level of letter-identity codes.

An alternative model to account for the letter-migration data has been suggested by Mozer (1984). This model relies on the notion of a hierarchy of detectors, starting at the lowest level with position-specific primitive-feature detectors, and progressing

---

[4] In the version of the model described in McClelland, 1985, feedback from the word to the letter level was excluded for simplicity, and subjects were assumed to generate their responses directly from the word level. The more recent version of the model described in McClelland (in press) includes the feedback mechanism and allows readout from the letter level, thereby permitting the generation of responses that might not be words. Both versions of the model can account for the basic migration effect, but the newer version is more easily extendable to cover nonword as well as word displays.

to a level composed of position-independent "letter cluster" detectors. Intervening levels register successively higher order features and collapse over local spatial regions of the level below, resulting in a decrease in position specificity at each successive level (Fukushima & Miyake, 1982). The higher order features are assumed to be familiarity dependent, having been learned through experience.

The letter cluster detectors at the highest level of the hierarchy represent patterns on the order of letter triples. Thus, presenting the word *SOLE* anywhere in the visual field would cause activation of units like **SO, SOL, S_LE, OLE, LE**, *E***, and partial activation of units like *SAL* and *ULE* (an asterisk indicates a blank space, and an underscore indicates "don't care" in a given position). Note that the set of units activated by a word, though unordered, is generally sufficient to reconstruct the ordered components of the word (Wickelgren, 1969).

The essence of this model is the assumption that when two letter strings are presented, the appropriate letter-cluster units for both strings are simultaneously activated to some degree. Because all positional information has been discarded at the letter-cluster level, the letter-cluster units do not explicitly code to which word they belong. Thus, the main problem faced by the perceptual mechanism is to disentangle activations produced by one string from activations produced by the other.

This process of "disentangling activations" is achieved by a competitive network called the *pull-out net*, which suppresses all but one of the consistent patterns active among the letter-cluster units. The pull-out net is composed of units in one-to-one correspondence with the letter-cluster units. Each letter-cluster unit excites its corresponding unit in the pull-out net; thus, the pattern of activity in the letter-cluster units is copied to the pull-out net. Within the pull-out net, units representing letter clusters that may fit together within a single string (e.g., *SOL* and *OLE*) are mutually excitatory, and units representing letter clusters that are potentially inconsistent (e.g., *SOL* and *SOF*) are mutually inhibitory. In addition, "semantic" units that represent higher level knowledge may come into play to help support sets of units in the pull-out net that form words. From these interactions, an internally consistent collection of activations emerges that represents the perceived string.

If attention is focused on one string in the display, activations from that string are enhanced at the primitive-feature level. As these activations work their way through the hierarchy of detectors, the letter-cluster units appropriate for the attended string will tend to become the most active as well. Consequently, these letter-cluster units will tend to dominate in the pull-out net competition, causing the attended string to be read out by the pull-out net. If attention is focused serially on each string in the display, the strings can be read out in succession. However, if attention is not focused, activations from the two strings are about equally strong, and random factors will strongly influence which string will win the competition. Under these circumstances, migration errors can result, and they should be more frequent when the stimulus strings share letters in common than when they do not: When the strings share no letters, for example, *SOLE CAMP*, the patterns of activation produced by the strings have little overlap, and hence separation of the strings is not difficult. However, when the strings share letters, for example, *SOLE SAME*, the activated letter cluster units will include **SO, SOL, OLE, LE**, **SA, SAM, SA_E, AME*, and *ME**, and, in addition, both strings will partially activate units like *SOM* and *ALE*. Due to the overlap in activation, the "migration" words *SOME* and *SALE* will be fairly consistent with the overall pattern of activation, and thus may be read out of the net accidentally in place of one of the presented strings.

The output of the pull-out net specifies a string, but not the spatial location of the string. Location information is supplied by the attentional system, because the current focus of attention is presumed to indicate the location from which activations are arising. The fact that subjects rarely report the context word indicates that the position postcue is able to direct their attention to the correct spatial region before activations arising from that region have been completely erased by the mask.

Now consider what will happen with LID strings. The pattern of activation produced by a LID string at the letter-cluster level will be fairly sparse. There may be units that detect single digits, small clusters of digits, and single letters, but there would be few, if any, that directly detect combinations of letters and digits; for example, the string 2X22 may activate units such as **_*X* (*X* in the second position of a string), **2, and 22*, but there is no 2X2 unit to become activated. When two LID strings are presented, whether same or different surround, the patterns of activation produced by the two strings should not overlap much, at least insofar as the critical letter is concerned. Without this overlap, surround similarity will have little effect on migration rates.

As for why any migrations should occur with LID strings, consider what might happen when the pair 2X22 3B33 is presented and attention is unfocused. The units **_*X* and **_*B* will both become activated, but it will be impossible to correctly determine which letter appeared in which location, or for that matter, which letter appeared with which surround. Consequently, migrations will result. According to this account, LID migration rates were lower than LIW rates in Experiment 2 because stimulus exposure durations were longer for LID, allowing subjects more time to focus attention on one or both strings.[5] In fact, this account predicts that LID migrations should be as frequent as LIW migrations when LID and LIW conditions are equated on exposure duration, rather than on accuracy of target report. We have not yet carried out a completely satisfactory test of this prediction. In one pilot study for Experiment 2, the results were apparently in accord with this prediction, but there were ambiguities in the interpretation of these results that could not be resolved without further experimentation.

The model shows a preference for words over pseudowords, for the following reason. As mentioned above, the pull-out network receives assistance from semantic-knowledge units that act to "hold together" the letter clusters composing meaningful words. If these units were removed from the model, the model would behave identically with words and pseudowords. However, with the semantic units in place, the model shows a preference for reading out letter clusters that compose words over those that

---

[5] The same explanation seems adequate to account for the lower LID rate in Experiment 1. Although the difference in exposure durations was not as dramatic in Experiment 1 as in Experiment 2, note that, in addition to the longer exposure durations for LID stimuli, the accuracy of target report was also somewhat higher for LID, affording fewer opportunities for LID migrations.

compose pseudowords, and, in fact, has a somewhat easier time with words by virtue of the additional knowledge. Consequently, the model predicts that performance should be better for word targets over pseudoword targets. With poorer performance for pseudoword targets, more opportunities for migrations would arise, and it is therefore not surprising that more migrations did actually occur in that condition. The semantic units also explain why word migrations are more frequent than pseudoword migrations: Components of a word migration are held together better than components of a pseudoword migration, making word migrations relatively more likely, especially when the target is a pseudoword.

As with the McClelland model, the Mozer model assumes that letter-cluster units represent letter-identity information abstracted from specific visual details. It thus follows that the surround-similarity effect should not depend on the physical similarity of the surrounds used in each of the display items, but on the fact that they are similar at the level of letter-identity codes.

Conclusions

It remains to be seen whether either of the models described above will ultimately provide a completely adequate account of the results obtained in our experiments, Treisman and Souther's recent experiments, and a wide range of other studies of the perceptual processing of letters in words and of familiar objects in general. Whether or not the models are correct in their details, they both make the central claim, supported by the data in the present experiments, that items in a multi-item display make simultaneous contact with processing structures embodying learned information about familiar objects and that letter migrations fall out of the resulting interactions.

In interpreting the surround-similarity effect, Treisman and Souther have proposed an elaboration of their model that essentially agrees with this claim. By the same token, both of our models have incorporated one of the central claims of Treisman and Souther, namely, that attention serves to limit access to central processing structures so as to reduce cross talk among multiple display items. At this point, it appears as if we have begun to see some agreement emerging among the models; we hope that this trend will continue as further research enriches our understanding of perceptual interactions in complex displays.

References

Adams, M. J. (1979). Models of word recognition. *Cognitive Psychology, 11,* 133–176.

Allport, D. A. (1977). On knowing the meaning of words we are unable to report: The effects of visual masking. In S. Dornic (Ed.), *Attention and performance VI.* Hillsdale, NJ: Erlbaum.

Baron, J., & Thurston, I. (1973). An analysis of the word-superiority effect. *Cognitive Psychology, 4,* 207–228.

Bradshaw, J. L. (1974). Peripherally presented and unreported words may bias the perceived meaning of a centrally fixated homograph. *Journal of Experimental Psychology, 6,* 1200–1202.

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review, 70,* 80–90.

Estes, W. (1975). The locus of inferential and perceptual processes in letter identification. *Journal of Experimental Psychology: General, 1,* 122–145.

Fukushima, K., & Miyake, S. (1982). Neocognition: A new algorithm for pattern recognition tolerant of deformations and shifts in position. *Pattern Recognition, 15,* 455–469.

Johnston, J. C., & McClelland, J. L. (1973). Visual factors in word perception. *Perception & Psychophysics, 14,* 365–370.

Johnston, J. C., & McClelland, J. L. (1974). Perception of letters in words: Seek not and ye shall find. *Science, 184,* 1192–1194.

Johnston, J. C., & McClelland, J. L. (1980). Experimental tests of a hierarchical model of word identification. *Journal of Verbal Learning and Verbal Behavior, 19,* 503–524.

Kahneman, D., Treisman, A., & Burkell, J. (1983). The cost of visual filtering. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 510–522.

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

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6,* 293–323.

Manelis, L. (1974). The effect of meaningfulness in tachistoscopic word perception. *Perception & Psychophysics, 16,* 182–192.

Massaro, D. W., & Klitzke, D. (1979). The role of lateral masking and orthographic structure in letter and word recognition. *Acta Psychologica, 43,* 413–426.

McClelland, J. L. (1976). Preliminary letter identification in the perception of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 80–91.

McClelland, J. L. (1978). Perception and masking of wholes and parts. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 210–223.

McClelland, J. L. (1985). Putting knowledge in its place: A scheme for programming parallel processing structures on the fly. *Cognitive Science, 9,* 113–146.

McClelland, J. L. (in press). The programmable blackboard model of reading. In J. L. McClelland, D. E. Rumelhart, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 2.* Cambridge, MA: Bradford Books.

McClelland, J. L., & Johnston, J. C. (1977). The role of familiar units in perception of words and nonwords. *Perception & Psychophysics, 22,* 249–261.

McClelland, J. L., & Miller, J. O. (1979). Structural factors in figure perception. *Perception & Psychophysics, 26,* 221–229.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part I. An account of basic findings. *Psychological Review, 88,* 375–407.

Morrison, I. (1983). *Computational model of visual information processing.* Unpublished doctoral thesis, Queens University, Kingston, Ontario.

Mozer, M. C. (1983). Letter migration in word perception. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 531–546.

Mozer, M. C. (1984). *The perception of multiple objects: A parallel, distributed processing approach.* Unpublished thesis proposal.

Norman, D. A. (1968). Towards a theory of memory and attention. *Psychological Review, 75,* 522–536.

Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation–verification model for letter and word recognition: The word-superiority effect. *Psychological Review, 89,* 573–594.

Palmer, S. E. (1975). Visual perception and world knowledge: Notes on a model of sensory-cognitive interaction. In D. A. Norman & D. E. Rumelhart (Eds.), *Explorations in cognition* (pp. 279–307). San Francisco: Freeman.

Ratcliff, R. (1981). A theory of order relations in perceptual matching. *Psychological Review, 88,* 552–572.

Rayner, K. (1978). Foveal and parafoveal cues in reading. In J. Requin (Ed.), *Attention and performance VII* (pp. 149–162). Hillsdale, NJ: Erlbaum.

Rayner, K., McConkie, G. W., & Ehrlich, S. (1978). Eye movements and integrating information across fixations. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 529-544.

Rayner, K., McConkie, G. W., & Zola, D. (1980). Integrating information across eye movements. *Cognitive Psychology, 12,* 206-226.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology, 81,* 274-280.

Rumelhart, D. E. (1977). Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and performance VI.* Hillsdale, NJ: Erlbaum.

Rumelhart, D. E., & McClelland, J. L. (1981). Interactive processing through spreading activation. In C. Perfetti & A. Lesgold (Eds.), *Interactive processes in reading* (pp. 37-60). Hillsdale, NJ: Erlbaum.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: Part II. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review, 89,* 60-84.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84,* 1-66.

Shallice, T., & McGill, J. (1978). The origins of mixed errors. In J. Requin (Ed.), *Attention and performance VII* (pp. 193-208). Hillsdale, NJ: Erlbaum.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127-190.

Spoehr, K., & Smith, E. (1975). The role of orthographic and phonotactic rules in perceiving letter patterns. *Journal of Experimental Psychology: Human Perception and Performance, 1,* 21-34.

Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12,* 97-136.

Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology, 14,* 107-141.

Treisman, A., & Souther, J. (1986). Illusory words: The roles of attention and of top-down constraints in conjoining letters to form words. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 3-17.

Treisman, A., Sykes, M., & Gelade, G. (1977). Selective attention and stimulus integration. In S. Dornic (Ed.), *Attention and performance VI.* Hillsdale, NJ: Erlbaum.

Virzi, R. A., & Egeth, H. E. (1984). Is meaning implicated in illusory conjunctions? *Journal of Experimental Psychology: Human Perception and Performance, 10,* 573-580.

Weisstein, N., & Harris, C. S. (1974). Visual detection of line segments: An object superiority effect. *Science, 186,* 752-755.

Wheeler, D. D. (1970). Processes in word recognition. *Cognitive Psychology, 1,* 59-85.

Wickelgren, W. A. (1969). Context-sensitive coding, associative memory, and serial order in (speech) behavior. *Psychological Review, 76,* 1-15.

Williams, A., & Weisstein, N. (1978). Line segments are perceived better in a coherent context than alone: An object-line effect in visual perception. *Memory and Cognition, 6,* 85-90.

Willows, D. M., & MacKinnon, G. E. (1973). Selective reading: Attention to the "unattended" lines. *Canadian Journal of Psychology, 27,* 292-304.

Wolford, G. (1975). Perturbation model for letter identification. *Psychological Review, 82,* 184-199.

Woodworth, R. S. (1938). *Experimental psychology.* New York: Holt.

Received July 22, 1985
Revision received October 11, 1985 ■

New Look for the APA Journals in 1986 and Change in Frequency for *JEP: Perception*

Beginning in 1986, the APA journals will have a new look. All the journals will be 8¼ × 11 inches—a little larger than the *American Psychologist* is now. This change in trim size will help reduce the costs of producing the journals, both because more type can be printed on the larger page (reducing the number of pages and amount of paper needed) and because the larger size allows for more efficient printing by many of the presses in use today. In addition, the type size of the text will be slightly smaller for most of the journals, which will contribute to the most efficient use of each printed page.

Also beginning in 1986, *JEP: Human Perception and Performance* will be published as a quarterly rather than a bimonthly. This change is a result of the change in trim size and consequent reduction in the absolute number of printed pages per issue and is not an indication that fewer articles are being published. It will also bring *Perception* in line with the other three *JEPs*, which are all quarterlies.

These changes are part of continuing efforts to keep the costs of producing the APA journals down, to offset the escalating costs of paper and mailing, and to minimize as much as possible increases in the prices of subscriptions to the APA journals.