Journal of Experimental Psychology:
Human Perception and Performance
1994, Vol. 20, No. 3, 537-554

0096-1523/94/$3.00

Interacting Sources of Information in Word Naming: A Study of Individual Differences

Patrick Brown, Stephen J. Lupker, and Lucia Colombo

Hypothetically, words can be named by spelling–sound translation rules or by looking up a phonological code in a lexicon. Following J. Baron and C. Strawson (1976), naming performance was measured as a function of skill with each route, using sets of stimuli varying in reliance on either route. "Phoenicians" were defined to be better with rules than with look-up; "Chinese" were better at look-up than with rules. As predicted by J. Baron and C. Strawson, Phoenicians named low-frequency regular words and nonwords faster than Chinese. Contrary to predictions, Phoenicians were also faster at naming irregular words of various frequencies. Implications of these results for various dual-route models versus single-route models are discussed.

The dual-route model of word naming (Coltheart, 1978) is still accepted in some form by most theorists, although there have been challenges (Glushko, 1979; Humphreys & Evett, 1985; Seidenberg & McClelland, 1989; Van Orden, Pennington, & Stone, 1990). According to the standard version of the model, words can be named in either of two ways. One, referred to as the assembly route, is based on the idea that readers use knowledge of common spelling-to-sound relationships to derive a likely set of sounds that are then assembled into a complete phonological code. The second, referred to as the lexical route, is based on use of word-specific association mechanisms. Readers identify the word as a pattern that allows them to access the appropriate location in lexical memory and then retrieve the corresponding phonological code.

There has been considerable debate about the nature of the assembly route (Brown & Besner, 1987; Carr & Pollatsek, 1985; Paap & Noel, 1991; Patterson & Coltheart, 1987; Seidenberg & McClelland, 1989; Van Orden, Johnston, & Hale, 1988). The general assumption is that this route allows the naming of nonwords and contributes to the naming of words that follow common spelling-to-sound rules. The lexical route, on the other hand, is assumed to be required when the words do not follow these rules (i.e., irregular words such as *have*).

Two major pieces of evidence from normal subjects support this model. The first, which we shall refer to as the "lexicality effect," is that low-frequency words that have regular spelling-to-sound correspondences are named faster than nonwords, regardless of how well the orthographic and phonological characteristics of the two kinds of stimuli are matched (Glushko, 1979; McCann & Besner, 1987; Stanhope & Parkin, 1987). The explanation is based on the idea that phonological codes for nonwords can only come from one source (the assembly route), whereas phonological codes for words can come from either route. Thus, to the extent that phonological codes provided by the lexical route tend to become available faster than those provided by the assembly route, words will be named faster than nonwords (Derouesné & Beauvois, 1985; McCann & Besner, 1987).

The second major piece of support for this model is the regularity effect: Readers are slower to name irregular words than regular words (Baron & Strawson, 1976; Glushko, 1979; Paap & Noel, 1991; Seidenberg, Waters, Barnes, & Tanenhaus, 1984; Stanovich & Bauer, 1978). This effect is taken to indicate that both routes are engaged by each word. For irregular words, the outputs of the two routes will tend not to match, creating interference and prolonging naming latency. More recently, Seidenberg et al. have also noted that the regularity effect obtains only with low-frequency words. Their account is that for high-frequency words, the lexical route is fast enough that a naming response can be made before a mismatch is registered.

Seidenberg (1985c) went on to argue that a similar logic applies to the subjects who are the fastest with the lexical route. That is, because the lexical route is so fast for those individuals, they should show smaller regularity effects even with low-frequency words. (In other words, fast subjects are fast because many ostensibly low-frequency words are nonetheless quite familiar to them, allowing these words to be named quickly through the lexical route.) In fact, a tertile split of the subjects from Seidenberg et al. on the

Patrick Brown and Stephen J. Lupker, Department of Psychology, University of Western Ontario, London, Ontario, Canada; Lucia Colombo, Dipartimento di Psicologia Generale, Universita' degli Studi di Padova, Padova, Italy.

This research was supported by Natural Sciences and Engineering Research Council of Canada Grants A38873 to Patrick Brown and A6333 to Stephen J. Lupker and by grants from the Centro Nazionale Ricerche in 1989 and the Human Frontiers Science Program Organization to Lucia Colombo. Major portions of this article were presented at the Second Annual Meeting of the Canadian Society for Brain, Behaviour, and Cognitive Science, Quebec City, June 1992, and the Third Meeting of the European Society for Cognitive Psychology, Paris, September 1992.

We are grateful to Carolynn Racicot, Susan McCorquodale, and Mel Mottram for testing the subjects in this study.

Correspondence concerning this article should be addressed to Patrick Brown or Stephen J. Lupker, Department of Psychology, University of Western Ontario, London, Ontario, Canada N6A 5C2.

basis of overall reaction time (RT) shows exactly this pattern. The regularity effect for low-frequency words is smallest for the subjects with the shortest overall RT and largest for the subjects with the longest overall RT. These results seem to argue quite strongly that the size of any regularity effect is crucially dependent on the time relationship between the two routes.

Seidenberg (1985b, 1985c) chose to interpret these results in terms of a somewhat different type of dual-route model, referred to as the time-course model. In this model, the two "routes" do not represent separate, independent processes but derive from a single interactive process that yields both orthographic and phonological information. Orthographic information starts to accumulate first, whereas phonological information, being parasitic on the orthographic information, starts to accumulate slightly later. If the orthographic information accumulates fast enough to allow word-specific activation and retrieval of the word's name, no regularity effect arises. If not, phonological information will enter the process, which will give an advantage to regular words. What is important here, however, is that two processes are being carried out in parallel and the size of the regularity effect is crucially dependent on the time relationship between them. This time relationship is dependent on (among other things) a word's frequency and a subject's ability with the two processes.

Whereas the time-course model is still somewhat in the spirit of the dual-route model, more recent challenges to the dual-route model have come from single-route models. In this type of model, all letter strings are named through the same process of translating orthographic codes into phonological codes that can then be passed to further (possibly, articulatory) processors. Early models of this sort did not adequately explain how subjects name irregular words (Gough, 1972) or how they ultimately synthesize a pronunciation (Glushko, 1979); however, the more recent models may be able to explain naming of all types of letter strings even as they account for the two major pieces of evidence favoring dual-route explanations, the lexicality and regularity effects (Brown & Besner, 1987; Seidenberg & McClelland, 1989; Van Orden et al., 1990; see Besner, Twilley, McCann, & Seergobin, 1990, for a dissenting view on the sufficiency of single-route models).

In these later models, the activation of graphemes and/or sets of graphemes generates candidate phonologies, and the computation of the final phonology is based on interactions involving these candidates. For example, Brown and Besner (1987) have suggested that candidate pronunciations are generated by the operation of "flexible rules." The candidates in a given set share a consonant frame but differ in pronunciation of vowels. Selection from among candidates is a function of lexical status, determined by interrogation of the lexicon using the phonological codes and, potentially, of contextual variables. If the candidates are ordered in terms of an experience variable such as probability of each vowel reading (see Berndt, Reggia, & Mitchum, 1987), the basic regularity effect follows. Nonwords are given a default reading after other candidates are eliminated by the selection process. The lexicality effect, on this view, arises from the same mechanism that produces the regularity effect: order of testing.

In the parallel distributed processing model advanced by Seidenberg and McClelland (1989), the orthographic and phonological representations of words are patterns of activation across orthographic and phonological units. The strength of activation is determined by the weights on the connections between units, where the weights are determined by the amount of experience with each connection. Thus, processing at the orthographic and phonological levels should be sensitive to the frequency with which that processing has been done previously. As such, any frequency effects would be explained by features intrinsic to the model.

What characterizes Van Orden et al.'s (1990) model is its emphasis on the role of phonological information in all aspects of word processing, including the retrieval of semantics. The main reason for this emphasis is that the covariance of orthography and phonology is very consistent in comparison to (for example) the covariance of orthography and semantics. According to Van Orden et al., some set of phonological codes are first derived and then these codes are subjected to a "cleaning-up" mechanism, through which the final complete phonological code is derived. Given the way in which the model is framed, frequency of experience plays an important role in both generation of the initial codes and the cleaning-up process. Most important, both the Seidenberg and McClelland (1989) model and Van Orden et al.'s model can also explain both the regularity and lexicality effects (essentially because of the major role frequency plays in the models), and they can do so without appealing to a second, qualitatively different route for naming.

One line of evidence that can be brought to bear on the single versus interactive versus dual (independent) routes debate comes from studies of individual differences in reading. Particularly relevant to this issue is the argument by Baron and Strawson (1976) that if the dual-route model is correct, it should be possible to classify individuals according to their abilities with each route and then to identify a group of subjects who are good with one route and bad with the other. Those who are good with the lexical route but bad with the assembly route should show the smallest regularity effects, whereas those who are bad with the lexical route but good with the assembly route should show the largest regularity effects. As will be described later, Baron and Strawson's data essentially supported this conceptualization.

In Baron and Strawson's (1976) study, subjects were classified by their scores on two pencil and paper tasks. To measure their ability on the assembly route, subjects were given a pseudohomophone detection task; they were shown a list of 22 nonwords and were asked to select those which, when pronounced according to the normal rules of English, were homophonic with an English word (e.g., *sope*). A simple count of the number of errors was used to indicate a subject's ability with the assembly route. To measure a subject's ability on the lexical route, a spelling-discrimination task was used. The subject began the experiment by spelling 25 often misspelled words, followed by a two-choice, forced-choice task for those same words. If a subject's original spelling of a given word was correct, the foil for that word was the most common misspelling of the word. If the original spelling was incorrect, the original misspelling and the correct spelling were presented. The score a subject received was the difference between the number correct on the forced-choice test and the number correct on the original spelling test. A high score represents a measure of the subject's ability to respond to the visual pattern of the word and, in doing so, correct the original misspellings.

The scores from the two tests were then added together. A low score (defined as 0 or 1) would indicate an individual who was quite good at detecting pseudohomophones (good assembly route) but not very good at responding to visual patterns (bad lexical route). Baron and Strawson (1976) labeled these individuals "Phoenicians." A high score (defined as 9 or more) would indicate a subject who was good at responding to visual patterns but not at detecting pseudohomophones. Baron and Strawson labeled these individuals "Chinese."

The 11 Phoenicians and 8 Chinese were then timed while they named 10-word lists of either regular or irregular words in lowercase, uppercase, or mixed case. Unfortunately, Baron and Strawson's (1976) irregular words were somewhat more frequent than their regular words and, thus, actually tended to be named slightly faster than the regular words in both uppercase and lowercase. Nonetheless, the Chinese showed a significantly stronger tendency than the Phoenicians to name the irregular words faster than the regular words, indicating that they were less affected by the words' irregularity, exactly as predicted.

Baron and Strawson's (1976) ability to predict subject performance in this way not only supports the existence of the lexical and assembly routes (and, hence, the dual-route model), but also suggests that the two routes are independent. Unfortunately, a number of questions arise with respect to their data that should be addressed before accepting their conclusions. At a purely methodological level, the task of naming lists of words rather than individual words introduces extraneous factors into the response process that could interact with the group factor. One problem is that words in a given list might affect one another through either facilitatory or inhibitory priming or through difficulties in selecting the next word to be pronounced, effects that are less likely in individual-word naming (see also Henderson, 1985). A second problem with this technique is that some variance will due to processes of finishing with one word and preparing for the next word. All of the time, and the variance in the time, associated with such between-word processes will be part of the list-reading time.

A second question concerns the lack of control over word frequency. Since the irregular words were higher in frequency than the regular words, frequency was totally confounded with regularity. Thus, the possibility exists that Baron and Strawson (1976) actually were observing a Groups × Frequency interaction rather than a Groups × Regularity interaction.

Finally, and most important, in order to support this categorization of subjects, it is necessary to show that the Phoenicians actually perform better in situations where the assembly route is supposed to be more important. For example, Phoenicians, whose assembly route is supposedly more efficient than their lexical route, should be quite good at naming nonwords and should name most regular words using their assembly route. Thus, they should show small lexicality effects. This point is especially relevant because, as Seidenberg (1985c) pointed out, in Baron and Strawson's (1976) study the Chinese were not only less affected by irregular words, they also had overall shorter naming latencies. The time-course model explicitly predicts that faster subjects will have smaller regularity effects. Similarly, all the single-route models mentioned previously predict a quantitative difference among subjects in their ability to use the single route and so are quite consistent with the result of both shorter naming latencies and smaller effects for more skilled subjects. As such, it is essential to demonstrate that Phoenicians and Chinese are qualitatively, rather than quantitatively, different by showing that each group is superior when using its preferred route.

Follow-ups to the original Baron and Strawson (1976) study, mainly with beginning readers and dyslexics, have provided only minimal support for the proposition that there exist Chinese and Phoenicians types that are qualitatively different. Freebody and Byrne (1988), for example, suggested that 22% of their Grades 2 and 3 students could be classified as Chinese and 16% could be classified as Phoenicians on the basis of their performance on reading irregular words and nonwords (the remainder were classified simply as either good readers or bad readers). The data indicate, however, that the Chinese were actually substantially better than the bad readers at reading nonwords (the assembly route task) and substantially worse than the good readers at reading irregular words (the lexical route task). Similarly, the Phoenicians were substantially better than the poor readers at reading irregular words and substantially worse than the good readers at reading nonwords. Furthermore, in neither grade did the Phoenicians show significantly better performance than the Chinese on the one test most relevant to their allegedly strong assembly route, a phonemic awareness test; at the same time, they did show better performance on comprehension measures. Finally, when these children were tested a year later, the performance differences between groups on both nonwords and irregular words had diminished substantially (Byrne, Freebody, & Gates, 1992). The authors concluded that if true Chinese readers (i.e., those who are good at using word-specific associations but not at using spelling-to-sound relationships) exist at all, they do not exist beyond the very early grades (see also Gough & Walsh, 1992).

Baron and colleagues (Baron, 1979; Baron & Treiman, 1980; Baron, Treiman, Wilf, & Kellman, 1980; Treiman, 1984) have, nonetheless, continued to argue that reading by spelling-to-sound rules and reading by word-specific associations do represent two independent abilities, suggesting that it should be possible to find true Chinese and Phoenician readers. Recently, arguments for the existence of two independent abilities have been based on the fact that correlations between performance on non-word-naming tasks and irregular word-naming tasks (measured in terms of errors) tend to be smaller than correlations between nonword-naming performance and regular word-naming performance, a result that would follow if qualitatively different routes were being used for naming nonwords and irregular words.

Several things should be noted about these results, however. First, the correlation between nonword performance and irregular word performance is often quite large in absolute terms (Baron, 1979; Baron and Treiman, 1980). Thus, although these data are consistent with the dual-route model, we note that they are also consistent with the position that irregular-word naming and nonword naming are actually based on qualitatively similar processes, a position taken by the single-route models (Brown & Besner, 1987; Seidenberg & McClelland, 1989; Van Orden et al., 1990). Second, the differences between the irregular word–nonword correlation and the regular word–nonword correlation are often significant, but they are also often quite small. For example, the two correlations were .90 and .95, respectively, in Baron and Treiman (1980), and .79 and .90, respectively, in Baron (1979, Experiment 1). (In fact, in one group of subjects in Baron, 1979, the difference was reversed.) Finally, because these correlations involve errors rather than RTs, a number of subjects had to be thrown out because they had too few errors. Their inclusion would certainly diminish the already small differences between the correlations.

We conducted the present study to reexamine and extend Baron and Strawson's (1976) original results with adult (i.e., university student) readers. The basic experimental procedure involved the subjects first taking the oral spelling test with an extended set of words (45 in total). The subjects were next run through a set of naming tasks, two of which should be easier for the Phoenicians. One involved naming a set of one- and two-syllable nonwords (the nonword task). Presumably, the Phoenicians should be much better at assembling these novel pronunciations, and they should have shorter overall RTs and show a smaller effect of number of syllables. This should be true regardless of whether the "syllable" effect is actually based on the difference in number of syllables or simply the difference in nonword length.

In the second task (the lexicality task), one- and two-syllable, low-frequency, regular words were mixed with one- and two-syllable nonwords. Again, the Phoenicians should show faster naming times and smaller syllable effects for the nonwords than the Chinese. They also should show smaller lexicality effects. The reasoning is that the Chinese are more likely to make use of the lexical route to facilitate naming of the words, whereas their nonword naming through the assembly route should be relatively slower.

In the other tasks, the Chinese would be expected to show better performance. The third task involved the naming of irregular words of various frequencies. The Chinese should have the faster overall RTs. They should also show smaller frequency effects, because, for them, a larger portion of the low-frequency words should be named fast enough, through the lexical route, to escape the interference from the assembly route.

The fourth task involved the stimuli used by Seidenberg et al. (1984) in their analysis of the Regularity × Frequency interaction. Following their line of argument, neither group would be expected to show a regularity effect with the high-frequency words. The Chinese, however, would be expected to show a much smaller regularity effect than the Phoenicians with the low-frequency words. The fifth task was analogous to the fourth task, with regularity defined slightly differently, that is, in terms of the stress pattern in two-syllable words. For English two-syllable words, the more common (i.e., regular) stress assignment is to the first syllable. Using this definition, Monsell, Doyle, and Haggard (1989) obtained a Regularity × Frequency interaction very similar to the interaction obtained by Seidenberg et al. (i.e., high-frequency words show no stress effect, whereas low-frequency words show a noticeable stress effect). The explanation they provide is also somewhat similar. That is, high-frequency words can be named through the lexical route before a phonological code can be assembled with a regular (i.e., first syllable) stress pattern. Thus, an irregular stress pattern does not slow naming. For low-frequency words, there is a mismatch in stress assignment for irregularly stressed words, leading to a stress effect. In this task, the expectation is that, with low-frequency words, the Chinese should show a smaller effect of stress than the Phoenicians.

At the end of the naming tasks, subjects performed a pseudohomophone-detection task involving 81 nonwords (30 of which were pseudohomophones) and finally the forced-choice spelling-discrimination task. It should be noted that both of the pencil and paper tasks used here involved more stimuli than in the original study. Thus, the cutoff scores to define our groups were higher. As Baron and Strawson (1976) pointed out, however, their scores were arbitrary anyway, and the feeling was that a larger set of stimuli would provide more protection against floor effects in these tasks.

Method

Subjects

The original pool of subjects was comprised of 73 students from the University of Western Ontario; each received $6.00 for participating.

Stimuli

Five sets of stimuli were used in the naming tasks, two of which were drawn from previously published papers (Monsell et al., 1989; Seidenberg et al., 1984). Of the stimulus sets created for this study, one contained only nonwords. A second set contained low-frequency regular words and nonwords. A third set contained only irregular words, in four frequency ranges. The stimuli for all these tasks are listed in Appendixes A through E.

The stimuli for the nonword task consisted of 40 one-syllable and 40 two-syllable items matched on distributions of first phonemes. The one-syllable stimuli were created by changing the first phoneme of one-syllable words that were not used in any other conditions of the experiment. The two-syllable stimuli were formed by randomly pairing first and second syllables from two-syllable English words that were not used in any other conditions of the experiment. The criteria for inclusion were that the resultant stimulus not be a word, a nonword stimulus in the lexicality task, or unpronounceable. The one-syllable stimuli averaged 4.4 letters in length, whereas the two-syllable stimuli averaged 7.0 letters (see Appendix A).

The stimuli for the lexicality task consisted of 80 words and 80 nonwords. The nonwords were created by breaking the words into first consonant (or consonant cluster) and remainder and then recombining the pieces. Half of each type of stimulus were one syllable, and half were two syllables in length. Thirty-five of the 40 one-syllable words were of Frequency 2 (Kucera & Francis, 1967), and the remaining 5 were of Frequency 1. Thirty-four of the 40 two-syllable words were of Kucera–Francis Frequency 2, and the remaining 6 were of Frequency 1. The one-syllable words and one-syllable nonwords both averaged 4.5 letters in length, whereas the two-syllable words and two-syllable nonwords both averaged 6.2 letters (see Appendix B).

The stimuli for the irregular word task consisted of 52 words divided into four frequency ranges: very high, high, low, and very low. Each set contained 13 words. The mean Kucera–Francis frequency for the very high set was 431.3 (SD = 261.1); for high, 109.9 (SD = 20.7); for low, 44.8 (SD = 14.4); and for very low, 9.2 (SD = 5.1). Very high frequency words averaged 4.8 letters in length; high words averaged 5.2 letters; low words averaged 5.0 letters; and very low words averaged 4.6 letters (see Appendix C).

The stimulus set drawn from Seidenberg et al. (1984) contained 52 words, 13 in each of 4 sets (high frequency regular, high frequency irregular, low frequency regular, and low frequency irregular). Since frequencies for the high frequency sets were highly variable (see Appendix D), we report only medians here: 715 for the high frequency regular set; 665 for the high frequency irregular set; 9 for the low frequency regular set; and 12 for the low frequency irregular set. These four sets were 4.1, 4.0, 4.2, and 4.4 letters long, respectively.

The stimulus set drawn from Monsell et al. (1989) contained 80 words, 20 in each of four sets defined by frequency and stress position. (See Appendix E.) The mean frequency for the high frequency, final stress words was 77.0 (SD = 44.5, median = 57.5). The mean frequency for the high frequency, initial stress words was 71.5 (SD = 20.5, median = 61.5); 2.8 for the low frequency, final stress words (SD = 2.5, median = 2), and 2.5 for low frequency, initial stress words (SD = 2.3, median = 2). High/final words averaged 5.6 letters in length; high/initial, 5.3 letters; low/final 5.4 letters; and low/initial 4.8 letters.

The stimuli for the pseudohomophone detection task (30 pseudohomophones and 51 nonhomophonic nonwords) were randomly selected from the set used by McCann and Besner (1987). (See Appendix F.) There was no overlap between these stimuli and those used by Baron and Strawson (1976). To create the stimuli for the spelling test, we selected 50 words from the Bad Speller's Dictionary and asked undergraduate students in psychology classes to spell them. We used the 35 words that had the highest misspelling rates (see Appendix G). Baron and Strawson did not report the words they used.

Apparatus

All stimulus sets were presented for naming on a Zenith Data Systems ZCM 1490 Flat Screen Technology color monitor. The experiment was controlled by a Zenith Low Profile 286 computer, using Psychology Software Tools Micro Experimental Laboratory software. Reaction timing began with the onset of a stimulus and ended when the microphone picked up the subject's voice and closed a switch connected to the computer. Trials on which extraneous noises set off the microphone too early were noted by the experimenter and those RTs were not analyzed.

All stimuli were presented centrally on the screen in lowercase letters. Each letter was approximately 0.7 cm high × 0.4 cm wide. Subjects were seated approximately 70 cm from the screen. Thus, the visual angle of a four-letter word would be approximately 1.3°.

Procedure

Testing sessions always began with the spelling test. Each of the 35 words were dictated to a subject, always in the same order. Subjects spelled the words on forms provided to them. They then entered the naming part of the study, always starting with the nonword task. This task always came first in an attempt to minimize lexical influences on performance that may have arisen if the subjects' lexical systems had been "primed" by participating in any of the tasks containing real words. Following this, subjects received the lexicality task (low-frequency words and nonwords), the irregular word task (irregular words in four frequency ranges), the Monsell et al. (1989) task, and the Seidenberg et al. (1984) task, in that order.

When the naming part of the experiment was finished, subjects were given the list of 81 nonwords, of which 30 were pseudohomophones, and asked to circle the pseudohomophones. While they were doing this, the experimenter customized the spelling test form by substituting any misspellings the subject made for the corresponding default misspelling. When subjects finished the pseudohomophone detection task, they were given the forced-choice spelling-discrimination task as the last procedure of the experiment. They were asked to circle the correctly spelled form of each word.

Results

The number correct on the initial spelling test ranged from 12 to 34 out of 35 items. Improvement scores on the spelling-discrimination task ranged from 0 to 12. Misses on the pseudohomophone detection task ranged from 0 to 20, with false alarms ranging from 0 to 10. The result was that the error totals on the pseudohomophone detection task ranged from 1 to 20. Totals for performance on the two tasks ranged from 2 to 36. Subjects with total scores of 6 or less were classified as Phoenicians (n = 15), and those with total scores of 19 or more were classified as Chinese (n = 14). For the 15 Phoenicians, the average number correct on the spelling test was 30.1, the average improvement score was 1.4, and the average number of errors on the pseudohomophone detection task was 2.3. For the 14 Chinese, the average number of correctly spelled words on the spelling test was 20.0, the average improvement score was 7.7, and the average number of errors on the pseudohomophone detection task was 12.9.

The five naming tasks will be presented separately. In all instances, an overall analysis of variance (ANOVA) based on correct RTs was first performed for the 73 subjects, followed by the ANOVA of subjects who were classified as Phoenicians or Chinese.

The overall RTs for the Phoenicians were substantially shorter than those for the Chinese. Thus, the results from a second classification scheme also will be presented in each instance. In this classification scheme, the categories were expanded so that Phoenicians were defined as those having a score of 9 or less, and Chinese were defined as those having a score of 15 or more. Within each group, we selected 12 subjects by matching overall RTs of a Phoenician and a Chinese subject. (These overall RTs were based only on RTs to word stimuli across the four tasks involving word stimuli.) This gave two groups of subjects with nearly identical mean RTs. Five members of each of these groups had been among the subjects selected for the initial analysis. For the 12 Phoenicians, the average number correct on the spelling test was 28.8, the average improvement score was 1.9, and the average number of errors on the pseudohomophone detection task was 4.9. For the 12 Chinese, the average number of correctly spelled words on the spelling test was 24.0, the average improvement score was 5.5, and the average number of errors on the pseudohomophone detection task was 11.5. The analysis of the naming task results from these individuals (i.e., the "matched groups" analysis) will have implications for an evaluation of the time-course model.

As in most naming tasks, error rates were low, approximately 4% overall. Nonetheless, similar ANOVAs were carried out on the error data. As these results were virtually the same as the results from the RT analysis, they are not reported here, although error rates are reported in the tables. Unlike words, nonwords do not necessarily have a single pronunciation. The pronunciation of a nonword was only scored as an error when the experimenter decided that there was no set of English spelling-to-sound rules that would support the pronunciation given.

Nonword Task

The results for all three of these analyses are contained in Table 1. The overall analysis indicates that subjects were much faster naming one-syllable nonwords than two-syllable nonwords, $F(1, 72) = 249.99$, $p < .001$, $MS_e = 2{,}621.72$. This effect was also significant when considering only the (unmatched) Chinese and Phoenicians, $F(1, 27) = 131.70$, $p < .001$, $MS_e = 1{,}897.58$, as was the groups effect, $F(1, 27) = 40.98$, $p < .001$, $MS_e = 34{,}468.09$, and the Groups × Number of Syllables interaction, $F(1, 27) = 9.58$, $p < .005$, $MS_e = 1{,}897.58$. Phoenicians were faster than the Chinese and, as predicted, showed a smaller syllable effect. In the matched groups analysis, however, only the number of syllables effect was significant, $F(1, 22) = 83.80$, $p < .001$, $MS_e = 2{,}472.96$. In particular, the interaction with groups did not approach significance, $F(1, 22) = 1.20$, $p > .25$, $MS_e = 2{,}472.96$.

Table 1
Mean Reaction Times (RTs) and Error Rates (ERs) in the Nonword Naming Task

	Stimuli				
	One-syllable		Two-syllable		
Group	RT	ER	RT	ER	Increase
Baron & Strawson split					
Chinese	810	5.5	977	15.9	167
Phoenician	533	1.7	629	8.0	96
Matched groups					
Chinese	619	4.6	734	13.3	115
Phoenician	595	1.1	742	6.4	147
Overall	647	3.4	781	9.4	134

Note. Correlation of syllable effect with overall RT = +.585 at $p < .001$.

Lexicality Task

The results for all three of these analyses are contained in Table 2. The overall analysis indicated an effect of lexicality, $F(1, 72) = 222.99$, $p < .001$, $MS_e = 1{,}227.17$, an effect of number of syllables, $F(1, 72) = 85.97$, $p < .001$, $MS_e = 713.09$, and an interaction of these factors, $F(1, 72) = 27.20$, $p < .001$, $MS_e = 456.97$. These results indicate that the word–nonword difference is larger for two-syllable letter strings than for one-syllable letter strings.

The same pattern of results was found when considering only the Chinese and Phoenicians: lexicality, $F(1, 27) = 198.88$, $p < .001$, $MS_e = 703.29$; number of syllables, $F(1, 27) = 51.01$, $p < .001$, $MS_e = 707.29$; and the interaction, $F(1, 27) = 6.00$, $p < .025$, $MS_e = 535.60$, were significant. Also significant was the groups effect, $F(1, 27) = 34.95$, $p < .001$, $MS_e = 34{,}103.47$. The groups effect again indicated that the Phoenicians were faster than the Chinese and the interactions with groups indicated that, as expected, the Phoenicians showed a smaller effect of both lexicality, $F(1, 27) = 29.34$, $p < .001$, $MS_e = 703.29$, and number of syllables, $F(1, 27) = 12.25$, $p < .005$, $MS_e = 707.73$. The three-factor interaction was nonsignificant, $F(1, 27) = 2.70$, $p > .10$, $MS_e = 535.60$.

In the matched groups analysis the lexicality effect, $F(1, 22) = 115.25$, $p < .001$, $MS_e = 663.28$, the number of syllables effect, $F(1, 22) = 56.97$, $p < .001$, $MS_e = 443.26$, and their interaction, $F(1, 22) = 4.31$, $p < .05$, $MS_e = 499.85$, were significant, but no effect involving groups was significant, all other $ps > .20$.

Irregular Word Task

The results for all three of these analyses are contained in Table 3. The effect of frequency was significant both in the overall analysis, $F(3, 216) = 52.39$, $p < .001$, $MS_e = 809.99$, and in the analysis considering only the Chinese and Phoenicians, $F(3, 81) = 54.09$, $p < .001$, $MS_e = 594.22$. The groups effect, $F(1, 27) = 25.37$, $p < .001$, $MS_e = 16{,}117.32$, and the Frequency × Groups interaction were also significant, $F(3, 81) = 6.04$, $p < .005$, $MS_e = 594.22$, in this second analysis. These two effects were again due to the Phoeni-

Table 2
Mean Reaction Times (RTs) and Error Rates (ERs) in the Assembly Task

Group	One-syllable RT	One-syllable ER	Two-syllable RT	Two-syllable ER	Increase
	Baron & Strawson split				
Chinese					
Words	667	0.9	702	1.8	35
Nonwords	746	9.6	816	10.2	70
Difference	+79		+114		
Phoenician					
Words	501	1.0	516	1.5	15
Nonwords	541	3.3	562	4.8	21
Difference	+40		+46		
	Matched groups				
Chinese					
Words	550	1.5	578	1.5	28
Nonwords	602	7.1	637	9.8	35
Difference	+52		+59		
Phoenician					
Words	553	0.9	572	1.1	19
Nonwords	595	3.0	644	3.4	49
Difference	+42		+72		
	Overall				
Words	569	0.9	585	1.4	16
Nonwords	618	5.2	660	6.7	42
Difference	+49		+75		

Note. Correlations with overall RT: syllable effect (words): $r = +.234$ ($p < .05$); syllable effect (nonwords): $r = +.620$ ($p < .001$); lexicality effect (1 syllable): $r = +.407$ ($p < .005$); lexicality effect (2 syllables): $r = +.682$ ($p < .005$).

cians being faster than the Chinese and, contrary to predictions, showing a smaller frequency effect for these irregular words. In the matched groups analysis, however, only the effect of frequency was significant, $F(3, 66) = 25.23$, $p < .001$, $MS_e = 592.71$. In particular, the interaction did not approach significance, $F(3, 66) = 1.38$, $p > .25$, $MS_e = 592.71$.

Seidenberg et al. (1984) Task

The results for all three of these analyses are contained in Table 4. The overall analysis indicated an effect of regularity, $F(1, 72) = 5.40$, $p < .025$, $MS_e = 1{,}215.94$, an effect of frequency, $F(1, 72) = 27.01$, $p < .001$, $MS_e = 1{,}111.84$, and an interaction, $F(1, 72) = 6.51$, $p < .025$, $MS_e = 1{,}145.15$. As in the original article, the interaction indicates that the regularity effect only exists for the low-frequency words.

When considering only the Chinese and Phoenicians, only the frequency effect was significant, $F(1, 27) = 20.62$, $p < .001$, $MS_e = 934.74$, although the interaction with regularity was marginal, $F(1, 27) = 3.66$, $p < .07$, $MS_e = 915.48$. The groups effect was again significant, $F(1, 27) = 24.45$, $p < .001$, $MS_e = 20{,}356.46$, as were the Groups × Frequency interaction, $F(1, 27) = 6.22$, $p < .025$, $MS_e = 934.74$, and the three-factor interaction, $F(1, 27) = 5.12$, $p < .05$, $MS_e = 915.48$. The groups effect again indicates that the Phoenicians were faster than the Chinese. The two interactions indicate that, contrary to predictions, it was the Chinese who showed the larger frequency effect and the larger Frequency × Regularity interaction. In particular, the only suggestion of a regularity effect was found for the Chinese with low-frequency words.

In the matched groups analysis, the two main effects regularity and frequency were significant, $F(1, 22) = 4.49$, $p < .05$, $MS_e = 685.97$, and $F(1, 22) = 7.01$, $p < .025$, $MS_e = 604.79$, respectively. Even though the groups had been matched on overall RT, this analysis also indicated that, again contrary to expectations, the Chinese still showed a significantly larger frequency effect, $F(1, 22) = 4.95$, $p < .05$, $MS_e = 604.79$, and a marginally larger regularity effect, $F(1, 22) = 2.97$, $p < .10$, $MS_e = 685.97$. Neither the Frequency × Regularity interaction nor the three-factor interaction approached significance (both ps > .25).

Table 3
Mean Reaction Times (RTs) and Error Rates (ERs) in the Irregular Word-Naming Task

Group	UH RT	UH ER	High RT	High ER	Low RT	Low ER	UL RT	UL ER	UL–UH
				Baron & Strawson split					
Chinese	537	0.6	558	5.0	564	2.2	639	11.0	+101
Phoenicians	435	0.0	452	3.1	448	1.0	488	7.7	+53
				Matched groups					
Chinese	476	0.0	479	5.1	481	2.6	540	10.9	+64
Phoenicians	481	0.0	485	2.1	489	0.7	523	5.6	+42
Overall	493	0.1	506	4.5	505	1.5	548	9.1	+55

Note. UH = ultra-high; UL = ultra-low. UL–UH correlation with overall RT = +.269 ($p < .05$).

Table 4
Mean Reaction Times (RTs) and Error Rates (ERs) in the Seidenberg, Waters, and Tannenhaus (1984) Task

	Words				
	Regular		Irregular		
Group	RT	ER	RT	ER	Effect
Baron & Strawson split					
Chinese					
High	560	0.6	550	1.8	−10
Low	576	1.2	613	8.9	+37
Difference	+16		+63		
Phoenicians					
High	437	0.0	439	3.3	+2
Low	450	0.0	449	8.3	−1
Difference	+13		+10		
Matched groups					
Chinese					
High	480	0.0	494	4.2	+14
Low	498	0.7	525	9.0	+27
Difference	+18		+31		
Phoenicians					
High	485	0.0	491	2.3	+6
Low	491	0.0	490	8.3	−1
Difference	+6		−1		
Overall					
High	500	0.2	499	2.1	−1
Low	510	0.8	529	7.5	+19
Difference	+10		+30		

Note. High = high frequency; low = low frequency. Correlation of overall RT with regularity effect (high-frequency words) = −.158, *ns*; with regularity effect (low-frequency words) = +.478 ($p < .001$); with frequency effect (regular words) = −.148, *ns*; and with frequency effect (irregular words) = +.526 ($p < .001$).

Monsell et al. (1989) Task

The results for all three of these analyses are contained in Table 5. The overall analysis indicated that both main effects, stress and frequency, were significant, $F(1, 72) = 10.31$, $p < .005$, $MS_e = 1{,}603.34$, and $F(1, 72) = 145.91$, $p < .001$, $MS_e = 2{,}163.60$, respectively, whereas the interaction was marginal, $F(1, 72) = 3.51$, $p < .07$, $MS_e = 787.58$. Thus, the general pattern observed in the original study was also observed here: The stress effect seems to be somewhat larger for low-frequency words.

When considering only the Chinese and Phoenicians, the two main effects, stress and frequency, were again significant, $F(1, 27) = 5.78$, $p < .025$, $MS_e = 2{,}475.20$, and $F(1, 27) = 104.02$, $p < .001$, $MS_e = 1{,}658.43$, respectively, as was the interaction, $F(1, 27) = 4.88$, $p < .05$, $MS_e = 1{,}213.88$. The groups factor was again significant, $F(1, 27) = 26.63$, $p < .001$, $MS_e = 27{,}739.29$, indicating that the Phoenicians were faster than the Chinese. Also significant was the Groups × Frequency interaction, $F(1, 27) = 15.76$, $p < .001$, $MS_e = 1{,}658.43$, indicating that, again contrary to predictions, the Chinese showed larger frequency effects. There was no hint, however, that the stress effect varied for the two groups, $F(1, 27) = .03$, $p > .85$, $MS_e = 2{,}475.20$. This lack of a Groups × Stress interaction indicates that the stress effect seems to behave differently than the regularity effect in the Seidenberg et al. (1984) task. This suggestion will be reinforced by the patterns of correlations with overall RT (reported in Table 5), which will be discussed below. Thus, it seems unlikely that these two "regularity" effects reflect similar processes. This point will be returned to in the Discussion section.

In the matched groups analysis, only the frequency effect was significant, $F(1, 22) = 83.52$, $p < .001$, $MS_e = 1{,}077.29$. No other effects even approached significance (all $ps > .25$).

Discussion

The data reported here provide very little evidence that it is possible to divide subjects meaningfully into Phoenicians and Chinese on the basis of the Baron and Strawson (1976) tasks. Although the Phoenicians showed smaller effects in tasks emphasizing the assembly route, the Chinese did not show smaller effects in tasks emphasizing the lexical route. In fact, in the lexical route task most comparable to that originally reported by Baron and Strawson (the Seidenberg et al., 1984, task), it was the Chinese and not the Phoenicians who showed evidence of being affected by irregular

Table 5
Mean Reaction Times (RTs) and Error Rates (ERs) in the Monsell, Doyle, and Haggard (1989) Task

	Initial		Final		
Group	RT	ER	RT	ER	Increase
Baron & Strawson split					
Chinese					
High	612	1.1	617	1.1	+5
Low	703	7.1	740	9.6	+37
Difference	+91		+123		
Phoenicians					
High	479	0.0	490	0.7	+11
Low	514	5.0	550	6.3	+36
Difference	+35		+60		
Matched groups					
Chinese					
High	520	0.0	520	1.7	0
Low	583	7.9	591	9.2	+8
Difference	+63		+71		
Phoenicians					
High	541	0.0	557	1.8	+16
Low	597	4.1	612	8.2	+15
Difference	+56		+55		
Overall					
High	541	0.3	549	1.1	+8
Low	600	5.7	621	7.7	+21
Difference	+59		+72		

Note. High = high frequency; low = low frequency. Correlation of overall RT with stress effect (high frequency) = −.046, *ns*; with stress effect (low frequency) = +.017, *ns*; with frequency effect (initial) = +.433 ($p < .001$); and with frequency effect (final) = +.387 ($p < .001$).

spelling-to-sound correspondences (for low-frequency words).

The matched groups analysis provides further support for this conclusion. One could argue that the Phoenician–Chinese differences were obscured because of the large difference in overall RTs between the two groups. If so, the anticipated differences should have emerged in the matched groups analysis. On the contrary, this analysis indicated that the newly defined Chinese and Phoenicians behaved almost identically. The only hint of an interaction with groups was in the Seidenberg et al. (1984) task and it was in the wrong direction: the Chinese showed slightly larger frequency and regularity effects.

These results raise the question of whether there really exist individuals who fit a Chinese–Phoenician categorization scheme. To answer this question, we evaluated individual differences in the two effects most central to the dual-route conceptualization, the lexicality effect in the lexicality task (averaged over both one- and two-syllable words) and the low-frequency regularity effect in the Seidenberg et al. (1984) task. If subjects are relatively better in the process of assembling subword segments, they should show smaller lexicality effects and larger regularity effects. If they are relatively better in the use of word-specific associations, they should show larger lexicality effects and smaller regularity effects. On the basis of the size of these effects, a tertile split was done on the 73 subjects in order to classify those subjects showing small lexicality effects and large regularity effects as Phoenicians and those subjects showing the opposite as Chinese. This split provided only 3 Phoenicians and 5 Chinese, too few for any further analysis. Instead, median splits were done on these effects, producing 15 Phoenicians and 14 Chinese. (Only 8 of these subjects had been selected by the original, Baron and Strawson, 1976, classification scheme, 4 of whom had been classified into the opposite group.)

Performance of these individuals was then examined on the nonword task and the irregular word task. The prediction was that subjects showing smaller lexicality effects and larger regularity effects (thus classified as Phoenicians) should also show smaller syllable effects, but larger frequency effects. As predicted by the classification scheme, the newly defined Phoenicians did show smaller syllable effects on the nonword task, as well as overall smaller RTs. However, they also showed smaller frequency effects in the irregular word task. Thus, a division of the groups according to a criterion based on lexicality and regularity effects was also not successful at creating a meaningful Chinese–Phoenician classification scheme.

Our failure to identify a set of Chinese and Phoenician subjects either in terms of the Baron and Strawson (1976) measures or through this more empirically based technique does suggest that these types of people do not exist (at least in our sample). This poses a considerable challenge for the independence assumption of the dual-route model. That is, if the two routes do represent even quasi-independent sources of information that work together as hypothesized, one of these techniques should have allowed us to identify individuals who show the predicted patterns across our various tasks. Certainly, there would be other ways of creating a Chinese–Phoenician classification scheme and the possibility exists that one of them would be successful at dividing subjects in a way that would yield the expected performance in the naming tasks. The techniques we used, however, would both seem to be quite central to the nature of the two routes.

One other possible reason for this failure, however, could be that we simply had few, if any, of the true Chinese and Phoenician types among our group of 73 subjects. It is quite possible, as Baron and Strawson (1976) suggest, that in the general population, abilities with the two routes are somewhat correlated. Thus, one could argue that a larger sample would have been required in order to produce a set of true Chinese and Phoenicians. This argument, of course, begs the question of why Baron and Strawson were able to locate 11 Phoenicians and 8 Chinese among their 60 subjects; nevertheless, the argument can not be dismissed out of hand.

Even if there were too few subjects to define true Chinese and Phoenicians, however, there is one very clear prediction of the dual-route model that the present data allow us to evaluate. The standard interpretation of the regularity effect for low-frequency words is in terms of the assembly route producing an incorrect phonological code either simultaneously with or before the lexical route can produce the correct code. Those individuals whose lexical route is, relatively, much faster than their assembly route should show smaller regularity effects. The lexicality effect is due to the fact that for regular words, a phonological code can be produced by either route, whereas a code for nonwords can only be produced by the assembly route. Thus, those same individuals whose lexical route is, relatively, much faster than their assembly route should show larger lexicality effects. The implication is that even among people who do not represent the extremes, there should be a negative correlation between the size of their regularity effect and the size of their lexicality effect. In the present sample, the correlation between the size of the low-frequency regularity effect in the Seidenberg et al. (1984) task and the size of the lexicality effect averaged over both one- and two-syllable words and nonwords was significant but positive, $r = +.315$, $t(71) = 2.80$, $p < .01$. The implication is that either the basic dual-route explanation of these two effects is incorrect, or the independence assumption is untenable.

These conclusions are, of course, limited to the population being sampled here, that is, university students who are reasonably skilled in reading. It is possible that other populations would not show similar patterns. For example, the basic dual-route explanation may describe beginning readers, as argued by Baron and colleagues (Baron, 1979; Baron & Treiman, 1980; see also Bryant & Impey, 1986), although other results suggest that this may only be true for the very initial stages of reading (Byrne et al., 1992). Alternatively, the basic dual-route explanation could describe dyslexic readers. For example, it could be suggested that Baron and Strawson's (1976) definition of Chinese and Phoenician readers may correspond to two of the proposed subtypes of dyslexic readers (Boder, 1973). Recent research, however, provides only minimal support for this position. For example, Baddeley, Logie, and Ellis (1988) reported nonsignificant correlations between regularity and lexicality effects for 15 dyslexic boys ($r = -.10$) as well as for 16 reading-age matched controls ($r = +.40$). Ben-Dror, Pollatsek, and Scarpati (1991) reported a highly significant but, as in the present study, positive correlation between the size of the lexicality and regularity effects for 20 dyslexic college students ($r = +.58$). Thus, results for this population would also appear to be incompatible with the basic dual-route conceptualization.

An Alternative Dual-Route Conceptualization

This inability to find the hypothesized negative relationship between the size of the regularity and lexicality effects appears to be the specific result most problematic for the standard version of the dual-route model. The reason that the model makes this prediction is that both of these effects are assumed to be due to the relative speed of the two routes, and the independence assumption, specifically, the assumption that no interactions occur until after phonological outputs from the two routes have been computed. The key question, then, is how the dual-route account must be modified to allow it to explain the positive relationship between the size of the regularity and lexicality effects as well as the rest of the present data.

In order to do this, it is necessary to abandon the idea that the two routes are independent in any way that would make it possible to dissociate the two processes on the basis of the subjects' performance. Rather, a useful operative distinction among subjects could be in terms of naming speed, on the assumption that this variable reflects, to some extent, the skills involved in word pronunciation. Faster subjects simply are more skilled in reading. The positive correlation between the lexicality and regularity effects (and, as discussed below, between these effects and overall naming times) can be interpreted simply as meaning that the faster subjects have more control over their use of the assembly route and, hence, can adjust the onset of pronunciation on the basis of its output, rather than on the lexical route's output, when it is to their advantage to do so. That is, it is assumed that more skilled subjects are better able to strategically use the output of the two processes adaptively in response to different contextual conditions.

The notion that subjects have some control over use of the output from the two routes is consistent with data from Baluch and Besner (1991), who showed that effects dependent on the use of the lexical route, like frequency effects and semantic priming effects, are reduced or disappear when nonwords are included in the experimental list (see also Carr, Davidson, & Hawkins, 1978; Colombo & Tabossi, 1992; Davelaar, Coltheart, Besner, & Jonasson, 1978; Monsell, Patterson, Graham, Hughes, & Milroy, 1992; and Tabossi & Laghi, 1992, for similar arguments). This notion is also consistent with the bypass hypothesis (Doctor & Coltheart, 1980), which claims that a reader's ability evolves from a more extended use of the assembly route to a more extended use of the lexical route. The point is simply that although skilled readers have acquired the ability to read through the lexical route, they do not lose their ability to use assembled phonological codes when it is to their advantage to do so. The explanation of the present data framed in these terms would then be as follows.

Although both lexical and sublexical phonological correspondences emerge automatically, subjects may have some control over the initiation of articulation, such that it can be anticipated or delayed with respect to the process of word identification. Fast subjects, who presumably are also more skilled, are more likely to initiate pronunciation independent of word identification. According to this idea, faster readers would be expected to show smaller lexicality effects, because when the stimulus set contains only nonwords and regular words it is more efficient (in speed terms) to use the assembly route to start the pronunciation without a complete identification of the presented stimulus. In Seidenberg et al.'s (1984) task or the irregular word task, this procedure may be counterproductive because of the irregular words. In these tasks, however, the regularity and frequency effects should be smaller for the faster subjects because of their more efficient lexical route.

This hypothesis, which is essentially based on the abilities of the faster subjects, appears to raise one further issue: Why do the slower subjects (who according to this conceptualization have little control over their routes) show both large regularity effects, indicating that the two routes are fairly equivalent in efficiency, and at the same time, large lexicality effects, indicating that their lexical routes are much more efficient than their assembly routes? In fact, among the 42 subjects with the longest overall RTs, the correlation between the lexicality and regularity effects was still significant and positive, $r = +.288$, $p < .05$, one-tailed, and, in fact, was essentially the same as the correlation over all 73 subjects, $r = +.315$.

This problem can be handled by appealing to the notion of a phonological output lexicon that speeds up the naming of familiar letter strings through a checking process (Derouesné & Beauvois, 1985; McCann and Besner, 1987; Morton & Patterson, 1980; Patterson & Coltheart, 1987). It must also be assumed that this lexical checking process would be most beneficial to the slower readers. This assumption, while post hoc, is consistent with the finding that poor readers benefit from context more than good readers (Perfetti & Roth, 1981). In both instances, the argument is that stored world knowledge plays a much larger role in the naming process for poor readers than for good readers. If so, the large lexicality effects for slower readers would be explained. Their large regularity effects would reflect the fact that the two routes are both slow and, thus, incongruent pronunciations for irregular words would often be available simultaneously. The result, as claimed in the standard dual-route explanation, is interference and, hence, a large regularity effect.

In summary, if it is assumed that the processes underlying the two effects are not independent, and therefore that the subjects' abilities cannot be dissociated (as opposed to Baron and Strawson's, 1976, original claim), if it is assumed that naming speed is an indication of how much control subjects have over the release of the output of their two routes, and if it is assumed that slower readers rely more on a phonological output lexicon, the dual-route hypothesis can be preserved. As we argue below, however, that our data are handled much more easily by virtually all the other models under consideration, without having to appeal to ad hoc assumptions.

The Time-Course Model

The time-course model was proposed by Seidenberg (1985c) as an alternative to the dual-route model. In the time-course model (as in the dual-route model), early word-processing can involve both an orthographically based process and a phonologically based process. Seidenberg does not assume, however, that these processes are independent but rather that they interact. In addition, the time-course model differs in its assumption that words are always named through access to the lexicon and retrieval of a lexically based phonological code (Seidenberg, 1985a), whereas according to the standard dual-route model, the assembly route can bypass the lexicon. Because of these differences, one could argue that: (a) the tasks we and Baron and Strawson (1976) used to divide subjects into Chinese and Phoenicians were inappropriate, and (b) a distinction of this sort is not possible in any case. Nonetheless, the model's explanation of the regularity effect is somewhat similar to that of the dual-route model. That is, recognition of low-frequency words is so slow through strictly orthographically based processing that phonologically based processing can contribute to the final output. The result is that regular words are named faster than irregular words. Further, according to Seidenberg (1985c), a fast overall RT indicates that an individual's orthographically based processing is fast enough that it will avoid the effects of phonology most of the time. Thus, the faster people should show the smallest regularity effects, as reported by Seidenberg et al. (1984).

An implication of this argument is that the observed patterns of data with our Chinese and Phoenicians may be explained in terms of the much longer RTs for the Chinese. The general expectation would be that in a matched groups analysis, no group differences would be observed. This is essentially what we found. At a more fine-grained level, however, this argument further suggests that the best way to evaluate the time-course model would be to examine the correlations of the effects with overall RT as done by Seidenberg (1985c). To this end, correlations between all the effect sizes and overall RT (based only on the word conditions) were calculated for each task. These correlations were based on all 73 subjects and each is shown in the corresponding table in the Results section. The one- and two-tailed cutoffs (α = .05) for these correlations are r = .195 and r = .231, respectively.

In terms of the time-course model, the explanation for the lexicality effect in the lexicality task could go as follows. Words are named through the retrieval of a phonological code after lexical access, whereas nonwords must have their pronunciations synthesized. Processing the nonwords causes the lexical representations of orthographically similar words to become partially activated. Consequently, the phonological codes of these words would also be partially activated. The synthesis process then uses these partially activated codes to produce a name for the nonword. As noted in Table 2, there is a positive correlation between the size of the lexicality effects and overall RT. In order to explain this positive correlation, one must assume that subjects who are fast are more highly skilled in general and, thus, should be able to synthesize a nonword pronunciation much more effectively than the slower, less-skilled subjects. Faster subjects will, therefore, be faster at reading nonwords. They are, of course, also faster at reading words. One could further assume, however, that the relative disadvantage of slower readers in pronouncing novel letter strings is larger than their relative disadvantage in pronouncing more familiar strings. In other words, slow readers are particularly bad with unfamiliar letter strings.

The syllable effects in the nonword and lexicality tasks could simply be explained as length effects. Presumably, both the process of accessing the lexicon, for words, and the process of synthesizing the pronunciation from partially activated phonological representations, for nonwords, could be sensitive to the size of the letter string. Again, the faster the subjects are, the better they may be at these processes. Thus, the faster subjects should have smaller syllable effects, as was found.

In the irregular word task, the model would predict a larger frequency effect for the slower subjects because the low-frequency words would tend to engage the phonologically based lexical access processes more for those subjects. In Seidenberg et al.'s (1984) task, the model would predict a correlation between overall RT and the regularity effect for the low-frequency words, but not for the high-frequency words, which should escape the effects of the phonologically based lexical activation for all subjects. Similarly, the model would predict a correlation between overall RT and the size of the frequency effect for irregular words but not necessarily for regular words. These results were all as predicted. (It should be noted that the nonsignificant correlations in two of these cells cannot be attributed to a range restriction. The variances of the four effects that were correlated with overall RT in this task were virtually identical.)

In Monsell et al.'s (1989) task, a similar set of predictions would follow on the basis of the assumption that stress irregularity (which is at the suprasegmental, or prosodic, level) involves similar processes as those involved in irregularity at the spelling-to-sound translation level. That is, there should be significant positive correlations between overall RT and the frequency effect only for irregularly stressed words and between overall RT and the stress effect only for low-frequency words. The results indicate a somewhat different pattern than predicted, however. Overall RT correlated with the frequency effects for both regular and irregular stressed words but did not correlate with the stress effect for either high- or low-frequency words. Again, this lack of correlation can not be attributed to a range restriction, because there was essentially no difference between the variances of the effects that correlated significantly and those that did not.

The results in the Seidenberg et al. (1984) task and in the irregular word task are quite consistent with the time-course model, whereas the results in the nonword and lexicality tasks can certainly be accommodated by it. Thus, although this model cannot explain all the present data, it certainly does a better job than the standard dual-route model does.

Single-Route Models

We wish to evaluate a strong prediction of the dual-route model, in particular, that the independence of the two routes would allow us to find individuals in whom the two abilities were dissociated, that is, to find groups of subjects corresponding to what Baron and Strawson (1976) characterized as Chinese and Phoenicians. As we have seen, this enterprise was quite unsuccessful. Instead, the only meaningful way to characterize individual differences was in terms of overall relative speed. As noted, the finding that overall speed is the only aspect of individual differences that seems to matter is generally consistent with the time-course model. It is also quite consistent with the single-route models under consideration, as will be explained in the following discussion.

A reasonable way to interpret the data in terms of this type of model would be to build on the suggestions of Van Orden et al. (1990), Seidenberg and McClelland (1989), and Brown and Besner (1987) about the processes involved in constructing a phonological code. In Van Orden et al.'s terminology, the process of creating a phonological code can be thought of as determining a point in phonological space. This process is frequency sensitive, but it may be so automatic that reasonably skilled readers (i.e., university students) differ only minimally in the speed with which it can be completed. Where readers of differing abilities can presumably be distinguished, however, is in the process of converging on the correct phonological code (the "attractor point"). Better readers may have initial encodings that are closer to the appropriate attractor points and, thus, for them, the converging process will be much faster. Distance from the attractor point may only be an issue at all, however, when there is some potential ambiguity in deriving the initial encoding and, hence, some potential for that encoding to be distant from the attractor point (i.e., as with irregular words, nonwords, and two-syllable words). When the letter string is a single-syllable, regular word, the potential for ambiguity is small and the encoding may be inevitably close to its attractor point for all readers. Thus, for these words, the "cleaning-up" process may be very rapid for all readers.

The basic prediction of all the single-route type of models is that the better subjects are at the process of deriving spelling–sound correspondences and selecting among them, the faster they will be at naming less common letter strings and, hence, the smaller should be the effects of the various factors investigated here. As noted, this was the case in almost all situations. What is interesting, then, is to consider the effects that did not correlate with overall RT because, presumably, these effects would need a different interpretation.

The most notable exceptions to the trend are the results with the two-syllable words in the Monsell et al. (1989) task, which showed a strong correlation between overall RT and the size of the frequency effect, on one hand, and a lack of correlation between overall RT and the size of the stress effect, on the other hand. The overall pattern of results was quite consistent in all the analyses, namely, a small stress effect for high-frequency words and a relatively larger stress effect for low-frequency words. This pattern is what is expected on the basis of the analogy of the irregularly stressed words to words that are irregular at the segmental level (e.g., the words used in the Seidenberg et al., 1984, task). However, what is different in the Monsell et al. task is that even fast subjects show this pattern, whereas in the Seidenberg et al. task the low-frequency regularity effect is negligible for faster readers.

What these results suggest is that the locus of the stress effect is different from the locus of the other effects. As Colombo (1992) argued, the stress assignment process can be biased by cues derived from subword units, but it may ultimately have to operate on whole-word representations. Furthermore, when these types of (neighborhood) cues are not available (as with the stimuli used in the present experiment, which were originally selected by Monsell et al., 1989, for precisely that reason), the correct assignment of stress must be based on knowledge of whole-word representations. The present results support Colombo's argument that stress assignment does occur somewhat later in the processing sequence and, in particular, after the process that distinguishes between good and poor readers.

Two other results showed a lack of correlation with overall RTs: the regularity effect for high-frequency words and the frequency effect for regular words in the Seidenberg et al. (1984) task. In the case of the regularity effect for high-frequency words, the explanation for the lack of a correlation is quite simple. Although the variance of this effect is as large as the variance of the regularity effect for low-frequency words, both Seidenberg's (1985c) results and the present results indicate that this effect does not actually exist. Thus, this variance would appear to be just random variation about a mean of zero, suggesting that there should be no correlation.

The explanation for the frequency effect for regular, one-syllable words is not so simple, however. Although our effect was small (only 10 ms), it was significant, $t(71) = 1.87$, $p < .05$, one-tailed, and not at all out of the range of what is typically reported in the literature (Paap & Noel, 1991; Seidenberg, 1985c; Waters & Seidenberg, 1985). It is interesting that Seidenberg's (1985c) data also appear to show no correlation between this effect and overall RT. That article reports naming data for these same stimuli with a tertile split of the subjects based on overall RT. The frequency effects for the regular words were +25 ms, +7 ms, and +20 ms for the fastest, medium, and slowest subjects, respectively. Thus, this frequency effect, though small, appears to be reliable, but does not vary as a function of reading speed; whatever process produces it may be different from the process that produces the other effects.

There would seem to be two possible explanations for this effect. One possibility, phrased in the terms of the Van Orden et al. (1990) model, would be to argue that this effect is mainly due to the initial process, that of establishing a point in phonological space, although it could also have some basis in the second process, the "cleaning-up" process. The point to realize, however, is that regular words are regular precisely because the spelling-to-sound correspondences of their components are found in many letter strings. As such, both processes would be so practiced even for relatively slow readers (keeping in mind that these are university readers) that only a small effect would be expected, and the size of the effect would not distinguish slow readers from fast readers.

An alternative possibility would be to ascribe a later locus to this effect and to argue that this particular frequency effect is an articulatory-motor effect (Balota & Chumbley, 1985; Theios & Muise, 1977). Balota and Chumbley reported that frequency effects in their standard naming task were only slightly larger than in a delayed-naming task, where the cue to respond came 400 ms after the word's presentation. Furthermore, even at the very longest delays, there was still a frequency effect of approximately 20 ms. On the basis of the argument that 400 ms should be sufficient time to retrieve a phonological code, the implication is that at least some component of frequency effects must be due to articulatory-motor factors, factors that must also be at work in the standard naming task.

The attribution of frequency effects to the articulatory-motor process (see also Andrews, 1989; Colombo, 1992) is not without its critics (McCann & Besner, 1987; McRae, Jared, & Seidenberg, 1990; Savage, Bradley, & Foster, 1990; see also Monsell, 1990). In particular, these authors have argued that the frequency effect in delayed naming is an artifactual consequence of the delayed-naming methodology. At present, however, it appears that there is less than unequivocal support for the contribution of the motor-articulatory process to frequency effects, but the possibility of their existence cannot be dismissed.

In terms of the present data, it can be argued that the major locus of the small (10 ms) frequency effect for regular words is the articulatory-motor process rather than any earlier process that may be responsible for most of the other effects. If one is further willing to assume that the articulatory-motor process is not one that differentiates university-level readers, one would not expect the size of this particular frequency effect to correlate with overall RT. Whether this effect actually is due to articulatory-motor processes or to other processes, as discussed earlier, it appears that the process that plays the major role in distinguishing poor from good readers seems to play a very small role in this effect.

We have discussed our results in terms of notions shared by various single-route models: Brown and Besner (1987), Seidenberg and McClelland (1989), and Van Orden et al. (1990), with an emphasis on the Van Orden et al. model. The present data do not, in fact, allow us to discriminate between these models (nor really between them and the time-course model). In addition, the data do not actually allow us to reject all versions of the dual-route model. Rather, with sufficient additional assumptions, some version of this model can also offer a plausible explanation of the data. As other authors (e.g., Humphreys & Evett, 1985; Seidenberg, 1985a) have repeatedly pointed out, however, it is exactly this ability to add assumptions in order to explain any new set of data that limits this model's usefulness by making it unfalsifiable. What the present data do is to provide a rather strong challenge to a basic assumption of the model, the independence of the two routes. The onus now appears to be on proponents of dual-route models to produce data necessitating the independence assumption as an essential component in any model of the word-naming process.

References

Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 802–814.

Baddeley, A. D., Logie, R. H., & Ellis, N. C. (1988). Characteristics of developmental dyslexia. *Cognition, 29,* 197–228.

Balota, D. A., & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language, 24,* 89–106.

Baluch, B., & Besner, D. (1991). Strategic use of lexical and nonlexical routines in visual word recognition: Evidence from oral reading in Persian. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 644–652.

Baron, J. (1979). Orthographic and word-specific mechanisms in children's reading of words. *Child Development, 50,* 60–72.

Baron, J., & Strawson, C. (1976). Orthographic and word-specific mechanisms in reading words aloud. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 386–393.

Baron, J., & Treiman, R. (1980). Use of orthography in reading and learning to read. In J. F. Kavanagh & R. L. Venezky (Eds.), *Orthography, reading, and dyslexia* (pp. 171–189). Baltimore: University Park Press.

Baron, J., Treiman, R., Wilf, J. F., & Kellman, P. (1980). Spelling and reading by rules. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 159–194). London: Academic Press.

Ben-Dror, I., Pollatsek, A., & Scarpati, S. (1991). Word identification in isolation and in context by college dyslexic students. *Brain & Language, 40,* 471–490.

Berndt, R. S., Reggia, J. A., & Mitchum, C. (1987). Empirically derived probabilities for grapheme-to-phoneme correspondences in English. *Behavior Research Methods, Instruments, and Computers, 19,* 1–9.

Besner, D., Twilley, L., McCann, R. S., & Seergobin, K. (1990). On the association between connectionism and data: Are a few words necessary? *Psychological Review, 97,* 432–446.

Boder, E. (1973). Developmental dyslexia: A diagnostic approach based on three atypical reading-spelling patterns. *Developmental Medicine and Child Neurology, 21,* 663–687.

Brown, P., & Besner, D. (1987). The assembly of phonology in oral reading: A new model. In M. Coltheart (Ed.), *Attention & performance* (Vol. 12, pp. 471–489). Hillsdale, NJ: Erlbaum.

Bryant, P., & Impey, L. (1986). The similarities between normal readers and developmental and acquired dyslexics. *Cognition, 24,* 121–137.

Byrne, B., Freebody, P., & Gates, A. (1992). Longitudinal data on the relations of word-reading strategies to comprehension, reading time, and phonemic awareness. *Reading Research Quarterly, 27,* 141–151.

Carr, T. H., Davidson, B. J., & Hawkins, H. L. (1978). Perceptual flexibility in word recognition: Strategies affect orthographic computation but not lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 674–690.

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

Colombo, L. (1992). Lexical stress effect and its interaction with frequency in word pronunciation. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 987–1003.

Colombo, L., & Tabossi, P. (1992). Strategies and stress assignment: Evidence from a shallow orthography. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology and meaning* (pp. 319–340). Amsterdam: Elsevier.

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.

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

Derouesné, J., & Beauvois, M. F. (1985). The "phonemic" stage in the non-lexical reading process: Evidence from a case of phonological alexia. In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia* (pp. 399–457). Hillsdale, NJ: Erlbaum.

Doctor, E. A., & Coltheart, M. (1980). Children's use of phonological encoding when reading for meaning. *Memory and Cognition, 8,* 195–209.

Freebody, P., & Byrne, B. (1988). Word-reading strategies in elementary school children: Relations to comprehension, reading time, and phonemic awareness. *Reading Research Quarterly, 23,* 441–453.

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

Gough, P. B. (1972). One second of reading. *Visible Language, 6,* 291–320.

Gough, P. B., & Walsh, M. A. (1992). Chinese, Phoenicians and the orthographic cipher of English. In S. Brady & D. Shankweiler (Eds.), *Phonological processes in literacy* (pp. 199–209). Hillsdale, NJ: Erlbaum.

Henderson, L. (1985). Issues in the modelling of pronunciation assembly in normal reading. In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading* (pp. 459–508). London: Erlbaum.

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

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

McCann, R. S., & Besner, D. (1987). Reading pseudohomophones: Implications for models of pronunciation assembly and the locus of word-frequency effects in naming. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 14–24.

McRae, K., Jared, D., & Seidenberg, M. S. (1990). On the roles of frequency and lexical access in word naming. *Journal of Memory and Language, 29*(4), 43–65.

Monsell, S. (1990). Frequency effects in lexical tasks: Reply to Balota & Chumbley. *Journal of Experimental Psychology: General, 119,* 335–339.

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General, 118,* 43–71.

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

Morton, J., & Patterson, K. (1980). A new attempt at an interpretation or an attempt at a new interpretation. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia* (pp. 91–118). London: Routledge, Chapman, & Hall.

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

Patterson, K. E., & Coltheart, M. (1987). Phonological processes in reading: A tutorial review. In M. Coltheart (Ed.), *Attention and performance* (Vol. 12, pp. 421–447). Hillsdale, NJ: Erlbaum.

Perfetti, C. A., & Roth, S. (1981). Some of the interactive processes in reading and their role in reading skill. In A. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading* (pp. 269–297). Hillsdale, NJ: Erlbaum.

Savage, G. R., Bradley, D. C., & Foster, K. I. (1990). Word frequency and the pronunciation task: The contribution of articulatory fluency. *Language and Cognitive Processes, 5,* 203–236.

Seidenberg, M. S. (1985a). Constraining models of word recognition. *Cognition, 20,* 169–190.

Seidenberg, M. S. (1985b). The time course of information activation and utilization in visual word recognition. In D. Besner, T. G. Waller, & G. E. MacKinnon (Eds.), *Reading research: Advances in theory and practice* (Vol. 5, pp. 199–252). San Diego, CA: Academic Press.

Seidenberg, M. S. (1985c). The time course of phonological code activation in two writing systems. *Cognition, 19,* 1–30.

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

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

Stanhope, N., & Parkin, A. J. (1987). Further explorations of the consistency effect in word and nonword pronunciation. *Memory and Cognition, 15,* 169–179.

Stanovich, K. E., & Bauer, D. W. (1978). Experiments on the spelling-to-sound regularity effect in word recognition. *Memory and Cognition, 6,* 410–415.

Tabossi, P., & Laghi, L. (1992). Semantic priming in the pronunciation of words in two writing systems: Italian and English. *Memory and Cognition, 20,* 303–313.

Theios, J., & Muise, J. G. (1977). The word identification process in reading. In N. J. Castellan, Jr. & D. Pisoni (Eds.), *Cognitive theory* (Vol. 2, pp. 289–327). Hillsdale, NJ: Erlbaum.

Treiman, R. (1984). Individual differences among children in spelling and reading styles. *Journal of Experimental Child Psychology, 37,* 463–477.

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

Appendix A

One- and Two-Syllable Nonwords in the Nonword Naming Task

1 syllable	2 syllable	1 syllable	2 syllable	1 syllable	2 syllable
bule	bectale	glap	gloiner	roke	rachoze
blit	blurneal	grost	groldor	sen	sutect
broap	broptly	heam	hampune	shud	shadess
cact	coltace	jind	jingope	skal	skorpary
chone	churtaid	kade	kentute	sleaf	slartful
climp	cleshel	kruve	kroakath	speve	spamtle
creld	crasted	leat	lormoid	starn	stipty
delt	dunkide	mim	muckish	straik	strumpness
droif	driskum	nonk	nerly	tain	turmy
feck	fandike	phoad	phaitle	thert	thistny
flipe	flithin	plurm	plurdish	trosk	treapter
frant	fronat	prot	prodery	vup	vuskie
gice	gilpome	pung	pocktor	wame	wurlner
				zulp	zictor

Appendix B

Low-Frequency Words and Nonwords Used in the Assembly Task

			1-syllable words				
bait	clan	flick	hulk	mend	roach	sleek	trash
bleed	crate	frock	jerk	nope	scuff	spike	thrush
brig	deft	gale	kelp	peep	sect	stave	vow
cain	drawl	glide	kraft	pluck	sheen	strap	wade
chunk	fern	grunt	lark	prick	skeet	tempt	zip
			2-syllable words				
batter	clatter	flimsy	hotly	meekly	relay	slimmer	trotter
blemish	crumble	frolic	jumpy	negate	scatter	spooky	throaty
bracket	dotted	gambit	kazoo	padlock	sedan	stocky	vainly
canny	drummer	glossy	krishna	plunder	shatter	streamer	windy
chubby	fable	grubby	leaky	prostate	skidded	tempest	zebra
			1-syllable nonwords				
bip	clow	flade	hend	meed	reft	slale	trelp
blempt	creen	freet	jope	naft	scark	spawl	thrain
breek	dap	gect	kuck	pulk	sern	stide	vig
cush	dreep	glick	kruff	plerk	shan	strick	wunk
chash	fike	groach	lave	prunt	skock	tate	zait
			2-syllable nonwords				
bebra	clempest	flimmer	hadlock	meaky	rossy	slimsy	tratter
blindy	crocky	fridded	junder	nishna	scumpy	spable	thrubby
brainly	deamer	gedan	kegate	pazoo	sambit	stotted	vanny
cotter	dratter	glostate	kratter	plotly	shummer	strumble	wemish
choaty	fooky	grelay	leekly	prubby	skolic	tacket	zatter

Note. All words have a frequency of either 1 or 2 in the Kucera–Francis frequency count.

(Appendixes continue on next page)

Appendix C

Irregular Words in the Irregular Word Naming Task

Ultra-high		High		Low		Ultra-low	
Word	Frequency	Word	Frequency	Word	Frequency	Word	Frequency
among	320	answer	152	bother	22	aisle	7
become	361	blood	121	calm	35	canoe	7
enough	430	built	103	foot	70	comb	6
front	220	column	71	guard	48	dare	21
good	807	corps	109	guide	36	debt	12
group	390	doubt	114	laugh	28	facade	7
look	399	eight	104	league	69	ghost	11
love	232	friend	133	lose	58	limb	5
most	1,160	scene	106	ocean	34	obey	8
often	368	sigh	94	prove	53	pearl	9
school	492	touch	87	shoes	44	soot	1
talk	154	volume	135	tongue	35	stein	18
word	274	walk	100	worse	50	sword	7

Note. Numbers are Kucera–Francis frequency counts in parts per million.

Appendix D

Words Used in the Replication of the Seidenberg, Waters, Barnes, and Tanenhaus (1984) Study

Word	Frequency	Word	Frequency	Word	Frequency	Word	Frequency
			Regular high-frequency words				
each	877	held	264	help	311	life	715
still	782	stop	120	take	611	these	1,573
name	294	this	5,146	not	4,609	turn	233
			Regular low-frequency words				
carve	3	cub	0	disk	25	gate	37
hunt	10	lent	5	mill	11	mode	21
plump	4	rink	2	struck	59	wail	3
			Irregular high-frequency words				
are	4,393	both	730	does	485	done	320
give	391	great	665	have	3,941	put	437
said	1,961	says	200	shall	267	some	1,617
were	3,284						
			Irregular low-frequency words				
broad	84	caste	3	deaf	12	doll	10
gross	66	lose	58	phase	72	pint	13
root	30	sew	6	spook	0	steak	10
wool	10						

Note. Numbers are Kucera–Francis frequency counts in parts per million.

Appendix E

Words Used in the Replication of the Monsell, Doyle, and Haggard (1989) Study

Word	Frequency	Word	Frequency	Word	Frequency	Word	Frequency
			Initial stress				
High-frequency							
bottom	88	busy	58	cover	88	dinner	91
dozen	52	follow	97	forest	66	garden	60
happen	63	image	119	index	81	jury	67
minor	58	motor	56	novel	59	royal	48
rural	54	seven	113	vital	56	yellow	55
Low-frequency							
arid	2	bogus	3	comet	2	coral	5
covet	1	edit	2	exit	7	heron	1
icon	0	laser	0	modal	3	navel	2
relic	6	rivet	0	sequel	1	suet	0
super	8	tacit	2	vigil	1	wager	3
			Final stress				
High-frequency							
accept	72	alive	57	apply	56	attend	54
avoid	58	begin	84	deny	47	device	55
divide	14	effect	213	enjoy	44	event	81
exist	59	express	42	extent	110	forget	54
hotel	126	machine	103	marine	55	police	155
Low-frequency							
annoy	2	attest	2	banal	2	cadet	4
cigar	10	depict	3	dissect	1	divan	6
effete	1	emit	1	inane	1	inept	2
invent	7	irate	1	lapel	1	omit	1
opine	0	parole	5	pecan	1	propel	4

Note. Numbers are Kucera–Francis frequency counts in parts per million.

Appendix F

Pseudohomophone Test Stimuli

		Nonwords			
baip	feap	hane	merse	serm	veeze
bleef	freest	heest	mong	tolph	vole
breem	frooze	jaul	merf	trufe	wirch
bruve	furve	jinje	neech	tude	wote
clud	gead	jurt	nung	tunce	yait
dawp	gool	keer	pern	turle	yome
deece	gurst	kerth	plew	turth	zoal
dight	haik	klite	preet	vawl	zupe
				vawx	zute
		Pseudohomophones			
berd	feal	grean	hokes	phocks	thret
boan	fownd	groe	mait	pirl	trax
coad	fues	groop	ment	plaie	turse
coph	gess	hazz	muel	raize	waije
dowt	gide	hoap	perge	sope	wirth

(*Appendix G follows on next page*)

Appendix G

Spelling Test Stimuli

Incorrect	Correct	Incorrect	Correct	Incorrect	Correct
asisstant	assistant	hemesphere	hemisphere	overwelm	overwhelm
atempt	attempt	honorible	honorable	perenial	perennial
beseige	besiege	inoculate	innoculate	primative	primitive
boistrous	boisterous	insistant	insistent	putrify	putrefy
comission	commission	interpilate	interpolate	quarentine	quarantine
curteous	courteous	invoise	invoice	renevate	renovate
crape	crepe	laserate	lacerate	sesession	secession
currancy	currency	ligiment	ligament	testiment	testament
despare	despair	loveable	lovable	terpitude	turpitude
elagent	elegant	militery	military	unanamous	unanimous
facinate	fascinate	narative	narrative	whisle	whistle
gorgous	gorgeous	oscilate	oscillate		

Note. Incorrect spellings appeared in either column, randomly, in test given to subjects.

Received July 20, 1992
Revision received June 30, 1993
Accepted July 30, 1993 ■