Journal of Experimental Psychology:
Human Perception and Performance
1999, Vol. 25, No. 2, 482-503

0096-1523/99/$3.00

Serial and Strategic Effects in Reading Aloud

Kathleen Rastle and Max Coltheart
Macquarie University

M. Coltheart and K. Rastle (1994) reported that the size of the regularity effect on word-naming latency decreases across position of irregularity, implicating a serial process in reading aloud. In response to criticism by D. C. Plaut, J. L. McClelland, M. S. Seidenberg, and K. Patterson (1996), these results were replicated with monosyllabic words that had been controlled for consistency. In a second experiment, participants named nonword- or regular-word targets mixed with either first-position irregular fillers or third-position irregular fillers. Target naming was slowed when first-position irregular fillers were present, compared with target naming when third-position irregular fillers were present. These data suggest that participants can slow use of the nonlexical route if faced with very costly exception words. Simulations using the dual-route cascaded model (M. Coltheart, B. Curtis, P. Atkins, & M. Haller, 1993) are presented.

The term *dual-route theory* refers to a particular class of theories of visual word recognition and reading aloud. The defining feature of such theories is the postulate that there are two different procedures for converting print to speech: a dictionary-lookup, or lexical, procedure and a rule-based, or nonlexical, procedure (expositions of dual-route theories can be found in, e.g., Baron & Strawson, 1976; Coltheart, 1978, 1985; Ellis & Young, 1988; Forster & Chambers, 1973; Gough & Cosky, 1977; Morton & Patterson, 1980; Ogden, 1996; Paap & Noel, 1991; Patterson & Morton, 1985; Patterson & Shewell, 1987).

Coltheart, Curtis, Atkins, and Haller (1993; see also Coltheart, Langdon, & Haller, 1996; Coltheart & Rastle, 1994; Rastle & Coltheart, 1998, 1999) described a computational realization of dual-route theory: the dual-route cascaded (DRC) model. The model described in these studies is computational in the sense that it exists as a computer program that can perform the tasks typically used in research on reading, such as lexical decision and reading aloud. The number of processing cycles needed to perform any such task with a particular stimulus is an analogue of the number of milliseconds needed by human beings, so direct simulations of experiments that yield human reaction times (RTs) is straightforward.

Four other models of reading, which are computational in this sense, currently exist: the parallel-distributed-processing (PDP) implementations described by Plaut, McClelland, Seidenberg, and Patterson (1996), the multiple-levels model of Norris (1994), the connectionist dual-process model of Zorzi, Houghton, and Butterworth (1998), and the multiple readout model of Grainger and Jacobs (1996). The DRC model differs from these models in that it is applicable both to the simulation of lexical decision and to the simulation of reading aloud (i.e., it is a model of both visual word recognition and reading aloud). The simulations contained in the Norris, Zorzi et al., and Plaut et al. implementations are of reading aloud only; those models, in their present form, have no procedure for making lexical decisions. Similarly, the simulations in Grainger and Jacob's model are of lexical decision only; the multiple readout model, in its current form, does not read aloud.

Because we are concerned in this article with reading aloud rather than with lexical decision, we focus on comparing the DRC model with the multiple-levels model, the connectionist dual-process model, and the PDP implementations; these are the only current computational models of reading aloud.

The DRC Model

The overall architecture of the DRC model is shown in Figure 1. As is evident, the model has both a dictionary-lookup (lexical) procedure for converting print to speech and a rule-based (nonlexical) procedure for such conversion.

The Lexical Route of the DRC Model

The lexical route consists of a sequence of five processing components or levels: feature detection, letter identification, orthographic lexicon, phonological lexicon, and phoneme activation. The first three of these levels are simply a generalization of the interactive activation (IA) model of visual word recognition (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982). However, instead of being restricted just to four-letter monosyllabic monomorphemic

Kathleen Rastle and Max Coltheart, Department of Psychology, Macquarie University, Sydney, New South Wales, Australia.

This work was supported by a National Science Foundation Graduate Research Fellowship awarded to Kathleen Rastle. Portions of this work were presented at the 37th Annual Meeting of the Psychonomic Society, Chicago, November 1996.

We are grateful to Conrad Perry and Jo Ziegler for helpful discussion.

Correspondence concerning this article should be addressed to Kathleen Rastle, who is now at the Centre for Speech and Language, Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, England. Electronic mail may be sent to kathy@currawong.bhs.mq.edu.au.

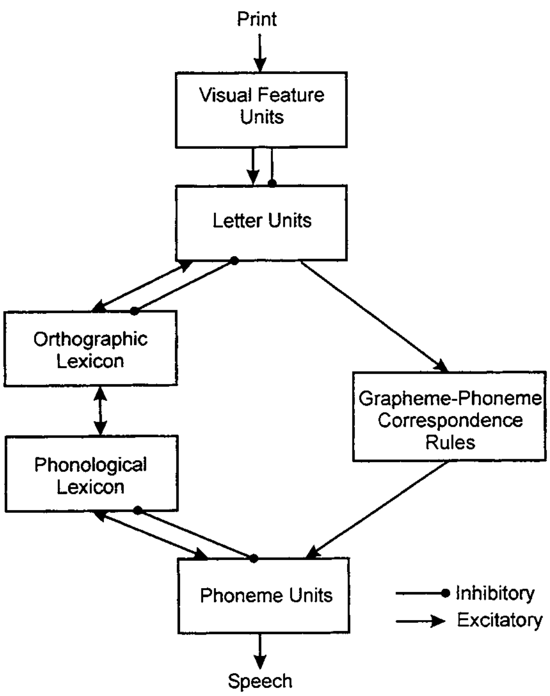

Figure 1. The architecture of the dual-route cascaded model.

words (as was the case with the IA model), the DRC model operates with words of any length up to seven letters[1] and is not restricted to monomorphemic words (though is restricted to monosyllabic words). Apart from these differences, the architecture, connectivity, and mode of operation of these three levels are identical to those in the IA model.

The feature level consists of eight sets of feature units, one for each of the eight possible letter positions of an input string. Each of these eight sets of feature units contains 16 feature-present units and 16 feature-absent units (the number is 16 because the system operates with the 16-stroke font used by Rumelhart & Siple, 1974, exactly as in the IA model).

The letter level also consists of eight sets of units. Each of these sets of letter units contains 27 units, one for each letter of the alphabet and one coding the absence of any letter in that position in the input string.

Each unit at the feature level for a particular input position has an excitatory connection to all of the letters for that input position that possess that feature and an inhibitory connection to all of the letters that do not. There are no connections from the letter level back to the feature level, nor are there inhibitory connections within each set of feature units. At the letter level, however, there is within-level inhibition: for each letter position, all letter units have inhibitory connections to all other letter units.

The orthographic lexicon contains 7,980 units, one unit for each monosyllabic word in the CELEX English database (Baayen, Piepenbrock, & van Rijn, 1993), except that infrequent words of foreign origin such as *kvass* or *lakh* have been culled. Each word unit in this lexicon has inhibitory connections to all other word units in the orthographic lexicon. Each word unit also has a parameter representing its frequency; as in the IA model, the value of this parameter ranges from −0.05 (for the least frequent word) to 0.00 (for the most frequent word). These frequency parameters cause the rate at which activation rises to be positively related to word frequency.

For each position at the letter level, every letter unit has excitatory connections to every entry in the orthographic lexicon representing a word that possesses that letter in that position and inhibitory connections to all other word units. Furthermore, each word unit has excitatory connections back to all of the letter units that represent its spelling and inhibitory connections to all other letter units.

The phonological lexicon contains 7,117 units, with each unit representing the pronunciation (in Australian English) of one of the entries in the orthographic lexicon. Each word unit in this lexicon has inhibitory connections to all other word units in the phonological lexicon. These units are frequency coded in the same way as are the entries in the orthographic lexicon. Heterographic homophones such as *so* and *sew* have separate entries in the orthographic lexicon but activate the same entry in the phonological lexicon. Homographic heterophones such as *lead* have a single entry in the orthographic lexicon, which activates two different entries in the phonological lexicon. Apart from these cases, connections from the orthographic lexicon to the phonological lexicon are one to one.

The phoneme level consists of eight sets of phoneme units. Each of these sets contains 44 phoneme units, one for each of the 43 phonemes in the DRC model's phonemic vocabulary (see the list of these phonemes in Appendix A) and one that codes the absence of a phoneme in that position in the output string. Each entry in the phonological lexicon has excitatory connections to all of its constituent phonemes and inhibitory connections to all other phonemes. In turn, each phoneme unit has excitatory connections back to all word units in the phonological lexicon that contain that phoneme in that position and inhibitory connections back to all other units in the phonological lexicon.

The DRC Model's Nonlexical Route

The nonlexical route in the DRC model is a sequence of four processing components or levels: feature detection, letter identification, grapheme–phoneme conversion, and phoneme activation. The feature, letter, and phoneme levels are shared with the lexical route and have been described. The rules used for grapheme–phoneme conversion are described in the next section.

The nonlexical route operates as follows. For the first N processing cycles (N is set to 10 cycles in the standard set of parameters), the route is inoperative. At Cycle 11, grapheme–phoneme correspondence (GPC) rules are applied to the first

[1] The model has eight sets of letter units, but its reading is nevertheless restricted to strings of up to seven letters in length only; the reasons for this are discussed in Coltheart, Rastle, Perry, Langdon, and Ziegler (1998).

letter of the input string. After a further *M* processing cycles have elapsed (*M* is set to 17 cycles in the standard set of parameters), the next letter of the string becomes available for nonlexical translation: The input to the GPC rules is now a two-letter string. Thus, the string becomes available for translation serially, letter by letter, from left to right; the translation process then operates on the entire available string in parallel.

Nonwords are not processed entirely by the nonlexical route, however. They activate word neighbors in the orthographic lexicon, which then activate phonological representations of words and their phonemes. Figure 2 displays activation in the four phoneme units in the stimulus item *flot.* Note that whereas activation for each of the phoneme units rises in serial order, activation for each of the phonemes does not rise at exactly the same rate. These activations would rise at the same rate only if the nonlexical route provided these units the sole source of activation. These phoneme units also receive activation from the lexical route by means of their word neighbors, however, which include *flat, flow,* and *foot.* The amount of lexical activation enjoyed by each phoneme unit will thus depend on what the word neighbors of the input string happen to be.

The Rules Used by the DRC Model's Nonlexical Route

The DRC model's nonlexical route uses four types of position-specific rules, which are applied to every letter string on the basis of the strength of the parameter controlling GPC activation: single-letter rules, multiletter rules, context-sensitive rules, and output rules. For a given portion of a letter string (that portion which has become available for translation), multiletter rules are applied before single-letter rules, so that, for example, if the letters *che* in the item *chean* were available for translation, the rule *ch* → /J/ would be applied before the rules *c* → /k/ and *h* → /h/. The full set of rules used by the nonlexical route is given in Appendix B.

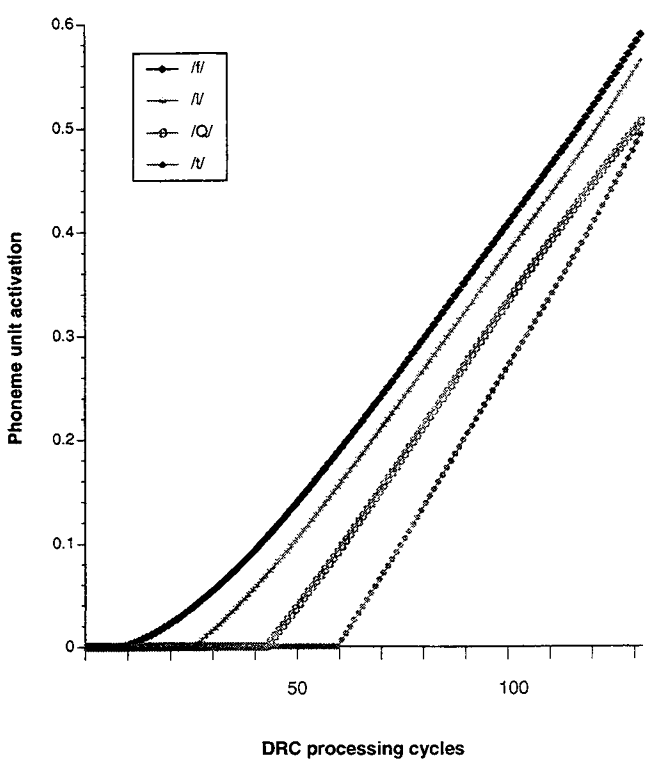

Figure 2. The rise of activation in phonemes /f/, /l/, /Q/, and /t/ in the naming of the word *flot.* DRC = dual-route cascaded.

In early work on the DRC model, Coltheart et al., 1993, explored a learning algorithm that was capable of discovering GPC rules when exposed to a corpus of word spellings and pronunciations. We soon abandoned the idea that the DRC model's GPC rules should be derived by means of a learning algorithm, however, because unless the learning algorithm itself was psychologically real (i.e., operated in just the same way as children operate when they are learning GPC rules), there would be no reason to expect that the set of algorithm-derived GPC rules would have any relation to the set of GPC rules the research participants would possess. Because there is no agreement on the question of how children learn GPC rules as they learn to read, there is no way of guaranteeing the psychological reality of any learning algorithm that might be used and hence no justification for adopting as the model's particular set of GPC rules those that the algorithm learned.

It is therefore important to see this rule set not as the outcome of the application of a learning algorithm but as a set of hypotheses about what GPC rules skilled readers possess. To the extent to which these hypotheses are false, the actual pronunciations of nonwords yielded by the DRC model will differ from the actual pronunciations that human readers will produce. No doubt different readers will have slightly different sets of GPC rules, because it is certainly not always the case that all people assign exactly the same pronunciation to a particular nonword (Masterson, 1985). Thus, the nonword pronunciations assigned by the DRC model will not be the pronunciations assigned by every person, because people do not even agree among themselves on these pronunciations. All we seek to achieve is that for all nonwords, the DRC model's pronunciation is the one that the majority of readers assign.

One refinement of dual-route modeling that goes beyond the DRC model in its current form is the idea that different GPC rules might have different strengths, with the strength of correspondence being a function of, for example, the proportion of words in which the correspondence occurs. We have not explored the notion of rule strength in the DRC model, although it would be simple to implement, because we are not aware of any work demonstrating that any kind of rule-strength variable has effects on naming latencies when other variables known to affect such latencies, such as neighborhood size (e.g., Andrews, 1992) and string length (e.g., Weekes, 1997), are controlled.

What Is an Exception Word?

Although some authors (e.g., Plaut et al., 1996) have striven to dissolve the distinction between consistency and regularity by claiming that exception words are simply those that are maximally inconsistent, whereas regular words are those that are maximally consistent, it was recognized as long ago as 1979 in the first article on the effects of consistency on naming latency (Glushko, 1979; see his Footnote 3, p. 684) that the two concepts are distinct.

Almost all of the empirical work on the effects of consistency on reading aloud has used the same definition of consistency: A word is consistent if the pronunciation of its orthographic body (the phonological rime) is the same in all words that share its orthographic body. No doubt one could seek to define other forms of consistency based on other orthographic units, but there is not adequate empirical evidence that consistency in relation to any other orthographic unit has any effect on reading performance.

An exception word is one whose pronunciation as derived from the application of a set of GPC rules differs from its dictionary pronunciation; a regular word is one for which these two pronunciations are the same (Baron & Strawson, 1976; Coltheart, 1978). Words defined as regular by a set of hypotheses regarding GPC rules may be consistent (e.g., *deep*) or inconsistent (e.g., *save*). Exception words defined by such a set of hypotheses also may be consistent (e.g., *calf*) or inconsistent (e.g., *have*). As long as regularity is defined by a clear set of spelling–sound rules, and consistency is defined as based on a particular orthographic unit, these variables can be orthogonally manipulated, and if they can be orthogonally manipulated then the argument that regularity and consistency are the same thing is fallacious.

There may be disputes about whether a particular word is exceptional or regular. However, such disputes are not disputes about the definition of regularity; they are disputes about the GPC rules being used. We have provided in Appendix B the complete set of the DRC model's GPC rules, and our classification of words as regular or exceptional is based on these rules. Given this set of GPC rules, the decision about whether a particular monosyllabic word is regular or exceptional is automatic.

Lexical Decision

The lexical decision task is performed by the DRC model according to the decision procedures proposed by Coltheart, Davelaar, Jonasson, and Besner (1977). The *yes* decision is made if any entry in the orthographic lexicon reaches a criterial activation level (typically .60). The *no* decision is made if a time deadline (expressed in terms of number of processing cycles) elapses before the *yes* decision has been made. This deadline is flexible in the sense that it is computed on every trial on the basis of the summed activation of all units in the orthographic lexicon measured at some stage in processing that is earlier than any *yes* decision time. The larger this sum is, the longer the computed deadline will be, reflecting the fact that the likelihood that the input string is a word will be positively correlated with the value of this summed activation.

When the DRC model carries out lexical decisions according to this procedure, its *yes* latencies are inversely related to word frequency, and they are positively related to the neighborhood size of the word but only when the word is of low frequency. Both of these effects have been reported in studies of human lexical decision (e.g., Andrews, 1989, 1992). Its *no* latencies are directly related to the neighborhood size of the nonword and are longer for pseudohomophones than for nonpseudohomophones, as is true for people (e.g., Coltheart et al., 1977).

Grainger and Jacobs (1996) proposed the addition of a third decision criterion to the two originally proposed by Coltheart et al. (1977), namely, a "fast-guess" mechanism by which a *yes* decision may be made on the basis of the summed activation of the orthographic lexicon. Because this summed activation is a datum yielded by the DRC model, there is no obstacle to using the three-criteria procedure with the DRC model if the human data indicate that this is required.

Reading Aloud

The model is deemed to have generated a reading-aloud response when a sufficient level of activation is present at the phoneme level for each of the relevant phonemes and a null phoneme representing the end of the string. Specifically, within each of these phoneme sets there must be a phoneme unit that has reached a criterial level of activation (we used .43 for this level). The processing cycle at which this is achieved by the model is the model's naming latency. In rare cases, it might be possible for two or more phoneme units in the same phoneme set to rise above criterial activation; if this multiple activation occurs, the phoneme with the highest activation is selected for the naming response when the other relevant phonemes have reached criterial level.

By manipulating the activation-criterion parameter in the model, we were able to simulate reading aloud under speeded and unspeeded conditions. Given sufficient time (a high-activation criterion), the model should respond correctly to all stimuli, because people read exception words correctly when reading at their leisure. When a stricter activation criterion is used, however, the model should make regularization errors, because people make such errors under speeded naming conditions. Consider the activation function shown in Figure 3, which represents the naming of the exception word *wholes*. Here, the model will make a regularization error at any activation criterion less than .47.

The Parameters of the DRC Model

The various parameters of the DRC model and their *standard values* are listed in Table 1. We arrived at these values after several years of work with the DRC model, simulating published data from a large number of studies of human lexical decision and reading aloud. We do not mean to imply that all simulations from the DRC model would use these exact values, because we provide evidence below that strategic effects in human reading experiments can be simulated if the parameters of the DRC model are thought of as being strategically variable.

This modeling strategy might seem to render the model immune from falsification, but that is not the case. There are many logically possible empirical results that could not be

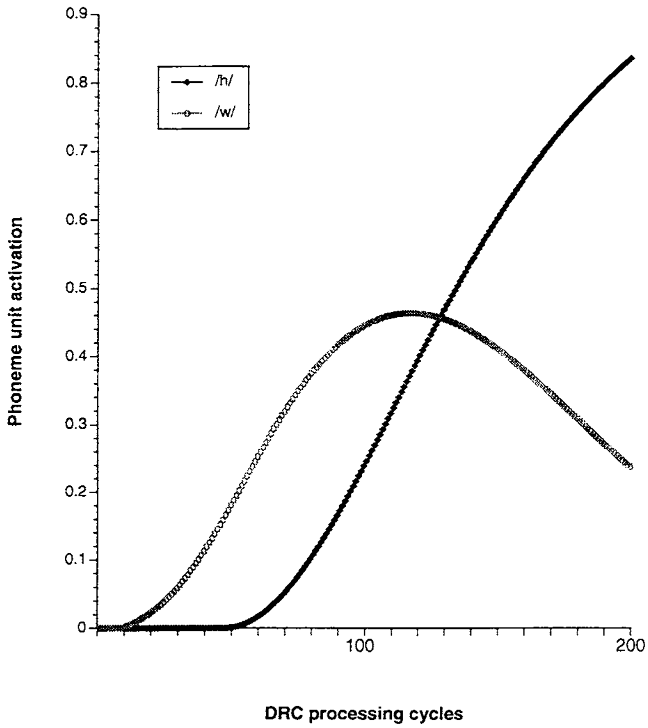

Figure 3. The rise of activation in phonemes /w/ and /h/ in the naming of the word *wholes.* DRC = dual-route cascaded.

simulated by the model no matter what parameters are used—for example, faster reading of nonwords than regular words, or faster reading of exception words than regular words. These may seem crude examples because there are no models that could predict such effects; but there are subtler effects for which this is not the case, one of which is the subject of this article. If there is an effect of irregularity on reading words aloud that is independent of the position in the word of the irregular GPC, then the DRC model, by virtue of its serial GPC procedure, is false.

The Parallel Models

And, we argue, if the effect of irregularity on reading words aloud is in fact dependent on the position in the word of the irregular correspondence between a grapheme and a phoneme, declining monotonically as this position moves from left to right, then models that operate solely in parallel, such as the PDP implementations of Plaut et al. (1996), the connectionist dual-process model (Zorzi et al., 1998), and the multiple-levels model (Norris, 1994), are false.

Plaut et al. (1996), Zorzi et al. (1998), and Norris (1994) have all described models that, though operating under extremely different architectures and processing assumptions, have simulated exception word and nonword reading and the well-known Regularity × Frequency interaction (e.g., Paap & Noel, 1991; Seidenberg, Waters, Barnes, & Tanenhaus, 1984). This interaction—that exception words yield longer naming latencies than regular words but only if they are of low frequency—was once interpreted as very strong evidence for dual-route theories of reading. If a model that does not have a dual-route architecture also yields these effects in simulations, however, then clearly it is wrong to claim that such effects require that the human reading system has a dual-route architecture. Something subtler is now needed to adjudicate between these models.

An Approach to Model Comparison and Evaluation

There are, of course, a number of different approaches to the task of adjudicating between competing models (see, e.g., Grainger & Jacobs, 1996; Jacobs & Grainger, 1994). Our approach is the one advocated by Coltheart and Coltheart (1972). Referring to models of visual word recognition (Rumelhart, 1970) and concept learning (Bower

Table 1
Simulation Parameter Set

Parameter	Value	Parameter	Value
Feature noise	0.000	Lat: Letters to letters	0.000
Letter noise	0.000	Inh: Letters to words	−0.435
Orthographic noise	0.000	Exc: Letters to words	0.070
Phonological noise	0.000	Exc: Phonological to words	0.200
Phoneme noise	0.000	Lat: Words to words	−0.060
Activation rate	0.200	Exc: Words to phonological	0.200
Letter decay	0.000	Inh: Phoneme to phonological	−0.160
Orthographic decay	0.000	Exc: Phoneme to phonological	0.040
Phonological decay	0.000	Lat: Phonological to phonological	−0.070
Phoneme decay	0.000	Inh: Phonological to phoneme	0.000
Frequency scale	1.000	Exc: Phonological to phoneme	0.140
Inh : Features to letters	−0.150	Lat: Phoneme to phoneme	−0.15
Exc: Features to letters	0.005	Exc: GPC to phoneme	0.055
Inh : Words to letters	0.000	GPC: Activation offset	10.000
Exc: Words to letters	0.300	GPC: Left-to-right interval	17.000
		Pronunciation latency: Minimum activation	0.430

Note. Inh = inhibition; Exc = excitation; Lat = lateral inhibition; GPC = grapheme–phoneme correspondence.

& Trabasso, 1964), they argued that definitive adjudication between competing models occurs as a result of experimental work that directly investigates basic postulates of the models. In support of their argument, they cited examples of models (e.g., Bower & Trabasso, 1964) that provided close numerical fits to a variety of data sets but that were subsequently rejected on the basis of experimental work designed to test their fundamental assumptions or postulates (Trabasso & Bower, 1966).

Broadbent (1958, p. 309) argued for exactly this approach to theory adjudication, and Schweickert and Boggs (1984) endorsed the approach thus:

> If a detailed theory is falsified, in itself this is not progress unless the set of remaining theories is notably smaller as a result. [Broadbent] suggests that the optimal strategy is to ask questions so that each answer reduces the number of remaining theories by half. (p. 272)

According to this approach, then, what should be done is to seek some rather general proposition—a Proposition X—about reading whose truth is asserted by one of the models being considered (and indeed by a whole class of models to which this model belongs) and whose truth is denied by the other model we are considering (and indeed by another whole class of models to which this other model belongs). Once such a proposition is identified, an experiment that seeks to determine the truth or falsity of this proposition should not only allow us to adjudicate between the two particular models we are specifically considering but also allow us to achieve Broadbent's (1958) desideratum of halving the number of defensible models.

The Position-of-Irregularity Effect

This adjudication strategy was exactly the one that we had adopted earlier (Coltheart & Rastle, 1994). In our earlier study, X was the following claim: All of the processes by which print is converted to speech are parallel processes. The truth of this claim is asserted by the PDP model and the multiple-levels model; its falsity is asserted by the DRC model (because in that model GPC conversion is a serial left-to-right process).

Following the dual-route explanation of the Regularity × Frequency interaction, we (Coltheart & Rastle, 1994) predicted that for low-frequency exception words, the cost of irregularity would be modulated by the position of the first spelling–sound irregularity in any word. Because nonlexical information about exception words is delivered to the phoneme system serially, words with irregularities in the first position should suffer more than words with irregularities in the fifth position. For example, when the low-frequency word *chef* is read, conflicting nonlexical information about the first phoneme /S/ may arrive at the phoneme system before lexical processing is completed, causing a latency, or accuracy, cost to the reader. Alternatively, when the low-frequency word *swap* is read, lexical processing will most likely finish before conflicting nonlexical information about the third phoneme /Q/ arrives at the phoneme system.

In our earlier study (Coltheart & Rastle, 1994), we examined two-syllable words with irregularities in the first through fifth positions. Controlling for neighborhood size, $\log_{10}$ word frequency, and number of letters, we found that the cost of irregularity (as determined by the irregular target's RT compared with that of a matched regular control) decreased monotonically and linearly over the five positions of irregularity, so that words with irregularities at the fifth phoneme position showed very little cost, whereas words with irregularities at the first position showed nearly a 60-ms cost. We went on to simulate this position-of-irregularity effect with the words *chef, tomb,* and *glow.* The simulation data supported our human data, in that the number of cycles required to name each word decreased over each position of irregularity.

We (Coltheart & Rastle, 1994) concluded that the monotonic decrease in the size of the regularity effect as a function of position of irregularity in the human and simulation data supported the dual-route model of reading and its computational version, the DRC model, and refuted any model that operates solely in parallel. Plaut et al. (1996), however, countered this conclusion by suggesting that a confound may have existed between position of irregularity and consistency, which may have driven the effect. Specifically, they suggested that those words with first-position irregularities may have generally been less consistent than those words with later position irregularities and that the decrease in latency cost over position of irregularity might, therefore, have had nothing to do with the serial nature of reading, but may instead have simply been an artefact of this confound.

One difficulty with Plaut et al.'s (1996) claim is that it is not at all clear how one should go about measuring consistency. As discussed, consistency has usually been defined with respect to the body of a word. This definition is, of course, troublesome for words like *chute,* which contain consistently pronounced bodies but are clearly irregular and inconsistent in the first GPC.

Plaut et al. (1996) realized the difficulty with their claim given the definition of consistency, and so they examined the consistency of each word in our (Coltheart & Rastle, 1994) small simulation at the irregular GPC instead of at the body. They examined the consistency of *chef* by calculating the ratio of friends to enemies at the *ch* segment, the consistency of *tomb* by examining the ratio of friends to enemies at the *to* and the *omb* segments, and the consistency of *glow* by examining the ratio of friends to enemies at the *ow* segment. On the basis of these calculations, they suggested that words with first-position irregularities were generally less consistent than words with second- or third-position irregularities.

As they suspected, Plaut et al. (1996), using these three words, were able to simulate the position-of-irregularity effect on their single-route model of reading, suggesting that the effect may not have arisen from a serial procedure. Unfortunately, in our 1994 study we used a disyllabic stimulus set, and as such, it is virtually impossible to do a post hoc analysis of our data to assess the validity of Plaut et al.'s claim regarding consistency. Consistency is difficult, if not impossible, to define for polysyllabic words because segmenting the word into orthographic components requires syllabification. In a word like *bandage,* for example, it is not clear whether the body of the first syllable is *-and* or *-an.*

Similarly, it is not clear in this example whether the coda in the first syllable is *-nd* or *-n*. Thus, even if one definition of consistency could be agreed upon, performing a post hoc analysis of our 1994 data seems a rather daunting task.

There is a second disadvantage to using disyllabic stimuli: Neither the parallel models nor the DRC model can deal with such stimuli, and therefore model simulations using the actual experimental stimuli, which are of course highly desirable, cannot be carried out. For these reasons, we designed the following experiment, in which we used monosyllabic regular and exception words to investigate the effect of position of irregularity.

Experiment 1

In Experiment 1 we sought to replicate the position-of-irregularity effect originally reported in our earlier study (Coltheart & Rastle, 1994) with the use of a monosyllabic stimulus set controlled for consistency across position of irregularity. Even when consistency across position of irregularity is controlled, models with a serial procedure such as the DRC model predict a position-of-irregularity effect, with first-position irregular words showing a greater cost of irregularity than words with later irregularities. Models that operate solely in parallel do not predict this serial effect.

Our success in demonstrating that the position-of-irregularity effect persists despite controlling for consistency across position of irregularity depends, in part, on the adequacy of the consistency measure that we adopt. As discussed, traditional measures of consistency as based on the orthographic body will not capture the inconsistencies in many first-position irregular words and are thus unsuitable. However, the means by which Plaut et al. (1996) analyzed the words *chef, tomb,* and *glow* for consistency (by examining the ratio of friends to enemies at the inconsistent segment only) seems rather arbitrary; in our assessment of Plaut et al.'s claim, we thus endeavored to adopt a more rigorous approach to the measurement of consistency than they had suggested.

Several such measures already exist (e.g., Berndt, Reggia, & Mitchum, 1987; Rosson, 1985; Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995; Venezky & Massaro, 1987), but for various reasons all of these measures are unsuitable for our purposes. All of these measures provide estimates of the statistical regularity of spelling–sound correspondences by counting the number of times a given orthography maps to one phonology relative to how many times that orthography maps to a different phonology in a sample of English words.

A measure used to examine the statistical regularity of a given orthography should maximize reference to regularities within the language in order to obtain realistic and uninflated measurements of inconsistency. These regularities can be captured by using position-specific information, by using context-sensitive information, and by considering as graphemes those items that are separated by an intervening letter (e.g., *a.e* in *shape*). A reference to position specificity is useful in maximizing regularities in the language because the pronunciation of an orthographic segment can often be predicted by its position in a letter string. The grapheme *y*, for example, is most often pronounced /j/, as in *yarn,* at the beginning of a word; in the middle of a word, it is most often pronounced /I/, as in *gym*; at the end of a word, it is most often pronounced /2/, as in *fly.* Context-sensitive information is useful in the same sense. For example, the statistical regularity of the $c \rightarrow$ /s/ correspondence is increased dramatically if the identity of the following letter is considered. Context-sensitive information can be derived systematically by examining for consistency orthographic units larger than the grapheme. We defined a grapheme as the spelling of a phoneme and therefore considered as graphemes those items with intervening letters. This classification also maximizes an assessment of regularities in the language and is thus desirable in a measurement of overall word consistency. Furthermore, a sound measure of word consistency should make reference to all monosyllabic items in the language and—for our purposes—must make reference to an Australian-English lexicon.

Berndt et al. (1987) provided probabilities of all GPCs based on the analyses of Hanna, Hanna, Hodges, and Rudorf (1966). Although Berndt et al.'s measurement adopts a desirable definition of grapheme—including as graphemes those items with an intervening letter—it is not position specific, it is based on monosyllabic *and* polysyllabic correspondences, and it is based on American English. Another approach was taken by Venezky and Massaro (1987), who designed a fluency measure similar that used by Rosson (1985) but that eliminated the confound between spelling–sound regularity and grapheme frequency in Rosson's measure. Their second-order fluency measure examined the statistical regularity of all GPCs; it was position specific and made reference to all monosyllabic words in English, but it did not recognize as graphemes those items with an intervening letter. Given that such a large number of the items that we used in Experiment 1 contain such graphemes, we deemed this measure unsuitable. Treiman et al. (1995) designed a consistency measure that deals effectively with most of these problems. Their *H* statistic considers the number of pronunciations for a given orthographic unit and the probability that those pronunciations will occur for that orthographic unit. The measure is position specific and adopts a desirable definition of grapheme. However, Treiman et al. calculated this statistic only for monosyllabic, monomorphemic consonant–vowel–consonant words, a sample numbering only 1,329. None of these measures is based on Australian English.

Recognizing the inadequacies of all of these measures, we endeavored to design a consistency calculation for the stimuli used here that takes into account the best features of all of these measurements and is based on Australian English. Our calculation is position specific, captures context-sensitive information, adopts a desirable definition of grapheme, and makes reference to all monosyllabic words in the DRC model's Australian-English database—a healthy sample numbering 7,980. The measure calculates a consistency index for each of five orthographic segments on the basis of the number of friends of each segment and the number of enemies of each segment and is detailed below.

Method

Participants. Participants were 20 first-year psychology students from Macquarie University, Sydney, Australia. All had normal or corrected-to-normal vision and were native Australian-English speakers. They were given course credit for their participation.

Materials. Eighty-eight words with irregular GPCs were chosen from the CELEX English database (Baayen et al., 1993). All target words were monosyllabic, had between three and six letters, and had Kučera and Francis (1967) frequencies between 0 and 22.

All words had irregular GPCs in either the first position, the second position, or the third position and were divided into three lists on that basis. There were 20 words with first-position irregularities, 39 words with second-position irregularities, and 29 words with third-position irregularities.

On the basis of Plaut et al.'s (1996) criticisms regarding a possible confound between consistency and position of irregularity, we examined each irregular target word for consistency at five levels: the head, the body, the antibody, the nucleus, and the coda. The head consists of all consonants before the first vowel, the body consists of the first vowel plus the rest of the word, the antibody consists of the head plus the first vowel, the nucleus consists of the vowel only, and the coda consists of the consonants following the nucleus. In this way, the head plus the body represents the entire word, as does the antibody plus the coda.

To identify each of these orthographic segments, however, we first examined the phonological form of the word. We have termed the five relevant phonological segments here as the *onset* (head), the *rime* (body), the *antirime* (antibody), the *torso* (nucleus), and the *tail* (coda). Once each of these components was isolated, we derived the spelling of each segment.

For each of these orthographic segments, we calculated a consistency index, *C*, based on the number of friends and enemies of that segment:

$$C = [(\text{friends} - \text{enemies})/(\text{friends} + \text{enemies})].$$

These calculations were averaged across segments, such that a value of -1.0 indicated perfect inconsistency and a value of $+1.0$ indicated perfect consistency. The consistency calculation for the word *chef* is shown in Table 2.

The average consistency index of first-position irregular words was .392, the average of second-position irregular words was .181, and the average of third-position irregular words was .219. Thus, consistency was confounded across position of irregularity, as Plaut et al. (1996) suspected, but in a direction that works against our hypothesis: The first-position irregular words were *more* consistent than were words with later irregularities. The entire consistency analysis is contained in Appendix C. First-position irregular words had an average of 2.6 neighbors. Second-position irregular words had 7.44 neighbors, and third-position irregular words had 3.45 neighbors. Matching on the basis of consistency and neighborhood is difficult, if not impossible, when dealing with irregular words. These factors were thus controlled by means of an analysis of covariance (ANCOVA).

Each irregular target word was then matched with a regular control word. Regular words were matched to irregular words on number of letters, initial phoneme, and frequency. In cases in which initial-phoneme matching was not possible, regular words were matched on the basis of phonetic class. The same consistency analysis was performed on the regular words. Results of this analysis showed that the first-position regular words averaged .885, the second-position regular words averaged .859, and the third-position regular words averaged .896. First-position regular words had an average of 5.25 neighbors. Second-position regular words had 8.10 neighbors, and third-position regular words had 4.24 neighbors.

One hundred and eighty-eight monosyllabic nonword fillers were generated. All fillers were orthographically legal and pronounceable.

Apparatus and procedure. Stimulus presentation and data recording were controlled by the DMASTR software (Forster & Forster, 1990), which was run on a DeltaCom 486 personal computer (PC). Responses were timed by means of a voice-key headset that was fitted to the participant to ensure that the microphone was kept at a constant distance from the mouth.

Participants were seated approximately 16 in. (40.64 cm) from the computer monitor. They were instructed to read words presented to them as quickly and as accurately as possible. Stimuli appeared on the screen continuously, one word after another, preceded only by fixation brackets lasting 900 ms, spaced eight characters wide. Pronunciation of the word then triggered immediate presentation of the fixation brackets. If a participant could not pronounce the word, it remained on the screen for 4,000 ms and then was replaced by fixation brackets for the next word. Participants were given 10 practice trials and then received the 370 experimental trials. The experimenter recorded errors by hand.

Table 2
The Calculation of Consistency Index C for the Word Chef

Segment	No. of friends (#F)	No. of enemies (#E)	C (#F − #E)/(#F + #E)
Head	6	151	−.924
Body	3	0	1.000
Antibody	2	9	−.636
Nucleus	516	30	.890
Coda	32	0	1.000
Average			.266

Results

Reaction times for target and control words were collected, and those for spoiled trials (because of voice-key failure) and errors were discarded. If a participant mispronounced an irregular target word, the RT for its matched control was also discarded; conversely, if the participant mispronounced a regular word, its matched irregular target word was discarded. Reaction times for fillers were discarded, and the remainder of the data points were winsorized to the second standard deviation boundary. Item data are contained in Appendix D.

Data were analyzed both by participants and by items. Data by participants were analyzed with an analysis of variance (ANOVA), with regularity and position as factors; data by items were analyzed with a two-way ANCOVA, with two factors (position and regularity) and two covariates (neighborhood size and consistency). Participant and adjusted item means are shown in Table 3. Critically, the interaction between regularity and position of irregularity was significant in the latency data both by participants, F_1(2,

Table 3
Naming Latency (in Milliseconds) and Percentage Error as a Function of Regularity and Position of Irregularity by Participants and by Items (Adjusted for Covariates) in Experiment 1

Regularity	Position of irregularity		
	Position 1	Position 2	Position 3
	Participants		
Irregular targets			
RT	553	506	509
% error	19.0	8.2	7.9
Regular controls			
RT	501	497	510
% error	0.3	1.3	0.9
	Items		
Irregular targets			
RT	552	516	510
% error	17.9	8.9	7.9
Regular controls			
RT	492	502	509
% error	0.0	1.8	0.9

Note. RT = reaction time.

38) = 28.94, $p < .0001$, $MSE = 281.39$, and by items, $F_2(2, 83) = 11.11$, $p < .0001$, $MSE = 921.79$; regularity had a greater effect at early positions of irregularity than at late positions. Position was significant by participants, $F_1(2, 38) = 15.14$, $p < .0001$, $MSE = 426.87$, but not by items, $F_2(2, 83) = 1.75$. Similarly, regularity was significant by participants, $F_1(1, 19) = 27.66$, $p < .0001$, $MSE = 422.38$, but not by items, $F_2(1, 83) = 2.19$.

We carried out planned comparisons using randomization tests[2] to investigate the regularity effect at each position of irregularity. The regularity effect was significant both by participants ($p < .00001$) and by items ($p < .001$) at Position 1. Similarly, Position 2 produced a significant regularity effect both by participants ($p < .01$) and by items ($p < .05$). The regularity effect was not significant at Position 3, either by participants or by items ($ps > .5$).

A regression analysis that examined the relationship between cost of irregularity and position of irregularity while holding the neighborhood and consistency covariates constant confirmed that the relationship between these variables was linear, $F(1, 83) = 21.21$, $p < .0001$, $MSE = 1,843.58$, and did not depart significantly from linearity, $F(1, 83) = 3.03$, $p = .09$. The line describing these data had a y-intercept of 71.88 ms and a regression coefficient of -27.86 ms ($r^2 = .18$).

Errors were tabulated and converted to percents and then were analyzed in the same way as were RTs. Position was significant by participants, $F_1(2, 38) = 18.25$, $p < .001$, $MSE = 0.0018$, but not by items, $F_2(2, 83) = 1.55$. Regularity was significant both by participants, $F_1(1, 19) = 103.75$, $p < .0001$, $MSE = 0.0037$, and by items, $F_2(1, 83) = 5.65$, $p < .05$, $MSE = 0.0176$; more errors occurred for irregular words than for regular words. The interaction between position and irregularity was significant by participants, $F_1(2, 38) = 24.46$, $p < .0001$, $MSE = 0.0018$, and there was a trend toward significance in the item data, $F_2(2, 83) = 2.36$, $p = .10$, $MSE = 0.0176$.

Because the interaction between position of irregularity and regularity was significant by participants and nearly significant by items in the error data, we carried out randomization tests to investigate the size of the regularity effect at each position of irregularity. The regularity effect was significant at all positions, both by participants and by items (all $ps < .05$).

Simulation

Stimuli

The irregular words and the matched regular controls used in the experiment were submitted to the DRC model operating under the standard set of parameters for naming. We expected that the simulation would show a monotonic decrease in the cost of irregularity across position of irregularity.

Results

Response times in cycles were collected for each target word and matched control. There were three errors in Position 1, and none in the other positions. Only one of these errors (/w5lz/ for *wholes*) was a regularization error. The other two errors, *isle* and *tsar,* were the result of lexical and nonlexical blending. Whereas the first phoneme of *isle* was regularized, lexical information accounted for the second phoneme; the resulting pronunciation was /Il/. Similarly, although the first two phonemes of *tsar* were regularized, the third phoneme—the null phoneme—was produced by the lexical procedure; the resulting pronunciation was /ts/. We discarded latencies for these errors and for their matched controls.

We analyzed latency data in the same way as in the experiment. Response times were analyzed in a mixed-design ANOVA, with two factors (regularity and position of irregularity) and two covariates (neighborhood size and consistency). Adjusted means are shown in Table 4, and item data are contained in Appendix D.

[2] Where possible, randomization tests have been carried out in place of other methods of statistical inference. The randomization method is a versatile, precise procedure that is resistant to nonnormal data and nonrandom sampling procedures. Statisticians have asserted that "the randomization test is the truly correct one and that the corresponding parametric test is valid only to the extent that it results in the same statistical decision" (Bradley, 1968, as cited in Edgington, 1995, p. 11). In addition, Cotton (1973, as cited in Edgington, 1995, p. 11) asserted that "randomization tests permit us to drop the most implausible assumption of typical psychological research—random sampling from a specified population . . ." and that "random sampling occurs infrequently in behavioral research and . . . therefore, any statistical tests making that assumption are questionable unless otherwise justified." Whereas randomization programs are available (Edgington, 1995) for both between-groups and within-groups one-way comparisons, programs for full-factorial and mixed-design procedures are not. In the data presented here, we therefore used the randomization procedure for one-way comparisons only. All tests in this work were carried out with 100,000 random permutations.

Table 4
Naming Latency (in DRC Processing Cycles, Adjusted for Covariates) as a Function of Regularity and Position of Irregularity

Regularity	Position of irregularity		
	Position 1	Position 2	Position 3
Irregular targets	97.25	87.83	79.84
Regular controls	78.61	77.83	78.12

Note. DRC = dual-route cascaded.

Critically, the interaction between regularity and position of irregularity was significant, $F(2, 80) = 33.82$, $p < .001$, $MSE = 22.60$, as the regularity effect was larger for words with irregularities in early positions than for words with irregularities in late positions. Both regularity, $F(1, 80) = 43.46$, $p < .001$, $MSE = 22.60$, and position of irregularity, $F(2, 80) = 47.14$, $p < .001$, $MSE = 17.93$, were significant.

We carried out planned comparisons between irregular words and regular matched controls at each position by means of randomization tests. Randomization tests showed significant regularity effects at Positions 1 and 2 ($ps < .001$). The regularity effect was not significant at the third position ($p > .05$).

A regression analysis designed to assess the relationship between cost of irregularity and position of irregularity while holding the neighborhood and consistency covariates constant confirmed that the relationship between these variables had a significant linear component, $F(1, 80) = 65.23$, $p < .0001$, $MSE = 45.20$; the residual nonlinear regression was not significant ($F < 1$). The regression equation describing these data had a y-intercept of 28.70 cycles and a regression coefficient of -8.43 cycles ($r^2 = .46$).

We compared the DRC data with human latency data to assess how much of the variance in human RT was captured by the DRC model. Because variance that cannot be explained by the DRC model—variance due to the phonetic class of the onset—is controlled through pairwise matching across regularity condition in our stimuli, we eliminated it by comparing the cost of irregularity produced by people with the cost of irregularity produced in the DRC model for each item pair. The resulting correlational analysis produced an r^2 of .244; thus 24.4% of the variance in the human data was accounted for by the DRC model.

How should this estimate of the variance accounted for by the DRC model be evaluated? That is, how much of the total variance in the human data is due to systematic linguistic variables that *could* be accounted for by the model and how much of this variance can be thought of as error due to individual differences in participants' performance? One method of estimating the contribution of variance from these systematic linguistic variables to the total variance is to calculate the overall reliability of the cost-of-irregularity measure within our participant sample. By correlating the cost of irregularity for each item pair for one random half of our sample with the cost of irregularity for each item pair for the other half of our sample, we should arrive at a reliability figure that does not include participant error variance; this correlation is $+.604$, $r^2 = .365$. On the basis of this reliability measure, then, we may think of 36.5% of the variance in cost of irregularity as being due to factors other than individual participant error variance. This is the largest amount of variance for which the DRC model should account. Thus, when evaluating the adequacy of the 24.4% of the variance in cost of irregularity accounted for by the DRC model, this 24.4% figure should not be measured against a total 100% of the variance but rather against the reliability figure of 36.5%.[3]

Discussion of Experiment 1 and the Simulation Data

The experimental and simulation data presented here suggest that the position of irregularity in an exception word modulates the regularity effect. Words with early irregularities (in the first or second positions) show a greater regularity effect than do words with later irregularities, even when differences in consistency across position of irregularity are controlled. Human data suggest that the function relating cost of irregularity to position of irregularity decreases monotonically and is linear; DRC data show the same monotonic and linear trend. Moreover, fine-grained analyses between human data and DRC data established that DRC data account for a portion of the variance produced in human data that we believe is satisfactory at this stage of modeling.

Although both human data and DRC data show a linear trend, the human data *appear* to be nonlinear; moreover, because the cost of irregularity declines from 60 ms to 14 ms to 1 ms over position of irregularity in the human data, the position-of-irregularity effect *appears* to be restricted to, or driven by, the exception words with first-position irregularities. This is clearly not the case, however. The human data show a significant regularity effect at Position 2, as did our earlier (Coltheart & Rastle, 1994) data; moreover, the human data fit the linear trend extremely well, with no significant departure from linearity.

The error data are more puzzling than the latency data. Participants showed no interaction between regularity and position of irregularity in the error data, although the human means do indicate that words with first-position irregularities produced a greater accuracy cost than did words with later position irregularities. The nature of the items was such that many items produced extremely high error rates, and other items produced no errors. As such, there was tremen-

[3] It could be argued that if model adequacy is to be assessed relative to a split-half reliability estimate in the way that we have advocated here, then only half of the model item data should be entered into the calculation of the r^2 statistic. To address this potential objection, we split at random the item pairs and recalculated the correlations between human cost of irregularity and model cost of irregularity. The resulting r^2 statistics for each half of the item data separately reflected what we observed when the entire sample was used: The DRC model accounted for a relatively high percentage of the variance in the human cost of irregularity ($r^2 = .28$ and $.22$) compared with the split-half reliability estimates derived from the human sample.

dous variability in the scores, decreasing the possibility of finding a statistically significant interaction.

As we predicted would be the case in the human data, the DRC model showed a position-of-irregularity effect in the error data, producing errors only in words with first-position irregularities. This result replicates the general pattern of error data produced by people, although the human data did show an error effect in both the second and third positions as well as in the first position. If the error effect at Positions 2 and 3 in the human data is reliable, then further simulation work with the DRC model may be necessary to simulate both the latency data and the error data perfectly.

In summary, Experiment 1 replicated our earlier (Coltheart & Rastle, 1994) findings but controlled for the possibility of a confound between position of irregularity and consistency; the cost of irregularity decreases monotonically and linearly as a function of the position of the irregular spelling–sound correspondence in an exception word. Together, the human and DRC data strongly suggest that a serial process is at work in reading aloud.

Experiment 2

Experiment 1 replicated our earlier finding that the position of irregularity in an exception word modulates the regularity effect (see Coltheart & Rastle, 1994). This effect remained even when controlling for consistency across position of irregularity.

We (Coltheart & Rastle, 1994) also reported a failure to find evidence of strategic manipulation of the lexical and nonlexical routes. Following Baluch and Besner (1991), Tabossi and Laghi (1992), and Monsell, Patterson, Graham, Hughes, and Milroy (1992), we had hypothesized that if lexical and nonlexical procedures both function in reading, readers may be able to strategically alter their use of one or both of the routes, depending on task demands. If one were reading only exception words, reading would proceed more quickly and accurately if the nonlexical route were slowed, because that route produces the wrong response for every stimulus. Alternately, when faced with only nonwords, a reader might find it helpful either to make greater use of the nonlexical route or to make less use of the lexical route. If evidence were found that implicated this sort of strategic processing in reading, it could be used as strong support for a dual-route framework.

After demonstrating a position-of-irregularity effect with nonword fillers, we (Coltheart & Rastle, 1994) predicted that changing the nature of the fillers would change the cost of irregularity at each position: If the fillers were changed to exception words, perhaps participants would make less use of the nonlexical route. Because we theorized that the position-of-irregularity effect was due to the serial processing of the nonlexical route, decreasing its use should diminish the position effect and, perhaps, the regularity effect altogether. Against prediction, there was no evidence in our 1994 study that participants had strategically altered their use of either of the routes. In fact, the Position × Regularity interaction was exactly the same in both filler conditions.

We offered several possible reasons why we did not find the hypothesized strategy effect; however, none made the issue clear (see Coltheart & Rastle, 1994). What we failed to take into account were the results of our first experiment.

In Experiment 1 of our 1994 study, we showed that the regularity effect is modulated by the position of the irregular phoneme in an exception word; first-position irregular words show a large regularity effect; words with fourth- and fifth-position irregularities do not show a regularity effect. Similarly, in the first experiment of the present article, we demonstrated that monosyllabic words with irregularities in the first position show a large regularity effect, and words with irregularities in the third position do not show a regularity effect. If words with late irregularities do not show a regularity effect (even when mixed with nonwords), reading them in pure blocks should not elicit a strategic slowing of the nonlexical route.

In our earlier strategy manipulation, we used nonword fillers in one condition and exception-word fillers in the other condition (see Coltheart & Rastle, 1994). However, further analysis of our exception-word fillers shows that although most of these words had second-position irregularities, only 38 had first-position irregularities, and 51 had irregularities in the third or later positions; in fact, the average point of irregularity was at 2.3 phonemes. If we had wanted participants to alter the use of the nonlexical route, we should have used exception-word fillers with irregularities in the first position, because those exception words are the most costly in reading aloud.

If all exception-word fillers had irregularities in the first position, participants may have been more inclined to slow their use of the nonlexical route (there were no nonwords to read in this condition), thus diminishing the position effect and the regularity effect entirely. Instead, we used exception-word fillers with late irregularities, which may not have forced participants to alter their use of the nonlexical route (see Coltheart & Rastle, 1994).

We therefore designed a strategy experiment that used first-position irregular fillers in one condition and third-position irregular fillers in another condition. Targets were monosyllabic nonwords and monosyllabic regular words. Of many possible outcomes, three would be compatible with the DRC model. Because the DRC model does not make any explicit predictions about whether strategic effects occur in reading at all, the absence of a strategy effect would not be incompatible with the DRC model. If strategy effects do occur in reading, however, the DRC model predicts that first-position irregular fillers will either speed the lexical route or slow the nonlexical route. If use of the lexical route is increased, then regular-word naming should become faster, and nonword naming should become slower. If the nonlexical route is slowed when naming first-position irregular words, both nonword naming and regular-word naming should become slower. Regular-word naming times will thus reveal which route is under strategic control in this manipulation.

Method

Participants. Participants were 24 first-year students from Macquarie University. All had normal or corrected-to-normal

	Nonword Targets	Regular Word Targets
Pos1 Fillers	Subject Group A: Targets 1-25, Fillers 1-25	Subject Group A: Targets 51-75, Fillers 26-50
	Subject Group B: Targets 26-50, Fillers 26-50	Subject Group B: Targets 76-100, Fillers 1-25
Pos3 Fillers	Subject Group A: Targets 26-50, Fillers 51-75	Subject Group A: Targets 76-100, Fillers 76-100
	Subject Group B; Targets 1-25, Fillers 76-100	Subject Group B: Targets 51-75, Fillers 51-75

Figure 4. The design of Experiment 2. Pos = position.

vision and were native Australian-English speakers. They received course credit for their participation.

Materials. Two lists of targets and two lists of fillers were created. One list of targets contained 50 monosyllabic regular words. The other list of targets contained 50 monosyllabic nonwords that were phonotactically legal. One list of fillers comprised 50 monosyllabic exception words, irregular in the first position; the other list of fillers comprised 50 monosyllabic exception words, irregular in the third position. Each list of targets and fillers was then divided randomly into two separate lists of 25 items each. Lists of targets and fillers were paired together on the basis of participant group. The design of the experiment is shown in Figure 4.

Words in each list were randomized. Each list was preceded by 12 practice items. Practice items were exception words with irregularities either in the first or third position, depending on the composition of the list.

Stimulus presentation and data recording were controlled by the DMASTR software package (Forster & Forster, 1990), which was run on a DeltaCom 486 PC. Latencies were measured by means of a voice key that fit to the participant's head, thus keeping the microphone at a constant distance from the mouth.

Procedure. Participants were divided into two equal groups and randomly assigned to a counterbalancing condition. Every participant took part in all four conditions. Every participant saw every target and every filler but saw each target and each filler only once. Additionally, the order in which the four lists were presented to each participant was counterbalanced in two ways: Order of filler condition was controlled so that participants saw first-position irregular fillers first as often as they saw third-position irregular fillers first; additionally, order of lexical condition was controlled so that each participant began with nonword targets as often as they began with regular-word targets. To maximize any strategic manipulations, we always presented filler conditions together; participants either started with two first-position filler conditions or started with two third-position filler conditions. Targets and fillers within each list were presented randomly.

Participants were asked to name each word as quickly and as accurately as possible. Before being presented with each list, they were given the 12 practice items. Each list contained 50 items; participants therefore named 200 items in total plus the practice items.

Results

Naming latencies were recorded by means of the DMASTR software, and errors were recorded by hand. Naming latencies for errors and spoiled trials were discarded, and the remaining RTs were winsorized to the second standard deviation boundary. Participant data are shown in Table 5, and item data are contained in Appendix E.[4]

A mixed-design ANOVA with three factors (participant block or item list, target lexicality, and filler condition) was carried out on the latency and error data. In the participant analysis, lexicality and filler condition were treated as within-groups factors, and participant block was treated as a between-groups factor; in the item analysis, filler condition was treated as a within-groups factor, and lexicality and item list were treated as between-groups factors.

Main effects of lexicality and filler condition emerged in the latency data in both the by-participants and the by-items analyses. Nonwords were read aloud more slowly than regular words by participants, $F_1(1, 22) = 58.23, p < .0001$, $MSE = 1{,}348.80$, and by items, $F_2(1, 96) = 58.79, p < .0001$, $MSE = 2{,}683.25$. In addition, targets paired with first-position irregular-word fillers were read aloud more slowly than targets paired with third-position irregular-word fillers by participants, $F_1(1, 22) = 7.10, p < .015, MSE = 906.16$, and by items, $F_2(1, 96) = 35.61, p < .0001, MSE = 410.15$. There was no interaction between lexicality and filler condition, however, either by participants, $F_1(1, 22) =$

[4] Because item means were not adjusted in this analysis, they were very similar to the participant means. Therefore, only participant means are reported here; complete item data are contained in Appendix E.

Table 5
Naming Latency (in Milliseconds) and Percentage Error as a Function of Target Type and Filler Condition by Participants

Target type	Position 1 fillers	Position 3 fillers
Nonword		
RT	549	529
% error	2.8	2.5
Regular word		
RT	488	476
% error	0.7	0.3

Note. RT = reaction time.

0.97, or by items, $F_2(1, 96) = 3.28$, $p = .07$. There were no effects of participant block or item list.

The error data were analyzed in the same way as were the latency data. Only the effect of lexicality emerged both in the participant, $F_1(1, 22) = 5.38$, $p < .05$, $MSE = 1.31$, and the item, $F_2(1, 96) = 9.35$, $p < .01$, $MSE = 0.36$, analyses, as nonwords attracted more errors than regular words. There were no effects of filler condition in either analysis, $F_1(1, 22) = 0.58$, $F_2(1, 96) = 0.43$. Finally, there were no interactions between lexicality and filler condition in either analysis, $F_1(1, 22) = 0.00$; $F_2(1, 96) = 0.00$.

Discussion

These data demonstrate that a target's naming latency is affected by the nature of the fillers present. Whether targets are regular words or nonwords, they are named more slowly when they are mixed with first-position irregular fillers than when they are mixed with third-position irregular fillers, indicating a general slowing of the nonlexical route, not a speeding of the lexical route.

These data lend support to the dual-route account of reading, and paired with Monsell et al.'s (1992) findings, they suggest that strategic effects do occur in languages like English with deep orthography, even though most of the evidence for these effects resides in languages with shallow orthography (e.g., Baluch & Besner, 1991; Tabossi & Laghi, 1992).

Lupker, Brown, and Colombo (1997), however, have recently reported data that suggest an alternative interpretation to the one we have suggested. They posited that the sorts of strategy effects reported in Monsell et al. (1992) reflect not a de-emphasis of routes but rather a more a general principle influencing speeded performance in the naming task: Simply put, when fast things are mixed with slow things, those fast things slow down; when slow things are mixed with fast things, those slow things speed up. Lupker et al. (1997) have found this general principle at work across several different stimulus types, and though this principle does not necessarily discount the possibility of route shifting, it may account for at least some of the strategy data contained in the literature.

In this experiment, regular words and nonwords were mixed either with first-position irregular fillers or with third-position irregular fillers. As shown in the first experiment, first-position irregular words are named more slowly than third-position irregular words. According to Lupker et al. (1997), because first-position irregular words are named so slowly, they drag down the naming latencies of both the regular-word targets and the nonword targets. Of course, this is the same prediction given by the dual-route account and is what we have reported. Thus, it is difficult using these data to disentangle the two theories. One possibility for disentangling these theories might be to examine the position-of-irregularity effect as a function of filler condition, using fillers with first- or third-position irregularities. Although the Lupker et al. account would predict a main effect of target latency driven by filler condition, it would not predict that filler condition would modulate the slope of the function relating cost of irregularity to position of irregularity, as a route de-emphasis account would predict. Unfortunately, there are not enough words that contain first-position irregularities to carry out this experiment. Clearly, because of the generality of Lupker et al.'s principle, it may be challenging to adjudicate between that account and the route-shifting account of strategy effects.

A post hoc analysis of the strategy data we had obtained in our earlier study (Coltheart & Rastle, 1994) may be helpful, however, in beginning to evaluate and compare these theories. In our attempt to modulate the size of the regularity effect by altering the nature of filler items, we reported no effects of filler condition that reached significance both by participants and by items in our 1994 study. Against our prediction, filler condition (nonwords or exception words) did not affect either general naming latencies of targets and controls across position of irregularity or interact with regularity across position of irregularity.

In reexamining our (Coltheart & Rastle, 1994) data, however, a randomization test showed a significant difference between nonword filler latency ($M = 577$) and exception-word filler latency ($M = 479$), $p < .0001$. Because one type of filler was much faster than the other type, and because these fillers were paired with the same set of targets, Lupker et al.'s (1997) account would predict that a main effect of filler condition would emerge in the target set: The very fast filler set would decrease target latencies, and the very slow filler set would increase target latencies. However, there was no difference in the general RTs of targets by filler condition. Thus, Lupker et al.'s account—although it may explain some other strategy data—does not seem to account for our earlier data (see Coltheart & Rastle, 1994).

In summary, the strategy effect we report in the present experiment may be explained within a route-shifting account or within the account posed by Lupker et al. (1997). Our explanation for this effect, based on the notion that participants can alter use of the nonlexical route, also accounts for our 1994 failure to find a strategy effect; the Lupker et al. account does not. Of course, although the route-shifting account seems better able to explain the data discussed here, further simulation work is clearly required to assess whether the DRC route-shifting account can explain the data reported by Lupker et al.

Simulation

Stimuli and Parameter Set

Although the DRC model has no means by which to adjust the speed of the nonlexical route in response to the composition of the stimulus set, it is not difficult to manually adjust one of the three nonlexical parameters to simulate a slowing of the nonlexical route. We simulated this decrease in processing speed by increasing the parameter that controls the speed of serial nonlexical assembly in the model. This parameter is normally set to 17 cycles. For the simulation of these data, the parameter value was increased to 22 cycles.

Fifty regular words and 50 nonwords used in Experiment 2 were submitted to the DRC model for naming. Except for the adjustment to the interletter interval parameter, the parameter set shown in Table 1 was used again here. We expected that slowing the nonlexical route would slow both regular words and nonwords, as was the case in the human data.

Results

We collected DRC naming latencies (in cycles), which are displayed in Table 6. Item data are shown in Appendix E. The DRC model named all of the regular words correctly in both the standard naming condition and the condition in which the speed of serial processing was decreased. The DRC model made a number of lexicalization errors when naming the nonword items: It produced two lexicalization errors in the standard naming condition and produced five lexicalization errors in the condition in which the nonlexical route was manipulated. These items were removed from the analyses.

The DRC simulation produced a lexicality effect, as nonword targets were named more slowly than regular-word targets: $F(1, 93) = 937.97, p < .0001, MSE = 395.42$. The simulation also revealed a main effect of strategic manipulation, as both nonword targets and regular-word targets were slowed when the speed of the nonlexical route was decreased, $F(1, 93) = 197.54, p < .0001, MSE = 31.85$. Target lexicality and strategic manipulation interacted, because nonword naming was slowed more by the slowing of the nonlexical route than was regular word naming, $F(1, 93) = 178.81$, $p < .0001$, $MSE = 31.85$. Randomization tests showed, however, that slowing the nonlexical route slowed both nonword naming ($p < .0001$) and regular-word naming ($p = .05$).

Table 6
Naming Latency (in DRC Processing Cycles) as a Function of Target Type and Strategy Manipulation

	Manipulation	
Target type	Before	After
Nonword	155.42	177.91
Regular word	77.90	78.46

Note. DRC = dual-route cascaded.

Discussion

In Experiment 2, first-position irregular fillers were accompanied by slowed regular-word and nonword naming. We believe that this effect was due to a strategic slowing of the nonlexical route. When the nonlexical route was slowed manually in the DRC model, the results were the same: both regular-word naming and nonword naming were slowed.

Although both regular words and nonwords were slowed in the simulation, they were not slowed equally. Slowing the rate of the serial operation in the nonlexical route had far greater consequences for nonwords than it had for regular words, which were affected only slightly by the slowing of the serial process in the nonlexical route. This interaction in the simulation data may prove problematic for the DRC model or its standard parameter set, given that the human data did not show a statistically significant interaction between word type and filler condition, though the means showed a trend in the appropriate direction. Further simulation work and model development may have to consider more precisely the contribution of lexical and nonlexical procedures to regular-word and nonword naming to simulate the null interaction produced by people.

A close inspection of the item data contained in Appendix E suggests that the nonlexical contribution to regular-word naming in the DRC model is not as straightforward as we originally thought; whereas slowing the serial procedure generally hurt regular-word naming, in some cases it actually speeded regular-word naming, and in other cases it had no effect at all. Simulation work must now focus on identifying the conditions under which nonlexical processing harms and helps regular-word naming in the DRC model; this work will most likely involve investigating how length and the presence of whammies[5] (Rastle & Coltheart, 1998) affect word naming in human readers and in the DRC model. In addition, further simulation work may consider alternative means of simulating the types of strategic modulations reported here, perhaps by decreasing the GPC excitation parameter or by increasing the GPC activation offset parameter.

Whether or not these data pose a problem for the DRC model or its standard parameter set, the basic strategy effect—that naming is slowed in the face of exceptional fillers with first-position irregularities—remains good evidence for the dual-route theory generally. More important, because strategic adjustment occurred as a result of the introduction of first-position irregular fillers, models that attempt to explain these results must contain some

[5] When a word contains a GPC in which the phoneme is represented by two or more letters, such as *ch* or *ee*, the left-to-right nonlexical procedure of the DRC model will initially translate only the first letter of these multiletter graphemes to an incorrect phoneme. For example, the multiletter grapheme *ch* will initially activate the incorrect phoneme /k/, which will compete with the subsequently generated correct phoneme /J/. We termed this inhibitory influence on the correct phoneme a *whammy* (Rastle & Coltheart, 1998), and we report data that suggest that not only the DRC model but also human readers are susceptible to this effect when reading nonwords.

process—we believe a serial process—that explains why first-position irregular fillers are accompanied by slower target naming than are third-position irregular fillers. Because it is unclear how the parallel models would capture the distinction between first-position irregular word fillers and third-position irregular word fillers, we do not believe that these data are compatible with those models.

General Discussion

Experiment 1 confirmed the existence of the position-of-irregularity effect in naming that we had reported earlier (Coltheart & Rastle, 1994). Even when controlling for consistency across position of irregularity, a larger cost of irregularity was present for first-position irregular words than was present for third-position irregular words. From these data, we conclude that the regularity effect is modulated not only by frequency (e.g., Seidenberg et al., 1984) but also by the position of irregularity in an exception word. Only exception words that are of low frequency and contain irregularities in early positions should show a latency or accuracy cost in naming.

From these results, we predicted that Experiment 2 would reveal a strategy effect in nonword and regular-word reading if the filler items were varied by position of irregularity. Indeed, those fillers that caused readers the greatest difficulty (first-position irregular fillers) led to slower target naming than fillers that caused readers no difficulty (third-position irregular fillers). These results thus represent strong evidence for two routes in the reading system, and we suspect that these effects come about through participants' explicit or implicit control over their use.

These experiments as a whole suggest that a new variable—position of irregularity—modulates the regularity effect and consequently should be examined and controlled in experiments investigating the naming of irregular words.

References

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

Andrews, S. (1992). Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 234–254.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database* [CD-ROM]. Philadelphia: University of Pennsylvania, Linguistic Data Consortium.

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, 2,* 386–393.

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

Bower, G., & Trabasso, T. (1964). Concept identification. In R. C. Atkinson (Ed.), *Studies in mathematical psychology* (pp. 32–94). Palo Alto, CA: Stanford University Press.

Broadbent, D. E. (1958). *Perception and communication.* Elmsford, NY: Pergamon Press.

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

Coltheart, M. (1985). Cognitive neuropsychology and the study of reading. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and performance XI* (pp. 3–37). Hillsdale, NJ: Erlbaum.

Coltheart, M., & Coltheart, V. (1972). On Rumelhart's model of visual information processing. *Canadian Journal of Psychology, 26,* 292–295.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review, 100,* 589–608.

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

Coltheart, M., Langdon, R., & Haller, M. (1996). Computational cognitive neuropsychology and acquired dyslexia. In B. Dodd, L. Worrell, & R. Campbell (Eds.), *Models of language: Illumination from impairment* (pp. 9–36). London: Whurr.

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.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (1998). *DRC: A computational dual-route model of visual word recognition and reading aloud.* Manuscript submitted for publication.

Edgington, E. S. (1995). *Randomization tests.* New York: Marcel Dekker.

Ellis, A. W., & Young, A. W. (1988). *Human cognitive neuropsychology.* East Sussex, England: Psychology Press.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12,* 627–635.

Forster, K. I., & Forster, J. C. (1990). *User's guide to the DMASTR display system: Laboratory software for mental chronometry.* Tucson: University of Arizona.

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

Gough, P. B., & Cosky, M. J. (1977). One second of reading again. In N. J. Castellan, D. B. Pisoni, & G. R. Potts (Eds.), *Cognitive theory* (Vol. 2, pp. 271–288). Hillsdale, NJ: Erlbaum.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103,* 518–565.

Hanna, P. R., Hanna, J. S., Hodges, R. E., & Rudorf, E. H. (1966). *Phoneme–grapheme correspondences as cues to spelling improvement.* Washington, D.C.: U.S. Department of Health, Education, and Welfare.

Jacobs, A. M., & Grainger, J. (1994). Models of visual word recognition—Sampling the state of the art. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 1311–1334.

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

Lupker, S. J., Brown, P., & Colombo, L. (1997). Strategic control in a naming task: Changing routes or changing deadlines? *Journal* *of Experimental Psychology: Learning, Memory, and Cognition, 23,* 570–590.

Masterson, J. (1985). Nonword reading in different populations. In K. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading* (pp. 289–299). London: Erlbaum.

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

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

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

Norris, D. (1994). A quantitative multiple-levels model of reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 1212–1232.

Ogden, J. A. (1996). Phonological dyslexia and phonological dysgraphia following left and right hemispherectomy. *Neuropsychologia, 34,* 905–918.

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

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

Patterson, K. E., & Shewell, C. (1987). Speak and spell: Dissociations and word-class effects. In M. Coltheart, G. Sartori, & R. Job (Eds.), *The cognitive neuropsychology of language* (pp. 273–294). London: Erlbaum.

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

Rastle, K., & Coltheart, M. (1998). Whammies and double whammies: The effect of length on nonword reading. *Psychonomic Bulletin & Review, 5,* 277–282.

Rastle, K., & Coltheart, M. (1999). Lexical and nonlexical phonological priming in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 461–481.

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

Rumelhart, D. E. (1970). A multicomponent theory of the perception of brief visual displays. *Journal of Mathematical Psychology, 7,* 191–218.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: Part II. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review, 89,* 60–94.

Rumelhart, D. E., & Siple, P. (1974). Process of recognizing tachistoscopically presented words. *Psychological Review, 81,* 99–118.

Schweickert, R., & Boggs, G. J. (1984). Models of central capacity and concurrency. *Journal of Mathematical Psychology, 28,* 223–281.

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.

Tabossi, P., & Laghi, L. (1992). Semantic priming in the pronunciation of words in two writing systems: Italian and English. *Memory & Cognition, 20,* 315–328.

Trabasso, T., & Bower, G. (1966). Presolution dimensional shifts in concept identification: A test of the sampling with replacement axiom in all-or-none models. *Journal of Mathematical Psychology, 3,* 163–173.

Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General, 124,* 107–136.

Venezky, R. L., & Massaro, D. W. (1987). Orthographic structure and spelling–sound regularity in reading English words. In D. A. Allport, D. G. MacKay, W. Prinz, & E. Scheerer (Eds.), *Language perception and production: Shared mechanisms in listening, speaking, reading, and writing* (pp. 159–175). San Diego, CA: Academic Press.

Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latency. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 50A,* 439–456.

Zorzi, M., Houghton, G., & Butterworth, B. (1998). Two routes or one in reading aloud? A connectionist dual-process model. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1131–1161.

(Appendixes follow)

Appendix A
The Phonemic Vocabulary of the Dual-Route Cascaded Model

Symbol	Example	Symbol	Example	Symbol	Example
1	bay	S	sheep	n	nat
2	buy	T	thin	p	pat
3	burn	U	put	r	rat
4	boy	V	putt	s	sap
5	no	Z	measure	t	tack
6	brow	b	bad	u	boon
7	peer	d	dad	v	vat
8	pair	f	fat	w	why
9	poor	g	game	z	zap
D	then	h	had	#	barn
E	pet	i	bean	{	pat
I	pit	j	yank	_	jeep
J	cheap	k	cad		
N	bang	l	lad		
Q	pot	m	mad		

Appendix B
Nonlexical Rules of the Dual-Route Cascaded Model

Grapheme	Phoneme	Position
Multiletter rules		
arre	#	A
augh	9	A
eare	7	A
eere	7	A
eigh	1	A
ough	9	A
tsch	J	e
urre	3	A
ai.e	1	A
ar.e	#	A
aw.e	9	A
ea.e	i	A
ee.e	i	A
er.e	3	A
ie.e	i	A
o.ue	5	A
oa.e	5	A
oi.e	4	A
oo.e	u	A
or.e	9	A
ou.e	6	A
ow.e	6	A
oy.e	4	A
ur.e	3	A
air	8	A
are	8	A
arr	#	A
awe	9	A
aye	2	A
ear	7	A
eer	7	A
ere	7	A
err	3	A
ewe	ju	A
ier	7	A
igh	2	A
irr	3	A
oar	9	A
oor	9	A
ore	9	A
our	9	A
ure	9	A
urr	3	A
Multiletter rules (*continued*)		
che	S	e
dge	_	m,e
gue	g	m,e
lle	l	m,e
que	k	m,e
tch	J	m,e
the	D	m,e
a.e	1	A
e.e	i	A
i.e	2	A
o.e	5	A
u.e	ju	A
y.e	2	m,e
aa	#	A
ah	#	A
ai	1	A
ar	#	A
au	9	A
aw	9	A
ay	1	A
bb	b	m,e
cc	k	m,e
ch	J	A
ce	s	m,e
ck	k	m,e
dd	d	m,e
de	d	m,e
ea	i	A
ee	i	A
ei	1	A
er	3	A
eu	u	A
ew	ju	A
ey	1	A
ff	f	m,e
ge	_	m,e
gg	g	m,e
gh	g	b
gn	n	A
ie	2	A
ir	3	A
je	_	m,e

Appendix B (*continued*)

Grapheme	Phoneme	Position
Multiletter rules (*continued*)		
jj	_	m,e
kh	k	A
kk	k	m,e
kn	n	b
le	l	m,e
ll	l	m,e
mb	m	m,e
mm	m	m,e
mn	m	m,e
ng	N	m,e
nn	n	m,e
oa	5	A
oe	5	A
oh	5	A
oi	4	A
oo	u	A
or	9	A
ou	6	A
ow	6	A
oy	4	A
ph	f	A
pp	p	m,e
ps	s	b
re	r	m,e
rh	r	b
se	s	m,e
sh	S	A
ss	s	m,e
te	t	m,e
th	T	A
tt	t	A
ue	ju	A
ui	u	A
ur	3	A
uy	2	m,e
ve	v	m,e
vv	v	m,e
wh	w	b
wr	r	b
x	ks	A
ye	2	m,e
ze	z	m,e
zz	z	m,e
Context-sensitive rules[a]		
gu[V]	g	b
n[k]	N	A
[q]u	w	m
g[e]	_	b
c[e]	s	A
c[i]	s	A
c[y]	s	A
[V][C]ed	d	e
[V][C][C]	d	e
[V][C][C][C]ed	d	e
Single-letter rules		
a	{	b,m
a	#	e
b	b	A
c	k	A
d	d	A
e	E	b,m
e	i	e
Single-letter rules (*continued*)		
f	f	A
g	g	A
h	h	A
i	I	b,m
i	2	e
j	_	A
l	l	A
k	k	A
m	m	A
n	n	A
o	Q	b,m
o	5	e
p	p	A
q	k	A
r	r	A
s	s	A
t	t	A
u	V	b,m
u	u	e
v	v	A
w	w	A
y	j	b
y	I	m
y	2	e
z	z	A
Output (phonotactic) rules[b]		
#s	#z	e
1s	1z	e
2s	2z	e
3s	3z	e
4s	4z	e
5s	5z	e
6s	6z	e
7s	7z	e
8s	8z	e
9s	9z	e
Ds	Dz	e
Ns	Nz	e
bs	bz	e
ds	dz	e
gs	gz	e
is	iz	e
ls	lz	e
ms	mz	e
ns	nz	e
us	uz	e
vs	vz	e
dT	tT	m,e
nk	Nk	m,e
pd	pt	m,e
kd	kt	m,e
Sd	St	e
Jd	Jt	e
rju	ru	A
Sju	Su	A
_ju	_u	A
lju	lu	A
Jju	Ju	A
sd	st	A
tz	ts	A
Td	Tt	A
fd	ft	A

Note. A = all; b = beginning; m = middle; e = end.
[a]The grapheme outside the brackets in column 1 is converted to the phoneme in column 2 in the presence of the preceding or following context specified in the brackets. [V] = any vowel; [C] = any consonant.
[b]The phoneme string in the first column is converted to the phoneme string in the second column.

(*Appendixes continue*)

Appendix C
Consistency Analysis

Word	Head	Body	Antibody	Nucleus	Coda	Total	Word	Head	Body	Antibody	Nucleus	Coda	Total
			Irregular words							Regular words			
aft		1.000	−.269	−.671	1.000	.265	ape		1.000	.800	.929	1.000	.932
aisle		1.000	−.750	−.953	1.000	.074	hutch	.981	1.000	1.000	−.412	1.000	.714
asked		1.000	−.391	−.671	1.000	.234	oust		1.000	.714	.257	1.000	.743
aunt		−.750	−.200	−.836	1.000	−.196	arch		1.000	1.000	.688	.805	.873
chef	−.924	1.000	−.636	.890	1.000	.266	shed	1.000	1.000	.800	.967	1.000	.953
chord	−.949	.667	1.000	.672	1.000	.478	crept	1.000	1.000	.600	.967	1.000	.913
chrome	1.000	.500	1.000	.702	1.000	.840	crunch	1.000	1.000	1.000	.889	.944	.967
chute	−.924	.428	1.000	−.662	1.000	.169	shrub	1.000	1.000	1.000	.889	1.000	.978
earls		1.000	.600	−.324	1.000	.569	hitch	.981	1.000	.882	.848	1.000	.942
earned		1.000	.600	−.324	1.000	.569	hoist	.981	1.000	1.000	1.000	1.000	.996
gaol	−.685	1.000	1.000	.200	1.000	.503	jest	1.000	1.000	1.000	.967	1.000	.993
heir		1.000	1.000	1.000		1.000	haze	.981	1.000	.556	.929	1.000	.893
isle		1.000	1.000	1.000	1.000	1.000	itch		1.000	.818	.848	1.000	.917
ones		−.857	−.556	−.984	1.000	−.349	oath		1.000	1.000	.941	.814	.939
owned		−.750	.000	−.242	1.000	.002	ounce		1.000	.714	.257	1.000	.743
thai	−.921	1.000	1.000	−.953		.031	tame	1.000	1.000	1.000	.929	1.000	.986
thyme	−.921	1.000	1.000	1.000	1.000	.616	tempt	1.000	1.000	1.000	.967	1.000	.993
tsar	1.000	.857	1.000	.713		.893	zoom	1.000	1.000	1.000	.648	1.000	.930
wholes	−.833	1.000	−.200	.702	1.000	.334	hoarse	.981	1.000	1.000	1.000	−.353	.726
whore	−.833	1.000	1.000	1.000		.542	hatch	.981	.667	.024	.396	1.000	.614
bears	1.000	−.412	.000	−.706	.713	.119	belch	1.000	.000	.926	.967	.429	.664
bind	1.000	−.053	−.771	−.888	1.000	.058	bait	1.000	.667	1.000	.886	.990	.908
books	1.000	.778	−.739	−.679	1.000	.272	beech	1.000	1.000	1.000	.983	.869	.970
bowled	1.000	−.714	.429	−.242	1.000	.294	breech	1.000	1.000	1.000	.983	.805	.958
bull	1.000	−.500	−.774	−.932	1.000	−.041	bang	1.000	1.000	−.129	.396	1.000	.653
bush	1.000	−.625	−.774	−.932	1.000	−.066	bark	1.000	1.000	1.000	.688	1.000	.938
butch	1.000	−.600	−.774	−.932	1.000	−.061	batch	1.000	.667	.080	.396	1.000	.629
chalk	.873	1.000	−.727	−.879	−.111	.031	cheat	.885	.667	1.000	.652	.990	.839
combs	.866	−.600	−.650	−.665	1.000	−.010	carve	.866	1.000	1.000	.688	1.000	.911
cooks	.866	.778	−.571	−.679	1.000	.279	crane	1.000	1.000	1.000	.929	1.000	.986
cough	.866	−.692	−.667	−.933	−.391	−.363	craps	1.000	.909	.846	.396	.983	.827
hearts	1.000	1.000	.000	−.882	1.000	.424	hound	.981	.800	1.000	.257	1.000	.808
hind	1.000	−.053	−.630	−.888	1.000	.086	hath	.981	−.429	.268	.396	.814	.406
hood	1.000	−.273	−.286	−.679	1.000	.152	hemp	.981	1.000	.769	.967	1.000	.944
hook	1.000	.667	−.286	−.679	1.000	.340	hive	.981	.714	1.000	.945	1.000	.928
knows	1.000	.259	1.000	−.242	.713	.546	knead	1.000	−.250	1.000	.652	1.000	.680
laughs	1.000	1.000	−.143	−.835	−.222	.160	loathe	1.000	1.000	1.000	.941	1.000	.988
looks	1.000	.778	−.647	−.679	1.000	.290	leach	1.000	1.000	1.000	.941	1.000	.988
loved	1.000	−.200	−.600	−.882	1.000	.064	lance	1.000	1.000	.609	.396	1.000	.801
mild	1.000	.500	−.778	−.888	.742	.115	mare	1.000	1.000	1.000	.959		.990
minds	1.000	.800	−.778	−.882	1.000	.227	mirth	1.000	1.000	1.000	.971	.814	.957
monk	1.000	−.500	−.733	−.936	1.000	−.033	mist	1.000	.818	.765	.848	1.000	.886
mousse	1.000	1.000	−.867	−.665	1.000	.294	morgue	1.000	1.000	1.000	.692	1.000	.939
mown	1.000	.000	1.000	−.242	1.000	.552	mash	1.000	.826	.366	.396	1.000	.718
ninth	1.000	.000	−.474	−.888	1.000	.128	nerve	1.000	1.000	1.000	.823	1.000	.965
pear	1.000	−.444	−.200	−.706		−.088	pave	1.000	.778	1.000	.938	1.000	.943
pearl	1.000	1.000	.200	−.324	1.000	.575	pants	1.000	−.111	.211	.396	1.000	.499
pint	1.000	−.846	−.805	−.888	1.000	−.108	peck	1.000	1.000	1.000	.967	1.000	.993
quay	1.000	1.000	1.000	1.000		1.000	kelp	1.000	1.000	1.000	.967	1.000	.993
rouge	1.000	.000	−.200	−.665	−.579	−.088	rinse	1.000	1.000	.889	.848	.714	.890
routes	1.000	1.000	−.200	−.665	1.000	.427	roach	1.000	1.000	.818	.832	.805	.891
sew	1.000	−.857	1.000	−.793		.087	sip	1.000	1.000	.622	.848	.984	.891
shoe	1.000	−.778	1.000	−.667		.139	sail	1.000	.882	.429	.886	1.000	.839
shove	1.000	−.600	−.333	−.882	1.000	.037	shrug	1.000	1.000	1.000	.889	1.000	.978
shows	1.000	.259	1.000	−.242	.713	.546	shave	1.000	.778	1.000	.929	1.000	.941
soup	1.000	.667	−.333	−.665	1.000	.334	sank	1.000	1.000	.846	.396	1.000	.848
wolf	1.000	.000	−.467	−.981	.800	.070	weed	1.000	1.000	1.000	.983	1.000	.997
womb	1.000	−.200	−.733	−.963	1.000	.021	wick	1.000	1.000	.750	.848	1.000	.920
yacht	1.000	1.000	−.857	−.918	1.000	.245	yeast	1.000	.667	−.333	.652	1.000	.597
blinds	1.000	.800	−.600	−.888	1.000	.262	blotch	1.000	1.000	1.000	.549	1.000	.910
blown	1.000	.000	1.000	−.242	1.000	.552	blaze	1.000	1.000	1.000	.929	1.000	.986
breaks	1.000	−.667	−.750	−.946	1.000	−.072	bilge	1.000	1.000	.852	.848	1.000	.940
brooch	1.000	−.600	−.750	−.992	.756	−.117	blanch	1.000	1.000	1.000	.396	.944	.868
brook	1.000	.667	−.250	−.679	1.000	.348	bland	1.000	.846	1.000	.396	1.000	.848
climb	1.000	.000	−.700	−.888	1.000	.082	cliff	1.000	1.000	.250	.848	1.000	.820
crepe	1.000	1.000	.333	−.825	1.000	.502	cloak	1.000	1.000	1.000	.941	1.000	.988
crook	1.000	.667	.000	−.679	1.000	.398	cleat	1.000	.667	.600	.652	.990	.782

Appendix C (*continued*)

Word	Head	Body	Antibody	Nucleus	Coda	Total	Word	Head	Body	Antibody	Nucleus	Coda	Total
		Irregular words (*continued*)							Regular words (*continued*)				
crow	1.000	.133	−.429	−.242		.116	cage	.866	.778	1.000	.929	.579	.830
flown	1.000	.000	1.000	−.242	1.000	.552	farce	1.000	.000	1.000	.688	1.000	.738
fronts	1.000	.000	−.818	−.938	1.000	.049	fierce	1.000	1.000	1.000	1.000	1.000	1.000
glove	1.000	−.600	−.200	−.882	1.000	.064	glide	1.000	1.000	1.000	.945	1.000	.989
glow	1.000	.133	1.000	−.242		.473	gaze	.968	1.000	1.000	.929	1.000	.980
grasp	1.000	.667	.000	−.671	1.000	.399	gland	1.000	.846	1.000	.396	1.000	.848
grind	1.000	−.053	−.800	−.888	1.000	.052	girth	.968	1.000	1.000	.971	.814	.951
groups	1.000	.500	−.333	−.665	1.000	.300	gauze	.968	1.000	.714	.765	1.000	.890
hose	1.000	.333	1.000	.702	−.324	.542	heap	.981	1.000	.625	.652	.984	.848
plaid	1.000	−.750	−.111	−.953	1.000	.037	peach	1.000	1.000	1.000	.652	.805	.891
proved	1.000	−.200	−.400	−.897	1.000	.101	pierce	1.000	1.000	1.000	1.000	1.000	1.000
scarce	.844	.000	−.778	−.991	1.000	.015	snatch	1.000	.667	1.000	.396	1.000	.813
sponge	.954	1.000	−.333	−.938	.905	.318	sneeze	1.000	1.000	1.000	.983	1.000	.997
steak	1.000	−.667	−.556	−.946	1.000	−.034	slate	1.000	1.000	1.000	.929	1.000	.986
swab	.979	−.875	.172	−.918	1.000	.072	sack	1.000	1.000	.846	.396	1.000	.848
swamps	.979	−.857	.172	−.918	1.000	.075	scribe	1.000	1.000	1.000	.945	1.000	.989
swap	.979	−.913	.172	−.918	1.000	.064	scab	.844	.875	.790	.396	1.000	.781
throws	1.000	.259	1.000	−.242	.713	.546	thatch	.421	.667	.429	.396	1.000	.582
tread	1.000	.250	.000	−.721	1.000	.306	tenth	1.000	1.000	1.000	.967	1.000	.993
troupe	1.000	1.000	−.500	−.665	1.000	.367	tights	1.000	1.000	1.000	1.000	1.000	1.000
truths	1.000	1.000	−.667	−.932	−.357	.009	trance	1.000	1.000	1.000	.396	1.000	.879

Appendix D
Human and Simulation Data: Experiment 1

Exc word	RT (ms)	RT (cycles)	Reg word	RT (ms)	RT (cycles)	Exc word	RT (ms)	RT (cycles)	Reg word	RT (ms)	RT (cycles)
aft	531	108	ape	485	73	ninth	524	92	nerve	476	76
aisle	560	100	hutch	453	83	pear	507	88	pave	511	83
asked	478	93	oust	569	81	pearl	500	93	pants	485	74
aunt	547	93	arch	501	77	pint	545	90	peck	470	78
chef	555	100	shed	474	77	quay	623	95	kelp	490	78
chord	569	97	crept	514	77	rouge	572	92	rinse	484	77
chrome	504	88	crunch	522	77	routes	533	85	roach	503	78
chute	551	101	shrub	476	78	sew	479	92	sip	498	71
earls	521	98	hitch	462	80	shoe	479	84	sail	448	76
earned	499	94	hoist	494	80	shove	465	86	shrug	466	79
gaol	621	112	jest	470	77	shows	466	77	shave	477	82
heir	578	115	haze	500	78	soup	481	84	sank	475	77
isle	533	—	itch	518	82	wolf	500	91	weed	473	77
ones	520	91	oath	464	78	womb	512	90	wick	491	77
owned	564	97	ounce	545	78	yacht	532	94	yeast	514	76
thai	605	86	tame	493	79	blinds	514	81	blotch	506	82
thyme	644	77	tempt	504	78	blown	487	77	blaze	482	79
tsar	648	—	zoom	487	80	breaks	517	77	bilge	565	78
wholes	563	—	hoarse	509	76	brooch	589	81	blanch	520	81
whore	520	108	hatch	448	79	brook	508	80	bland	507	76
bears	523	84	belch	530	78	climb	487	82	cliff	502	74
bind	548	95	bait	499	79	crepe	559	82	cloak	534	76
books	508	79	beech	513	76	crook	504	79	cleat	569	78
bowled	554	91	breech	535	77	crow	518	81	cage	515	78
bull	512	87	bang	471	82	flown	497	79	farce	520	78
bush	527	86	bark	491	77	fronts	550	80	fierce	556	76
butch	561	102	batch	497	79	glove	521	80	glide	505	82
chalk	517	82	cheat	534	80	glow	521	78	gaze	506	77
combs	509	96	carve	492	80	grasp	495	79	gland	517	77
cooks	474	88	crane	483	80	grind	522	81	girth	515	80
cough	519	75	craps	498	80	groups	538	73	gauze	584	81
hearts	504	84	hound	504	79	hose	486	85	heap	466	77
hind	501	91	hath	538	76	plaid	533	76	peach	516	79
hood	473	88	hemp	474	77	proved	512	78	pierce	542	82
hook	441	83	hive	497	80	scarce	527	78	snatch	489	77

(*Appendixes continue*)

Appendix D (*continued*)

Exc word	RT (ms)	RT (cycles)	Reg word	RT (ms)	RT (cycles)	Exc word	RT (ms)	RT (cycles)	Reg word	RT (ms)	RT (cycles)
knows	497	77	knead	543	77	sponge	458	80	sneeze	445	79
laughs	476	86	loathe	496	82	steak	499	79	slate	464	79
looks	464	80	leach	508	80	swab	474	84	sack	450	74
loved	465	86	lance	483	78	swamps	477	81	scribe	463	79
mild	483	87	mare	521	77	swap	468	83	scab	492	78
minds	469	85	mirth	515	79	throws	527	78	thatch	574	80
monk	496	94	mist	483	75	tread	501	77	tenth	539	75
mousse	616	91	morgue	525	80	troupe	525	82	tights	496	81
mown	550	90	mash	490	77	truths	539	79	trance	503	76

Note. Exc = exception; Reg = regular; RT = reaction time. Dashes signify errors made by the dual-route cascaded model.

Appendix E
Strategic Effects in Reading Aloud: Human and Dual-Route Cascaded (DRC) Data

Target	Human data: Position 1 fillers	Human data: Position 3 fillers	DRC data: Standard	DRC data: Strategic	Target	Human data: Position 1 fillers	Human data: Position 3 fillers	DRC data: Standard	DRC data: Strategic
Nonword targets					Regular-word targets				
frant	556	569	129	137	jest	490	460	77	77
woize	606	544	156	178	crunch	469	459	77	77
sleam	520	440	153	173	belch	526	500	78	79
stoab	498	495	156	178	bait	507	499	79	80
dreeb	524	516	156	178	breech	517	479	77	77
greel	555	534	143	—	batch	501	482	79	78
kneam	699	597	156	178	carve	531	452	80	80
droze	540	507	179	216	crane	491	462	80	79
scaid	548	481	156	177	craps	547	497	80	80
clite	640	585	202	270	hound	536	521	79	79
murse	533	490	151	171	hive	508	487	80	85
cheen	553	523	151	170	knead	506	501	77	77
knafe	685	554	—	—	leech	482	495	79	79
wheeb	532	553	151	166	lance	506	492	78	78
gurch	574	528	168	196	mash	468	449	77	78
deest	566	589	156	178	nerve	463	466	76	76
kirth	576	525	159	185	pave	483	463	83	89
reeze	499	473	156	178	pants	478	455	74	74
heece	590	544	164	190	peck	484	471	78	78
poote	534	538	156	178	kelp	529	488	78	78
spant	484	444	145	152	rinse	507	491	77	77
treen	533	499	140	157	roach	463	496	78	78
taize	547	550	155	175	sip	480	487	71	76
skart	516	497	146	165	weed	477	480	77	77
speen	463	449	143	161	wick	512	493	77	77
thurn	611	588	155	175	yeast	470	500	76	76
brank	536	544	162	158	blotch	516	470	82	81
sanse	526	516	151	166	blaze	459	466	79	77
plews	544	570	154	175	cliff	516	479	74	74
merch	492	498	156	177	cloak	494	454	76	76
soize	511	501	156	177	cleat	554	565	78	78
pleap	546	560	152	170	cage	481	496	78	82
snerk	462	460	156	177	glide	496	507	82	80
borch	546	556	163	189	gaze	474	475	77	81
jelp	564	517	130	142	heap	462	467	77	76
nent	501	496	119	131	peach	463	476	79	78
whert	561	557	140	—	sneeze	444	423	79	79
flar	579	521	146	—	slate	476	446	79	77
sperk	495	446	151	173	sack	473	438	74	74
crean	536	518	148	169	scribe	450	429	79	79
frim	617	552	—	—	scab	476	432	78	78
shest	613	553	149	168	tenth	494	457	75	75
joor	527	544	141	162	ape	503	516	73	79
phize	629	612	182	231	shed	443	422	77	78
trufe	541	521	194	239	shrub	414	448	78	79

Appendix E (*continued*)

	Human data		DRC data			Human data		DRC data	
Target	Position 1 fillers	Position 3 fillers	Standard	Strategic	Target	Position 1 fillers	Position 3 fillers	Standard	Strategic
tolph	598	637	169	198	hoist	504	503	80	80
pait	520	496	129	142	itch	462	489	82	85
zupe	495	504	177	220	oath	473	482	78	79
snocks	466	450	163	185	zoom	451	459	80	82
goph	569	601	153	175	hatch	479	469	79	77

Note. Dashes signify errors made by the DRC model.

Received December 23, 1996
Revision received January 6, 1998
Accepted March 23, 1998 ■