915

Journal of Experimental Psychology:
Human Perception and Performance
1994, Vol. 20, No. 4, 807-825

0096-1523/94/$3.00

Evidence of Flexible Coding in Visual Word Recognition

Kenneth R. Pugh, Karl Rexer, and Leonard Katz

In 3 visual word recognition experiments, the authors examined Ss' differential dependence on phonological versus orthographic information in accessing the lexicon. The critical manipulation was the presence or absence of pseudohomophones in the nonword context of a lexical decision task. Ss received a list with either no pseudohomophones (NPsH group) or 17%–30% pseudohomophones among the nonwords (PsH group). In the first 2 experiments Ss in the PsH group were faster and no less accurate on word trials than Ss in the NPsH group. Furthermore, performance in the NPsH group was adversely affected by phonological inconsistency in the target's orthographic neighborhood. In the final experiment, a double lexical decision paradigm was used, and performance on orthographically similar but phonologically dissimilar pairs differed in the 2 conditions.

Whether access to the mental lexicon during reading is mediated by phonological codes, visual codes, or both has been researched extensively in the field of cognitive psychology (Besner & Smith, 1992; Carello, Turvey, & Lukatela, 1992; Carr & Pollatsek, 1985; Coltheart, Davelaar, Jonasson, & Besner, 1977; Humphreys & Evett, 1985; Rayner & Pollatsek, 1989; Seidenberg & McClelland, 1989; Van Orden, Pennington, & Stone, 1990). Whereas it is widely accepted that both types of codes can be computed by a reader, a major issue concerns whether one or the other type of coding will dominate the process, depending on factors such as word frequency, spelling regularity, reading experience, and type of orthography (Katz & Frost, 1992; Seidenberg, 1992; Van Orden et al., 1990; Waters & Seidenberg, 1985). A related issue is whether the relative contribution of each of these codes can be modulated by task demands. In several studies subjects have demonstrated an apparent flexibility in their degree of dependence on phonological or visual codes as a function of changes in experimental context; depending on task demands, phonological coding could be made either advantageous or disadvantageous, and subjects appeared to vary their behavior accordingly (Andrews, 1982; Hanson & Fowler, 1987; Hawkins, Reicher, Rogers, & Peterson, 1976; McQuade, 1981, 1983; Monsell, Patterson, Graham, Hughes, & Milroy, 1992; Paap & Noel, 1991; Shulman, Hornak, & Sanders, 1978). If adult readers do, in fact, possess the ability to change their dependence readily between phonological and visual codes, it is a matter of interest because it suggests that this flexibility is useful for their normal everyday reading; otherwise, why would a readiness for strategic variation exist at all? The current research addresses this question of coding flexibility.

Dual Route Debate

Research on letter string pronunciation has been strongly influenced by the idea that more than one way of generating a phonological output must exist in order to account for people's ability to pronounce both words that the reader has never seen before (including pseudowords, such as BINT) and words with exceptional or unconventional spelling-to-sound relations (e.g., AISLE and PINT). The speed with which subjects can name novel words or pseudowords suggests a process of early and efficient conversion from graphemic to phonologic codes; this compiled or assembled phonology may play a role in skilled reading of familiar words as well. The ability to correctly pronounce words that violate typical grapheme-to-phoneme conversion rules (e.g., PINT) suggests a lexical constraint on phonological output and has been interpreted as evidence that phonological information can be recovered from lexicon; this information is called addressed phonology.

By far, the most popular account that incorporates all these considerations has been the so-called dual-route theory of reading (Coltheart et al., 1977). Two routes to pronunciation are posited: a phonologic route and a direct access route (note that we use the term *coding* to describe the cognitive operations within these routes or pathways). The phonologic route is said to consist of two stages. In the first stage, orthographic representations such as letters and letter clusters are converted into appropriate phonological representations such as phonemes (assembled phonology). In a second stage, these phonological representations are matched to their appropriate lexical entries or, in the case of naming, to an appropriate articulation. The direct access

---

Kenneth R. Pugh, Department of Psychology, Dartmouth College, and Haskins Laboratories, New Haven; Karl Rexer (now at Department of Psychology, Northeastern University, and Haskins Laboratories, New Haven) and Leonard Katz, Department of Psychology, University of Connecticut, and Haskins Laboratories, New Haven.

This research was supported by National Institute of Child Health and Human Development Grant HD-01994 to Haskins Laboratories. Mira Peter assisted us in testing subjects, along with students in Kenneth R. Pugh's Seminar in Cognition.

Correspondence concerning this article should be addressed to Kenneth R. Pugh, Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511. Electronic mail may be sent via Bitnet to Haskins@yalevm.

route, on the other hand, is thought to involve direct mapping from orthographic representations to lexical entries. Although specific versions of dual route theories may differ on some point or another, the following assumptions are usually made explicitly or implicitly. First, the two routes to lexicon, direct and phonologic, operate independently of one another. Second, given that the phonologic process logically requires an extra step, it will, on average, take longer to finish than direct access (Coltheart, 1978; Waters & Seidenberg, 1985; but see Stone & Van Orden, 1993). Third, it is also assumed that as reading ability develops (or in the case of specific words, as familiarity increases), subjects will tend to bypass the phonological route and rely on the direct route for lexical access. (See Van Orden et al., 1990, for a detailed criticism of each of these assumptions.)

Although the dual-route concept has continued to frame much of the experimental work in the word recognition field, all of its major tenets have been challenged in recent years (see Humphreys & Evett, 1985, and Van Orden et al., 1990, for reviews). The existence of context-independent grapheme-to-phoneme conversion (GPC) rules has been challenged (Glushko, 1979; Humphreys & Evett, 1985). Empirical challenges come from what have been called consistency effects, wherein two words, both of which follow GPC rules, behave differently in a naming task if one of them has a neighbor that shares the target's orthographic rime but whose pronunciation of this rime is different than the target's (e.g., words like PINT and LINT). This effect suggests a lexical constraint on phonological mapping; pronunciation is strongly influenced by lexically stored information. However, a GPC process that is sensitive to frequency of occurrence and number of alternatives could be seen as consistent with these effects. In fact, Rosson (1985) has obtained evidence that words and nonwords with stronger rules (as indexed by the frequency of occurrence of their GPC mappings relative to others) are named more quickly than words with weaker rules, even when controlling for consistency effects. This finding, although consonant with the GPC view, suggests that the process is sensitive to what Van Orden and his colleagues have termed *statistical regularity* between print and pronunciation (Van Orden et al., 1990).

Dual-route accounts usually are based on the assumption that phonological information builds up too slowly to be relevant to the processing of all but very low-frequency words. That assumption has been challenged by data suggesting that phonological masking benefits target processing relative to orthographic masking even when the target is masked very shortly after presentation (Perfetti & Bell, 1991; Perfetti, Bell, & Delaney, 1988). Furthermore, the idea that phonological processing can be bypassed is challenged by Van Orden's categorization experiments (Van Orden, 1987; Van Orden, Johnston, & Hale, 1988). In these studies false-positive error rates to homophone and pseudohomophone foils are much higher than to orthographically matched controls, suggesting an early influence of phonology on access to meaning. However, Jared and Seidenberg (1991) were able to replicate Van Orden's results only for low-frequency words, and this finding is broadly consistent with dual-route accounts. In these models the visual route will tend to be slower for low-frequency words, and therefore greater phonological influences are possible on these items.

Lukatela and Turvey (1993) contrasted the naming latencies of high- and low-frequency words with their pseudohomophone counterparts (e.g., DOOR vs. FOAL and DORE vs. FOLE) under different levels of attentional load. They found that words and their pseudohomophones were similarly influenced by load, suggesting that they were processed in the same way. According to dual-route theory, words (especially high-frequency ones) should be processed by the automatic direct route; therefore, this theory predicted interactions between lexical status and level of memory load. Along similar lines these authors also found strong pseudohomophone associative priming, both with pseudohomophones as primes and as targets (Lukatela & Turvey, 1991). The authors concluded that their results suggested that lexical access is primarily phonological.

Alternatives to the dual-route model have been suggested. Van Orden (1987; Van Orden et al., 1990) also proposed an account in which phonological coding always mediates lexical access, although it is conceived within the framework of a connectionist system. Van Orden pointed out that researchers tend to assume that direct access is a given and that it is the role of phonology that is debated. In fact, he argues, it is possible to question all of the existing data purporting to show direct access and consequently to treat direct access as the suspect construct. Other challenges to dual-route theory have ranged from proposals of a visual basis for access (e.g., Glushko's, 1979, analogy theory) to attempts to create modified dual-route accounts, wherein GPC mapping occurs at several levels of structure or is in some way sensitive to the statistical regularities in the mapping (Carr & Pollatsek, 1985). Although dual-route theory stands challenged in several ways, it still provides a useful framework within which to organize research questions, and the notion of more than one pathway to lexicon has not been made implausible by any of these results.

The experiments reported here were motivated by the idea that clear evidence suggesting a variable reliance on phonological or visual information would, among other things, obviously pose a challenge to any model that assumes a single route to lexicon. It might be possible to induce subjects to modify which type of coding they rely on in a word recognition task. Such evidence would not only be generally relevant to the study of reading but would suggest a very fine degree of attentional control over relatively low-level cognitive processes, and therefore would also be relevant to other areas of cognitive psychology. In the following section, we review some data relevant to the question of processing flexibility.

Evidence of Flexible Coding Processes

As noted above, dual-route theories usually assume that with increased reading skill or word familiarity, reliance on orthographic information for accessing the lexicon should also increase. Such a developmental shift in reliance on type of code would constitute important support for dual-route theory. Seidenberg and his colleagues (Seidenberg, Waters, Barnes, & Tannenhaus, 1984; Waters & Seidenberg, 1985; Waters, Seidenberg, & Bruck, 1984) found that regularity effects in the lexical decision task (faster response latencies to words that conform to GPC rules than to words that violate these rules), which would appear to implicate prelexical phonological processing, diminish both with increasing reading ability and with increasing word frequency within a given reading level. This has been taken to suggest a shift to reliance on direct access (but see Van Orden et al., 1990).

Evidence suggesting experimentally induced shifts in reliance on phonological information in several different types of word recognition tasks has been reported (Andrews, 1982; Hanson & Fowler, 1987; Hawkins et al., 1976; McQuade, 1981, 1983; Monsell et al., 1992; Paap & Noel, 1991; Shulman et al., 1978). Using a two-alternative forced-choice recognition task with masked stimuli, Hawkins, Reicher, Rogers, and Peterson (1976) found that when subjects knew that the proportion of homophone pairs was high, they were no worse at choosing the correct target in homophone pairs than they were at choosing the correct target in nonhomophone pairs. However, when the proportion of homophones was low, subjects were significantly less accurate on homophone pairs than on nonhomophone pairs. The authors argued that subjects were able to strategically control the extent to which they used phonological coding.

Shulman and colleagues (1978) used a paired lexical decision task, wherein the subject decides if two letter strings, simultaneously presented, are both words. In one experiment, subjects received pronounceable nonwords, and in a second experiment they received illegal (i.e., nonpronounceable) nonwords. When subjects received pronounceable nonwords, latencies on pairs of words that were orthographically similar but phonologically dissimilar (e.g., *COUCH–TOUCH*) were inhibited relative to a control condition. However, when illegal nonwords were used performance on this type of pair was actually facilitated relative to control. They interpreted this as evidence that subjects in the former condition used phonological codes, whereas subjects in the latter condition did not. Hanson and Fowler (1987) essentially replicated this finding. However, Van Orden and his colleagues (1990) argued that this result can be interpreted within a phonologically oriented model if it is assumed that subjects in the illegal nonword context can rely on "noisy" as opposed to "cleaned-up" phonological codes. This issue is addressed in our third experiment.

Davelaar, Coltheart, Besner, and Jonasson (1978) conducted several experiments whose outcomes can be interpreted as suggesting strategic flexibility in reliance on phonological coding. In a lexical decision task, they manipulated whether the nonword context contained pseudohomophones (nonwords which, when pronounced, sound like real words; e.g., *BRANE* and *BOTE*). Words were either homophones (e.g., *SALE*) or matched nonhomophonic controls. In an initial condition with no pseudohomophones among the nonwords, low-frequency homophonic words were responded to more slowly than their controls. However, in a second condition where the nonword context contained many pseudohomophones, no homophony effect was observed. These authors concluded that subjects can strategically control whether they use phonological coding.

McQuade (1981, 1983) also manipulated the proportion of pseudohomophones used in a lexical decision experiment. One group of subjects received a high proportion of pseudohomophones, whereas a second group received a low proportion of these items. Performance on a common set of pseudohomophone targets was compared to performance on a set of matched nonword controls. Previous studies had shown that subjects tend to respond more slowly to pseudohomophones than to nonpseudohomophones; this has been referred to as the pseudohomophone effect. In one study (McQuade, 1981) the high-proportion pseudohomophone group showed no pseudohomophone effect, whereas the low-proportion group did. McQuade surmised that the high-proportion group had suppressed phonological coding, because phonological codes would be misleading and disadvantageous on a large proportion of trials. Presumably, these subjects relied on visual access coding and, therefore, were not slower on the critical pseudohomophones than on the nonpseudohomophones. This finding, although suggestive, speaks primarily to nonword processing and does not necessarily provide insight into the processing of words.

Andrews (1982) also manipulated nonword context in a lexical decision experiment. Two groups of subjects received a common set of words, but for one group half of the nonwords were pseudohomophones, whereas for the second group no pseudohomophones were included among the nonwords. The pseudohomophone group was significantly faster on word trials than the nonpseudohomophone group. Andrews suggested that subjects in the pseudohomophone group bypassed the phonological route and, relying on the direct access route, were faster than the no-pseudohomophone subjects who were waiting for the output from the phonological route. However, a possible speed–accuracy trade-off was present in these data. Andrews also manipulated other characteristics of the words. She crossed regularity (regular vs. exception word) with consistency (absence or presence of neighbors with different rime pronunciations) and found that consistency was more reliably associated with latencies than was regularity. However, there were consistency effects for both groups, and no interactions between the group variable and consistency were reported. On the view that subjects in the pseudohomophone group were, in some way, bypassing the phonological route, whereas the no-pseudohomophone subjects were not, differences in the magnitude of phonological effects in the two conditions would have been expected, and this outcome was not obtained. In the current Experiments 1 and 2, which involved similar manipulations of nonword context, we attempted to determine whether a pseudohomophone-induced speed difference coupled with a difference in the magnitude of phonological effects between the groups can be obtained.

In contrast to Andrews's (1982) results, Stone and Van Orden (1993) found a word response latency difference favoring a group that received no pseudohomophones over those who received 100% pseudohomophones in the nonword context of a lexical decision task (see James, 1975, for similar results). This result directly opposes the one obtained by Andrews. Furthermore, G. Stone (personal communication) reports that in some as-yet unpublished studies, a latency disadvantage was observed for a group receiving only 50% pseudohomophones, and that would constitute a failure to replicate the outcome obtained by Andrews (1982). However, the differences obtained in these experiments are far from reconciled at this point (Stone & Van Orden, 1993), and the current experiments provide further evidence along these lines.

Paap and Noel (1991) also manipulated context across groups using a naming task. One group of subjects was asked to pronounce a list composed exclusively of exception words, whereas a second group was given 50% exception words and 50% regular words. Performance on a common set of exception words was the variable of interest. Subjects who received all exception words were faster on the critical items than subjects in the mixed context. Paap and Noel argued that this finding is consistent with dual-route theory. They claimed that because phonological coding is not efficient for exception words, subjects in the all-exception word context bypassed assembled phonology and instead used addressed phonology to name the target. By relying on direct access, they processed words more quickly than subjects in the mixed list condition who, presumably, were engaged in a greater degree of assembled phonological coding. One problem for this interpretation is that subjects in the all-exception word context had, in effect, more naming practice with this kind of word and might have been faster on critical trials regardless of the route used in lexical access. In the same study Paap and Noel also looked at naming performance under dual-task conditions (concurrent memory-load task) and found that low-frequency exception words were actually named faster under the high memory-load condition rather than under the low memory-load condition. In contrast, low-frequency regular words and both high-frequency regular and exception words were all named more slowly under high load than under low load. They claimed that this effect came about because the assembled phonological route was more handicapped by concurrent attentional demands than the addressed phonological route; because the assembled phonological information is thought to inhibit primarily the naming of low-frequency exception words, slowing it down through the use of a heavier memory load reduced its negative influence (but see Lukatela and Turvey, 1993, for contrasting results using similar procedures).

Monsell et al. (1992) also used a naming task and contrasted conditions in which lists consisted of words only, of nonwords only, or both words and nonwords. All words were of the exception type. They found that words presented in the word-only list received fewer regularization errors than words presented in the mixed word–nonword context. They argued that because nonwords require the phonologic route in order to generate a pronunciation, subjects receiving the mixed word–nonword context relied more on this assembled phonology and hence made more regularization errors. Subjects receiving the exclusive word context, on the other hand, could rely on the lexically generated addressed phonology, and therefore regularization errors were less likely. The authors proposed that subjects can strategically disable the assembled route in a naming task when conditions make it useful to do so.

Taken as a whole, these studies seem consistent with the proposal that subjects are flexible in the degree to which they use phonological codes in word-recognition tasks. Furthermore, these results have been obtained with several different word-recognition tasks. In the current experiments, we explore further the nature and consequences of coding flexibility. We begin with a quasi-replication of the basic phenomenon of coding flexibility together with a demonstration that the effect is indeed a phonological one: The effect is shown to involve the phonological similarity of the lexical neighborhood. In the second experiment, evidence is presented that suggests that the use of assembled phonological information makes measurable demands on attentional resources. Finally, the results of the third experiment suggest that the extraction of meaning from an identified word is not affected by which route predominates in lexical access. This is consistent with the proposal that coding flexibility affects prelexical, not postlexical, processing.

A pilot study was conducted using a between-groups manipulation of pseudohomophony. One group of subjects in a lexical decision experiment received no pseudohomophones (NPsH group). The stimulus list for a second group was created by replacing 15% of the nonwords in the first list with pseudohomophones (PsH group). Both groups received identical word lists. Half of the 128 words were of low frequency and half were of high frequency. Results indicated that the word responses of subjects in the NPsH group were significantly slower than the responses of subjects in the PsH group (NPsH = 569 ms, PsH = 524 ms). Frequency was significant and there was a marginally significant Group × Frequency interaction, indicating that subjects in the NPsH group were more adversely affected by low-frequency words than subjects in the PsH group. An analysis of the accuracy data revealed no significant differences between conditions. Hence, subjects who received pseudohomophones produced faster and no less accurate word responses than subjects who received no pseudohomophones. This outcome conforms to Andrews (1982), who used a much higher proportion of pseudohomophones (50%) but also found a latency advantage for subjects in the PsH condition; however, the result conflicts with Stone and Van Orden's (1993) results in which a 100% PsH group was much slower than a NPsH group.

The latency advantage for subjects in the PsH condition does not appear to be attributable to a simple lowering of a response threshold criterion, because that should result in a lower accuracy rate for this condition: subjects in this group were actually slightly (though not significantly) more accurate than subjects in the NPsH group. The between-group differences obtained in the pilot study might be thought of as the consequence of the fact that subjects in the PsH group are in some way either disabling the phonologic route or, perhaps, are executing a response prior to its output.

A less interesting account of the results from this pilot study is that the speed advantage (without a corresponding increase in errors) comes about because subjects in the PsH group exert more cognitive effort (greater attention) due to the difficult homophony created by the PsH items. This attentional account would not require any assumptions about differences in type of coding between the two conditions. Experiment 1 was conducted to determine whether the observed between-group difference in latencies is also associated with differences in the magnitude of effects of phonological processing difficulty. That outcome would implicate processing type differences in the two conditions.

Experiment 1

As in the pilot experiment, a pseudohomophone manipulation was used in a lexical decision experiment. However, in Experiment 1 words were selected specifically to provide a broad range on two dimensions: frequency and phonological processing difficulty. Phonological processing difficulty was indexed for each target word by a count of the number of "unfriendly" neighbors, defined as the number of English words sharing the same orthographic rime (the same spelling) as the target word but differing in rime pronunciation (e.g., BOOT and FOOT). A target word's number of unfriendly neighbors can also be considered as an index of phonological inconsistency for that word. Some words contained no "unfriendly" neighbors, whereas others contained many. This continuous measure of phonological processing difficulty is correlated with whether the word is regular or exceptional with regard to GPC rules (and many words of both types were contained in the list). However, there were several regular words with unfriendly neighbors and a few exception words with none. As noted above, the number of words that either share or do not share rime pronunciation with the target would be psychologically important in any dual-route theory in which grapheme-to-phoneme mapping is sensitive to the statistical characteristics (such as frequency of occurrence) of these transforms (Rosson, 1985; Van Orden et al., 1990). By any such account, generating the appropriate GPC mapping for the target word will be more difficult with the increase in the number of words in the lexicon possessing the same orthographic structure but a different phonological realization. In any case, without theoretical commitment as to how consistency and regularity might differ, we noted that several studies of lexical decision have found that indexes of phonological processing complexity that are based on an examination of the target's phonological neighborhood are associated with performance (Andrews, 1982; Jared, McRae, & Seidenberg, 1990; Perfetti & Bell, 1991). We predicted that subjects relying on phonological information during lexical access (NPsH condition) would be more sensitive to phonological processing difficulty than subjects engaged in direct access (PsH condition).

A recognition memory test was also conducted to determine whether subjects in these conditions differed in their depth of processing. For instance, if subjects in the PsH condition are faster as a consequence of failing to process the targets through to meaning, whereas subjects in the NPsH condition are processing through to meaning, then episodic memory differences would be expected, because semantic processing is associated in recognition memory with superior performance (Craik & Lockhart, 1972).

Method

Subjects

Forty-nine undergraduate students from the College of the Holy Cross participated for partial course credit.

Stimuli

Two lists, each containing 128 monosyllabic words and 128 monosyllabic pronounceable nonwords, were constructed. The only difference between the two was that List 1 contained no pseudohomophones among the nonwords whereas List 2 contained 22 pseudohomophones (17%). Sixty nonhomophonic words were the critical experimental items chosen to provide a broad range of frequency (Kucera & Francis, 1967, range = 2–1,617) and phonological consistency values, and 68 words were fillers (consisting of 30 homophones and 38 nonhomophones and included for use in a subsequent recognition memory test). Phonological inconsistency was indexed as the number of monosyllabic words that share the target's orthographic rime but that pronounce the rime differently than the target (range = 0–26). We called these words "unfriendly neighbors." The log of the number of unfriendly neighbors (NU) was used in the analysis. Length (number of letters) was included as a control variable (range = 3–6).

A surprise recognition memory test also was used, consisting of a 140-word list. The list included 70 previously viewed items (15 of which were homophonic filler words) and 15 words that the subjects had not seen but that were homophonic to words used in the lexical decision task.

Procedure

Subjects were randomly assigned to either the PsH group or the NPsH group. A standard lexical decision procedure was followed. Items were presented in a different random order to each subject in uppercase letters on a Macintosh 512K computer screen. Targets were preceded by a 500-ms fixation point (asterisk) in the middle of the screen and a 500-ms blank. Target presentation continued until the subject responded or until 1,600 ms had elapsed. Latencies shorter than 150 ms or longer than 1,600 ms were recorded as errors. "Word" responses were made with the dominant hand and "nonword" responses were made with the nondominant hand on two telegraph keys. Reaction time (RT) was measured with an accuracy of ±2 ms. Subjects were given 40 practice trials. Following the 256 lexical decision trials subjects were given a surprise recognition test consisting of 140 words. They were informed that half of these words were from the lexical decision list and half were not, and they were told to indicate the items that had been presented in the task by circling them. They were also instructed to work at a fairly quick pace. Each subject's participation lasted approximately 35 min.

Results

For each subject, mean latencies were calculated for the correct and incorrect word and nonword responses. Within each of these categories, trials with latencies greater than two standard deviations from the subject's own mean (calculated independently for each category) were treated as errors. Mean latencies were computed, averaged over subjects, for each experimental word and for each nonword that appeared in both the NPsH and PsH conditions. Accuracy for each item was calculated as the proportion of subjects responding correctly to it. One of the 60 experimental words failed to be displayed due to a programming error, and 3 other words had error rates greater than 60%; these 3 items were excluded from the latency analysis but not the accuracy analysis.

Standard (simultaneous) multiple regression analyses were performed on word latency and accuracy with items as the unit of analysis. The following regressors were used: the log number of phonologically unfriendly neighbors (NU), log word frequency, the interaction between these two, word length, and PsH group (as a repeated measure). The interactions between the repeated measure and each of the other regressors were also included in the analyses but were removed from an analysis if they were nonsignificant. This procedure was followed in all subsequent analyses. The categorical variable regular–exception and the proportion of neighbors that were unfriendly to the target (NU/total number of neighbors) was tried as well, but only NU was significantly associated with performance; hence, all subsequent analyses use NU as the index of the phonological inconsistency of a target word's neighborhood.

For word latency there was a significant effect of group, $F(1, 54) = 9.96$, $p < .01$, $MS_e = 424.25$, indicating that mean latencies were faster in the PsH condition (513 ms) than in the NPsH condition (535 ms). There was a significant effect of NU, $F(1. 51) = 4.08$, $p < .05$, $MS_e = 1.706.07$, and its positive regression coefficient (31.70) indicates that latencies increased as NU increased. Whereas the main effect of frequency was not significant in this model, it should be noted that with the term representing the interaction between frequency and NU removed from the model, frequency was significant, $F(1. 52) = 15.78$, $p < .001$, $MS_e = 1.830.27$. The NU × Frequency interaction was, however, significant, $F(1, 51) = 4.79$, $p < .05$, $MS_e = 1.706.07$. To examine this interaction we split the words into two roughly evenly sized frequency categories: low (including items with frequencies between 1 and 20) and high (including items with frequencies between 22 and 1,617). The positive coefficient for NU, reliable in the overall analysis, was present only for low-frequency words (14.04); the high-frequency coefficient was actually negative (−2.73). Thus, the inhibitory effect of NU appears to have been largely carried by the lower frequency words. The Group × NU interaction was also significant, $F(1, 51) = 4.25$, $p < .05$, $MS_e = 405.22$. Given the Group × NU interaction, data from the two groups were analyzed separately.

Table 1 provides a summary of the separate latency analyses of the data from the PsH and NPsH groups. Of critical interest is the fact that NU, as well as the NU × Frequency interaction, was significant only for the NPsH group: for NU, $F(1, 51) = 5.98$, $p < .05$, $MS_e = 1,304.56$; for the NU × Frequency, $F(1, 51) = 5.75$, $p < .05$, $MS_e = 1,304.56$. The positive regression coefficient for the NU effect indicates that latencies increased as the number of unfriendly phonological neighbors increased. Like the omnibus analysis, the interaction revealed that this was especially true for the low-frequency items (regression coefficients were as follows: low frequency = 25.55, high frequency = −.607). In the PsH condition neither of these terms was significant.

The omnibus analysis of the accuracy data revealed that only NU was significant, $F(1, 54) = 5.02$, $p < .05$, $MS_e = .033$; as NU increased, accuracy decreased (the coefficient for NU = −.145). Without the Frequency × NU interaction term in the model, frequency was significant, $F(1, 55) = 11.93$, $p < .001$, $MS_e = .035$. There was no group difference (NPsH = 87.2%, PsH = 87.4%). However, in keeping with the latency analysis, the accuracy data from the two groups were also separately examined. Table 1 also summarizes the word accuracy data for the two groups. As with the latency analysis, NU and the NU × Frequency interaction were significant for the NPsH group: for NU, $F(1, 54) = 6.82$, $p < .05$, $MS_e = .016$; for the interaction, $F(1, 54) = 5.30$, $p < .05$, $MS_e = .016$, but not for the PsH group. The negative regression coefficient for the NU effect indicates that as the number of unfriendly neighbors increased, accuracy decreased. As with the latency data, the words were divided into lower and higher frequency sets to examine the NU × Frequency interaction. For lower frequency words the coefficient was negative (−.143), whereas for

Table 1
Experiment 1: Regression Analyses by Condition

Variable	Pseudo-homophone Coefficient	Pseudo-homophone *F*	Nonpseudo-homophone Coefficient	Nonpseudo-homophone *F*
	Latency			
NU	15.93	1.09	47.46	5.98*
Frequency	−9.73	2.45	−9.47	1.44
NU × Frequency	−13.11	2.10	−27.57	5.75*
Length	—	<1.00	—	<1.00
	Accuracy			
NU	−.12	2.89	−.17	6.82*
Frequency	.04	1.70	.03	1.23
NU × Frequency	.06	1.71	.09	5.30*
Length	—	<1.00	—	<1.00

Note. Latency values for the pseudohomophone group: $R^2 = .24$. $MS_{res} = 807$; for the nonpseudohomophone group: $R^2 = .30$, $MS_{res} = 1.305$. Accuracy values for the pseudohomophone group: $R^2 = .18$. $MS_{res} = 0.02$; for the nonpseudohomophone group, $R^2 = .28$. $MS_{res} = 0.02$. Dashes refer to nonsignificant coefficients. NU = number of unfriendly neighbors.
* $p < .05$.

higher frequency words the slope was actually slightly positive (.016). Thus, the accuracy results parallel the latency results in this regard.

An analysis was also conducted on the 106 nonwords that subjects in both conditions had received (not including the pseudohomophones or corresponding nonpseudohomophones that were unique to one or the other condition). As with words, correct rejection latencies were faster in the PsH condition (mean = 569 ms) than in the NPsH condition (mean = 596 ms), $F(1, 105) = 61.59$, $p < .001$, $MS_e = 629.08$. No significant accuracy difference was obtained.

Signal detection analysis was used for the memory data from the recognition test that was administered after the lexical decision trials. Mean d' for the NPsH group was 2.56, and it was 2.57 for the PsH group; this difference was not significant ($F < 1.0$). An analysis of performance on just the homophonic targets and the foils also failed to reveal a significant difference between the NPsH and PsH groups.

Discussion

Experiment 1 revealed a latency advantage for words and nonwords in the PsH condition over the NPsH condition, even though the inclusion of pseudohomophones in the nonword context might have made the former condition more difficult, not less. This result replicates the results of the pilot study as well as the results reported by Andrews (1982), who used 50% pseudohomophones in the PsH condition. In the present experiment the latency advantage for the PsH group cannot be attributed to a speed–accuracy trade-off because the PsH group was slightly more accurate than the NPsH group (although the difference between the two groups was not statistically significant). However, the latency disadvantage for a group receiving 100% pseudohomophones reported by Stone and Van Orden (1993) stands in contrast to the current results. Furthermore, these authors argue that eliminating a slower route will not necessarily speed latencies, especially if a horse race process is assumed (Paap & Noel, 1991). The current results might be taken to indicate either that disabling the phonological route frees attentional resources, thereby producing more efficient orthographic processing, or alternatively, that subjects who do not disable the phonological route in some sense wait for this information to build up and that this produces the longer latencies observed in the NPsH context. Although the conflicts within the literature are not resolved, this experiment establishes, for the first time, a link between faster responding and diminished phonological influences.

There was additional evidence that phonology had been used for word recognition in the NPsH group but not in the PsH group (or, at least, not to the same degree). In the NPsH condition, the difficulty of phonological processing, as indexed by the number of phonologically unfriendly neighbors, had significant adverse effects on latencies and accuracy, whereas in the faster PsH condition this variable had no influence. This finding lends support to the claim made by Andrews (1982) that in a pseudohomophonic context subjects strategically inhibit phonological processing (because it tends to generate false positives), thereby shifting resources to the faster direct route. This is consistent with the basic architecture of dual-route theories but is problematic for most single-route models. Connectionist accounts (Lukatela & Turvey, 1990, 1993; Seidenberg & McClelland, 1989; Van Orden et al., 1990) might cope with the results of Experiment 1 by recourse to a short-term, context-induced adjustment of network dynamics; however, only actual simulations can inform such speculation.

Even for dual-route theories, the precise locus of strategic flexibility is not clear. It is possible that subjects disable or attenuate GPC level processing (Monsell et al., 1992). It is also possible that the strategic change occurs late, and that subjects simply ignore the output from GPC level processing. It does seem unlikely that the flexibility demonstrated in Experiment 1 occurs at an even later postlexical checking stage. If subjects in the NPsH group tended to engage in an extra step of "sounding out" an already recognized lexical representation, then effects of the number of phonologically inconsistent neighbors should either not be of any consequence to this group or, alternatively, should affect both groups equally. Furthermore, it is unlikely that the small latency advantage for the PsH group would be attributable to the elimination of what should be a relatively time-demanding postlexical check. It also seems that if the subjects in the NPsH group had adopted a more stringent criterion than the PsH group in checking the response, they would have had higher accuracy rates than subjects in the PsH group—but they did not. As noted earlier, Monsell and his associates (1992) and Paap and Noel (1991) both claim that subjects can ignore or bypass assembled phonological information when it is advantageous to do so in a naming task; the current results suggest a similar flexibility in lexical decision.

Finally, the recognition memory test was included to explore the possibility that subjects in the PsH group obtained a speed advantage by initiating their responses before processing the targets fully. Several recent lexical decision studies have suggested that subjects can use an orthographic familiarity bias, wherein a lexical decision is made without fully discriminating the target from its active neighbors (Johnson & Pugh, in press; Pugh, Rexer, Peter, & Katz, 1994). It seemed reasonable that if subjects in the NPsH condition were more fully processing the target than subjects in the PsH condition (perhaps using target semantic information in making the decision), then episodic memory for the lexical decision stimuli would be superior in the former group. Had group differences been obtained the results would have been provocative; the failure to show group differences, however, should not be considered overly informative.

Experiment 2

The explanation that subjects in the PsH condition were simply more attentionally focused than subjects in the NPsH condition, as a consequence of the difficult PsH context, seems at odds with the fact that different neighborhood phonological inconsistency effects were also found for the two groups. This suggests a difference in the kind of processing, not simply a difference in the efficiency of processing. Nonetheless, an attentional account cannot necessarily be ruled out because greater attentional effort, if its influence reached down to the level of lexical processing, might suppress phonological ambiguity effects to some extent; of course, attentional consequences of this sort would have relevance to the dynamics of word recognition. To examine the attentional characteristics of performance in the NPsH and PsH conditions, we used a dual-task paradigm in Experiment 2 (Lukatela & Turvey, 1993; Paap & Noel, 1991; Posner & Boies, 1971). As noted in the introduction, Paap and Noel (1991) found that naming latencies were influenced by the difficulty of a concurrent memory task. A similar manipulation was used in the current experiment. Subjects were required to make lexical decisions while holding either one digit (low load) or four digits (high load) in short-term memory. Immediately after making the lexical decision, a target probe digit appeared on the screen, and subjects decided whether the probe matched, or did not match, an element in the memory set. This memory load manipulation allowed us to examine lexical decision performance while attentional demands on the subject were either low or high.

If subjects in a PsH condition simply exert greater attentional effort in lexical decision, then increasing attentional resource demands by increasing memory load should cut into the available resource more for these subjects than for subjects in an NPsH condition. However, if the pattern of results obtained in Experiment 1 resulted from a selective disabling of the phonologic route because of the PsH context, then PsH subjects in Experiment 2, unencumbered by the presumably greater attentional demands of phonological processing, might not only perform the lexical decision task more efficiently but might also perform the secondary memory task more efficiently.

Method

Subjects

Thirty undergraduate students from the University of Connecticut participated for partial fulfillment of a course requirement.

Stimuli

Ninety-six experimental words, possessing a broad range of values on the dimensions of frequency (1–1,617) and NUs (0–26), were used. Along with these, 36 filler words (32 of them homophones) were also included, as in Experiment 1. The nonwords for the NPsH group were 132 pronounceable nonpseudohomophones; in the PsH condition, 30% of these nonwords were replaced with pseudohomophones. All of the word and nonword stimuli are presented in Appendix A. Half of the memory sets consisted of one digit (low load) and half consisted of four digits (high load). Within each of these, half of the target probes matched an element in the set, whereas for half there was no match. Each stimulus was viewed by half of the subjects under the one-digit memory load and by the other subjects under the four-digit memory load.

Procedure

Each trial began with the presentation of a fixation point (an asterisk) for 500 ms. After a 500-ms pause the memory set was presented for 1,500 ms. Following a 1,500-ms interval, the lexical decision target was then presented. "Word" responses were made with the dominant hand and "nonword" responses were made with the nondominant hand on two telegraph keys. Subjects had 1,500 ms from the onset of the stimulus to make their lexical decisions. The probe digit appeared on the screen 1,700 ms after the offset of the lexical decision target; it lasted for 600 ms and subjects had up to 1,500 ms to respond positively or negatively using the same telegraph keys. Lexical decision and probe response latencies shorter than 150 ms or longer than 1,500 ms were recorded as errors. Latencies were measured with an accuracy of ± 2 ms. Subjects received 40 practice trials and then the experiment's 264 trials, which were presented in a different random order to each subject. Subjects were instructed to respond as quickly and as accurately as possible to both the lexical decision and the memory probe judgment and were told that the word–nonword task was inserted into the retention interval in order to make the memory task more challenging. However, subjects were not explicitly told that the memory task was the primary task. Each subject's participation lasted approximately 45 min.

Results

Mean lexical decision and memory probe response latencies were computed for each item, averaging over subjects following the same procedure that was used in Experiment 1. Accuracy values were also calculated, as in Experiment 1.

Lexical Decision Analyses

Standard (simultaneous) multiple regression analyses were conducted on the latency and accuracy of the lexical decision word and nonword responses, using items as the unit of analysis, as in Experiment 1. The between-item variables in the analyses were log frequency, log NUs, the interaction between these two terms, and word length. Group (NPsH vs. PsH) and load (low vs. high) were within-item variables. Separate regression analyses were also performed on the NPsH data and the PsH data, as in Experiment 1.

Words: Latency. The omnibus analysis of the word latency data revealed a significant Group × Length interaction, $F(1, 91) = 8.65$, $p < .01$, $MS_e = 880.65$. As Figure 1 illustrates, the latency advantage for the PsH over the NPsH condition is larger for longer words (5–6 letters) than for shorter words (3–4 letters). The only other term to reach significance in this analysis was frequency, $F(1, 91) = 17.06$, $p < .001$, $MS_e = 5{,}971.12$; latencies decreased with increased frequency. It should be noted that with the Group × Length interaction removed from the model the main effect of group did obtain significance, $F(1, 95) = 112.52$, $p < .0001$, $MS_e = 947.69$; NPsH (633 ms) and PsH (599 ms). The NU did not reliably affect response latency,

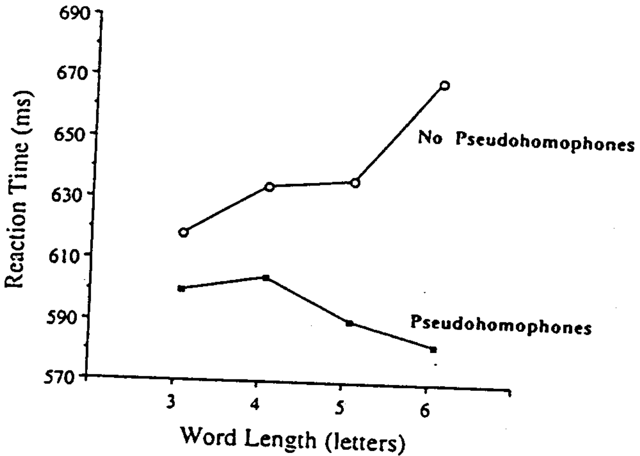

Figure 1. Word response latency in the dual-task procedure as a function of group and word length.

nor did it interact with any of the other variables. The separate regressions performed on each group revealed no additional effects.

Words: Accuracy. Neither the omnibus nor the separate regression analyses of the word accuracy data revealed any significant effects. The range of mean percentage correct for the two groups across the two load conditions was only 2% (88%–90%).

Nonwords: Latency. The omnibus analysis of the latency data from the subset of nonwords that had been presented to subjects in both groups revealed a significant effect of group, $F(1, 90) = 132.64$, $p < .001$, $MS_e = 861.01$; correct rejection latencies were faster in the PsH condition (674 ms) than in the NPsH condition (709 ms). There was also a significant effect of load (low = 685 ms, high = 698 ms), $F(1, 90) = 14.15$, $p < .001$, $MS_e = 1{,}100.89$, but no significant interaction between these two factors. Length was significant, $F(1, 90) = 3.98$, $p < .05$, $MS_e = 5{,}351.46$, with longer nonwords yielding slower rejection latencies. A Load × Length interaction was also obtained, $F(1, 90) = 8.14$, $p < .01$, $MS_e = 1{,}100.89$; the means were 672 ms, 698 ms, 695 ms, and 701 ms for low (load)–shorter (length), low–longer, high–shorter, and high–longer conditions, respectively. Thus, the length effect was considerably larger under a low memory load (26 ms) than under a high load (6 ms). Finally, the three-way Group × Load × Length interaction approached significance, $F(1, 90) = 3.15$, $p < .08$, $MS_e = 1{,}809.11$. The length effect for the NPsH group under low load was nearly twice as large as in any other cell (35 ms).

Nonwords: Accuracy. The omnibus analysis of the accuracy data revealed a significant effect of group, $F(1, 90) = 10.21$, $p < .01$, $MS_e = .0072$, indicating greater accuracy in the PsH condition (92%) than in the NPsH condition (89%). No other terms were significant.

Probe Task Analyses

Standard multiple regression analyses were also conducted on the latency and accuracy of the probe task responses following words and nonwords, using items as the unit of analysis. These analyses used the same variables as the analyses of the lexical decision responses, and included the additional within-item variable of match or mismatch between probe and memory set (match), and its interactions with the other variables. Separate analyses were also performed on the NPsH and PsH data.

Words: Probe latency. The omnibus analysis of the latency of probe responses following words revealed main effects of load and match, $F(1, 180) = 119.11$, $p < .001$, $MS_e = 2{,}632.35$, and $F(1, 180) = 38.37$, $p < .001$, $MS_e = 2{,}632.35$, respectively. The interaction between these terms was also significant, $F(1, 180) = 12.34$, $p < .001$, $MS_e = 2{,}632.35$. The means were 525 ms, 433 ms, 643 ms, and 587 ms for the low–mismatch, low–match, high–mismatch, and high–match conditions, respectively. Thus, the interaction indicates that the advantage on match trials was larger under low load (92 ms difference) than under high load (55 ms difference). Group (NPsH vs. PsH) was also significant, $F(1, 190) = 87.90$, $p < .001$, $MS_e = 2{,}000.96$, revealing that probe task performance was faster in the PsH condition (526 ms) than in the NPsH condition (569 ms). The Group × Load interaction was also significant, $F(1, 190) = 6.22$, $p < .05$, $MS_e = 2{,}000.96$. The means were 495 ms, 642 ms, 464 ms, and 588 ms for the NPsH–low, NPsH–high, PsH–low, and PsH–high conditions, respectively. Thus, the latency advantage of low load over high load was 23 ms larger in the NPsH condition than in the PsH condition.

In the separate analyses of the PsH and NPsH groups, the effect of NU was marginally significant in the NPsH condition, $F(1, 184) = 2.92$, $p < .10$, $MS_e = 2{,}463.52$. The regression coefficient for NU was positive (18.36); thus, as the number of phonologically unfriendly neighbors increased so did latencies. Also, the NU × Frequency interaction was marginally significant in the NPsH condition, $F(1, 184) = 3.80$, $p = .05$, $MS_e = 2{,}463.52$. As in Experiment 1, the effect of NU in the NPsH condition was examined separately for high- and low-frequency words. As expected, the regression coefficient for NU was positive for the lower frequency items (2.90), whereas for higher frequency items the coefficient was negative (−26.67). Neither of these terms approached significance in the PsH condition.

Words: Probe accuracy. The omnibus analysis of the accuracy of probe responses following words also revealed a main effect of load, $F(1, 180) = 8.78$, $p < .01$, $MS_e = .004$; the means were 97% and 94% correct for the low- and high-load conditions, respectively. There was also a significant Load × Match interaction, $F(1, 180) = 7.10$, $p < .01$, $MS_e = .004$. The means were 97%, 96%, 95%, and 92% for the low–mismatch, low–match, high–mismatch, and high–match conditions, respectively. The interaction appears to come from the fact there were relatively more errors on high-load trials when the probe matched an item in the memory set than in any of the other conditions. Group was also significant, $F(1, 191) = 7.10$, $p < .01$, $MS_e = .004$, indicating somewhat greater accuracy on the probe task for subjects in the PsH condition (96%) than for subjects in the

NPsH condition (94%). No other effects or interactions were reliable, although NU approached significance, $F(1, 180) = 2.81, p < .10, MS_e = .004$. An examination of the separate NPsH and PsH analyses revealed that NU was significantly related to accuracy only in the NPsH analysis, $F(1, 180) = 4.10, p < .05, MS_e = .005$, and $F < 1.0$ for NPsH and PsH, respectively; its negative regression coefficient (−.869) indicated that as the NUs increased, accuracy decreased. As with latencies, subjects in the PsH condition were not influenced by this variable.

Nonwords: Probe latency. The omnibus analysis of the latency of probe responses following nonwords revealed main effects of load, $F(1, 154) = 527.74, p < .001, MS_e = 2{,}441.11$, and match, $F(1, 154) = 200.23, p < .001, MS_e = 2{,}441.11$. As with word trials latency advantages for both low-load and match trials were obtained. The interaction between these terms was also significant, $F(1, 154) = 9.55, p < .01, MS_e = 2{,}441.11$, and it reflects that the match advantage was larger (96 ms) under low load than high load (61 ms; the means were 565 ms, 469 ms, 675 ms, and 614 ms for the low–mismatch, low–match, high–mismatch, and high–match conditions, respectively). Group was, once again, significant, $F(1, 158) = 60.29, p < .001, MS_e = 3{,}234.60$, revealing that performance on the probe task was faster in the PsH group (556 ms) than the NPsH group (605 ms). None of the interactions with group was reliable.

Nonwords: Probe accuracy. The omnibus analysis of the accuracy of probe responses following nonwords indicated a main effect of load, $F(1, 154) = 6.71, p < .05, MS_e = .004$; accuracy was greater in the low-load condition (95%) than in the high-load condition (93%). Group was also significant, $F(1, 155) = 13.71, p < .001, MS_e = .004$, with greater accuracy in the PsH group (95%) than the NPsH group (93%). No interactions were reliable.

Discussion

The main focus of Experiment 2 was to determine whether lexical access made greater demands on attention for the NPsH group than for the PsH group (the latter condition is hypothesized to be less dependent on phonology). The results supported the hypothesis but in a manner that was less direct than expected. The memory load affected the NPsH group adversely (as expected), but its major effect was on that group's memory probe recognition rather than on its lexical decision performance. Responses to the memory probe were slower and less accurate in the NPsH group than in the PsH group. Furthermore, for the NPsH group, increasing the memory load (from one to four digits held in memory) had a relatively greater deleterious effect on probe RT than for the PsH group. Thus, it was clear that subjects in the NPsH group were less able than those in the PsH group to perform the attention-demanding memory probe recognition.

Although the results of the lexical decision analyses were less straightforward, subjects in the NPsH group showed deficits in their performance on this task as well. The clearest results were for nonword lexical decisions, which were slower and less accurate for the NPsH group. With regard to lexical decisions on words, the only evidence of a NPsH–PsH group difference was a Group × Word Length interaction (see Figure 1); longer words were processed more slowly by the NPsH group, but the reverse was true for the PsH group. This interaction suggests differences in the kind of processing and not simply in the efficacy or efficiency of processing. Any explanation that does not posit a change in type of processing would find this disordinal interaction problematic. To account for the interaction, we speculated that if the NPsH subjects (who are hypothesized to be relatively dependent on phonological coding) were engaged in a series of grapheme-to-phoneme conversions while processing a word, then a longer word should require more conversions and, therefore, should take longer to process. PsH subjects, on the other hand, apparently engaged in visual (i.e., orthographic) processing; perhaps this visual processing is a parallel rather than a serial process. Under parallel convergence, longer words would be processed faster than shorter words because longer words have smaller numbers of competing items in their respective neighborhoods (longer words are more nearly unique).

As noted above, the Group × Load interaction in the probe recognition shows that subjects in the NPsH condition were more adversely affected by high load than PsH subjects. This is consistent with the idea that, due to extra phonologic processing in the NPsH condition, fewer attentional resources were available for the demanding memory condition. This interpretation is reinforced by the fact that as the neighborhood's phonological inconsistency (NU) increased, subjects in the NPsH condition were more likely to forget what they were holding in short-term memory (i.e., accuracy decreased). Additionally, as NU for low-frequency words increased, NPsH subjects were slower on probe judgments. PsH subjects, on the other hand, who were hypothesized not to be dependent on phonological coding, showed no hint of an influence of neighborhood phonological inconsistency on probe recognition. In summary, the results suggest that subjects under PsH conditions adjusted processing in some way so as to eliminate the disadvantageous influence of phonological processing.

The fact that subjects in the PsH condition were generally faster and more accurate on both tasks than subjects in the NPsH condition seems consistent with the idea that subjects in the former condition were processing words in a way that was not only more advantageous for lexical decision, but actually freed up attentional resources for the memory task in which the word recognition task was embedded. If subjects in the PsH condition were in some way disabling the presumably attentionally expensive assembled phonology route, then the results can be explained. An alternative account that preserves the notion that GPC processes are in some way disabled in the PsH context but does not borrow on the attentional resources explanation, might also be consistent with these data. Some residual interference between word recognition and memory probe tasks might result if the two share common code types (the memory set is likely held in short-term memory by a phonological code). In the current experiment, then, subjects in the PsH condition (who do not generate phonological codes) would suffer less interference than subjects in the NPsH condition (who engage in extra phonological processing during word recognition). In any event, the idea that the speed and accuracy advantages for PsH subjects in the first experiment could have come about due to greater attentional effort in lexical decision would seem to have predicted a trade-off between the two performances. Instead, superior performance as a function of the inclusion of pseudohomophones was found on both tasks. An attentional account might be constructed that can handle aspects of the current results, possibly one assuming a great deal more vigilance in the PsH context, but this approach would seem less consistent with the total pattern of data than the coding flexibility hypothesis.

A somewhat perplexing aspect of the current experiment is that, whereas neighborhood phonological inconsistency only influenced the performance of NPsH subjects, as in Experiment 1, in this experiment the influence revealed itself, not on the lexical decision, but on the subsequent memory probe. How is it that subjects in the NPsH condition, if engaged in more phonologically based reading, would not exhibit NU effects on the word recognition task itself, as well as on the subsequent memory judgment? The failure to obtain a stronger NU effect for the NPsH subjects on lexical decision trials in this experiment cannot likely be attributed to a lack of statistical power: in fact, the nonsignificant regression coefficients were 11.01 and 29.37 for the NPsH and PsH groups, respectively, and this actually reverses the pattern observed in Experiment 1.

One possible account of these results is predicated on the following set of assumptions. Subjects in the PsH condition might have actively disabled the phonologic route (because it signals false positive responses to pseudohomophones), and then relied exclusively on the direct route in making lexical decisions. This not only optimized lexical decision performance, it also freed attentional resources for the memory task, and hence performance was enhanced there as well. Subjects in the NPsH condition did not disable the phonologic route, and in the first experiment, they waited for the build-up of phonological information and used it in making lexical decisions (hence the influence of NU). In Experiment 2, on the other hand, they also retained the phonologic route, but under the demanding conditions of the memory load context, they tended to make the lexical decision before the phonological representation was fully generated (hence no influence of NU and only a small unreliable latency disadvantage). Thus, they actually relied on the direct route to read out the lexical decision response. Nonetheless, because the phonologic route was not actively suppressed by these NPsH subjects, it continued to operate, and for words with many NUs it was particularly resource demanding. This unsuppressed activity might then impair the subsequent memory judgment performance; thus, as NU increased accuracy on the probe task actually decreased. This speculation is ultimately grounded in the view that phonologic information builds up more slowly than direct access (but see Stone and Van Orden, 1993, for a contrasting view), and that even subjects in the NPsH condition can make a word–nonword decision prior to completion of the phonologic process. However, pseudohomophones in the nonword context drive subjects not simply to make a decision prior to the completion of phonological processing but, instead, to actively suppress it. Obviously, such an account is speculative and is contingent on the view that strategic control operates at several points early in processing. Nevertheless, it seems to provide the only account that can handle the results of the two experiments. The possibility that subjects can either ignore phonological information or disable it should be investigated further. Still, it remains the case that even in this experiment when NU had an influence on performance (in the probe task), it was only for subjects in the (slower) NPsH group; once again, the (faster) PsH subjects showed no sensitivity to this variable.

Experiment 3

The results of Experiments 1 and 2 are consistent with an account of lexical access that allows for strategic control over the extent to which phonological coding mediates lexical access. It is clear that there is a general performance advantage for subjects in the PsH condition, coupled with the failure to observe any evidence that these subjects were sensitive to the phonological inconsistency of its target's neighborhood (i.e., NU). However, the argument that the failure to observe effects of NU indicates the absence of phonological coding obviously hinges on the validity of NU as a diagnostic criterion. Although there is evidence from Experiments 1 and 2 that NU effects are, in fact, useful, they are somewhat variable in magnitude (and, therefore, in reliability) between experiments. Thus, it is important to supplement the evidence provided by NU in order to provide converging evidence that the PsH and NPsH groups differ in their degree of phonological coding. To this end, we decided to seek converging evidence of phonological processing flexibility in a double lexical decision study.

The basic task involves presenting the subject with two letter strings (one above the other), with the subject responding positively if both are words and negatively if one or both are not. The relations between the two words in a given pair can be varied to measure orthographic, phonological, or semantic processing. In an initial investigation Meyer and his colleagues (Meyer, Schvaneveldt, & Ruddy, 1974) examined performance on words that were either orthographically and phonologically similar (e.g., BRIBE–TRIBE, LOOK–BOOK) or orthographically similar but phonologically dissimilar (e.g., COUCH–TOUCH, LEMON–DEMON). Relative to control conditions consisting of the same words in unrelated pairings (e.g., BRIBE–BOOK, TOUCH–DEMON) they found a small facilitatory effect for BRIBE–TRIBE pairs, and an inhibitory effect for COUCH–TOUCH pairs. They interpreted this result as evidence that subjects were using phonological codes, whereby the consistency of the pronunciations of the rime in BRIBE–TRIBE-type pairs produced facilitatory priming, and the inconsistency of the two rime pronunciations in COUCH–TOUCH-type pairs produced inhibitory priming. Shulman and his colleagues (Shulman et al., 1978) replicated this finding when the nonwords were orthographically legal.

However, when illegal nonwords were used (either consonant strings or random letter strings that violated orthotactic rules) they found facilitatory effects for both BRIBE–TRIBE and COUCH–TOUCH pairs. It is possible that this result was obtained because subjects were making decisions at a prelexical level on the basis of orthographic familiarity, and so they included an associatively related condition (OCEAN–WATER). Because they observed associative facilitation in both conditions, they suggested that subjects were activating lexical representations. They interpreted their results as suggesting that subjects in the illegal nonword condition were getting to lexicon without mandatory phonological coding. However, they did not compare the magnitude of semantic priming in the two conditions, because the nonword manipulation was across experiments. Hanson and Fowler (1987) replicated the Shulman et al. (1978) finding that facilitatory priming with illegal nonwords was obtained with COUCH–TOUCH pairs, and this held for both hearing and deaf readers. However, they did not include OCEAN–WATER-type pairs in their study.

Recently, Van Orden and his colleagues (Van Orden et al., 1990) noted that whereas facilitation for COUCH–TOUCH stimuli could be seen as one of the few positive findings supporting the existence of a direct route, it is subject to an alternative interpretation. They argued that when the nonwords are illegal, subjects can rely on partial phonological information—"noisy" phonological codes—to recognize words, and because COUCH and TOUCH have a good deal of phonological overlap, they can still partially prime each other phonologically. When nonwords are legal and, therefore, the discrimination is more demanding, noisy coding is too error-prone and subjects rely on a "cleaned-up" phonological code; here the COUCH–TOUCH inhibition is found. In other words, they suggest that the differences are due to quantitative and not qualitative differences in processing.

The following experiment provides a test of Van Orden et al.'s (1990) hypothesis. Instead of using illegal nonwords to induce a shift away from phonological processing (as in previous experiments using this paradigm), a pseudohomophone manipulation was once again used (as in Experiments 1 and 2). All subjects received nonwords that were legal (thus, presumably difficult); the intention was for subjects in both groups to rely on cleaned up phonological codes and not to rely on orthographic representations. By this account, so far, one might predict COUCH–TOUCH inhibition in both the NPsH and PsH conditions. However, if subjects in the PsH condition disable or weaken the phonologic route (in spite of the difficulty of the nonwords), then COUCH–TOUCH orthographic facilitation should be obtained because there will be no basis for any phonological competition that would result in inhibition. In contrast, subjects in the NPsH group (who are presumably more dependent on phonological coding) should still show COUCH–TOUCH inhibition. Furthermore, if subjects in the PsH condition have been gaining a speed and accuracy advantage in Experiments 1 and 2 by somehow making lexical decisions without really achieving lexical access (for example, by means of an orthographic word familiarity check), they might show COUCH–TOUCH orthographic facilitation, but they would not be expected to show OCEAN–WATER semantic facilitation. Thus, the current experiment can converge with Experiments 1 and 2 to show that PsH subjects attenuated their phonological processing. It also provides a test of Van Orden's noisy-code account of the Shulman et al. (1978) data, and it tests for the possibility that subjects in the PsH condition are engaging in only shallow (nonsemantic) processing. Furthermore, in the first two experiments a number of the nonwords were orthographically unusual (small neighborhood items). In this experiment nonwords are, on average, more orthographically familiar patterns.

Method

Subjects

Forty-six undergraduate students at the College of the Holy Cross participated in the experiment for partial fulfillment of a course requirement.

Stimuli

Subjects received 96 word–word pairs (positive trials) and 90 word–nonword and 6 nonword–nonword pairs (96 negative trials). Six types of word–word pairs were prepared. Type 1 pairs were orthographically and phonologically similar (BRIBE–TRIBE, LOAD–TOAD). Type 2 pairs were controls that were orthographically and phonologically dissimilar. The control pairs were generated by pairing dissimilar Type 1 items (e.g., BRIBE–TOAD, LOAD–TRIBE). Type 3 pairs were orthographically similar and phonologically dissimilar pairs (COUCH–TOUCH, GONE–BONE). Type 4 pairs were controls for the Type 3 pairs, constructed in the same way as the Type 2 controls. These experimental pairs (Types 1 and 3) were the same as those used by Meyer et al. (1974) and Hanson and Fowler (1987), and Type 1 and Type 3 pairs were matched as closely as possible for length and frequency (see Meyer et al. for details). Type 5 pairs were semantically related pairs (OCEAN–WATER) chosen from the norms of Battig and Montague (1969), and Type 6 pairs were controls for the Type 5 pairs, again generated by rearranging the Type 5 pairs. Type 5 pairs were chosen from among the top five exemplars of each category in the norms; this was done to insure that each member of a related pair was a good category exemplar. However, this constraint did not allow for a matching of these pairs with Type 1 or Type 3 pairs on dimensions such as length and frequency. Thirty-two pairs of each word–word pair type were prepared (all stimulus pairs are presented in Appendix B).

Two stimulus lists were constructed. List A consisted of 16 Type 1 pairs, 16 Type 3 pairs, and 16 Type 5 pairs, with the words from the remaining 16 pairs of each type rearranged to serve as controls (Types 2, 4, 6); in List B the situation was reversed. Thirty-two of the 96 negative trials consisted of orthographically similar items (e.g., LOOK–DOOK); this matched the number of positive pairs that were orthographically similar (Types 1 and 3). Thus, orthographic similarity was not correlated with whether the pair was a positive or negative response type. Half of the word–nonword pairs were presented with the word as the upper display item, and half with the word as the lower display item. Subjects in the NPsH condition received either List A or List B as they are described above. Subjects in the PsH condition received one of these lists with a pseudohomophone substituted for the nonword item in 30% of its word–nonword pairs. Both NPsH and PsH subjects received the same 32 orthographically similar negative trials.

Procedure

The procedure was the same as in Experiment 1, except stimulus pairs instead of single letter strings were presented (they appeared one above the other in the center of the screen), and subjects had 1,500 ms to respond to the items. As in the other experiments, subjects received a practice list of 40 trials before the experimental trials. Each subject's participation lasted approximately 25 min.

Results

For each subject, mean latencies were calculated for the six types of word pairs and for correct responses to the two types of nonword pairs. Within each of these categories, trials with latencies greater than two standard deviations from the subject's own mean (calculated independently for each category) were treated as errors. The data for the subjects and items analyses were based on these data. The data from three subjects who made more than 30% errors in at least two response categories were excluded from further analyses.

Word Analyses

The primary analyses were conducted on the matched Type 1 and Type 3 pairs and their respective controls (Type 3 and Type 4). Type 5 and Type 6 pairs were not matched with the first four pair types (see stimuli section), and were examined separately, to determine whether there were group differences in the magnitude of semantic priming. Analyses of variance were conducted on the latency and accuracy data using both subjects (F_1) and items (F_2) as random factors. For the subjects analyses, mean latencies and proportions correct were computed for each subject for each of the six pair types. In the items analyses, mean latencies were computed for each of the experimental words, averaged over subjects, and accuracy was calculated as the proportion of subjects responding correctly to each item. In the subjects analyses, pair type was a within-subjects factor and group (NPsH vs. PsH) was a between-subjects factor. These designations were reversed in the items analyses. List (A vs. B) served as an additional control variable in these analyses.

The latency analysis on the data from the first four pair types yielded a significant main effect of pair type, $F_1(3, 117) = 19.30$, $p < .001$, $MS_e = 6{,}349.48$, and $F_2(3, 120) = 8.98$, $p < .001$, $MS_e = 23{,}811.03$, and a significant Pair Type × Group interaction, $F_1(3, 117) = 2.93$, $p < .05$, $MS_e = 6{,}349.48$, and $F_2(3, 120) = 3.98$, $p < .01$, $MS_e = 6{,}731.47$. Separate analyses on the two pair types and their respective controls indicated no significant Group × Pair Type (experimental vs. control) interaction in the orthographically and phonologically similar condition. However, as expected, the interaction was significant in the orthographically similar but phonologically dissimilar condition, $F_1(1, 39) = 5.42$, $p < .05$, $MS_e = 7{,}808.70$, and $F_2(1, 60) = 6.04$, $p < .025$, $MS_e = 9{,}231.01$. Subject means in the orthographically and phonologically similar condition were as follows: NPsH experimental pair type = 798 ms, NPsH control pair type =903 ms, PsH experimental pair type = 804 ms, and PsH control pair type = 912 ms. Subject means in the orthographically similar but phonologically dissimilar condition were as follows: NPsH experimental pair type = 945, NPsH control pair type = 895, PsH experimental pair type = 875, and PsH control pair type = 912.

Figure 2 shows the differences in response latency between the experimental pair types and their respective controls. Positive numbers indicate that the experimental pairs were faster than their controls (facilitatory effects), and negative numbers indicate that experimental pairs were slower than controls (inhibitory effects). Subjects in both groups showed facilitation to orthographically and phonologically similar pairs. However, for orthographically similar but phonologically dissimilar pairs NPsH subjects showed inhibitory effects, whereas PsH subjects were facilitated on these pairs relative to the control condition (hence the two-way interaction noted above). There was also a significant List × Pair Type interaction, $F_1(3, 117) = 5.38$, $p < .01$, $MS_e = 634.48$, and $F_2(3, 120) = 3.24$, $p < .05$, $MS_e = 23{,}811.03$. The cell means indicated that facilitatory effects for Type 1 pairs were larger for List B than List A, and that inhibitory effects on Type 3 pairs were smaller for List B than List A. However, the three-way List × Pair Type × Group interaction was not significant in either the subject or item analyses; the critical Group × Pair Type interaction was not qualified by list. Finally, the 10-ms latency advantage for PsH subjects was not significant in either analysis.

The latency analysis on the data from the semantically related pairs (Type 5) and their controls (Type 6) revealed a main effect of pair type, $F_1(1, 39) = 61.11$, $p < .001$, $MS_e = 3{,}156.72$, and $F_2(1, 60) = 35.10$, $p < .001$, $MS_e =$

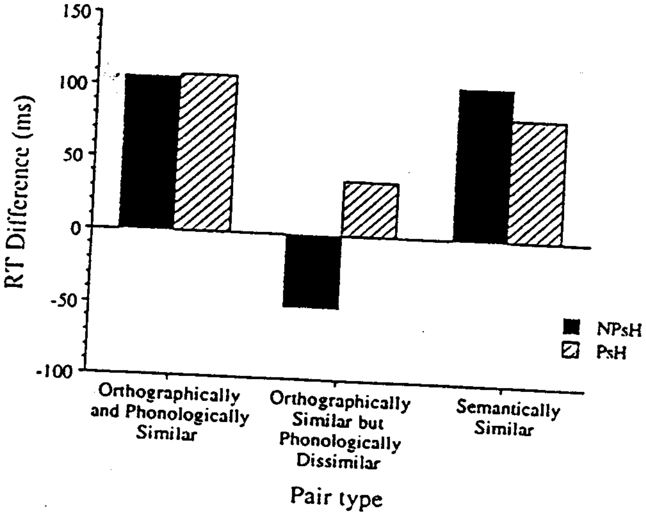

Figure 2. The difference between experimental and control response latencies in the double lexical decision task as a function of Group and pair type. Reaction time (RT) difference is the control group mean minus the experimental group mean. (NPsH = nonpseudohomophone group; PsH = pseudohomophone group.)

8,349.30. The semantically related pairs were responded to 94 ms faster than the control items (related mean = 718 ms, unrelated mean = 812 ms). Of critical interest, however, is that this variable did not interact with group in either the subject or item analyses ($ps > .25$). Figure 2 presents the differences between the experimental and the control latencies for the two groups (105 ms and 85 ms for the NPsH and PsH groups, respectively). Thus, the magnitude of semantic priming effects was quite large for both groups.

The analyses conducted on the accuracy data showed no significant effects of group, pair type, or their interaction.

Nonword Analyses

Analyses of the nonword data included the following variables: group and similarity (orthographically similar vs. dissimilar pairs). An effect of group on latencies was significant in the item analysis, $F_2(1, 66) = 5.98$, $p < .05$, $MS_e = 2{,}487.51$, but not in the subject analysis ($F < 1.0$). Subjects in the NPsH condition were faster on correct rejections than PsH subjects (956 ms and 976 ms, respectively).

Discussion

The results of this experiment are quite clear with regard to the question of whether phonological processing differences between NPsH and PsH subjects are present. Subjects in the NPsH condition showed inhibitory effects when pair members were orthographically similar but phonologically dissimilar (COUCH–TOUCH) along with facilitatory effects when pair members were similar on both dimensions (BRIBE–TRIBE). In contrast, PsH subjects showed facilitatory effects of both types of pairs. The hypothesis that subjects in the PsH condition curtail phonological processing is strongly supported by these data.

Semantic association facilitated the responding of subjects in both groups. Note that whereas the magnitude of the effect was slightly greater in the NPsH condition, the difference was not reliable. Subjects in the PsH group seem to have been able to get to lexicon without phonological representations of the target words. This was the same conclusion supported by Experiments 1 and 2, neither of which showed any influence of the number of the target word's phonologically unfriendly neighbors in this PsH condition.

It should be noted that whereas this experiment revealed clear differences in phonological processing between the PsH and NPsH groups, the latency advantage for PsH subjects obtained in four other experiments (Andrews, 1982; our pilot; and Experiments 1 and 2) was not reliable in the current experiment. Perhaps when the judgment involves two letter strings, the additional cognitive processing obscures the latency advantage that might be obtained due to attenuating or curtailing phonological processing. Furthermore, responses to nonwords were actually faster in the NPsH condition, although this was reliable only in the item analysis (subjects, $F < 1.0$). This reverses the latency advantages in the PsH condition obtained in the other experiments on nonword trials. However, this weak effect favoring the NPsH subjects on nonword latencies should not be taken as evidence that somehow PsH subjects engaged in more phonological processing. If this had been so then the COUCH–TOUCH trial inhibition should have been observed for these subjects as well. Furthermore, PsH subjects were not slower in rejecting pseudohomophones than regular nonwords (see the General Discussion section for a full discussion of nonword pseudohomophone effects across experiments). The cognitive differences between the single and double lexical decision merit further exploration.

The facilitatory effect of PsH subjects on COUCH–TOUCH pairs is not consistent with claims that the strategic effects that we have been documenting in this study are trivially postlexical in origin. If, for instance, phonological lexical access were mandatory, as several researchers have suggested (Lukatela & Turvey, 1990, 1993; Van Orden et al., 1990), and subjects in the NPsH group were simply engaging in an extra postlexical phonological check, which the PsH subjects suppressed, then COUCH–TOUCH facilitation for PsH subjects would not be expected. Instead, by their account some obligatory inhibition on these pairs would be predicted, albeit of possibly smaller magnitude than in the NPsH condition. The observed facilitation suggests lexical access that is orthographically based, which is consistent with dual-route theory. Furthermore, the claim that the PsH subjects in Experiments 1 and 2 were engaged in a kind of orthographic word familiarity judgment in lieu of actual lexical activation is not consistent with the large semantic priming effects observed in this experiment.

Following the hypothesis that PsH subjects disabled the phonologic route, BRIBE–TRIBE facilitation should be no different from COUCH–TOUCH facilitation for these subjects. However, that was not the case: for subjects in the PsH condition the magnitude of the facilitation for BRIBE–TRIBE pairs was more than twice as large as for COUCH–TOUCH pairs. A possible explanation for this difference is that most, but not all, subjects in the PsH condition showed facilitation on COUCH–TOUCH pairs, whereas nearly all showed BRIBE–TRIBE facilitation. Apparently, not all subjects responded to the PsH manipulation by disabling the phonologic route, although a significant proportion seem to have done so. These individual differences in response to contextual manipulations should be examined in subsequent investigations of strategic flexibility (see also Hanson & Fowler, 1987).

General Discussion

Our purpose was to demonstrate coding flexibility in lexical access. The three experiments varied the composition of the nonwords in a lexical decision task; in each experiment one group of subjects received pseudohomophones among its nonwords and the other group did not. The intention was to make dependence on phonological assembly counterproductive for the PsH group because the phonological realization of pseudohomophones falsely represents them as real words: this should lead to greater reliance on orthographic coding if this is, in fact, possible. In the first experiment subjects in the PsH condition performed faster and no less accurately than subjects in the

NPsH condition on both word and nonword trials, suggesting that the presence of pseudohomophones had the predicted effect. Moreover, the performance of PsH subjects on word trials was uninfluenced by the phonological inconsistency of a target's orthographic neighborhood, whereas the latency and accuracy of NPsH subjects' responses were inhibited by neighborhood phonological inconsistency. Together, these results suggest that subjects in the PsH group did not depend on assembled phonology to access lexicon (at least not to the extent of subjects in the NPsH group) but that they were, nevertheless, more efficient than the NPsH group.

In Experiment 2 we found similar latency and accuracy advantages for PsH subjects on the lexical decision and memory probe components of a dual-task procedure. However, unlike the outcome of Experiment 1, neighborhood phonological inconsistency did not affect lexical decisions in the NPsH group. Instead, inconsistency influenced performance on the memory probe that followed lexical decision for these NPsH subjects, with no corresponding influence on PsH subjects. Furthermore, subjects in the NPsH condition were more adversely affected on probe performance by high memory load than were PsH subjects. These effects can be interpreted as suggesting that processing for the NPsH subjects not only involved phonological coding but was also more attentionally demanding. The Group × Word Length interaction on lexical decision latency suggested differences in the type of processing and not merely differences in efficacy or efficiency.

In Experiment 3 we used a double lexical decision paradigm to examine the influence of this phonological coding flexibility on phonological consistency effects. Whereas subjects in both groups showed facilitation on phonologically consistent pairs (e.g., BRIBE–TRIBE), NPsH subjects showed inhibition on phonologically inconsistent pairs (e.g., COUCH–TOUCH). PsH subjects, however, were facilitated on these pairs, suggesting once again that they were not relying on phonological coding. Furthermore, facilitation on semantically related pairs (e.g., OCEAN–WATER) was equivalent in both conditions, suggesting that subjects in both groups were, in fact, activating lexical entries. Dependence on direct access does not appear to diminish activation of semantic information.

The clear implication from these studies is that, in the presence of pseudohomophones, a substantial proportion of subjects will process words so as to minimize phonological influences. The precise locus of this flexibility is unclear, but there are several reasons to suppose it is not trivially postlexical. If it is not, then this poses a problem for single-route theories in general, which would seem compelled to place coding flexibility—evidence for two kinds of processes—at some postlexical cognitive stage. As an often-proposed example of a postlexical mechanism, consider confirmatory postlexical phonological checking, performed after a word has already been selected in lexicon but before the response is made. The check uses a phonological representation of the target; if the representation does, indeed, "sound" identical to a word in the subject's speech lexicon, the original printed stimulus is confirmed to be a word. There are two possible sources for such a phonological representation: prelexical and lexical. It seems implausible that the former would ever be used when the latter is available: prelexical (i.e., assembled) phonology is typically incomplete—syllable stress for multisyllabic words is not indicated in the print and, therefore, cannot be present in the prelexical representation. Yet syllable stress is critical for the identification of spoken words. On the other hand, once a lexical entry has been activated (whether by assembled or direct processes), its complete phonological representation, including stress, is available. However, after lexical access, both conditions are identical with regard to access of lexical phonology. Thus, this assessment predicts (contrary to fact) that there will be no difference on words between conditions as a function of their postlexical processing, because processing should be identical for both PsH and NPsH at that point.

Nevertheless, there will be no such lexical phonology for pseudowords. Here, pseudohomophones will prove to be problematical for the PsH subjects. Suppose, therefore, that they completely eliminate the postlexical check. Because the NPsH subjects would not suppress the postlexical test, we would appear to have a possible explanation of the speed advantage of the PsH condition; the PsH subjects perform one less operation than the NPsH subjects (although we might expect the latency differences to be somewhat larger than they actually were). However, we should also see an elevated error rate for PsH subjects, because they are eliminating the check. In fact, the opposite result obtained; PsH subjects were slightly (though nonsignificantly) more accurate even while performing faster. There are other difficulties encountered by an explanation that is based solely on postlexical checking differences. If initial lexical access for both groups was phonologically mediated as suggested by several researchers (Lukatela & Turvey, 1993; Van Orden et al., 1990), then some mandatory inhibition on phonologically dissimilar pairs in the third experiment should have been seen in both groups, and that was not the case. If, on the other hand, prelexical processing was primarily orthographic for both groups, and phonological influences only occurred at a later stage, then it seems unlikely that NPsH subjects would show increased sensitivity to phonological neighborhood inconsistency on the subsequent memory probe judgment and not on the initial lexical decision in the second experiment. Certainly, however, additional experiments are needed to examine the precise mechanisms involved in coding flexibility. At present, however, it would appear that the most plausible account of the results of these three experiments is one that emphasizes context-induced differences at the level of lexical access.

Several findings suggest that subjects in the PsH condition disabled the assembled phonological route, and not that they simply read out responses prior to the completion of the phonological processing. First, as noted above, Experiment 2 probe-task performance was influenced by phonological neighborhood inconsistency for NPsH subjects, suggesting a spillover effect; if PsH subjects had merely been reading out quick responses and had not been disabling phonological processing, a similar spillover would still have been predicted in that condition. However, probe task performance was uninfluenced by NU for this group. Second, quick responses that are orthographically based should not be possible for nonword trials, and consequently nonword performance should not have indicated group differences. In two of the three experiments, PsH subjects were faster on nonword responses than were NPsH subjects. To examine this issue further, we compared performance on regular nonwords and pseudohomophones for the PsH subjects. Note that these two sets of items were not specifically equated on any dimensions. However, one dimension that they do differ on is the phonological dimension, and if subjects in this condition were processing nonwords in a phonologically sensitive fashion, then PsH rejection latencies should have been somewhat slower than regular nonwords. In all three experiments rejection latencies were actually somewhat faster for PsHs. The means for the regular nonwords and pseudohomophones were 569 ms vs. 566 ms in Experiment 1, 674 ms vs. 654 ms in Experiment 2, and 977 ms vs. 965 ms in Experiment 3. Thus, there was no hint of the standard PsH effect in any of these experiments. Such an outcome strongly implies that subjects in the PsH group were operating in a nonphonological mode even on nonword trials. There seems to be every indication in these data that PsH subjects performed the lexical decision task in a fundamentally different way than NPsH subjects.

It should be noted that the interpretations being considered here are grounded in the idea that subjects are, in a strategic sense, altering the word recognition process. However, selective inclusion or elimination of pathways in lexical processing is not the only possible means of strategic control over performance. In considering the results of their pseudohomophone context manipulation (see above), Stone and Van Orden (1993) contrast pathway selection accounts with accounts that are based on flexible criterion setting. They find both accounts lacking in some ways but are more inclined toward the latter approach. A criterion-setting account of the results of the experiments reported here does not appear to be very plausible. First, if subjects in the NPsH condition had simply set higher word and nonword response thresholds, resulting in longer latencies on word and nonword responses (Experiments 1 and 2), and this somehow amplified phonological influences, then correspondingly higher accuracy rates should have been observed in that condition. As noted, accuracy was slightly greater in the PsH group. Second, in Experiment 3 differential phonological sensitivity on COUCH–TOUCH trials was observed, without any corresponding group differences in latency or accuracy. Again, the notion of differential use of phonological coding seems most plausible.

Given that it took the "unnatural" presence of PsHs to force subjects to adopt an apparently nonphonological mode of processing, it might be tempting to conclude that NPsH subject performance is more representative of how normal word recognition. However, it is entirely possible that in a lexical decision task, in which half of the letter strings have no lexical representation, subjects occasionally adopt an inordinately phonological strategy. Given the plausibility of both arguments, it remains for experiments using more naturalistic reading tasks than lexical decision to resolve the issue of what constitutes normal phonological involvement for skilled readers (cf. Pollatsek, Lesch, Morris, & Rayner, 1992). We suggest, however, that the very existence of flexibility suggests that both phonological and direct processing are required in everyday reading, and that coding flexibility is, therefore, highly practiced. If not, why would the flexibility that we have demonstrated occur at all? If subjects always used only one strategy (at least since they became skilled readers), why should they be able to switch with apparent efficiency to another strategy (even if that switch is only partial, as from a single coding strategy to a mixed strategy)? In this regard, it is worth mentioning that our subjects may have had very little insight into the effect of the manipulation on their reading. Subjects appear to be both exquisitely sensitive to the pseudohomophone manipulation and unaware of what effect the pseudohomophones have on their process of word recognition. Anecdotally, we can report that when we queried a number of PsH subjects after the experiment, several were unaware that any of the nonwords were pseudohomophones. Those who were aware of the pseudohomophones claimed that they were forced to "slow down and be more careful." As we have seen, the opposite was the case: responses were faster in the PsH condition.

In conclusion, the current experiments suggest that subjects can control the extent to which they engage phonological processing in making lexical decisions. Furthermore, in conditions where they apparently disable or attenuate such processing, lexical decision and concurrent performance that is based on short-term memory are enhanced. That this adjustment does not come at the expense of lexical access is suggested by the facilitatory semantic priming evident for all subjects in Experiment 3. This set of results converges with prior studies in suggesting subject flexibility with regard to the use of assembled phonological processing. The results pose a serious challenge to single-route accounts in general. Most important, these data speak of remarkably fined-tuned strategic adjustments in performance and suggest caution in interpreting lexical decision results without first carefully examining the specific experimental context.

References

Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent? *Memory and Cognition, 10,* 565–575.

Battig, W., & Montague, W. (1969). Category norms for verbal items in 56 categories: A replication and extension of the University of Connecticut norms. *Journal of Experimental Psychology: Monograph, 80,* 1–46.

Besner, D., & Smith, M. (1992). Models of visual word recognition: When obscuring the stimulus yields a clearer view. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 468–482.

Carello, C., Turvey, M. T., & Lukatela, G. (1992). Can theories of word recognition remain stubbornly nonphonological? In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology and meaning* (pp. 221–226). Amsterdam: North-Holland.

Carr, T. H., & Pollatsek, A. (1985). Recognizing printed words: A look at current models. In D. Besner, T. Waller, & E. MacKinnon (Eds.), *Reading research: Advances in theory and practice* (Vol. 5, pp. 1–82). San Diego, CA: Academic Press.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151–216). San Diego, CA: Academic Press.

Coltheart, M., Davelaar, E., Jonasson, J., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance* (Vol. 6, pp. 535–555). Hillsdale, NJ: Earlbaum.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671–684.

Davelaar, E., Coltheart, M., Besner, D., & Jonasson, J. T. (1978). Phonological recoding and lexical access. *Memory and Cognition, 6,* 391–402.

Glushko, R. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 674–691.

Hanson, V. L., & Fowler, C. A. (1987). Phonological coding in word reading: Evidence from hearing and deaf readers. *Memory and Cognition, 15,* 199–207.

Hawkins, H. L., Reicher, G., Rogers, M., & Peterson, L. (1976). Flexible coding in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 235–242.

Humphreys, G. W., & Evett, L. J. (1985). Are there independent lexical and nonlexical routes in word processing? An evaluation of the dual-route theory of reading. *The Behavioral and Brain Sciences, 8,* 689–740.

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance, 1,* 130–136.

Jared, J., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in naming. *Journal of Memory and Language, 29,* 687–715.

Jared, J., & Seidenberg, M. S. (1991). Does word identification proceed from spelling to sound to meaning? *Journal of Experimental Psychology: General, 120,* 358–394.

Johnson, N. F., & Pugh, K. R. (in press). An examination of cohort effects in visual word recognition. *Cognitive Psychology, 26.*

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology and meaning* (pp. 67–84). Amsterdam: North-Holland.

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

Lukatela, G., & Turvey, M. T. (1990). Phonemic similarity effects and prelexical phonology. *Memory and Cognition, 18,* 128–152.

Lukatela, G., & Turvey, M. T. (1991). Phonological access of the lexicon: Evidence from associative priming with pseudohomophones. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 951–966.

Lukatela, G., & Turvey, M. T. (1993). Similar attentional, frequency and associative effects for pseudohomophones and words. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 166–178.

McQuade, D. V. (1981). Variable reliance on phonological information in visual word recognition. *Language and Speech, 24,* 99–109.

McQuade, D. V. (1983). Phonological recoding becomes automatic with stimulus repetition. *Language and Speech, 26,* 393–405.

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1974). Functions of graphemic and phonemic codes in visual word recognition. *Memory and Cognition, 2,* 309–321.

Monsell, S., Patterson, K., Graham, A., Hughes, C. H., & Milroy, R. (1992). Lexical and sublexical translations of spelling to sound: Strategic anticipation of lexical status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 452–467.

Paap, K. R., & Noel, R. W. (1991). Dual-route models of print to sound: Still a good horse race. *Psychological Research, 53,* 13–24.

Perfetti, C. A., & Bell, L. C. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and priming. *Journal of Memory and Language, 30,* 473–485.

Perfetti, C. A., Bell, L. C., & Delaney, S. M. (1988). Automatic (prelexical) phonetic activation in silent word reading: Evidence from backward masking. *Journal of Memory and Language, 27,* 59–70.

Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 148–162.

Posner, M., & Boies, S. (1971). Components of attention. *Psychological Review, 78,* 391–408.

Pugh, K. R., Rexer, K., Peter, M., & Katz, L. (1994). Neighborhood effects in visual word recognition: Effects of letter delay and nonword context difficulty. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 639–648.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading.* Englewood Cliffs, NJ: Prentice Hall.

Rosson, M. B. (1985). The interaction of pronunciation rules and lexical representations in reading aloud. *Memory and Cognition, 13,* 90–98.

Seidenberg, M. S. (1992). Beyond orthographic depth in reading: Equitable division of labor. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology and meaning* (pp. 85–118). Amsterdam: North-Holland.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 523–568.

Seidenberg, M., Waters, G., Barnes, M. A., & Tannenhaus, M. K. (1984). When does irregular spelling and pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior, 23,* 383–404.

Shulman, H. G., Hornak, R., & Sanders, E. (1978). The effect of graphemic, phonetic, and semantic relationships on access to lexical structures. *Memory and Cognition, 6,* 115–123.

Stone, G., & Van Orden, G. C. (1993). Strategic control of processing in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 744–774.

Van Orden, G. C. (1987). A ROWS is ROSE: Spelling, sound, and reading. *Memory and Cognition, 15,* 181–198.

Van Orden, G. C., Johnston, J. C., & Hale, B. L. (1988). Word identification in reading proceeds from spelling to sound to meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 371–386.

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review, 97,* 488–522.

Waters, G., & Seidenberg, M. S. (1985). Spelling–sound effects in reading: Time course and decision criteria. *Memory and Cognition, 13,* 557–572.

Waters, G., Seidenberg, M. S., & Bruck, M. (1984). Children's and adults' use of spelling–sound information in three reading tasks. *Memory and Cognition, 12,* 293–305.

(Appendix follows on next page)