

Regularity effects in word naming: What are they?

MICHAEL J. CORTESE

Morehead State University, Morehead, Kentucky

and

GREG B. SIMPSON

University of Kansas, Lawrence, Kansas

In a word-naming experiment, word-body consistency was crossed with grapheme-to-phoneme regularity to test predictions of current models of word recognition. In the latency and error data, a clear effect of consistency was observed, with the influence of regularity somewhat weaker. In addition, simulation data from three contemporary models of word recognition were obtained for the stimuli used in the experiment in order to compare the models' latencies with those of humans. The simulations showed that the human latency data are most consistent with the parallel-distributed-processing model of Plaut, McClelland, Seidenberg, and Patterson (1996), less so with the dual-process model (Zorzi, Houghton, & Butterworth, 1998), and least so with the dual-route-cascaded model (Coltheart & Rastle, 1994).

In experiments of reading aloud, irregular words (e.g., *pint*) take longer to name than regular words (e.g., *punt*). Usually, a word is considered to be irregular if it violates grapheme-to-phoneme correspondence (GPC) rules, which typically correspond to the most frequent pronunciations of graphemes. For example, the rule for *i* is /ɪ/, because this is the dominant pronunciation for *i*. Thus, according to this definition of regularity, *pint* is irregular because the *i* rule is violated. However, *pint* is also inconsistent. That is, there are similarly spelled words (e.g., *mint*, *lint*) with conflicting pronunciations of the word body (the vowel and following consonants). Many irregular words are also inconsistent, which has often led to a confounding of these variables, but they are separable dimensions. As a measure of consistency, we can examine the distribution of pronunciations associated with a particular word body. Words high on this measure of consistency are those that have many more friends (words that share the same body and a common pronunciation) than enemies (words containing the same body associated with a different pronunciation). For example, *storm* has *worm* as an enemy, but many more friends (*form*, *norm*, *dorm*, etc.). A word lower on this consistency dimension would be one with more enemies than friends, such as *pint*, which has the enemies *mint*, *hint*, *lint*, *tint*, and no friends.

Several studies have demonstrated an effect of consistency that is independent of GPC regularity (Glushko, 1979; Jared, McRae, & Seidenberg, 1990). Words that are regular as defined by GPC rules, but have many enemies

(e.g., *mood*) yield longer naming latencies than do regular consistent words (e.g., *moon*). In contrast, there is very little evidence that GPC regularity effects are independent of word-body consistency. However, in a recent study, Jared (1997, Experiment 4) examined consistency/regularity effects for high-frequency words, using regular- and irregular-inconsistent words that varied according to their friends-to-enemies ratio. This led to conditions in which regular and irregular words either had more friends than enemies ($F > E$), or more enemies than friends ($E > F$). Differences in naming latencies between experimental words and a set of regular control words were observed for $E > F$, but not for $F > E$ words.

Because current word-recognition models differ sharply in their explanations of regularity effects, the present study was designed for the purpose of further studying the relationship between regularity and word-body consistency. Participants named low-frequency words that varied in regularity and the measure of word-body consistency described above. Specifically, four types of words were presented: regular $F > E$ (e.g., *toad*), regular $E > F$ (e.g., *tone*), irregular $F > E$ (e.g., *toll*), and irregular $E > F$ (e.g., *tomb*). In addition, simulation data were gathered for the purpose of comparing the naming times for these words and the output of three implemented word-recognition models that make predictions about regularity effects. We turn now to brief descriptions of these models and their predictions regarding the roles of consistency and regularity in word recognition.

The dual-route cascade (DRC) model (Coltheart, Curtis, Atkins, & Haller, 1993) represents an implemented version of the dual-route theory of reading, containing separate lexical and sublexical routes. The lexical route is an interactive activation procedure (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982) that generates a pronunciation by activating a phonological code in a mental lexicon via a visual word representation. A sep-

Special thanks go to Max Coltheart and Kathy Rastle, David Plaut, and Marco Zorzi, who provided simulation data from their respective models. We also thank Maura Pilotti, who provided useful comments on an earlier draft of this manuscript. Correspondence should be addressed to M. J. Cortese, Department of Psychology, Morehead State University, 601 Ginger Hall, Morehead, KY 40351 (e-mail: m.cortese@morehead-st.edu).

arate sublexical procedure operates on the basis of GPC rules. The sublexical route generates pronunciations of orthographic inputs by applying categorical GPC rules that are based on the most frequent pronunciation of each grapheme in a word. The sublexical route processes orthographic stimuli one letter at a time from left to right (Rastle & Coltheart, 1998). During word recognition, these two processes work separately and simultaneously, and feed information to a common phoneme system. Irregular words (e.g., *pint*) take longer to name because the phoneme system receives conflicting information from the two processing routes. The incorrect code produced by the sublexical route competes with the correct code retrieved from the lexicon. For example, when the rules are applied to *pint*, the unit representing /ɪ/ at the second position in the phoneme system becomes active, and it competes with /ai/ activated by the lexical route. When two or more phoneme units are active, their activation serves to inhibit each other, and this increases reaction time (Coltheart & Rastle, 1994).

It is important to note that GPC rules are categorical (Seidenberg & McClelland, 1989). In other words, a grapheme corresponds to a single phoneme and no others. For example, although *-ood* is associated strongly with the pronunciation occurring in *good*, the sublexical route activates only /u/ when *-ood* words are processed. Therefore, potential interference delivered to the phoneme system from the sublexical route is, in general, context independent. The model contains context-sensitive rules, but these rules are applied rather narrowly.

Effects of word-body consistency, on the other hand, have been theorized to arise within the lexical route (Coltheart et al., 1993). Although localized pronunciation codes for words exist within this route, information at various levels (e.g., feature, letter, visual-word, phonological-word) is activated in cascade. That is, as soon as information becomes active at one level, it is shared with adjacent levels. Word-body units (e.g., *ood*) activate words that contain them (e.g., *mood*, *good*, *wood*, etc.). The activation of words with contrasting pronunciations results in competition in the phoneme system and leads to slower naming latencies. For example, *mood* activates *good*, *wood*, and so forth, and this generates competition at the second phoneme position (owing to the alternative pronunciations of *oo*). It is important to note that this within-route conflict is separate from the between-route competition responsible for regularity effects. That is, consistency effects are due solely to information activated within the lexical route, and regularity effects are due to information activated through both routes.

Therefore, in the present study, according to the DRC model, naming latency should be related to the number of sources in which competition occurs. Specifically, the model predicts that latencies should be longer for irregular words than for regular words because of competition occurring between processing routes, and latencies should be longer for inconsistent words than for consistent words because of competition within the lexical route.

In contrast to the DRC model, in which regularity and word-body consistency effects have different origins within the lexical and sublexical routes, the model offered by Plaut, McClelland, Seidenberg, and Patterson (1996) proposes that these effects share a common basis. This model (hereafter referred to as PMSP96) consists of a single network representing an orthographic input level, a phonological output level, and a semantic level representing the frequency of the orthographic input. In the implemented model, hidden units mediate connections between the levels in such a way that each orthographic unit connects to each hidden unit, which, in turn, is connected to each phonological unit. Each connection is associated with a weighted value that is adjusted during a learning phase, and knowledge of spelling-to-sound relationships is contained in these weighted connections. During learning, words that have consistent pronunciations (e.g., *punt*) affect weight adjustment similarly for each word in the neighborhood (e.g., *hunt*, *bunt*, *runt*, etc.), whereas weight adjustment for words containing inconsistent letter strings (e.g., *bough*, *cough*, *dough*, etc.) will vary, depending on the target output. For example, when *bough* is presented to the model, connection weights are adjusted so that *ough* yields /au/, whereas when *dough* is presented, the weights are adjusted so that *ough* yields /o/. After learning, the model yields longer settling times for inconsistent words regardless of GPC regularity, because the connection weights have been adjusted differently for each particular orthographic-phonological relationship. Because irregular words (e.g., *pint*) typically contain spelling patterns that are strongly associated with alternative pronunciations, they are likely to show longer settling times (which represents recognition speed in the model). Consequently, the PMSP96 approach does not distinguish between regular and irregular words on the basis of GPC rules, but rather in terms of degree of consistency; words are more or less regular, depending on the consistency of pronunciation of similarly spelled words.

In the present study, therefore, the PMSP96 model should predict a robust effect of consistency. Regularity, on the other hand, should have no effect that is independent of the consistency measure. That is, irregular words should be named slowly only if they have a high proportion of enemies.

The DP model (Zorzi, Houghton, & Butterworth, 1998) combines properties of both the DRC and PMSP96 models. Like the DRC model, the DP model contains separate lexical and sublexical routines that interact in a phonological decision system. As in the PMSP96 model, representations are distributed across units, pronunciation codes are learned via standard algorithms, and the DP model contains a hidden layer that mediates relationships between orthographic inputs and phonological outputs. In contrast to the PMSP96 model, however, the DP model has a separate two-layer phonological assembly route that maps orthographic inputs directly onto phonological units (i.e., the sublexical procedure). Because two-layer networks cannot learn the arbitrary mappings that

occur in irregular words (Hinton & Shallice, 1991), the pronunciation of irregular words is learned via the hidden units (i.e., the lexical procedure). The lexical route is represented as a three-layer network with hidden units mediating orthographic-to-phonological relationships. It is assumed that the two-layer side of the model is faster than the three-layer side because one less step is involved.

Regarding consistency and regularity effects, the DP model should behave similarly to the PMSP96 model. The three-layer network should operate like the PMSP96 model because they contain similar representations, and they are trained similarly. Thus, the model should show an effect of consistency. However, the two-layer network should be more sensitive to grapheme-to-phoneme mappings than orthographic-to-phonological body mappings and show a strong regularity effect. Therefore, the model should produce main effects of both variables. However, on the basis of simulations of Jared's (1997) high-frequency words (Zorzi, personal communication, July 1999), we expect the model to produce a larger effect of regularity than consistency.

In summary, regularity and consistency effects have different bases in the DRC and DP models, and a common basis within the PMSP96 model. In the DRC model, regularity effects arise from competition between the lexical and sublexical routes while consistency effects reflect a conflict within the lexical route. In the PMSP96 model, consistency effects and regularity effects derive from the same source, but the model is more sensitive to consistency than to regularity. Finally, the DP model should show sensitivity to consistency because it contains a three-layer network as does the PMSP96 model, and it should show sensitivity to regularity because it contains a two-layer network that is sensitive to grapheme-to-phoneme mappings.

METHOD

Subjects

Forty-nine undergraduates from the University of Kansas participated in the experiment for course credit. All were native speakers of English who reported to have normal or corrected-to-normal vision.

Design

The design for the experiment was a 2 × 2 within-subjects factorial. The factors were consistency (*F* > *E* vs. *E* > *F*) and GPC regularity (regular vs. irregular) of the targets.

Stimuli

The stimuli and their mean naming latencies, grapheme-to-phoneme (Berndt, Reggia, & Mitchum, 1987) and orthographic-to-phonological body (Ziegler, Stone, & Jacobs, 1997) conditional probabilities are presented in the Appendix, and summary statistics for these items are provided in Table 1. All of the stimuli used were monosyllabic words with printed frequency of less than 57 per million words (Carroll, Davies, & Richman, 1971). Of the 40 targets selected, 20 were regular in terms of GPC rules (Coltheart et al., 1993) and 20 were irregular. In addition, both the regular and irregular items were further classified into *F* > *E* and *E* > *F* groups. We defined word consistency in terms of the summed frequencies of each of the pronunciations of a particular word body.¹ A consis-

Table 1
Summary Statistics for the Items Used in the Experiment
(Mean Values are Provided)

Variable	Regular		Irregular	
	<i>F</i> > <i>E</i>	<i>E</i> > <i>F</i>	<i>F</i> > <i>E</i>	<i>E</i> > <i>F</i>
Frequency	16.8	10.8	13.5	11.7
Familiarity	6.25	6.41	6.41	6.44
Imageability	5.07	4.95	4.99	5.10
Length	4.5	4.5	4.7	4.9
Bigram-frequency	4,997	5,312	5,730	5,767
Coltheart's <i>N</i>	7.1	7.7	6.4	5.2
Mean summed frequency-friends	449.1	184.0	745.8	140.7
Mean summed frequency-enemies	125.5	873.3	175.5	569.0
Conditional probability-body	.84	.41	.70	.27
Conditional probability-grapheme	.57	.62	.12	.17

tent word was one for which the summed frequencies of its friends (i.e., similarly pronounced words) were greater than those of its enemies (i.e., differently pronounced words), and was labeled *F* > *E*. For inconsistent words, the summed frequencies of enemies was greater than those of its friends (labeled *E* > *F*). Varying consistency and GPC regularity (Coltheart et al., 1993) allowed us to define four stimulus categories: (1) regular *F* > *E* (e.g., *toad*), (2) regular *E* > *F* (e.g., *tone*), (3) irregular *F* > *E* (e.g., *toll*), and (4) irregular *E* > *F* (e.g., *tomb*). The orthographic-to-phonological conditional probabilities (i.e., the probability of a particular pronunciation, given a particular orthographic sequence) showed that the range and variation of consistency were approximately equal for these two units.² For example, the mean conditional probabilities of the *F* > *E* and *E* > *F* words were .770 and .342, respectively (mean range = .432, standard deviation = .305), and the mean conditional probabilities of the regular and irregular words were .595 and .147, respectively (mean range = .448, standard deviation = .285). Also, because of the recent controversy involving the criterion to initiate pronunciation (see Cortese, 1998; Cortese & Zevin, 1998; Kawamoto, Kello, Jones, & Bame, 1998; Rastle, Harrington, Coltheart, & Palethorpe, 2000), equal numbers of plosive and nonplosive initial phonemes occurred in each group. Other aspects of initial phoneme were controlled by analysis of covariance. Furthermore, no two words contained the same spelling pattern.

A series of regularity × consistency analyses of variance were performed to compare the items on frequency, familiarity (rated by a group of 30 subjects), word length, imageability (Cortese, Simpson, & Woolsey, 1997; Strain, Patterson, & Seidenberg, 1995; rated by another group of 30 subjects), neighborhood size (Coltheart, Davelaar, Jonasson, & Besner, 1977), and summed bigram frequency (Solso & Juel, 1980). None of these analyses yielded any significant main effects or interactions (all *ps* > .23).

Procedure

The Micro Experimental Laboratory (MEL) software (Schneider, 1990) was used to present the stimuli via a microcomputer. On each trial, a fixation mark (+) was presented for 1,000 msec. The word stimulus then appeared until a naming response was made. An experimenter seated next to the subject coded responses as correct, error, or noise (i.e., the microphone failed to record the response or if it recorded some extraneous noise). An intertrial interval of 1,000 msec was employed.

RESULTS

The responses coded as noise were removed from the analyses. Also, responses beyond 2.5 standard deviations of the mean for each condition were classified as outliers

and were removed from the analyses. Noise responses and outliers accounted for 4.4% of the data. The latency data and error rates for each condition are presented in Figure 1.

Analysis of variance by subjects and items (Clark, 1973) was performed on latencies and errors. Consistency and regularity were both within-subjects factors in the analyses by subjects (F_1), and between-words factors in the analyses by items (F_2). Voice, fricative, palatal, bilabial, and liquid were included as covariate factors in the analysis by items. Unless noted, all effects are significant at $p < .05$.

In the latency data, consistent ($F > E$) words were named in less time than were inconsistent ($E > F$) words [$F_1(1,48) = 82.80, MS_e = 445.41; F_2(1,31) = 10.91, MS_e = 1,025.42$]. Regularity was significant by subjects [$F_1(1,48) = 7.39, MS_e = 307.13$], but not by items [$F_2(1,31) = 1.42, MS_e = 1,025.42, p > .24$]. The interaction between consistency and regularity was also significant by subjects [$F_1(1,48) = 4.16, MS_e = 560.48$], but not by items ($F_2 < 1$). Simple effects show that consistency was significant for both regular words [$t_1(48) = 7.26, SE = 4.73$] and irregular words [$t_1(48) = 4.75, SE = 4.32$].

In the error data, more errors were made for inconsistent ($E > F$) than for consistent ($F > E$) words [$F_1(1,48) = 53.64, MS_e = .01; F_2(1,31) = 13.34, MS_e = .01$]. Regularity was significant by subjects [$F_1(1,48) = 22.63, MS_e = .01$], but not by items [$F_2(1,31) = 2.16, MS_e = .01, p > .13$]. The consistency \times regularity interaction was not significant (both $F_s < 1$). The vast majority of the errors produced could be seen as consistency driven (e.g., *bead* as rhyming with *head*, and *pint* rhyming with *mint*). One exception to this was that *suite* was com-

monly mispronounced as *suit*. There was no sign of a speed-accuracy tradeoff in the data.

The results of this experiment suggest that naming latencies and errors are more affected by word-body consistency than by GPC regularity. A clear effect of consistency was observed (35 msec), whereas the influence of GPC regularity was somewhat weaker (16 msec).

Simulations

We obtained measures of naming latencies from each of the models for the stimuli in the experiment in order to compare the performance of each model to the human data (see Figure 2).³ The DRC model showed an effect of regularity that was similar to the results observed with Jared's (1997) stimuli [$F(1,36) = 7.01, MS_e = 27.62$], but the effect of consistency was in the opposite direction to that observed in the experiment and was not significant [$F(1,36) = 2.64, MS_e = 27.62, p > .11$]. As predicted, the PMSP96 model showed a large effect of consistency [$F(1,36) = 10.90, MS_e = .01$], whereas regularity was not significant ($F < 1$). Also as predicted, the DP model produced a large effect of regularity [$F(1,35) = 27.53, MS_e = .62$] and a small effect of consistency [$F(1,35) = 4.30, MS_e = .62$]. The consistency \times regularity interaction was not significant for any of the models (all $F_s < 1$).

In addition, in order to test the models at the item level (e.g., Spieler & Balota, 1997), we calculated the correlation between the output of each of the models and the item means obtained from our subjects. The results of these analyses are consistent with the results obtained at the factor level. Only the latencies obtained from the

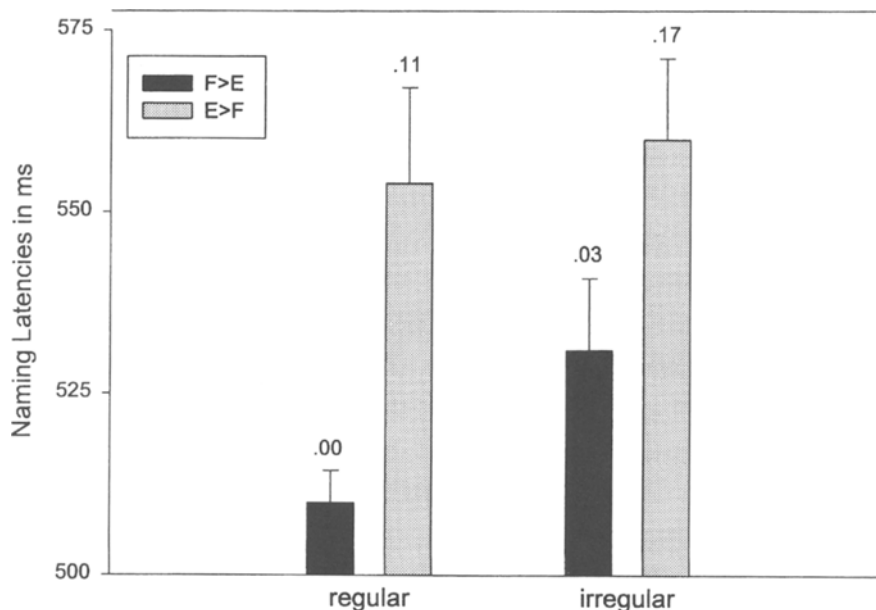


Figure 1. Mean latencies and proportion errors (above each bar) as a function of regularity and consistency.

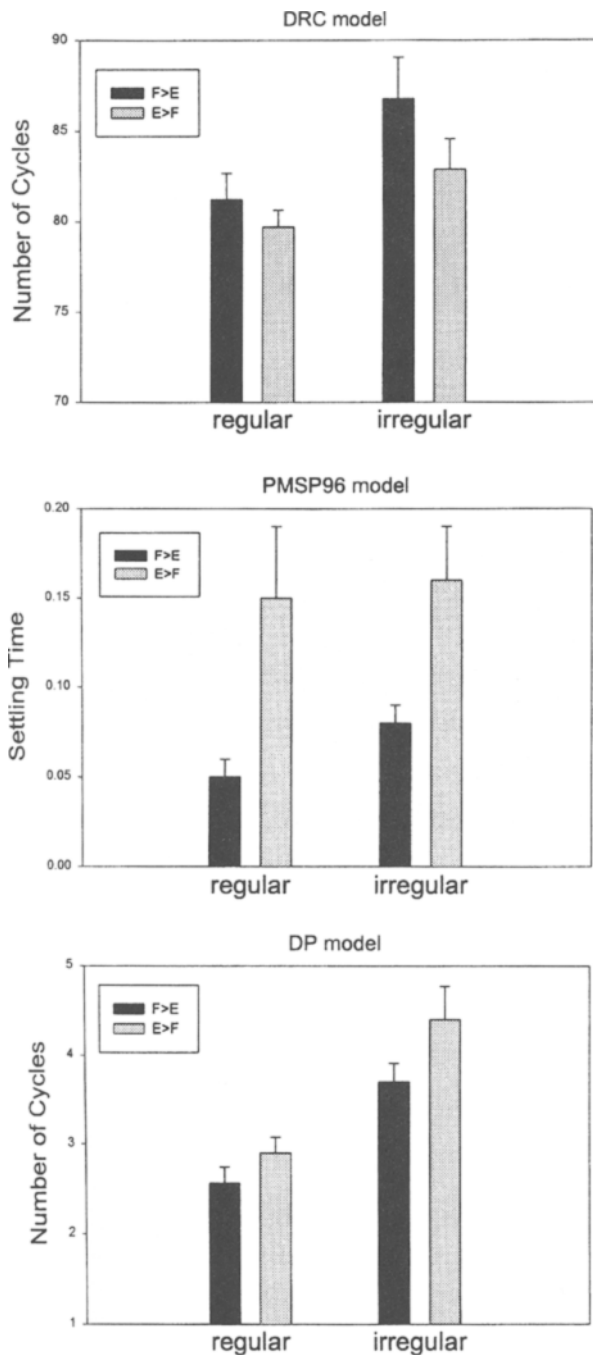


Figure 2. Simulation data from the DRC, PMSP96, and DP models.

PMSP96 model significantly correlated with the human latencies [$r(39) = .44$, DP model; $r(38) = .28$, $p < .09$, DRC model; $r(39) = -.06$, $p > .10$].

Overall, the pattern of reaction time data generated by our subjects resembled most closely that of the PMSP96 model, whereas the output of the DP model overestimated regularity effects and underestimates consistency effects,

and that of the DRC model failed to show the obtained pattern.

DISCUSSION

The results of the experiment reported here support the predictions made by the PMSP96 model. In fact, consistency as measured by word-body friends and enemies had a clear effect on naming performance, whereas GPC regularity did not. The DP model performed similarly to the PMSP96 model, but it did not produce as close of a match to the human data. The DRC model was not supported by the human data.

On the basis of these results, we would argue that conceptions of orthographic/phonological regularity in word recognition need to be considered more in terms of word-body consistency than of GPC regularity. Inconsistent words take more time to respond to and lead to more errors than do consistent words, irrespective of GPC regularity. Much of the previous research manipulating spelling-to-sound regularity has typically failed to take into account the fact that irregular words usually have many enemies, whereas regular words tend to have many friends. In other words, the research has frequently confounded regularity with consistency.

Our finding is consistent with other studies in which it has been found that word-body consistency plays an important role in word recognition (e.g., Glushko, 1979; Jared, 1997; Jared et al., 1990; Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995). For example, using a large-scale regression analysis, Treiman et al. found that for word body-units in three-phoneme words, the consistency of the body unit accounted for more variance ($p < .05$) than did vowel consistency ($p > .27$). In addition, Jared (1997), crossing the same variables as in the present study but using high-frequency words, found that consistency was stronger than regularity. Finally, studies done primarily with regular words have shown word-body consistency effects that are independent of regularity (Glusko, 1979; Jared, 1997; Jared et al., 1990). On the other hand, not one study involving the reading of words has shown a stronger effect of regularity than of consistency. Thus, we would argue that in experiments using only real words where the system is not biased in an unnatural way, consistency effects should be and always are larger than regularity effects. It is important to note that we are not arguing that grapheme-phoneme regularity has no effect in word naming. Rather, conceptions of spelling-to-sound effects in reading aloud should be thought of more in terms of word-body consistency than grapheme-to-phoneme regularity.

It may be argued that the consistency effect reported in the present study is a within-word context effect. That is, information about the grapheme may be overridden by the contextual information provided by the word body. For example, the most common way to pronounce *i* is /i/ (as in *lid*), but the *ind* body is most frequently pronounced

as in *kind*. The present results suggest that the irregularity of *i* is overridden by the consistency of *ind*. This finding contrasts sharply with previous notions regarding the supreme status of the grapheme in visual word recognition (Coltheart et al., 1993).

Regarding the theoretical implications of the present study, the simulation results show that the PMSP96 model most closely simulates the observed outcome. Thus, the model appears to be more sensitive to word-body consistency than to grapheme–phoneme consistency. The PMSP96 approach views regularity as a continuous variable based on the statistical properties of spelling-to-sound mapping of English. The irregularity of a word corresponds to the phonological characteristics of similarly spelled words. If the pronunciation of similarly spelled words is consistent with the target word (e.g., *toad*, *bind*), the adjustment of the weights during learning will be applied in a consistent manner (so that *ind* yields /aɪnd/, and *oad* yields /od/). However, if the pronunciation of similarly spelled words is inconsistent with the target word (e.g., *bead* and *pint*), the weights will be adjusted differently for the different pronunciations (e.g., so that *int* maps onto /aɪnt/ when *pint* is presented and /ɪnt/ when *hint*, *mint*, *tint*, *print*, *lint*, etc. are presented). When the target pronunciation is the least common, performance will suffer. It is important to note that the model, in principle, could handle an effect of regularity as long as it was always smaller than that of consistency (assuming that all other factors are equated). Consider *bind* again. We might expect the model to be affected by other *i* words in the sense that during learning when the model is presented with the grapheme *i*, the weights are being adjusted so that *i* yields the GPC-consistent pronunciation. However, when presented with *ind* words, more connections within the network are being adjusted so that *ind* yields the /aɪnd/ pronunciation rather than the /ɪnd/ pronunciation. The hidden layer that allows the model to learn the pronunciation of inconsistent words is more sensitive to larger orthographic units than it is to smaller ones (Seidenberg & McClelland, 1989). Consistent with this idea, the PMSP96 model shows a large consistency effect, and a small, nonsignificant effect of regularity. Presumably, if the number of items increased, a significant regularity effect could be found in the model, but the consistency effect would be greater.

The DP model operates in a similar fashion, but the sublexical route tends to emphasize grapheme units such that words containing inconsistent graphemes (e.g., *bind* and *pint*) will create equivalent interference in the phoneme decision system regardless of the body consistency. The lexical route will be sensitive to body consistency because it operates very similarly to the PMSP96 model. Thus, the model simulates both a regularity effect and a consistency effect. However, given the simulation data, the model in its current form tends to overemphasize the effect of regularity and underestimate the effect of consistency.

The DRC model produced a large effect of regularity and no effect of consistency; in fact, the phonological code took slightly more time for $F > E$ words than for $E > F$ words. Thus, the model greatly overemphasizes the effect of regularity. The problem with the DRC model is that it considers regularity a categorical variable. Thus *bind* is viewed as being as irregular as *pint* because the rule for medial *i* has been violated equally in both words. Interestingly, *bind* and *pint* are very similar phonologically. In fact, phonologically they differ only by the voicing of the first and last phonemes. However, the mean naming latency of *bind* was 560 msec, whereas the mean naming latency of *pint* was 615 msec. We claim that the difference in naming latencies of these two words can be traced to their differences in word-body consistency. It seems possible that the DRC model could account for these findings by making its rules more context sensitive. Currently, context-sensitive rules exist within the model, but they are applied rather narrowly. In terms of the present study, context sensitivity could be expanded, and this would result in making some of our irregular stimuli regular and vice versa. For example, rules such as *i(nd)*–/aɪ/ (due to the pronunciation of *find*, *kind*, *mind*, etc.) might be formed. Another possibility would involve changing parameter values so that there would be more of a reliance on the lexical route. If consistency effects reflect activity within the lexical route as Coltheart et al. (1993) state, then presumably they could be made greater by putting more emphasis on lexical processing. There is some recent evidence that subjects may be able to rely on lexical/semantic processing in some situations and sublexical/phonological in others (Zevin & Balota, 2000). However, the model did not show any hint of a consistency effect in the simulations presented here or in those presented by Jared (1997). Thus, in its present form, the DRC model does not appear to be compatible with the results reported here, and it is difficult to say whether a change in parameter values would be detrimental to the model's performance in other areas that the model has had success with.

It should be noted that the results of the present study are somewhat of a paradox when compared with the results of studies examining the role of orthographic units in nonword pronunciation. For example, in Andrews and Scarrett's (1998) recent study, although word-body consistency accounted for more variance than did grapheme-to-phoneme consistency in pronunciation variability of nonwords (i.e., the extent to which subjects gave different pronunciations for particular stimuli), subjects usually assigned pronunciations of nonwords according to GPC rules. Furthermore, the results of a large-scale study by Seidenberg, Plaut, Petersen, McClelland, and McRae (1994) on the pronunciation of 590 nonwords are consistent with the finding of Andrews and Scarrett. In Seidenberg et al.'s study, 24 participants named nonwords generated from 590 different word bodies. If we consider the stimuli used in the present study, 20 items

(the 10 regular $E > F$ words and the 10 irregular $F > E$ words) had pronunciations in which word-body consistency and GPC rules were in conflict with one another. Fortunately, the preferred pronunciations of these particular word bodies in nonwords containing them are available in Seidenberg et al. The GPC-generated pronunciation was favored for 14 of the critical nonwords, whereas the word-body generated pronunciation was favored for only 5 of the items (*choll*, *tearl*, *doup*, *grall*, and *skose*). The results from studies of nonword pronunciation suggest that the grapheme is more influential than the consistency of the word body when applied to the pronunciations of novel word-like stimuli. The paradox can be made clear by returning to the items *bind* and *pint*. Whereas it seems apparent that the difference in latencies reflects the consistency of these two items, in the Seidenberg et al. study, the GPC-based pronunciation was favored for both *jind* (by 23 out of 24 subjects) and *bint* (by all 24 subjects). Thus, it seems that consistency matters most for word pronunciation and favors the PMSP96 model, and that regularity matters most for nonword pronunciation and favors the DRC and DP models.

In addition, Andrews and Scarrett (1998) found that in terms of the word body, the number of neighbors accounted for more variance in a regression analysis than did the summed frequencies of those neighbors on several different measures. In contrast, Jared et al. (1990) found that with regard to word naming, the summed frequency of enemies rather than the number of enemies was responsible for increased naming latencies for inconsistent words (see note 1). Future research on consistency should focus on the basis of the discrepancy between the word and nonword studies.

In summary, the manipulation of GPC regularity and word-body consistency allowed for the examination of contrasting predictions made by the PMSP96, DRC, and DP models. The results are most consistent with the predictions made by the PMSP96 model and less so with the other two models. The conclusion from these studies is that during normal visual word recognition, the word body controls processing more than does the grapheme unit. Furthermore, a more accurate definition of phonological regularity should emphasize the consistency of pronunciation of similarly spelled words.

REFERENCES

ANDREWS, S., & SCARRETT D. R. (1998). Rule and analogy mechanisms in reading nonwords: Hough dou peapel rede gnew wirds? *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 1052-1086.

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.

CARROLL, J., DAVIES, P., & RICHMAN, B. (1971). *The American Heritage word frequency book*. Boston: Houghton Mifflin.

CLARK, H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning & Verbal Behavior*, **12**, 335-339.

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

CORTESE, M. J. (1998). Revisiting serial position effects in reading. *Journal of Memory & Language*, **39**, 652-665.

CORTESE, M. J., SIMPSON, G. B., & WOOLSEY, S. (1997). Effects of association and imageability on phonological mapping. *Psychonomic Bulletin & Review*, **4**, 226-231.

CORTESE, M. J., & ZEVIN, J. D. (1998). Lexicality effects in naming: Where in the word are they? *Abstracts of the Psychonomic Society*, **3**, 64.

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

HINTON, G. E., & SHALLICE, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review*, **98**, 74-95.

JARED, D. (1997). Spelling-sound consistency affects the naming of high-frequency words. *Journal of Memory & Language*, **36**, 505-529.

JARED, D., MCRAE, K., & SEIDENBERG, M. S. (1990). The basis of consistency effects in word naming. *Journal of Memory & Language*, **29**, 687-715.

KAWAMOTO, A., KELLO, C., JONES, R., & BAME, K. (1998). Initial phoneme versus whole word criterion to initiate pronunciation: Evidence based on response latency and initial phoneme duration. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 862-885.

MCCLELLAND, J. L., & RUMELHART, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, **88**, 375-407.

PLAUT, D. E., MCCLELLAND, J. L., SEIDENBERG, M. S., & PATTERSON, K. E. (1996). Understanding normal and impaired reading: Computational principles in quasi-regular domains. *Psychological Review*, **103**, 56-115.

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

RASTLE, K., HARRINGTON, J., COLTHEART, M., & PALETHORPE, S. (2000). Reading aloud begins when the computation of phonology is complete. *Journal of Experimental Psychology: Human Perception & Performance*, **26**, 1178-1191.

RUMELHART, D. E., & MCCLELLAND, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, **89**, 60-94.

SCHNEIDER, W. (1990). *Micro-experimental laboratory*. Pittsburgh, PA: Psychology Software Tools, Inc.

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

SEIDENBERG, M. S., PLAUT, D. C., PETERSEN, A. S., MCCLELLAND, J. L., & MCRAE, K. (1994). Nonword pronunciation and models of word recognition. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 1177-1196.

SOLSO, R. L., & JUEL, C. L. (1980). Positional frequency and versatility of bigrams for two-through nine-letter English words. *Behavior Research Methods & Instrumentation*, **12**, 297-343.

SPIELER, D. H., & BALOTA, D. A. (1997). Bringing computational models of word naming down to the items level. *Psychological Science*, **8**, 411-416.

STRAIN, E., PATTERSON, K., & SEIDENBERG, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 1140-1154.

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 Memory & Language*, **124**, 107-136.

- ZEVIN, J. D., & BALOTA, D. A. (2000). Priming and attentional control of lexical and sublexical pathways during naming. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 121-135.
- ZIEGLER, J. C., STONE, G. O., & JACOBS, A. M. (1997). What is the pronunciation for *-ough* and the spelling for /u/? A database for computing feedforward and feedback consistency in English. *Behavior Research Methods, Instruments, & Computers*, **29**, 600-618.
- 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 & Performance*, **24**, 1131-1161.

NOTES

1. Jared et al. (1990) have demonstrated that the consistency effect is driven by the summed frequencies of friends and enemies rather than their number. Jared et al. observed an equivalent consistency effect between conditions in which the ratio of the number friends and enemies varied while the summed frequencies remained constant. Specifically, one condition of words had a greater number of friends and a smaller

number of enemies than another, but the summed frequencies of friends and enemies was the same, and naming latencies were equivalent.

2. It should be noted that we selected our stimuli on the basis of the summed frequencies of friends and enemies due to the findings of Jared et al. (1990). Both the conditional probabilities of Berndt et al. (1987) and Ziegler et al. (1997) are based on the number of words that contain these orthographic-to-phonological relationships. Thus, there may be some minor discrepancies in stimulus classification if a type approach (based on the number of words) is employed instead of a token approach (based on the summed frequencies of words). However, regarding the words in the present study, the friends-to-enemies ratio of these two approaches is highly correlated.

3. The simulation data is based on an assumption that the naming response begins when the phonology of the whole word has been completed. Although this assumption has been challenged recently (see Cortese, 1998; Kawamoto et al., 1998), changing the assumption to an initial-phoneme criterion should not greatly affect the results of the models, because an equal number of plosive and nonplosive initial phonemes were employed in each condition. An initial-phoneme criterion assumption would decrease the consistency/regularity effects for nonplosives but not for plosives. Because equal numbers of each occur in each condition, the pattern of results would be maintained for either criterion.

APPENDIX

Words Used in the Experiment, Their Naming Latencies, Grapheme-to-Phoneme Conditional Probabilities, and Orthographic-to-Phonological Probabilities

Item	Latency	Conditional Probability		Item	Latency	Conditional Probability	
		Grapheme	Body			Grapheme	Body
Regular $F > E$				Regular $E > F$			
toad	515	.933	.800	tone	544	.785	.750
batch	526	.542	.889	bead	634	.576	.308
storm	508	.314	.750	spook	536	.570	.000
dice	496	.589	1.0	dose	520	.868	.182
shut	528	.417	.889	foul	529	.324	.333
shear	512	.576	.789	font	585	.261	.500
paste	495	.651	.800	dome	521	.785	.667
leaf	505	.576	.666	mood	522	.570	.333
drool	538	.570	.857	grease	539	.868	.571
lass	519	.542	1.0	wreath	614	.576	.500
Irregular $F > E$				Irregular $E > F$			
toll	539	.314	.714	tomb	542	.006	.500
bind	560	.074	.875	pint	615	.074	.100
scarce	558	.020	.500	shove	524	.055	.689
pearl	514	.056	1.0	plaid	568	.003	.125
stealth	590	.298	1.0	suite	617	.046	.000
soup	516	.041	.666	flown	568	.502	.429
pour	488	.041	.200	cough	530	.517	.250
mall	505	.002	.929	youth	537	.041	.250
climb	537	.074	.500	gross	529	.314	.111
host	505	.314	.142	height	530	.142	.250

(Manuscript received February 8, 1999;
revision accepted for publication February 25, 2000.)