Journal of Experimental Psychology:
Learning, Memory, and Cognition
1997, Vol. 23, No. 3, 570–590

Strategic Control in a Naming Task: Changing Routes or Changing Deadlines?

Stephen J. Lupker and Patrick Brown
The University of Western Ontario

Lucia Colombo
Università degli Studi di Padova

S. Monsell, K. E. Patterson, A. Graham, C. H. Hughes, and R. Milroy (1992) reported that high-frequency irregular words are named faster when presented in a "pure" block than when mixed with nonwords. They attributed this effect to a de-emphasis of an assembly route in the pure block. The current authors replicated this effect in Experiment 1. In Experiments 2 and 3, similar effects resulted from mixing high- and low-frequency regular words with nonwords and from mixing high- and low-frequency irregular words together. Further, in all cases, the more slowly named stimuli were named faster in mixed blocks than in pure blocks. An alternative to the de-emphasis account, which is based on strategic control of initiation of articulation, was supported in Experiment 4 by confirming the alternative account's novel prediction of a regularity effect for high-frequency words in pure blocks. Implications for the single- versus dual-route debate and for interpretations of strategy effects in general are discussed.

For fluent readers, reading appears to be an extremely automatized process. That is, while reading, both phonological and semantic information appear to become available to the reader essentially automatically. Experimental evidence validating this observation is relatively plentiful, with the most compelling evidence probably coming from interference tasks (Klein, 1964; Lupker, 1979; Rosinski, 1977; Stroop, 1935; see MacLeod, 1991, for a review) and masked priming tasks (Fowler, Wolford, Slade, & Tassinary, 1981; Hines, Czerwinski, Sawyer, & Dwyer, 1986; Marcel, 1983; Perfetti, Bell, & Delaney, 1988). In these tasks, the finding is that phonological and semantic information from an unattended and response-irrelevant word affects processing of the response-relevant aspects of a stimulus. In all instances, the argument is that because the word was either unattended or unavailable to consciousness, its influence must have been the result of automatic processing.

More recently, a subsidiary question has gained some attention: the question of what aspects of the reading process a reader does have control over and how that control can be strategically exercised (e.g., Dorfman & Glanzer, 1988; Gordon, 1983; Stone & Van Orden, 1993). For example, Glanzer and Ehrenreich (1979) have suggested that readers can alter how lexical access is accomplished on the basis of the nature of the words they see in an experimental condition (although see Dorfman & Glanzer, 1988, and Forster, 1981, for alternative views). Others (e.g., Becker, 1980; McKoon & Ratcliff, 1995; Shulman & Davison, 1977) have shown that the way in which a particular type of semantic context affects lexical access depends on the more general context in which it is embedded.

The focus of the present article is not the lexical-access process per se but is the process of producing a phonological code. Much of the existing work on this issue has been couched within the framework of the dual-route model (Coltheart, 1978; Patterson & Morton, 1985). According to this model, there are two ways to produce a phonological code. One way, referred to as the assembly route, involves assembling a pronunciation based on knowledge of spelling-to-sound mappings. This route can be successfully used whenever the letter string's spelling-to-sound mappings follow the standard rules of the language (the so-called "regular" words). Further, this route also allows the naming of unfamiliar letter strings and nonwords. The other route, referred to as the lexical route, involves accessing a lexical representation and retrieving the associated phonological code in an essentially holistic fashion. This route can only be used for naming letter strings that have a lexical representation (i.e., words) and must assume the dominant role whenever the word's spelling-to-sound mappings are not all standard (i.e., the so-called "irregular" words). Further, unlike the assembly route, this route is assumed to be frequency sensitive, with speed of processing being a direct function of the word's frequency.

In the dual-route model, both routes are presumed to work in parallel and relatively automatically (although, as discussed below, they may have different attentional demands).

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

This research was supported by Natural Sciences and Engineering Research Council of Canada Grant A6333. Major portions of this article were presented at the 6th Meeting of the European Society for Cognitive Psychology, Elsinore, Denmark, September 1993 and at the 34th Annual Meeting of the Psychonomic Society, Washington, DC, November 1993.

We thank Sachiko Kinoshita for sharing her data and thoughts about these issues with us and Carolynn Racicot for her help in the data collection and analyses.

Correspondence concerning this article should be addressed to Stephen J. Lupker, Department of Psychology, The University of Western Ontario, London, Ontario N6A 5C2, Canada. Electronic mail may be sent via Internet to lupker@sscl.uwo.ca.

Thus, when a word is being processed, both routes will produce a phonological code. In the case of regular words, the two codes should match, allowing the reader to produce a relatively rapid naming response. For irregular words, however, there will be a mismatch because the assembly route will produce the incorrect "regularized" code. The result is that a competition situation will be created that takes time to resolve, delaying the naming of irregular words (the "regularity effect"; Baron & Strawson, 1976; Glushko, 1979; Stanovich & Bauer, 1978).

In general, regularity effects appear to exist only for low-frequency words (Andrews, 1982; Brown, Lupker, & Colombo, 1994; Paap & Noel, 1991; Seidenberg, Waters, Barnes, & Tanenhaus, 1984; Taraban & McClelland, 1987; although see Content, 1991, and Jared, 1995). In terms of the dual-route model, the explanation is that high-frequency irregular words can apparently be processed fast enough via the lexical route that naming can occur before the assembly route can create a competing code. Thus, this Regularity × Frequency interaction is nicely explained in dual-route terms, and in fact, its existence has been taken as strong support for the dual-route model.

Although both routes are assumed to operate essentially automatically, some researchers working within this framework have suggested that processing on the lexical route is somewhat more automatized than processing on the assembly route (e.g., Herdman, 1992; Paap & Noel, 1991). That is, because the lexical route is assumed to be based on direct, automatic associations between orthographic whole word forms and their corresponding phonologies, whereas the assembly route involves the actual assembly of phonological segments, the assembly route is presumed to require more resources than the lexical route. Therefore, because the speed of retrieval of the phonological code via the lexical route for any given word depends on the frequency with which the word has been seen, and because most of the words typically encountered by readers are reasonably frequent, the lexical route is normally presumed to play the major role in reading.

One implication of the assumption that both routes operate essentially automatically is that the process of naming a word should be fairly impervious to strategies. Nonetheless, recent research has strongly suggested that readers do seem to be able to deploy different strategies in a naming task. In particular, assuming that the two routes are at least to some extent independent, one strategy that may be available to readers is to selectively emphasize or de-emphasize the output of one of the routes. For example, using Italian stimuli, Tabossi and Laghi (1992) reported that if nonwords were included as targets in a naming task (hence putting greater emphasis on the assembly route), word targets showed no associative priming effect (an effect attributed to processing on the lexical route). When those same targets were presented by themselves or when they were presented with irregularly stressed words (words that seem to require lexical involvement) in addition to the nonwords, reliable associative priming effects were observed.

It is worth noting that Tabossi and Laghi (1992) were not able to obtain parallel effects in English. That is, including nonwords did not eliminate the associative priming effect (see also Keefe & Neely, 1990; West & Stanovich, 1982). Tabossi and Laghi attributed this difference to the fact that spelling-to-sound mappings are much less regular in English than in Italian. That is, they claimed that in English, reading "can never be accomplished safely without reliance on lexical knowledge" (Tabossi & Laghi, 1992, p. 310).

Baluch and Besner (1991) reported a similar result with Persian words. In Persian, some words contain all the orthographic information necessary to generate a correct phonological code without consulting lexical knowledge (transparent words), whereas others do not (opaque words). For the transparent words, including nonwords in the experiment eliminated not only the associative priming effect but also the frequency effect, suggesting that their inclusion had induced readers to rely mainly on information from the assembly route. It is worth noting that these effects were obtained even when opaque words were also included in the experiment, that is, even when it was necessary to use the lexical route on some portion of the trials. For the opaque words, in fact, neither the priming effect nor the frequency effect was noticeably altered by the inclusion of nonwords.

Finally, Colombo and Tabossi (1992) observed a similar result using a slightly different manipulation. Although Italian has only regular spelling-to-sound mappings, stress assignment in multisyllabic words is not regular. Although the majority of multisyllabic words are stressed on the penultimate syllable (e.g., *lav-or'-o*), about 30% are stressed on the antepenultimate syllable (e.g., *tav'-ol-o*). The latter type of stress assignment can thus be labeled as irregular. As a result, it can be argued that under normal circumstances, these irregular words cannot be named correctly without lexical involvement because the specification of the correct stress pattern can come only from the stored pronunciation.

In Colombo and Tabossi's (1992) study, three-syllable words with their stress on the antepenultimate syllable (i.e., irregular-stress words) were presented either in a block with words having a regular stress pattern or in a block with other words with the same (irregular) stress pattern. Results indicated that an associative priming effect arose only in the mixed block. Apparently, when all the words have the same stress, even when it is an irregular one, it can be assigned by default. Thus, in that situation, it is only necessary to use information from the assembly route in order to generate a correct naming response.

These results all appear to suggest that lexical information can be virtually ignored if it is not necessary. The reader should note two things, however. First, although these results do suggest that lexical information can be ignored, they do not provide much evidence for the existence of a "lexical route" (i.e., a route that maps an orthographic representation to a lexical representation, which then allows retrieval of a phonological code in an essentially holistic fashion). As other researchers (e.g., Brown & Besner, 1987; Glushko, 1979; Humphreys & Evett, 1985) have suggested, there are a number of other ways in which lexical information could be used in the process of naming a word. Second, all of these studies involved languages with shallower orthographies than English, orthographies in which lexical information may play a less critical role than it does in English. As noted, Tabossi and Laghi (1992) were not able to obtain a similar effect with English stimuli. Thus, the question of whether it is possible for English readers to ignore lexical information is clearly an open one.

Investigations with English stimuli have, instead, centered on ways of manipulating the (presumably) more resource-demanding assembly route. Paap and Noel (1991), for example, demonstrated that requiring the participants to hold a set of digits in memory while naming words seemed to slow the assembly route. The result was the disappearance of the regularity effect for low-frequency words (although see Bernstein & Carr, 1996, and Pexman & Lupker, 1995, in press). The argument is that because the lexical route operates in a more automatic fashion, memory load should have a stronger impact on the resource-demanding assembly route. As a consequence, the assembly route would essentially play no role in naming when the memory load is high. Thus, no competition would arise, so that the only factor driving naming latency would be frequency, not regularity.

A more direct manipulation of readers' strategies was provided by Monsell, Patterson, Graham, Hughes, and Milroy (1992). As in the studies described above, Monsell et al. attempted to change the way in which readers named words by manipulating the nature of the context in which the words appeared. The idea was that if readers know that all the words to appear in a block will be irregular, they may be able to de-emphasize (e.g., slow down) the output from the assembly route. This would lead to less competition and, hence, to faster naming latencies in comparison with a condition in which the assembly route is necessary, in particular, when the irregular words are mixed with nonwords.

The results from Monsell et al.'s (1992) Experiment 2 do provide some support for this hypothesis. Of most importance for the present investigation, *high*-frequency irregular words were named faster when presented by themselves (the pure block) than when presented mixed with nonwords (the mixed block; hereinafter, this effect is referred to as the "pure-block response time [RT] advantage"). For the low-frequency irregular words, however, there was little evidence of a pure-block RT advantage, although there was an indication that pure blocks provided some benefit in terms of the number of regularization errors. That is, participants tended to make slightly fewer regularization errors with low-frequency words (a total of 15 fewer regularization errors over the 32 participants) in the pure blocks than in the mixed blocks. On the basis of these results, Monsell et al. offered as one of their two main conclusions that "skilled readers can (to some degree) reduce the impact of the sublexical assembly process on their naming performance; . . . it is functionally equivalent to slowing the sublexical process relative to the lexical process" (p. 463).

On the other hand, Monsell et al.'s (1992) results also raise some questions about the viability of their de-emphasis explanation. In particular, if readers really were able to de-emphasize the assembly route in pure blocks, it is surprising that there was so little evidence of blocking effects for the low-frequency words. It is not high-frequency irregular words, but low-frequency irregular words, that are assumed to be more affected by competition from the assembly route, because the lexical route is presumed to be slower at processing low-frequency words. Thus, if the assembly route really were de-emphasized in the pure blocks, it would seem that low-frequency irregular words should have benefitted at least as much as high-frequency irregular words. Nonetheless, as Monsell et al. noted, not only do their own results indicate that pure-block RT advantages tend to be limited to high-frequency words but the two previous articles that provided an evaluation of this issue (Andrews, 1982; Fredericksen & Kroll, 1976) also seemed to show the same pattern.[1]

---

[1] Monsell et al. (1992) offered an explanation of what "de-emphasizing the assembly route" might mean in terms of the distributions of processing times for words on the lexical and assembly routes (p. 463). Unfortunately, this explanation is also problematic. Their suggestion is that de-emphasizing the assembly route is equivalent to increasing the mean of the distribution of processing times on the assembly route (and possibly changing the variance). The result would be a reduction in the probability that the assembly route would produce an incorrect, competing phonological code before the lexical route produced the correct code for high-frequency words, speeding naming of high-frequency words in the de-emphasized situation.

To actually calculate the relevant probabilities of the assembly route providing an incorrect code before the lexical route can provide the correct code, one must consider the distribution of differences between processing times for the two routes. This distribution has a mean equal to the difference between means of the assembly and lexical distributions and a variance equal to the sum of their two variances. If one knew these values, one could calculate two Z scores (one in the case where the assembly route was de-emphasized and one in the case where it was not), based on the zero point in the distribution of differences, in order to determine the probabilities that the assembly route would produce a code first in each of the two situations.

The point to be made with respect to the high-frequency lexical distribution is simply that the effect this de-emphasizing could have would, of necessity, be a minor one. That is, because of the large difference between the mean processing time for the high-frequency words on the lexical route and the mean processing time for the assembly route either in the de-emphasized situation or in the normal situation, these calculations for the high-frequency words would both involve rather large absolute Z scores and, hence, rather small probabilities. Thus, the *difference* between these two probabilities would also, of necessity, be quite small. Nonetheless, it is possible that with a large enough interference effect from competing incorrect assembled codes, even a very small difference in probabilities could produce the observed 15-ms pure-block RT advantage.

Where things go awry is in the consideration of what must then happen for low-frequency irregular words. The relation between the distribution of processing times for low-frequency words on the lexical route and the distribution of processing times on the assembly route is quite different. In particular, the means of these two distributions are quite similar. Thus, the mean of the distribution of differences would be near zero. De-emphasizing the assembly route (i.e., shifting the assembly distribution to the right) would produce a change of Z scores from a small negative value to a small positive value in the distribution of differences.

More recent evidence (Coltheart & Rastle, 1994; Kinoshita & Woollams, 1996) also raised concerns about the viability of the route de-emphasis explanation. These investigators suggested that if readers really do place more emphasis on the assembly route when nonwords are present, as Monsell et al. (1992) claimed, competition in the naming of low-frequency irregular words should be stronger in blocks containing nonword fillers than in blocks containing only words. The result should be a larger regularity effect in blocks containing nonword fillers. What these investigators reported, however, was that the size of the low-frequency regularity effect was essentially unaffected by the presence of nonword fillers.

The purpose of the present article was to provide a further investigation of the proposal that readers can strategically de-emphasize their assembly route. Before proceeding, however, we should note that Monsell et al. (1992) offered a second conclusion about readers' ability to strategically control the naming process: "Skilled readers can (to some degree) adjust their readiness to initiate articulation of the currently available phonological description on the basis of the anticipated time course of the generation of phonology by means of the lexical process" (p. 463). This conclusion was based on the fact that nonwords were named faster in pure blocks than when mixed with either mixed-frequency or low-frequency irregular words. The idea is that when only nonwords are being named, whatever code the assembly route produces is generally acceptable and, hence, can be produced. When low-frequency irregular words are also contained in the block, however, participants delay producing an assembled code just in case there is a late-arriving, irregular code from the lexical route. As should become clear, our experiments do allow us to evaluate this conclusion as well, and our interpretation of this effect is a bit different from Monsell et al.'s. However, because this second conclusion does not relate to our basic question—the question of what control readers might have over supposedly automatic processes—it was not a focus of our initial experiments.

Apparently, Monsell et al.'s (1992) blocking manipulation induced a strategy change in their participants, at least in the blocks containing high-frequency words. The question is whether that effect can, in fact, be characterized as having been due to a de-emphasis of the assembly route. Experiment 1 was simply a replication of Monsell et al.'s Experiment 2, to confirm that the effect does tend to be limited to high-frequency words. The technique used was essentially identical to Monsell et al.'s. In particular, words and nonwords were named in 40-stimulus blocks that were either pure or mixed. Before each block, participants were told what type of stimuli they would be seeing in the block and were shown some example stimuli that were not used in the experiment proper. The first 16 stimuli in each block were regarded as warm-up stimuli to allow the participants time to settle into whatever strategy they found most useful for that block, with the data coming from the final 24 stimuli in the block. Among the word stimuli, high- and low-frequency words were always presented in separate blocks (i.e., the pure word blocks contained words from only one frequency class, and the mixed blocks contained only nonwords and words from one frequency class).

The only differences between Experiment 1 and Monsell et al.'s (1992) Experiment 2 were as follows: (a) Because Monsell et al. got cleaner data with monosyllables, only monosyllables were used here; (b) twice as many participants (and half as many stimuli) were used in the present experiment; (c) one set of nonwords was mixed only with high-frequency words, and another set of nonwords was mixed only with low-frequency words rather than counterbalancing nonword sets across word type; (d) there was no general practice block before the first block of experimental trials; and (e) no attempt was made to equate the stimuli on digram frequency because Monsell et al. reported that this factor had no bearing on their results.

Experiment 1

Method

Participants. The participants were 64 undergraduates from the University of Western Ontario who received course credit for their participation. All were native English speakers and had normal or corrected-to-normal vision.

Materials and design. Four sets of 80 letter strings were created. One set contained only high-frequency, irregular, monosyllabic words (mean frequency of 594.0, median frequency of 180.0, range from 43 to 9,816, according to Kučera & Francis, 1967; mean word length of 4.5 letters). These words (as well as the low-frequency irregular words) were compiled from sets of irregular words reported in the previous literature (e.g., Brown et al., 1994; Paap & Noel, 1991; Seidenberg et al., 1984). A second set contained only monosyllabic nonwords that were to be used with the high-frequency words. Each nonword was matched with a high-frequency word in terms of first phoneme. The third set contained only low-frequency, irregular, monosyllabic words (mean frequency of 6.1, median frequency of 6.0, range from 1 to 15, according to Kučera & Francis, 1967; mean word length of 4.7 letters). The fourth set contained only monosyllabic nonwords that were to be used with the low-frequency words. Each nonword was matched with a low-frequency word in terms of first phoneme. Each set of words was then divided into two half-sets of 40 stimuli so that half of the stimuli could be used in the pure blocks and half could be used in the mixed blocks. (This assignment of stimuli to block type was, of course, counterbalanced over participants such that each stimulus was seen only once by each participant.) A complete list of the stimuli is contained in Appendix A.

To create the pure blocks, 16 of the 40 stimuli in each half-set were selected to serve as warm-up stimuli, with the other 24 serving as the experimental stimuli. Thus, there were four pure

---

Because we are dealing with the middle of the distribution of differences, the probabilities involved here (and, hence, the change of probabilities that this involves) would be much more substantial. That is, de-emphasizing the assembly route would dramatically decrease the probability that the assembly route would provide an incorrect code before the lexical route could for low-frequency words. This, should, of course, cause a much larger pure-block RT advantage for low-frequency irregular words than for high-frequency irregular words, an effect that Monsell et al. (1992) did not observe.

blocks, one derived from each of the original stimulus sets (two containing words and two containing nonwords).

The mixed blocks were constructed by taking each of the other half-sets and designating 16 stimuli as warm-up stimuli and the other 24 as experimental stimuli. Half of each of these (8 warm-up and 12 experimental stimuli) were mixed with the appropriate other set, creating two mixed blocks of each type. Thus, there were also four mixed blocks, two containing high-frequency words and nonwords and two containing low-frequency words and nonwords. Each participant saw all eight blocks (i.e., four pure and four mixed blocks).

As previously noted, in order to counterbalance properly, the words and nonwords used in the mixed blocks also had to be used in the pure blocks (and vice versa). To accomplish this, these same procedures were carried out again with the roles of the half-sets being reversed (i.e., from mixed to pure). The stimuli that were originally assigned to be warm-up stimuli remained warm-up stimuli, and those originally assigned to be experimental stimuli remained experimental stimuli.

All participants saw either all the pure blocks first or all the mixed blocks first. For the pure blocks, there were four orders of presentation. For half the participants, the two word blocks came first, and for the other half of the participants, the two nonword blocks came first. Within each set of word and nonword pure blocks, the order of the two blocks was counterbalanced. For the mixed blocks, there were also four orders of presentation. For half the participants, the two low-frequency blocks came first, and for the other half of the participants, the two high-frequency blocks came first. Which of these two blocks came first within each frequency condition was also counterbalanced. Crossing these four orders with the factors of whether the pure or mixed blocks came first and with the assignment of half-sets to either the pure or mixed blocks created 16 different sequences of conditions. Four participants received each sequence.

Equipment. Stimuli were presented 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 (MEL; Schneider, 1988).

The stimuli were presented in the center of the screen in lowercase letters. Each letter was approximately 0.7 cm high and 0.4 cm wide. Participants were seated approximately 45 cm from the screen. Thus, the visual angle of a four-letter word was approximately 2.0 degrees.

Procedure. Each participant was tested individually. Participants were told that they would be seeing a number of letter strings to name and that they should name them as rapidly and accurately as possible. They then started the eight blocks of the experiment. Prior to each block, they were told what type of stimuli the block would contain and were verbally presented with examples drawn from the warm-up stimuli. The 40 trials in the block then commenced. Each trial begin with a 750-ms presentation of a fixation point. The letter string followed immediately and remained on the screen until the participant named it. The interval between the start of the naming response and the re-presentation of the fixation point was 750 ms. Participants were given no feedback on either naming latencies (RTs) or error rates during the experiment.

Results

A trial was considered an error and its RT was not entered in the RT analysis if (a) the pronunciation was incorrect, (b) the RT was longer than 1,500 ms or shorter than 150 ms, or (c) the participant spoke too softly to trigger the voice key. Only mispronunciations and long RTs were considered as errors for purposes of the error analysis of variance (ANOVA).[2] In all cases, the criterion for statistical significance was two-tailed, $p < .05$, unless otherwise stated.

Table 1

Mean Response Times (RTs) and Error Rates (ERs) in Experiment 1 as a Function of Stimulus Type and Blocking Condition

	Condition					
	Pure		Mixed		Effect	
Stimulus type	RT	ER	RT	ER	RT	ER
HF words	463	2.3	485	2.7	+22	+0.4
LF words	563	9.4	547	11.8	−16	+2.4
Nonwords with HF words	554	4.8	535	8.2	−19	+3.4
Nonwords with LF words	555	4.4	564	7.0	+9	+2.6

Note. The mixed conditions were created by mixing words from a particular frequency class with nonwords. Thus, words from different frequency classes were not mixed with one another. HF = high frequency; LF = low frequency.

In both the RT and error ANOVAs, word data were analyzed separately from nonword data.[3] In both situations, the design was a 2 (frequency) × 2 (block) within-subject design; however, it should be kept in mind that when discussing the nonword analyses, *frequency* refers only to the words used as context in the mixed blocks. The data from Experiment 1, both RTs and error rates, are shown in Table 1.

Word RTs. Both the frequency effect, $F(1, 63) = 267.63$, $MSE = 1{,}580.5$, and the Frequency × Block interaction, $F(1, 63) = 23.20$, $MSE = 1{,}055.6$, were significant. The frequency effect was due to participants responding 81 ms faster to high-frequency words. The interaction was due to a pure-block RT advantage (i.e., shorter RTs in the pure-block condition) for high-frequency words, but a mixed-block RT advantage (i.e., shorter RTs in the mixed-block condition) for low-frequency words. Simple main effects analyses showed that the 22-ms pure-block RT advantage for the high-frequency words was significant, $t(63) = 3.74$, as was the 16-ms mixed-block RT advantage for the low-frequency words, $t(63) = -2.15$.

---

[2] The percentage of trials that exceeded the 1,500-ms cutoff was approximately 1% in all experiments, ranging from a low of 0.98% in Experiment 4 to a high of 1.14% in Experiment 3. Because these percentages were so low, the distribution of percentages over conditions is not reported.

[3] Because the same targets appeared in the pure and mixed conditions for different participants, error variance due to items contributed to the expected mean squares for the conventional Subject × Block and Subject × Block × Frequency error terms. Thus, although only the conventional F values were calculated, any block main effects, any Frequency × Block interaction effects, or any simple main effect involving block should generalize over items. Although error variance due to items did not contribute to the conventional error terms for the frequency main effect or the other main effect analyzed in Experiment 4 (i.e., regularity), any issues about the generalizability of these variables over items are easily resolved by noting the replicability of these effects in the literature.

Word errors. Both the frequency effect, $F(1, 63) = 136.77$, $MSE = 30.7$, and the block effect, $F(1, 63) = 7.24$, $MSE = 19.0$, were significant, because the error rate was 8.1% higher for low-frequency words and 1.4% higher in the mixed blocks, respectively.

Table 2 contains the classifications for all the pronunciation errors as a function of the frequency and block variables. The categories and classification scheme are the same as those used by Monsell et al. (1992). As the reader can see, more regularization errors were made in the mixed block than in the pure block but only for the low-frequency words, essentially replicating Monsell et al.'s pattern of results.

Nonword RTs. Both the frequency effect, $F(1, 63) = 10.32$, $MSE = 1{,}385.3$, and the Frequency × Block interaction, $F(1, 63) = 6.77$, $MSE = 1{,}854.5$, were significant. The frequency effect was due to participants responding 15 ms faster to the nonwords that were mixed with the high-frequency irregular words. This effect was qualified, however, by the significant interaction, which indicates that the effect of mixing nonwords with high-frequency words and the effect of mixing nonwords with low-frequency words were different. Simple main effects analyses indicated that the 19-ms mixed-block RT advantage for the nonwords mixed with high-frequency words was significant, $t(63) = -2.27$; however, the 9-ms pure-block RT advantage for the nonwords mixed with low-frequency words was not significant, $t(63) = 0.93$, *ns.*

Nonword errors. Only the block effect was significant, $F(1, 63) = 21.36$, $MSE = 26.3$. This effect was due to the error rate being 3.0% higher in the mixed blocks.

Discussion

In its most important aspects, Experiment 1 produced results similar to the results of Monsell et al. (1992). In particular, high-frequency irregular words were pronounced significantly faster in the pure than in the mixed blocks. Further, as in Monsell et al.'s experiment, there was a tendency for more regularization errors for the low-frequency irregular words to occur in the mixed blocks than in the pure blocks. Finally, there was no tendency for the low-frequency words to be named faster in the pure-block condition. On the contrary, these words were named significantly faster in the mixed-block condition.

Table 2
Frequencies of Types of Pronunciation Errors in Experiment 1 for Words

	Block type			
	High frequency		Low frequency	
Type of error	Pure	Mixed	Pure	Mixed
Regularization	12	11	81	104
Confusion	16	6	28	35
Other fluent	0	0	9	15
Dysfluent	0	1	5	3
Total	28	18	123	157

Note. There were 1,536 trials of each block type in this experiment. Errors on 8 low-frequency trials could not be classified because the nature of the error was not recorded.

These data do differ from Monsell et al.'s (1992) results in three ways, however. To begin with, the pure-block RT advantage for the nonwords mixed with low-frequency words was not significant here. The size and the direction of the effects, however (9 ms in Experiment 1; 18 ms in Monsell et al.'s, 1992, Experiment 2), were reasonably similar. There was, however, a much larger discrepancy between the two experiments in the low-frequency word conditions and in the nonwords mixed with high-frequency words conditions. In both of these situations, there was actually a significant mixed-block RT advantage. There had been a trend in this direction in Monsell et al.'s results for the nonwords mixed with high-frequency words; however, there had been no hint of such an effect for the low-frequency words.

Independent of the question of replication, there would appear to be no obvious reason within the dual-route framework for observing a mixed-block RT advantage with any type of stimuli. Presumably, blocks that contain only one type of stimulus present the optimal opportunity for readers to balance their reliance on the two routes in a way that produces the shortest possible naming latencies. Assuming that readers can adopt strategies of this nature, introducing a second type of stimulus into those blocks can only upset that balance, which should lead to longer naming latencies in the mixed-block conditions. Thus, these particular results do not easily lend themselves to an explanation in terms of a de-emphasis of routes.

What should, of course, be noted is that these two mixed-block RT advantages were accompanied by pure-block advantages in error rate. Thus, what appears to have been at work here is some sort of speed–accuracy trade-off. The other point to note, however, is that the process that generated these trade-offs cannot have been a simple one. That is, these results cannot simply have been due to the fact that participants decided to trade accuracy for speed in the mixed blocks. If so, the other stimuli in the mixed blocks (the high-frequency irregular words and the nonwords mixed with low-frequency words) should also have shown the same pattern (i.e., a mixed-block RT advantage accompanied by a pure-block error advantage). Clearly, they did not. As we argue subsequently, however, the processes that were at work here do affect the way in which participants balance the trade-off between accuracy and speed.

Experiment 2

The purpose of Experiment 2 was to provide a more direct examination of the route de-emphasis explanation of the pure-block RT advantage for high-frequency words. The explanation is based on the idea that the assembly route provides incorrect information when naming irregular words. Thus, it is to a reader's advantage to de-emphasize its contributions whenever possible (e.g., in the pure blocks). Such would not be the case for regular words, however. For regular words, any information that the assembly route could provide would be accurate and, hence, helpful. As such, de-emphasizing the assembly route in pure blocks with regular words would appear to be a counterproductive strategy because useful information would be ignored. The effect would then be, if anything, to create a pure-block RT disadvantage (i.e., a mixed-block RT advantage).

To investigate this issue, Experiment 2 was directly analogous to Experiment 1, except that the words were all regular. According to the route de-emphasis explanation, the most straightforward expectation is that the assembly route would be kept maximally active in both the pure and mixed blocks, and thus, there should be no effect of the pure versus mixed manipulation. On the other hand, one could make the argument that de-emphasizing the assembly route in the pure word blocks might have some positive benefit even for regular words because it would allow a reallocation of resources. That is, if the assembly route were de-emphasized, it may free up resources that could then be used by the lexical route to more than counteract whatever is lost in terms of useful information from the now de-emphasized assembly route. Thus, it may be possible for a de-emphasis explanation to account for a pure-block RT advantage for high-frequency regular words if one is observed. We evaluate the viability of a reallocation-of-resources explanation in the *Discussion* section of this experiment.

Method

Participants. The participants were 32 undergraduates from the University of Western Ontario who received course credit for their participation. All were native English speakers and had normal or corrected-to-normal vision. None had participated in Experiment 1.

Materials, design, equipment, and procedure. Two sets of 80 monosyllabic regular words were selected for use in this experiment. These words were compiled from sets of regular words reported in the literature (e.g., Brown et al., 1994; Paap & Noel, 1991; Seidenberg et al., 1984). One set contained only high-frequency words (mean frequency of 499.2, median frequency of 150.0, range from 41 to 7,289, according to Kučera & Francis, 1967; mean word length of 4.2 letters). The other set contained only low-frequency words (mean frequency of 5.8, median frequency of 5.0, range from 1 to 14 according to Kučera & Francis, 1967; mean word length of 4.2 letters). These words were mixed with the nonwords used in Experiment 1 in the same way that the irregular word sets had been mixed in Experiment 1. A complete list of these words is contained in Appendix B.

In choosing the regular words, an attempt was made to select sets of words that matched the sets of irregular words as closely as possible in terms of frequency and length. An additional criterion for selecting the regular words was that each matched a nonword in the set of nonwords to be used with those words in terms of first phoneme. That is, each high-frequency word was matched to a nonword in the set of nonwords to be used with high-frequency words, and each low-frequency word was matched to a nonword in the set of nonwords to be used with low-frequency words. Because the same had been done when selecting words and nonwords in Experiment 1 (the high- and low-frequency words had been matched with their nonwords in terms of first phoneme), the result was that the two sets of high-frequency words and the two sets of low-frequency words had also been matched on first phoneme. Finally, in general, most of the words selected (64 in each set of 80) would be classified as "regular-consistent" words.

The counterbalancing was identical to that in Experiment 1 with the regular words replacing the irregular words. Thus, there were again 16 different sequences of conditions. Two participants received each sequence. The equipment and procedure were also the same as in Experiment 1.

Results

The error criteria, design, and ANOVAs were identical to those of Experiment 1. The data from Experiment 2, both RTs and error rates, are shown in Table 3. Again, it should be kept in mind that when discussing the nonword analyses, *frequency* refers only to the words used as context in the mixed blocks.

Word RTs. Both the frequency effect, $F(1, 31) = 63.45$, $MSE = 558.6$, and the block effect, $F(1, 31) = 15.90$, $MSE = 807.3$, were significant. The interaction was not significant, $F(1, 31) = 2.07$, $p > .15$, $MSE = 635.1$. The frequency effect was due to participants responding 34 ms more rapidly to high-frequency words. The block effect was due to a 21-ms pure-block RT advantage.

Word errors. Neither the main effects nor the interaction was significant.

Nonword RTs. Both the frequency effect, $F(1, 31) = 5.39$, $MSE = 756.0$, and the block effect, $F(1, 31) = 7.79$, $MSE = 1{,}832.8$, were significant. There was no hint of an interaction ($F < 1$). The frequency effect was due to participants responding 12 ms faster to the nonwords used with high-frequency regular words. The block effect was due to a significant 21-ms mixed-block RT advantage.

Nonword errors. Neither the main effects nor the interaction was significant.

Discussion

In the most important respect, the results of Experiments 1 and 2 were identical. That is, there was a 20+ ms pure-block RT advantage for the high-frequency words and an approximately 20-ms mixed-block RT advantage for the nonwords mixed with them. In contrast, unlike in Experiment 1, the

Table 3
Mean Response Times (RTs) and Error Rates (ERs) in Experiment 2 as a Function of Stimulus Type and Blocking Condition

	Condition					
	Pure		Mixed		Effect	
Stimulus type	RT	ER	RT	ER	RT	ER
HF words	442	1.4	469	1.4	+27	0.0
LF words	482	2.2	496	2.9	+14	+0.7
Nonwords with HF words	541	6.2	520	9.8	−21	+3.6
Nonwords with LF words	552	6.6	532	6.9	−20	+0.3

Note. The mixed conditions were created by mixing words from a particular frequency class with nonwords. Thus, words from different frequency classes were not mixed with one another. HF = high frequency; LF = low frequency.

low-frequency regular words and the nonwords mixed with them behaved exactly like the high-frequency words and their nonwords. As argued, a pure-block RT advantage for high-frequency words in Experiment 2 does not easily follow from the route de-emphasis account of the parallel effect in Experiment 1. That is, because all the words were regular, there would have been no reason for participants to de-emphasize the assembly route in the pure blocks in Experiment 2. Thus, there would be no reason to expect any effect of the blocking manipulation in Experiment 2. In fact, if, for some reason, participants actually did choose to de-emphasize the assembly route in the pure blocks, the expected result would not be a pure-block RT advantage. Rather, the more obvious prediction would be a mixed-block RT advantage because participants would no longer be getting potentially useful information from the assembly route in the pure blocks.

The following question then emerges: Can the route de-emphasis explanation provide some way of accounting not only for the pure-block RT advantage for the high-frequency words but also for the entire set of results observed in Experiment 2? Two possibilities suggest themselves. One possibility, suggested earlier, would be based on reallocation of attention. Because the assembly route is assumed to be somewhat attention demanding (Paap & Noel, 1991), de-emphasizing this route in the pure block could be useful because it would free up resources that the lexical route could then use. Whatever is saved in this way may make up for any disadvantage that may arise by de-emphasizing a useful source of information.

An account of this sort would have problems of its own, however. To begin with, the standard assumption has been that the lexical route demands few resources (e.g., Paap & Noel, 1991), especially for high-frequency words. Thus, it seems unlikely that the processing of high-frequency words on the lexical route could be facilitated to any large degree by whatever resources might have been freed up by de-emphasizing the assembly route.

Second, this account would also have a great deal of difficulty explaining why the low-frequency words also showed a pure-block RT advantage. For low-frequency words, the general assumption has been that the assembly route is essentially as fast as and, hence, as important as the lexical route. What necessitates this assumption is the standard result (e.g., Brown et al., 1994; Paap & Noel, 1991; Seidenberg, 1985; Seidenberg et al., 1984) that the frequency effect for regular words is quite small. The explanation has been that processing on the assembly route must be fast enough to make up for the relatively slow processing of these words on the lexical route. Thus, if the assembly route really had been de-emphasized in the pure blocks, it would be extremely unlikely that the naming of low-frequency regular words would actually have been faster in those blocks.

Finally, the large mixed-block RT advantages for the nonwords would remain unexplained. Making the assumption that the assembly route was de-emphasized in the pure blocks for words implies that this route played a somewhat larger role in the mixed blocks. What should also be true, however, is that the assembly route should have played at least as large a role in the pure blocks with nonwords because nonwords can be named only by the assembly route. If so, this account would predict that nonwords would be named at least as fast in the pure blocks as in the mixed blocks, which, of course, is not what was observed.

A second alternative account, in terms of the strategic adjustments of routes, could be based on the assumption that in Experiment 2, it was not the assembly route that was being de-emphasized but the lexical route. As noted, there are a number of studies suggesting that readers may have the ability to ignore lexical information (Baluch & Besner, 1991; Colombo & Tabossi, 1992; Tabossi & Laghi, 1992), although there is not as yet any evidence that this can be done in English. Nonetheless, the following account could be proposed. Because all the stimuli in the mixed-block conditions could be named by the assembly route but only half of them could be named by the lexical route, less emphasis may have been placed on the lexical route in the mixed-block conditions than in the pure-block conditions involving words. With less emphasis placed on the lexical route, the expectation would be that both high-frequency and low-frequency words would slow down in the mixed-block conditions (with the effect being somewhat greater for the high-frequency words), yielding the observed pure-block RT advantages.

A problem for this account is posed by the nonword data. Any strategy that could have been invoked to de-emphasize the lexical route in the mixed blocks could, presumably, also have been invoked for the pure blocks with nonwords. In fact, there would have been even more motivation to invoke a strategy of this sort in the pure blocks with nonwords than in any of the other conditions. The result would be that, at the very least, there should have been no difference between the pure and mixed blocks for nonwords and possibly that there should even have been a pure-block RT advantage. As such, the mixed-block RT advantages for nonwords would appear to provide a fairly strong argument against this possibility.

What appears to be a more fruitful way to look at the present data is based on noting a common trend. Whenever stimuli that were named relatively rapidly in pure blocks (e.g., regular words—hereinafter referred to as "fast stimuli") and stimuli that were named more slowly in pure blocks (e.g., nonwords—hereinafter referred to as "slow stimuli") were mixed in the same block, naming latencies for the fast stimuli increased, and naming latencies for the slow stimuli decreased. This description characterizes exactly the results of Experiment 2, as well as the results with high-frequency words in both Experiment 1 of the present article and in Monsell et al.'s (1992) Experiment 2. (Note, however, that this description does not characterize the circumstance of mixing low-frequency irregular words and nonwords in Experiment 1 of the present article or in Monsell et al.'s Experiment 2. That is, those conditions did not represent the mixing of fast and slow stimuli but rather represented the mixture of two different types of slow stimuli. A discussion of possible strategies that might have been invoked in those circumstances is presented in the General Discussion.)

A fairly straightforward interpretation of this type of pattern could be proposed by invoking the concept of a criterion that is used to control the timing of output (e.g., Patterson, Seidenberg, & McClelland, 1989). We start by making the assumption that the process that accepts a phonological code and turns it into an articulatory code is essentially an assembly–synthesis process that works in cascade with the phonological code generation process. Thus, it is not the case that an articulatory code simply becomes "available" at some point in time but that the viability of the articulatory code increases continually over time. As such, if there is some time pressure to respond, it would be possible to terminate this process prior to completion and to start articulation of a best guess response.

The more important assumption for our purposes would be that in order to maintain an acceptable level of accuracy in any naming task and at the same time to produce responses acceptably rapidly, participants set a time criterion for when articulation should begin. The position of this criterion would be determined mainly by the perceived difficulty of the stimuli being named. That is, the key factor would be the perceived average strengths of the stimulus–response (S-R) mappings (i.e., the orthography-to-phonology mappings) for the stimuli in the block. Thus, it would take some trials in each block before the position stabilizes. Once stabilized, however, the criterion would act as a flexible guide for the beginning of articulation for all subsequent responses in the block.

In the present tasks, we assume that the 16-trial warm-up in each block helped the participants to become familiar with the type of stimuli and, hence, with the average difficulty of the stimuli in the block. Thus, at the point that the 24 experimental trials began, the criterion had essentially stabilized at a position that participants felt was as appropriate as possible for all the stimuli in the block. If the stimuli were fairly homogeneous, as they would have been in pure blocks, the criterion setting would have been appropriate for most stimuli in the block. That is, for virtually all stimuli, the quality of the articulatory code would have been sufficient to support an acceptably accurate and rapid response. When easy and difficult (i.e., fast and slow) stimuli were mixed together, however, the criterion would have tended to stabilize at a point that was beyond the preferred responding point for the fast stimuli but prior to the preferred responding point for the slow stimuli. The result was that the processing of fast stimuli would have tended to continue for longer than it needed to, causing them to be named more slowly than in the pure block. The processing of slow stimuli, on the other hand, would have tended to be rushed, causing them to be named more rapidly than in the pure block.

The existence of a time criterion to guide the start of articulation in mixed blocks would not necessarily have produced equal naming latencies for all stimuli in the block, however. The articulatory code for a slow stimulus would not always have been sufficiently complete to start articulation when the criterion was reached. Similarly, if the articulatory code for a fast stimulus was ready substantially before the criterion was reached, the participant may have begun articulation at that point. Nonetheless, the major result would be a trend toward homogenization of naming latencies, as observed in these experiments.

This tendency toward homogenization should, of course, have had consequences for the error rates, as would be expected in any explanation based on setting a time criterion. That is, when articulation was started before it normally would have been, more errors would be expected, and when articulation was delayed, accuracy should have improved. There were, in fact, noticeable trends of this nature in the present experiments (although many were nonsignificant). A more detailed discussion of how different criterion placements may influence error rates in naming tasks in general is presented in the General Discussion.

We view our particular proposal as, basically, just a specific version of a general, criterion-setting framework proposed by investigators working in other realms (e.g., Krueger, 1985; Sperling & Dosher, 1986; Strayer & Kramer, 1994a, 1994b; Treisman & Williams, 1984). Empirically, Strayer and Kramer's (1994a, 1994b) results obtained in a visual-scanning and memory-scanning task, in fact, provide a rather interesting parallel to the present results. What those authors demonstrated is that the normal processing advantage enjoyed by stimuli in consistent-mapping (CM) conditions over stimuli in varied-mapping conditions in "pure blocks" decreased in "mixed blocks." Further, the shrinkage of the size of the effect was due both to an increase in latencies for CM targets and to a decrease in latencies for varied-mapping targets. In essence, latencies in the two conditions homogenized, just as demonstrated in the present experiments.[4]

According to Strayer and Kramer's (1994a, 1994b) interpretation, participants in the pure-block condition with CM stimuli could take advantage of the strength of the S-R mappings in order to decrease the criterion for responding. That is, the lower setting of the criterion in blocked-CM conditions (the pure blocks) was due to a response bias based on the knowledge that there would be a rapid buildup of perceptual evidence for *all* stimuli in the block (because, due to practice, their processing had become highly automatized).

---

[4] Strayer and Kramer's (1994a, 1994b) data also illustrate some other interesting points. First, they found that their participants had a clear inability to consciously alter criterion placements during a block of trials. For example, in the mixed blocks, cuing participants up to 1,500 ms before the trial as to whether it would involve CM stimuli or varied-mapping stimuli did not have any effect. Thus, participants had limited ability to make major changes in criterion placement on-line, at least within this time frame. On the other hand, the nature of the previous trial did seem to have a minor effect on criterion placement. In fact, a long run of CM trials in a mixed block produced performance on a subsequent CM trial that was similar to performance on CM trials in the pure-block condition (although the analogous effect did not occur following a long run of varied-mapping trials).

Second, although the mixing manipulation tended to homogenize RTs, there was still often a discernible advantage for CM trials in mixed blocks. Thus, the argument is that, although criterion placement homogenizes RTs, it does not dictate them. True processing differences can still emerge.

In the present situation, high-frequency words would represent the type of condition in which the S-R associations (the mapping between orthographic and phonological representations) have become highly practiced and should therefore be highly automatized. Thus, this high consistency may be exploited in a blocked condition where only high-frequency words are presented by setting a very strict criterion. For low-frequency irregular words or nonwords, however, the orthography-to-phonology mapping is weaker. Thus, the criterion setting in a pure block of low-frequency irregular words or nonwords should be higher. Finally, in the mixed condition, the S-R mappings would have an average strength intermediate to that found in the two pure blocks. The expectation would, therefore, be that an intermediate setting for the criterion would be established.

We are, of course, not the first to invoke the notion that a time criterion affects performance in a naming task. As noted, a time criterion was also a part of Patterson et al.'s (1989) model. The specific instantiation of the criterion concept was a bit different, however, and as Monsell et al. (1992) argued, having the concept in their model does not allow the model to do a particularly good job of accounting for Monsell et al.'s results (nor would the model do a particularly good job of accounting for the present results).

As also noted, Monsell et al.'s (1992) model invokes the idea of a criterion. In their model, the criterion provides a way of explaining why nonwords are named faster in pure blocks than when mixed with low-frequency irregular words in their experiments (an effect that we had only partial success in replicating in our Experiment 1). The idea is that readers can use the criterion to delay responding when it might be to their advantage to do so. It is unclear, however, whether the criterion plays any other role in their model. Thus, their view of the role of the criterion seems to be a bit different from the role we assign to it, that of a prime determinant of naming latencies that is strongly affected by the relative processing difficulty of the stimuli used. In some sense, then, what we call a criterion and what Monsell et al. called a criterion may actually be somewhat different concepts.

Experiment 3

Experiment 3 was an attempt to test directly a prediction of the criterion-setting account of the block effects. In a standard naming task, the typical finding is that there are fairly large frequency effects for irregular words when high- and low-frequency irregular words are mixed in the same block (e.g., Paap & Noel, 1991; Seidenberg et al., 1984). The criterion-setting account predicts that if those same words are presented in separate blocks, the frequency effect would be even larger because the latencies for high-frequency words would decrease and the latencies for low-frequency words would increase. That is, in the pure blocks, the participants would now be free to adopt different criterion settings for the two types of words. The setting for the high-frequency words should be somewhat more strict because processing of those words is easier than processing of the low-frequency words, whereas the setting for the low-frequency words should be somewhat less strict. The result should be a larger frequency effect in comparison with the condition in which a single criterion setting is applied to both types of words.

On the other hand, the dual-route framework, when expanded to allow the de-emphasis of either route, suggests that because all the stimuli are irregular and no nonwords are included, the most useful strategy would be to de-emphasize the assembly route as much as possible in all conditions. Thus, no effect of this type of block manipulation would be expected.

Method

Participants. The participants were 32 undergraduates from the University of Western Ontario who received course credit for their participation. All were native English speakers and had normal or correct-to-normal vision. None had participated in either of the previous experiments.

Materials, design, equipment, and procedure. The words used in this experiment were the high- and low-frequency irregular words from Experiment 1. They were divided into half-sets in the same way as in Experiment 1 with the stimuli maintaining their designation as warm-up or experimental stimuli. One half-set of each frequency class was used in the pure-block condition, and the other half-set of each frequency class was further divided in half (as in Experiment 1) to allow the creation of two mixed blocks.

The design was essentially the same as in the previous experiments. Each half-set of words was used in the pure blocks for half the participants and in the mixed blocks for the other half. Half the participants received the two pure blocks first, and half received the two mixed blocks first. Within both the pure- and the mixed-block conditions, the order of the two blocks was also counterbalanced. Thus, there were eight sequences of conditions with 4 participants receiving each sequence. The equipment and all other aspects of the procedure were identical to those in previous experiments.

Results

Error criteria were the same as in the first two experiments. In both the RT and error ANOVAs, the design was a 2 (frequency) × 2 (block) within-subject design. The data from Experiment 3, both RTs and error rates, are shown in Table 4.

RTs. Both the frequency effect, $F(1, 31) = 116.45$, $MSE = 1{,}372.4$, and the Frequency × Block interaction, $F(1, 31) = 8.75$, $MSE = 2{,}187.6$, were significant. The frequency effect was due to participants responding 71 ms more rapidly to high-frequency words. The interaction was

Table 4
Mean Response Times (RTs) and Error Rates (ERs) in Experiment 3 as a Function of Word Type and Blocking Condition

	Condition					
	Pure		Mixed		Effect	
Word type	RT	ER	RT	ER	RT	ER
High-frequency words	488	3.6	513	1.8	+25	−1.8
Low-frequency words	583	12.0	559	12.2	−24	+0.2

due to a pure-block RT advantage for high-frequency words, but a mixed-block RT advantage for low-frequency words. Planned comparisons showed that both the 25-ms pure-block RT advantage for the high-frequency words, $t(31) = 2.68$, $p < .01$ (one-tailed), and the 24-ms mixed-block RT advantage for the low-frequency words, $t(31) = -1.70$, $p < .05$ (one-tailed), were significant.

Errors. The only significant effect was the main effect of frequency, $F(1, 31) = 124.00$, $MSE = 22.7$. This effect was due to the error rate being 9.4% higher for the low-frequency words.

Discussion

The results came out exactly as predicted by the criterion-setting account. High-frequency irregular words were named faster in pure blocks, and low-frequency irregular words were named faster in mixed blocks. Thus, these data provide very clear support for the claim that participants' responding was in part controlled by time criteria in a way that led to a homogenization of RTs in mixed blocks.

The major implication of these results is that an explanation in terms of time criteria provides a much more parsimonious explanation of Monsell et al.'s (1992) pure-block RT advantage for high-frequency irregular words than an explanation in terms of de-emphasizing routes. What should also be noted, however, is that these results and this analysis should not be construed as evidence against the dual-route model per se (or against the version of it offered by Monsell et al., 1992). The explanation of the block effects that is being presented here is not based on how readers derive phonological codes from print but is based on the process of transcoding phonological codes into articulatory codes. The dual-route model makes few, if any, assumptions about this process, and the notion of a criterion for the start of articulation is not at all inconsistent with any aspect of the model.

In addition to providing support for the criterion-setting account, the results of Experiment 3 also make one very clear empirical point. That is, the frequency effect for irregular words was substantially smaller when high- and low-frequency words were mixed than when they were presented in separate blocks. If this empirical phenomenon generalizes to other effects, as the criterion-setting account predicts it should, an interesting empirical question arises: How many of the null effects reported in the literature were actually due to criterion placement in a mixed block and would, therefore, become significant if the stimuli were presented in pure blocks (so that participants could adopt a more suitable criterion placement for each type of stimulus)? In Experiment 4, we investigated the one effect of this type that is most central to the issue of modelling phonological-code retrieval: the lack of a regularity effect with high-frequency words.

Experiment 4

As noted earlier, the lack of a regularity effect with high-frequency words is a pervasive finding (Brown et al., 1994; Paap & Noel, 1991; Seidenberg, 1985; Seidenberg et al., 1984; but see Content, 1991, and Jared, 1995). The dual-route model explains this result by assuming that the lexical route is so fast at delivering the phonological code for high-frequency words that it almost always beats the assembly route. Hence, there is no competition and no regularity effect.

It should be noted, however, that in the dual-route framework, the issue of whether there would be a regularity effect for high-frequency words was originally an empirical question rather than a prediction of the model itself. If a regularity effect for high-frequency words had been the standard finding, it would simply have been assumed that the assembly route was fast enough to create competition even for high-frequency words, with no harm done to the model. However, given that whenever frequency has been explicitly manipulated there typically has been no regularity effect for high-frequency words in mixed-block conditions, there would be no obvious way for the model to account for an effect of this sort in pure-block conditions (without adding assumptions relating to time criteria).

Most parallel distributed processing (PDP) models (e.g., Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989), on the other hand, were designed so that they would specifically predict little or no regularity effect for high-frequency words for skilled readers. The way this type of model could explain a regularity effect for high-frequency words, if one were found in either pure or mixed blocks, would be to suggest that for the particular participants involved in the experiment, these words were not really highly familiar (i.e., they were not really high-frequency words for those participants). What the model could not do, however, would be to explain an interaction between block and regularity, that is, to explain how the size of this effect could differ between the pure and mixed blocks (again, without adding assumptions relating to time criteria).

In Experiment 4, regular and irregular high- and low-frequency words were presented either in pure blocks or mixed with words of similar frequency (i.e., high-frequency regular and irregular words were mixed together, and low-frequency regular and irregular words were mixed together). Considering first the high-frequency words, our analysis suggests the possibility that one reason that a regularity effect is generally not found with high-frequency words is that because of the placement of the criterion, RTs for regular and irregular high-frequency words are homogenized. If so, presenting the regular and irregular words in pure blocks may allow a regularity effect to emerge. On the other hand, if it is the case that high-frequency regular and irregular words are processed equally rapidly, as results of most experiments in the literature suggest, no regularity effect should emerge in either the mixed- or the pure-block conditions.

It should be noted that a comparison of Experiments 1 and 2 does suggest that we should obtain a high-frequency regularity effect in the pure blocks. That is, although the contrast is somewhat problematic because it is not a within-subject contrast, the pure-block naming latency for high-frequency regular words was 21 ms faster than the pure-block naming latency for high-frequency irregular words. Experiment 4 allowed us to determine: (a) whether this difference was a real one and (b) of equal importance, if the difference was real, whether it arose specifically because the words had been presented in pure blocks (i.e., whether it would disappear when our sets of high-frequency words were mixed together).

Considering now the low-frequency words, these words have inevitably shown a regularity effect. Thus, the prediction from the criterion-setting account for these words is simply that the effect would be larger in the pure blocks than in the mixed blocks. On the other hand, from a route de-emphasis perspective, the only prediction that can be made here is that mixing should hurt both the regular and irregular words. As noted earlier, pure blocks provide the optimal opportunity for readers to balance their reliance on the two routes in a way that produces the shortest possible naming latencies. Mixing can only serve to upset that balance, which, if anything, should harm performance for both word types. Depending on what balance of routes is established in the mixed blocks, in theory, this harm could be greater for the regular words, leading to a decrease in the regularity effect, or it could be greater for the irregular words, leading to an increase in the regularity effect. Thus, it is not possible to make a prediction about the relative sizes of the regularity effects in the two blocks. What is clear, however, is that this type of account could not predict, as the criterion-setting account could, that low-frequency irregular words would be named faster in the mixed block.

Method

Participants. The participants were 32 undergraduates from the University of Western Ontario who received course credit for participating in this experiment. All were native English speakers and had normal or corrected-to-normal vision. None had participated in any of the previous experiments.

Materials, design, equipment, and procedure. The stimuli were the two sets of 80 irregular words from Experiment 1 and the two sets of 80 regular words from Experiment 2. Half-sets from each of the four sets were presented in the pure blocks. The other half-sets were combined to form two blocks containing only low-frequency words and two blocks containing only high-frequency words. As in all previous experiments, each half-set was used in the pure-block conditions for half of the participants and in the mixed-block conditions for the other half.

The order and counterbalancing of conditions were essentially the same as those in Experiments 1 and 2. Thus, there were 16 sequences of conditions with 2 participants receiving each sequence. The equipment and all other aspects of the procedure were identical to those of the previous experiments.

Results

Error criteria were the same as in the other experiments. In both the RT and error ANOVAs, the design was a 2 (frequency) × 2 (regularity) × 2 (block) within-subject design. The data from Experiment 4, both RTs and error rates, are shown in Table 5.

RTs. The regularity effect, $F(1, 31) = 64.96$, $MSE = 1{,}051.2$; the frequency effect, $F(1, 31) = 121.40$, $MSE = 2{,}535.5$; and their interaction, $F(1, 31) = 72.98$, $MSE = 530.8$, were significant. More importantly, there were significant interactions of block and regularity, $F(1, 31) = 28.02$, $MSE = 756.1$, and of block, regularity, and frequency, $F(1, 31) = 6.06$, $MSE = 609.8$.

Table 5

Mean Response Times (RTs) and Error Rates (ERs) in Experiment 4 as a Function of Word Type and Blocking Condition

	Condition					
	Pure		Mixed		Effect	
Word type	RT	ER	RT	ER	RT	ER
HF regular words	453	2.3	455	1.7	+2	−0.6
HF irregular words	471	3.4	452	3.8	−19	+0.4
Regularity effect	+18	+1.1	−3	+2.1		
LF regular words	486	4.6	512	2.9	+26	−1.7
LF irregular words	569	10.5	543	13.3	−26	+2.8
Regularity effect	+83	+5.9	+31	+10.4		

Note. HF = high frequency; LF = low frequency.

A separate ANOVA done on only high-frequency words indicated a significant two-way interaction of block and regularity, $F(1, 31) = 6.60$, $MSE = 544.3$. Planned comparisons showed that the 18-ms regularity effect in the pure blocks was significant, $t(31) = 2.43$, $p < .05$ (one-tailed), whereas the −3-ms regularity effect in the mixed blocks was not, $t(31) = -.78$, *ns*.

A separate ANOVA done on only low-frequency words indicated both an effect of regularity, $F(1, 31) = 103.12$, $MSE = 1{,}017.6$, and a significant Block × Regularity interaction, $F(1, 31) = 25.92$, $MSE = 821.5$. As indicated by the Block × Regularity interaction, the regularity effect was larger in the pure blocks than in the mixed blocks. Planned comparisons, however, indicated that the regularity effect was nonetheless significant in both the pure blocks, $t(31) = 9.15$, $p < .001$ (one-tailed), and the mixed blocks, $t(31) = 5.51$, $p < .001$ (one-tailed).

Errors. Paralleling the RT data, the regularity effect, $F(1, 31) = 104.59$, $MSE = 15.0$; the frequency effect, $F(1, 31) = 55.67$, $MSE = 29.7$; and their interaction, $F(1, 31) = 36.21$, $MSE = 18.7$, were significant. Also significant was the interaction of block and regularity, $F(1, 31) = 8.71$, $MSE = 12.5$. This latter interaction was due to the regularity effect being 2.7% larger in the mixed blocks than in the pure blocks. As suggested by the marginal three-way interaction of block, regularity, and frequency, $F(1, 31) = 3.07$, $p < .10$, $MSE = 17.3$, however, the increase in the regularity effect in mixed blocks was found for low-frequency words (4.5%) but not for high-frequency words (1.0%).

Discussion

The results of Experiment 4 are quite clear. First of all, as expected on the basis of the results from earlier experiments, the regularity effect for low-frequency words was substantially larger when the words were presented in pure than in mixed blocks. The reader should note that this was due to naming latencies for low-frequency regular words showing a substantial pure-block RT advantage and to naming latencies for low-frequency irregular words showing a substantial mixed-block RT advantage. This second result, although predicted by the criterion-setting account, is directly opposite to the prediction that would be made by a route de-emphasis account. That is, mixing low-frequency irregular words with regular words can only have the effect of leading to more emphasis being placed on the assembly route than when those same irregular words are named in a pure block. The result should be, if anything, a pure-block RT advantage for the low-frequency irregular words rather than the mixed-block RT advantage observed in Experiment 4.

The more central issue in Experiment 4 concerned the regularity effect for high-frequency words. Although when high-frequency regular and irregular words were presented in mixed blocks there was virtually no regularity effect, when they were presented in pure blocks, a noticeable regularity effect emerged. This result, coupled with the pure-block data from Experiments 1 and 2, suggests that even for highly familiar words, it is easier to name a regular word than an irregular word. The lack of a regularity effect for these words in most previous reports can then be attributed to the fact that when they were mixed in the same block (and, generally, also with low-frequency words), the participants' use of a time criterion that remained fairly stable throughout the block pushed the mean RTs together.

General Discussion

The main purpose of the present set of experiments was to examine the claim (Monsell et al., 1992) that readers can strategically de-emphasize information from one of the two routes hypothesized by the dual-route model. In particular, we examined the argument that information from the assembly route can be de-emphasized when naming only irregular words, allowing readers to speed up the naming of high-frequency irregular words. Although we have provided further support for the reality of Monsell et al.'s phenomenon, the results of the current investigation suggest a quite different explanation. Rather than a strategic de-emphasis of the assembly route, this result appears to have been due to the influences of a time criterion.

Our criterion-setting account is based on the idea that in naming tasks, readers establish a time criterion to act as a flexible deadline for when articulation should start, a criterion that stays more or less constant over a block of trials.[5] The placement of this criterion is undoubtedly determined by a number of factors, with a very important one being the average strength of the S-R mappings of the stimuli in a block. When all the stimuli in a block are fairly homogeneous on this dimension, the criterion can be set at a point that seems most appropriate for that type of stimulus. Thus, articulatory codes for most of the stimuli in the block are at approximately the same level of completeness when articulation starts. When easy and difficult (i.e., fast and slow) stimuli are mixed together in the same block, however, the criterion tends to get set at a position that is intermediate to the positions used for the fast and slow stimuli (a position that, although not particularly appropriate for either type of stimulus, is nonetheless reasonably appropriate for the block). Thus, articulatory codes for rapidly processed stimuli are allowed to develop beyond the point where articulation could start, whereas the start of articulation for more slowly processed stimuli is initiated while the articulatory codes are still a bit incomplete. The main effect, however, is to homogenize RTs. That is, RTs for fast stimuli increase and RTs for slow stimuli decrease as participants strive to start articulation at approximately the same point in time for all stimuli in a block.

One additional effect of naming slow stimuli more rapidly and naming fast stimuli more slowly should be reciprocal changes in error rates, as observed by Strayer and Kramer (1994a, 1994b) in their memory-scanning and visual-scanning task. There were some trends in this direction in the present experiments. In particular, in all experiments, when slow stimuli were named more rapidly, numerically larger error rates were observed. However, these trends were often not significant (a notable exception being the significant block effect for the low-frequency irregular words in Experiment 4). Further, naming fast stimuli more slowly did not, in general, lead to any improvement in accuracy, although, because error rates for these stimuli tended to be rather low in the first place, there was very little room for improvement.

The implication of these patterns seems to be that speed–accuracy trade-off functions in a naming task have a particular form, a form that may be slightly different from those in other tasks (e.g., binary decision tasks). Specifically, in naming tasks, although the functions may approach asymptote at different rates for different types of stimuli, the functions for all types of stimuli have a long, shallow slope just below perfect accuracy. It is within this area of shallow slope rather than near the point of inflection (i.e., where the steep slope ends and the shallow slope begins) that participants tend to position their criterion.

The main reason that the function has this form is presumably because a naming response unfolds in time. Thus, one does not need to have a complete articulatory code available when pronunciation starts in order to make an acceptably accurate response. For example, as demonstrated in shadowing tasks (e.g., Marslen-Wilson & Tyler, 1980), if participants can gain information about initial phonemes, articulation can start and be successfully completed because the final phonemes will become available by the time they are needed. Drawing the parallel to the present situation implies that participants often had sufficient phonological–

---

[5] When one is considering only overall latency for a block as we have done here, our constant position hypothesis would be mathematically equivalent to the hypothesis that the position of the criterion changes to some extent after every trial by moving in the direction of the latency on the previous trial. A more fine-grained analysis of trial-by-trial effects would be necessary to see which of these hypotheses would be preferred. As noted, Strayer and Kramer (1994b) did find some evidence that criterion placement was affected by the nature of the previous trial. We thank Guy Van Orden for bringing this alternative hypothesis to our attention.

articulatory information to allow them to make an acceptably accurate response (i.e., they were on the shallow part of the function) well before they actually started responding.

If this argument is correct, participants should be able to show a large degree of flexibility in the placement of their criterion in a naming task without a high cost in terms of accuracy. Data from Colombo and Tabossi (1992) support this point quite nicely. In their Experiment 2, they induced participants to respond faster with a deadline procedure. Mean RTs dropped by over 60 ms whereas, if anything, error rates decreased slightly.

A major determinant of when participants actually do respond is, then, the nature of the context. In the present experiments, experimental context was altered by mixing fast stimuli with slow ones. This would not be the only way to effectively alter context, of course (e.g., the deadline technique used by Colombo & Tabossi, 1992). Further, in the present experiments, there may have been other aspects of the experimental procedure that may have made a mixing manipulation more or less potent (e.g., the length of the intertrial intervals, the stress put on minimizing errors at the expense of RTs, or even the ease with which a strategy of this sort can be adopted after only a few trials).[6] The important point, however, is that readers can and do respond to changes in context by exercising control over the point at which articulation is begun.

With only two exceptions, the entire set of data reported here follows the pattern of slow stimuli being named more rapidly and fast stimuli being named more slowly when the two types of stimuli were mixed. A brief summary of the effects that do follow this pattern is given: in Experiment 1, mixing rapidly named high-frequency irregular words with slowly named nonwords led to faster RTs for the nonwords and slower RTs for the words. In Experiment 2, mixing rapidly named regular words (both high- and low-frequency) with nonwords led to faster RTs for the nonwords and slower RTs for the words. In Experiment 3, mixing rapidly named high-frequency irregular words with more slowly named low-frequency irregular words led to faster RTs for the low-frequency words and slower RTs for the high-frequency words. Finally, in Experiment 4, mixing more rapidly named low-frequency regular words with more slowly named low-frequency irregular words led to faster RTs for the irregular words and slower RTs for the regular words.

The only exceptions to this pattern occurred when high-frequency regular and irregular words were mixed in Experiment 4 and when low-frequency irregular words were mixed with nonwords in Experiment 1. With respect to the former case, the results can be explained fairly straightforwardly if one assumes that the criterion setting used in the mixed blocks reflected a floor on what our participants viewed as acceptable performance. That is, although our participants may have been able to name words faster without a corresponding unacceptable increase in error rate in the pure block with high-frequency regular words, they may have been unwilling to attempt it. This hypothesis could be evaluated by motivating more rapid responding through, for example, the use of a deadline technique.

Mixing Low-Frequency Irregular Words and Nonwords

The second, more complicated exception concerns the results from Experiment 1 in which low-frequency irregular words were mixed with nonwords. From the criterion-setting perspective, the main thing to note here is that pure-block performance (both in Experiment 1 and in subsequent experiments) indicates that our low-frequency irregular words and our nonwords were quite similar in difficulty and, hence, invoked similar criterion placements in those pure blocks. Thus, unlike in all other situations, the mixed blocks did not constitute a mixture of slow and fast stimuli. The specific principle for criterion setting discussed above is that an adjustment of the criterion occurs when different stimulus types having somewhat different pure-block latencies are mixed. Thus, on the basis of this particular principle alone, there would be no reason to have expected any effect at all of the mixing manipulation.

There is, of course, no reason to claim that the *only* principle for moving a criterion is that fast and slow stimuli are being mixed. That is, as suggested earlier, many factors, including things like instructions or general expectations (e.g., Treisman & Williams, 1984), could affect criterion placement. The other important thing to note here, however, is that the effects of mixing were different for the two types of stimuli being mixed. That is, when mixed, the low-frequency irregular words were named more rapidly, and the nonwords were named more slowly. If the effect of the mixing manipulation was simply to cause a criterion shift, for whatever reason, the expectation would be that this shift would have had the same effect for both stimulus types. As such, these effects do indicate that a more complicated explanation of these results is required.[7]

As Monsell et al. (1992) suggested, one way to try to understand the nature of processing in this type of situation would be to consider the error data. As they also suggested, there appear to be fewer regularization errors for low-frequency words in pure blocks than in mixed blocks, a generalization that seems to characterize not only the results

---

[6] We thank Steve Monsell and Ken Paap for suggesting these possibilities.

[7] There is also a mixing effect in Experiment 2 of Paap and Noel (1991) that cannot be explained by our criterion-setting account. In this task, high- and low-frequency irregular words were presented in the context of either irregular words or regular words. The result was that the irregular words, as a whole, were named faster when presented in the context of other irregular words. (Unfortunately, the data broken down by word frequency are unavailable.) The main point to note here is that this manipulation does not constitute a mixture of a set of fast and slow stimuli, which means that there would be no reason for the criterion-setting account to predict the different performance observed in the pure and mixed blocks. That is, in the pure block, the irregular words were of varying frequencies, and thus, that block itself consisted of a mixture of slow and fast stimuli. As such, mixing in the regular words should not have altered the criterion setting in the mixed block to any large degree, and thus, no RT change for either high- or low-frequency irregular words would have been expected.

of our Experiment 1 but also the results of both of their experiments. Monsell et al.'s interpretation of this effect was that the assembly route was de-emphasized in the irregular-word pure blocks, placing relatively more emphasis on the lexical route. In addition to the fact that we have not been able to obtain any other evidence supporting a route de-emphasis explanation, there are two other important points to consider when evaluating this interpretation.

First, although there was a decrease in the number of regularization errors for low-frequency irregular words in the pure blocks, there is little evidence for a decrease in the *percentage* of errors that were regularization errors in the pure blocks. In Experiment 1, the percentage of errors that were regularization errors was 66% in both the pure and mixed blocks (see Table 2). In Monsell et al.'s (1992) Experiment 2, the percentage was 55% in both the pure and mixed blocks (see Monsell et al., 1992, Table 5). Only in Monsell et al.'s Experiment 1, in which frequency was not blocked, was there any evidence for a decrease in the percentage of regularization errors from the mixed blocks (66%) to the pure blocks (47%).

Second, even if participants' tendencies to regularize did decrease in pure blocks, that result would not be compelling support for the conclusion that this was due to the de-emphasis of an "assembly route." That is, as noted earlier, there are a number of ways that lexical information could be used at some point in the naming process other than by simply increasing *relative* reliance on a "lexical route" (Brown & Besner, 1987; Glushko, 1979; Humphreys & Evett, 1985). Within any of these alternate frameworks, a decrease in regularization errors would presumably follow from any increased use of lexical information. Within other frameworks (e.g., Kawamoto & Zemblidge, 1992; Plaut et al., 1996), a decrease in regularization errors could come about simply as a function of changes in the position of a time criterion. Thus, by itself, a decrease in regularization errors in the pure blocks would provide only minimal support for the route de-emphasis account.

A better way to understand the effect of mixing low-frequency irregular words and nonwords would be to extrapolate from Monsell et al.'s (1992) account of what happens with the nonwords. This account is based on two notions. First, Monsell et al. suggested that in the pure-block condition with nonwords, participants do not wait for output from the lexical route but, rather, start articulation as soon as a regularized code emerges. In contrast, in the mixed blocks, where half of the stimuli are low-frequency irregular words, participants tend to wait for output from the lexical route before producing a response. As such, naming latencies for nonwords should be longer in the mixed blocks than in the pure blocks, as observed in both of their experiments and, to a lesser extent, in our Experiment 1.

Monsell et al.'s (1992) suggestion that readers rely more on stored lexical information in the mixed blocks than in the pure blocks with nonwords does seem to be a reasonable one (although there is no need to couch it in a route de-emphasis framework). Our suggestion is that it be carried one step further and be extended to account for the effects for the low-frequency irregular words as well. That is, we suggest that the entire pattern of these data is consistent with the idea that participants choose to invoke a "lexical-checking" strategy on some portion of the trials.

The idea is that as the articulatory code builds up, readers can choose to consult an output lexicon to determine whether the phonological code generated by the phonological-code-generation process matches a code in their lexicon. The more low-frequency irregular words contained in a block, of course, the more useful this type of strategy would be, hence, the more often it should be invoked. That is, in the pure blocks with low-frequency irregular words, this strategy would be most useful because it is only in this situation that participants *must* be sure that the articulatory code they produce is a word code. In the mixed blocks, because half the stimuli are nonwords, the tendency would be to invoke this strategy less often. Thus, even though the criterion settings may have been approximately the same in the two blocks, naming latencies for low-frequency irregular words would tend to be a bit shorter and there would be an increased error rate in the mixed blocks.

On the other hand, in the pure blocks with nonwords, there would be no reason at all to invoke this type of strategy. Thus, without this extra process, nonword naming latencies should be shorter in the pure blocks than in the mixed blocks with irregular words, as was generally observed. Note also that error rates were significantly lower in pure blocks as well. This result provides additional support for the argument that the difference between the pure and mixed blocks for the nonwords was not simply due to a criterion shift. Instead, it is more consistent with the idea that, although this lexical-checking process may be beneficial for words, on those trials in the mixed block when it is used and the stimulus turns out to be a nonword, the effect is to actually increase the difficulty of coming up with a reasonable articulatory code.

Implications for Models of Pronunciation

As noted, the present results cause considerable problems for both the dual-route and PDP models if the models attempt to account for the data *solely* in terms of the phonological-code-generation process. Such is not the case, however, if it is assumed that these effects are not due to the phonological-coding process itself but rather to the demands placed on the system by using a time criterion to control the articulation process.

Consider first the dual-route model. According to the model, phonological codes are produced by each route. When a viable code is available, it is used by the system responsible for motor programming and articulation. The operations involved in all these processes would presumably be time demanding, and thus, the time criterion could have its impact during any of those operations.

Given that a difference between high-frequency regular and irregular words was observed in the pure blocks in Experiment 4, one more assumption should be added to the model, namely, that the assembly route sometimes does provide a competing response for high-frequency irregular words. Time is then taken to resolve the competition, which implies that on average, the output system receives a completely resolved code slightly later for high-frequency irregular words. In pure blocks, this difference translates into a regularity effect for high-frequency words. When high-frequency regular and irregular words are mixed, however, the irregular words are named faster (i.e., the system does not wait for a full resolution of the competition but uses the most viable code it has available). The result is that these words are then named as fast as the high-frequency regular words.

Consider next Plaut et al.'s (1996) PDP model. This model involves an initial computation of phonology, which is followed by a settling process in which the initial "noisy" code is "cleaned up." The clean-up process is time consuming, and hence, in theory, the criterion notion could be thought of as affecting either this process or the process of turning a cleaned-up code into an articulatory code.

The question is how would this model account for the results of Experiment 4. In their Simulation 2, Plaut et al. (1996) demonstrated that different assumptions about the impact of frequency of exposure during training can lead to different size "regularity effects" for high-frequency words in terms of the noisiness of the initial code (i.e., the cross-entropy scores). Of most importance here is that when the training regimen was based on the actual frequencies of the words in the language, the quality of the initial codes was virtually identical for the regular and irregular high-frequency words. As their Simulation 3 shows, however, in spite of this initial equality and in spite of the fact that the initial codes for high-frequency words are all quite good already, the model takes less time to clean up regular high-frequency words than to clean up irregular high-frequency words. That is, this simulation does predict a small regularity effect in terms of RT for high-frequency words.

A simple way for the model to explain the results of Experiment 4 would be to suggest that the small regularity effect shown in Simulation 3 represents the situation in pure blocks. When the word types are mixed together, however, the clean-up process for the irregular words may be terminated slightly prematurely. Nonetheless, because the initial code is so good, it would be quite easy for an accurate articulatory code to be derived from it. Thus, high-frequency irregular words could be pronounced both as accurately and as quickly as regular words.

Methodological Issues

The results of Experiment 4 also raise an important methodological–interpretation issue. That is, whenever a null effect has been reported with latency data in a mixed condition, the question could be raised as to whether the effect would have emerged (thus, changing the interpretation) if the conditions had been run in pure blocks. In particular, the question would be whether the null effect was due to the influence of a time criterion that diluted any difference between the two conditions. At the very least, the possibility that different results may emerge in mixed and pure conditions because of effects of time criteria is one that investigators need to be wary of in both designing and interpreting experiments.

In fact, there may be a number of null effects now in the literature that were due to the influence of time criteria (rather than to the lack of a processing difference between two conditions). As a possible example, consider the recent results of Buchanan and Besner (1993). In one experiment, they showed a significant 21-ms semantic/associative priming effect for Katakana target words named aloud by Japanese readers. (Katakana is a script used for writing Japanese words that is orthographically very shallow. Thus, in principle, readers could name these words just by applying spelling-to-sound rules.) In a second experiment, the authors showed that this effect disappeared when an equal number of unfamiliar Hiragana transcriptions of Katakana targets were also included in the stimulus set. (Hiragana is also a Japanese script that is orthographically very shallow. Japanese words can be written in either of these scripts, but they are usually written in one or the other. Thus, the Katakana words transcribed into Hiragana were familiar phonologically but were, in essence, pseudohomophones when written in Hiragana.) These Hiragana-transcribed words were much more difficult to name and did show a priming effect.

According to our analysis, one result of including the more difficult Hiragana targets along with the fairly rapidly named Katakana targets should have been to slow down naming of all Katakana targets, in effect, homogenizing their latencies with those of the Hiragana transcriptions. Such was the case, as RTs to Katakana targets were approximately 55 ms slower in the mixed block. In addition, however, a secondary consequence of having a time criterion set at a substantially later point than the normal point for pronunciation of Katakana targets may have been to homogenize the naming latencies in the related and unrelated conditions for those words. That is, the naming latencies for all Katakana targets would increase toward the level set by the criterion, a level that was near that needed for accurate performance on the Hiragana targets, thereby effectively eliminating the difference between related and unrelated targets (i.e., the priming effect).

Buchanan and Besner's (1993) studies are, of course, just one example of a situation in which investigators need to be wary. As noted, there are now a number of studies in the literature (e.g., Baluch & Besner, 1991; Tabossi & Laghi, 1992) in which the key manipulation was mixing nonwords with words, with the result being a change in the nature of the relation between word conditions. Because nonwords are typically harder to name than words, their inclusion very likely influenced criterion settings. The question that needs to be dealt with by these investigators is whether their effects were independent of or were a consequence of criterion settings.

The additional point should also be made, of course, that including nonword targets in a word-naming task will not inevitably lead to a full homogenization of word latencies. In fact, as noted, no one has yet demonstrated that including nonword targets in a priming task eliminates priming effects when naming English words (e.g., Keefe & Neely, 1990;

Tabossi & Laghi, 1992; West & Stanovich, 1982). How participants react to the inclusion of slow stimuli depends on a number of things, including the difficulty of those stimuli, the difficulty of the word stimuli, and the length of the intertrial interval. As suggested previously, this latter factor may be quite important because longer intertrial intervals may allow a greater opportunity for other contextual factors to manifest themselves. It should be noted that the shortest intertrial interval used in the three priming studies mentioned above was 3 s.

Although our analysis suggests that the best way to determine whether there is a regularity effect with high-frequency words is to present the words in pure blocks, we are not arguing that using pure blocks is always the best way to verify the presence of an effect. Pure blocks also can create a situation in which criteria can inappropriately affect the results (Sperling & Dosher, 1986). Consider, for example, a semantic-priming task. Having a block of pure related trials and a block of pure unrelated trials would very likely cause participants to adopt completely different strategies in the two blocks, preventing any meaningful comparison between conditions. Nonetheless, the use of pure blocks should be encouraged whenever it makes sense and whenever important theoretical points are being made about a null effect in a mixed-block condition.

Conclusion

In recent years, a number of investigators have argued that readers have some control over the way in which phonological codes are generated. The present results support a slightly different claim that there is a process, external to the phonological-code-generation process, that can have a major impact on naming latencies and that readers' strategies can affect that process. In particular, readers appear to be able to initiate a response somewhat prior to the point that a normally acceptable articulatory code is available or to delay the initiation of that response when the contextual constraints of the task dictate. Whether and to what extent readers' strategies can actually affect the process of phonological-code generation itself remains an open question.

References

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

Baluch, B., & Besner, D. (1991). Visual word recognition: Evidence for strategic control of lexical and nonlexical routines in oral reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 644–652.

Baron, J., & Strawson, C. (1976). Use of orthographic and word-specific knowledge in reading words aloud. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 207–214.

Becker, C. A. (1980). Semantic context effects in visual word recognition: An analysis of semantic strategies. *Memory & Cognition, 8,* 493–512.

Bernstein, S., & Carr, T. H. (1996). Dual-route theories of spelling to pronunciation: What can be learned from concurrent task performance? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 86–116.

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

Brown, P., Lupker, S. J., & Colombo, L. (1994). Interacting sources of information in word naming: A study of individual differences. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 537–554.

Buchanan, L., & Besner, D. (1993). Reading aloud: Evidence for the use of a whole word nonsemantic pathway. *Canadian Journal of Experimental Psychology, 47,* 133–152.

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. 112–174). New York: Academic Press.

Coltheart, M., & Rastle, K. (1994). Serial processing in reading aloud: Evidence for dual-route models of reading. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 1197–1211.

Content, A. (1991). The effect of spelling-to-sound regularity on naming in French. *Psychological Research/Psychologische Forschung, 53,* 3–12.

Dorfman, D., & Glanzer, M. (1988). List composition effects in lexical decision and recognition memory. *Journal of Memory and Language, 27,* 633–648.

Forster, K. I. (1981). Frequency blocking and lexical access: One mental lexicon or two? *Journal of Verbal Learning and Verbal Behavior, 20,* 190–203.

Fowler, C., Wolford, G., Slade, R., & Tassinary, L. (1981). Lexical access with and without awareness. *Journal of Experimental Psychology: General, 110,* 341–362.

Fredericksen, J. R., & Kroll, J. K. (1976). Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 361–379.

Glanzer, M., & Ehrenreich, S. L. (1979). Structure and search of the internal lexicon. *Journal of Verbal Learning and Verbal Behavior, 18,* 381–398.

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.

Gordon, B. (1983). Lexical access and lexical decisions: Mechanisms of frequency sensitivity. *Journal of Verbal Learning and Verbal Behavior, 22,* 24–44.

Herdman, C. M. (1992). Attentional resource demands of visual word recognition in naming and lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 124–132.

Hines, D., Czerwinski, M., Sawyer, P. K., & Dwyer, M. (1986). Automatic semantic priming: Effect of category exemplar level and word association level. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 370–379.

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

Jared, D. (1995, November). *Spelling-sound consistency affects the naming of high-frequency words.* Poster presented at the meeting of the Psychonomic Society, Los Angeles, CA.

Kawamoto, A. H., & Zemblidge, J. H. (1992). Pronunciation of homographs. *Journal of Memory and Language, 31,* 349–374.

Keefe, D. E., & Neely, J. H. (1990). Semantic priming in the pronunciation task: The role of prospective prime-generated expectancies. *Memory & Cognition, 18,* 289–298.

Kinoshita, S., & Woollams, A. (1996). *The left-to-right regularity effect and strategy effects in naming.* Manuscript submitted for publication.

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

Krueger, L. E. (1985). Effect of intermixed foveal and parafoveal presentation on same-different judgments: Evidence for a criterion-inertia model. *Perception & Psychophysics, 37,* 266–271.

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

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

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163–203.

Marcel, A. T. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology, 15,* 197–237.

Marslen-Wilson, W. D., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition, 8,* 1–71.

McKoon, G., & Ratcliff, R. (1995). Conceptual combinations and relational contexts in free association and in priming in lexical decision and naming. *Psychonomic Bulletin & Review, 2,* 527–533.

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.

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., & Morton, J. (1985). From orthography to phonology: An attempt at an old interpretation. In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading* (pp. 335–359). London: Erlbaum.

Patterson, K. E., Seidenberg, M., & McClelland, J. L. (1989). Connections and disconnections: Acquired dyslexia in a computational model of reading processes. In R. G. M. Morris (Ed.), *Parallel distributed processing: Implications for psychology and neurobiology* (pp. 131–181). Oxford, England: Oxford University Press.

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

Pexman, P. M., & Lupker, S. J. (1995). Effects of memory load in a word naming task: Five failures to replicate. *Memory & Cognition, 23,* 581–595.

Pexman, P. M., & Lupker, S. J. (in press). Word naming and memory load: Still searching for an individual differences explanation. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. E. (1996). Understanding normal and impaired reading: Computational principles in quasi-regular domains. *Psychological Review, 103,* 56–115.

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

Schneider, W. (1988). Micro Experimental Laboratory: An integrated system for IBM PC compatibles. *Behavior Research Methods, Instrumentation, and Computers, 20,* 206–217.

Seidenberg, M. S. (1985). 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.

Shulman, H. G., & Davison, T. C. B. (1977). Control properties of semantic coding in a lexical decision task. *Journal of Verbal Learning and Verbal Behavior, 16,* 91–98.

Sperling, G., & Dosher, B. (1986). Strategy and optimization in human information processing. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 1, pp. 2-1–2-65). New York: Wiley.

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

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

Strayer, D. L., & Kramer, A. F. (1994a). Strategies and automaticity: I. Basic findings and conceptual framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 318–341.

Strayer, D. L., & Kramer, A. F. (1994b). Strategies and automaticity: II. Dynamic aspects of strategy adjustment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 342–365.

Stroop, J. R. (1935). Studies of interference in serial reactions. *Journal of Experimental Psychology, 18,* 643–662.

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

Taraban, R., & McClelland, J. L. (1987). Conspiracy effects in word pronunciation. *Journal of Memory and Language, 26,* 608–631.

Treisman, M., & Williams, T. C. (1984). A theory of criteria setting with an application to sequential dependencies. *Psychological Review, 91,* 68–111.

West, R. F., & Stanovich, K. E. (1982). Source of inhibition in experiments on the effect of sentence context on word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 385–399.

(Appendixes follow)

Appendix A

Irregular Words and Nonwords Used in Experiment 1

HF irregular words		Nonwords used with HF words	
Set 1	Set 2	Set 1	Set 2
Warm-up stimuli			
are	broad	arn	broap
corps	does	cact	delt
gross	give	groach	gice
half	have	heam	hend
most	put	mim	pone
says	two	sern	tive
walk	weight	wame	woif
war	world	wunk	wup
break	breath	breek	bick
dead	course	deve	kale
eyes	guy	ert	gow
health	learn	hain	lund
one	our	wuff	osk
rough	search	roke	sile
touch	through	tash	thrart
warm	won	wuck	wung
Experimental stimuli			
bear	both	bosk	bule
blood	built	blit	bip
come	court	kade	keck
four	doubt	fipe	dap
friend	eight	freet	eaf
gone	front	gect	frant
great	group	grost	grulp
guess	guard	gan	greep
heard	heart	herk	heen
hour	house	hig	himp
lose	know	lave	naft
move	love	meed	leat
none	month	nonk	melp
once	ought	wuve	eld
phase	post	phoad	pulk
shall	scene	shud	sen
sign	school	seft	skal
son	some	sawl	surm
though	touch	thark	tate
view	tour	vide	tind
was	want	wike	wope
whole	were	haik	wot
work	word	wole	wurn
worth	worse	wunt	wempt

Appendix A (*continued*)

Irregular Words and Nonwords Used in Experiment 1

LF irregular words		Nonwords used with LF words	
Set 1	Set 2	Set 1	Set 2
Warm-up stimuli			
deaf	dwarf	doke	derk
blown	beau	blale	bame
chasm	chic	koif	shung
cough	choir	culk	kark
coupe	flown	kip	flurm
fjord	ghoul	fope	gup
guise	isle	gar	ock
lieu	moll	luff	mot
pi	pier	pode	pern
pique	reins	pait	reen
rogue	scourge	runk	skawl
sioux	sigh	sund	sost
suave	swap	swaik	swow
wan	warn	wibe	wule
spook	yacht	sim	yeat
pearl	plaid	pash	plind
Experimental stimuli			
aisle	ache	ard	erb
caste	calf	ceck	keek
chalk	chute	chig	shipe
chord	coup	kend	keld
comb	doll	kice	dag
debt	dough	dube	dant
dose	fierce	doach	fimp
feud	ghost	faft	gream
gauge	glove	goap	gade
guild	hearth	geed	hect
heir	hood	ide	hulp
hymn	leapt	haim	lelp
mosque	niche	meve	nuck
pint	plague	peaf	plick
quart	psalm	quate	sule
reign	rouge	rive	ral
sew	scarce	sule	scoad
shove	scent	shact	sem
sponge	soot	speep	seft
thigh	sword	thelt	sosk
ton	thread	tave	theet
veil	tomb	vit	tud
warp	wasp	wurt	wund
worm	weird	wonk	wemp

Note. These words were also used in Experiments 3 and 4. These nonwords were also used in Experiment 2. HF = high frequency; LF = low frequency.

(*Appendixes continue*)

Appendix B

Regular Words Used in Experiments 2 and 4

HF regular words		LF regular words	
Set 1	Set 2	Set 1	Set 2
	Warm-up stimuli		
bright	act	doom	dill
day	claim	bait	blend
game	green	shave	cain
him	held	cave	coil
past	made	flip	cove
time	side	gull	faze
wish	wide	itch	garb
way	will	mink	latch
broke	bright	pane	paw
case	dark	rack	punk
gate	east	skeet	roost
lord	horse	silk	sack
all	went	swoop	swerve
seek	raw	woke	wade
three	team	yoke	spout
wine	wheel	plod	poll
	Experimental stimuli		
back	bay	soy	fife
born	based	sheen	sheep
call	case	leak	coop
deep	felt	wisp	chess
aid	french	nudge	quell
free	gun	ape	weld
growth	ground	cart	wick
grand	gold	hike	cord
hot	hair	sill	sunk
help	hall	plead	inn
night	leg	scoop	cured
less	meet	sash	gall
march	news	hound	ant
out	went	rude	roam
pain	phone	soak	speck
serve	short	fake	vent
south	sight	gland	hint
same	sort	ditch	thaw
team	these	wept	merge
teach	voice	cult	duke
west	we	gram	gaze
with	hand	tide	dot
week	well	dice	par
wave	wait	theft	tub

Note. HF = high frequency; LF = low frequency.

Received July 14, 1995
Revision received August 7, 1996
Accepted August 22, 1996 ■