UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Neighborhood and Position Effects Interact in Naming Latency

Permalink

Journal

Authors

Milostan, Jeanne C.
Cottrell, Garrison W.
Ferreira, Victor

Publication Date

Peer reviewed

Neighborhood and Position Effects Interact in Naming Latency

Jeanne C. Milostan (JMILOSTA@CS.UCSD.EDU)
Garrison W. Cottrell (GARY@CS.UCSD.EDU)
Computer Science and Engineering Department 0114
University of California San Diego
La Jolla, CA 92093 USA

Victor Ferreira (FERREIRA@PSY.UCSD.EDU)
Department of Psychology
University of California San Diego
La Jolla, CA 92093 USA

Abstract

Naming latency studies have recently shown a position-of-irregularity effect (words with early irregularities seem slowed compared to those with late irregularities), for which Dual-Route models of reading can account. Milostan & Cottrell (1998) showed that the initial studies contained a confound between irregularity position and friend/enemy ratio, and that the statistical confound could be captured by connectionist networks which then show the supposed position effect. This paper presents work to disentangle the position/regularity confound through a subject study and additional connectionist explorations. The latency data show that, once friend/enemy ratios are controlled for, the supposed position effect is driven entirely by high-enemy words in the first position. Further, connectionist network simulations show that network error at the first phoneme position only is a better match for naming latency, while overall network error produces a better match to subject error counts.

Filler	Position 1	of 2	3	Irregular 4	5
Nonword					
Irregular	554	542	530	529	537
Regular	502	516	518	523	525
Difference	52	26	12	6	12
Exception					
Irregular	545	524	528	526	528
Regular	500	503	503	515	524
Difference	45	21	25	11	4
Avg. Diff.	48.5	23.5	18.5	8.5	8

Table 1: Naming Latency vs. Irregularity Position

Introduction

A major component of the task of learning to read is the development of a mapping from orthography to phonology. In a complete model of reading, message understanding must play a role, but many psycholinguistic phenomena can be explained in the context of this simple mapping task. A difficulty in learning this mapping is that in a language such as English, the mapping is *quasiregular* (Plaut et al., 1996); there are a wide range of exceptions to the general rules. As with nearly all psychological phenomena, more frequent stimuli are processed faster, leading to shorter naming latencies. The regularity of mapping interacts with this variable, a robust finding that is well-explained by connectionist accounts (Seidenberg and McClelland, 1989; Taraban and McClelland, 1987).

In this paper we continue consideration of a recent effect that seems difficult to account for in terms of the standard parallel network models. Coltheart & Rastle (1994) have shown that the amount of delay experienced in naming an exception word is related to the phonemic position of the irregularity in pronunciation. Specifically, the earlier the exception occurs in the word, the longer the latency to the onset of pronouncing the word. Table 1, adapted from (Coltheart and Rastle, 1994) shows the response latencies to two-syllable words by normal subjects. There is a clear left-to-right ranking of the latencies compared to controls in the last row of the Table. Coltheart *et al.* claim this delay ranking cannot be achieved by standard connectionist models. Earlier work (Milostan and Cottrell, 1998) showed that the origin of the effect seen in the Coltheart study lies in a statistical regularity of English, related to the number of "friends" and "enemies" of the pronunciation within the word's neighborhood. [1] The human subject study and network simulations presented in this paper attempt to tease apart the effects of phoneme position and neighborhood ratio.

Background

Computational modeling of the reading task has been approached from a number of different perspectives. Advocates of a dual-route model of oral reading claim that two separate routes, one lexical (a lexicon, often hypothesized to be an associative network) and one rule-based, are *required* to account for certain phenomena in reaction times and nonword pronunciation seen in human subjects (Coltheart et al., 1993). Connectionist modelers claim that the same phenomena can be captured in a single-route model which learns simply by exposure to a representative dataset (Seidenberg and McClelland, 1989).

In the Dual-Route Cascade model (DRC) (Coltheart et al., 1993), the lexical route is implemented as an Interactive Activation (McClelland and Rumelhart, 1981) system, while the non-lexical route is implemented by a set of grapheme-phoneme correspondence (GPC) rules learned from a dataset. Input at the letter identification layer is activated in a left-to-right sequential fashion to simulate the reading direction of English, and fed simultaneously to the two pathways in the

---

[1]Friends are words with the same pronunciations for the ambiguous letter-to-sound correspondence; enemies are words with different pronunciations.

model. Activation from both the GPC route and the lexicon route then begins to interact at the output (phoneme) level, starting with the phonemes at the beginning of the word. If the GPC and the lexicon agree on pronunciation, the correct phonemes will be activated quickly. For words with irregular pronunciation, the lexicon and GPC routes will activate different phonemes: the GPC route will try to activate the regular pronunciation while the lexical route will activate the irregular (correct) pronunciation. Inhibitory links between alternate phoneme pronunciations will slow down the rise in activation, causing words with inconsistencies to be pronounced more slowly than regular words. This slowing will not occur, however, when an irregularity appears late in a word since the lexicon will try to activate *all* of a word's phonemes as soon as the word's lexical node becomes active. If an irregularity is late in a word, the correct pronunciation will begin to be activated before the GPC route is able to vote against it. Hence late irregularities will not be as affected by the conflicting information. This result is validated by simulations with the one-syllable DRC model (Coltheart and Rastle, 1994).

Several connectionist systems have been developed to model the orthography to phonology process (Seidenberg and McClelland, 1989; Plaut et al., 1996). These connectionist models provide evidence that the task, with accompanying phenomena, can be learned through a single mechanism. In particular, Plaut *et al.* (henceforth PMSP) develop a recurrent network which duplicates the naming latencies appropriate to their data set, consisting of approximately 3000 one-syllable English words (monosyllabic words with frequency greater than zero in the Kuçera & Francis corpus (Kuçera and Francis, 1967)). Naming latencies are computed based on time-to-settle for the recurrent network, and based on mean squared error (MSE) for a feed-forward model used in some simulations. The structure of the feed-forward network is shown in Figure 1. In addition to duplicating frequency and regularity interactions displayed in previous subject work, this model also performs appropriately in providing pronunciation of pronounceable nonwords. This provides an improvement over, and a validation of, previous work with a strictly feed-forward network (Seidenberg and McClelland, 1989). (Milostan and Cottrell, 1998) then showed that the serial position effect proposed by Coltheart & Rastle could be accounted for by a statistical regularity in English, as measured by the Enemy Ratio (# of enemies in a word's neighborhood divided by the total size of the word's neighborhood). (Milostan and Cottrell, 1998) showed that, for the words used in (Coltheart and Rastle, 1994), words with earlier irregularities had higher enemy ratios, and that the parallel connectionist model of PMSP, exposed to the same statistical regularities, also shows the same left-to-right effect that (Coltheart and Rastle, 1994) claimed it would not.

Experiment

Intuition suggests that, since English is read from left to right, left-to-right phenomena such as the serial position effect might be seen, independent of statistical confounds. However, as with all assumptions, such effects must be verified through careful testing, and the source of such effects must be carefully delineated within the model hypothesized for the system at hand.
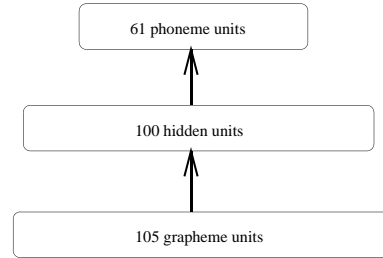

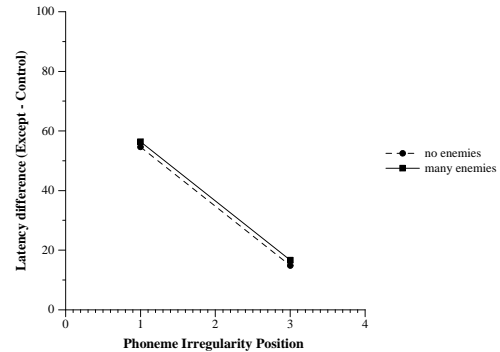Figure 1: Single Syllable Ortho-to-Phono Network

Figure 2: Hypothetical Position-Only Effect

In a serial system such as the DRC, which by design processes input orthography from left to right, any observed left-to-right irregularity effect is a direct result of the GPC operation. On the other hand, for a parallel model such as the PMSP system, which produces the output phonology all at once, effects of irregularity are driven by neighborhood enemy/friend measures, and serial effects should disappear once these enemy ratios are controlled.

The serial position effect seen by Coltheart & Rastle could be the result of a confound between the position of the irregularity and the statistics of English. Earlier positions appear to have more irregularities. It would be productive, then, to retest the Coltheart & Rastle hypothesis, this time controlling for amount of consistency. If the serial position effect does hold regardless of the enemy ratio of the test words, an effect similar to that shown in Figure 2 would be expected. If, however, the effect is due to enemy ratio alone, the results should be similar to that of Figure 3. The subject experiment and network simulation presented here are an attempt to adjudicate between these options, and stimuli will vary in both position of irregularity, and in enemy ratio, in order to determine the source of the effects.

Difficulties of GPC rules

One of the major discrepancies between the PMSP work and DRC model is the latter's assumption of the existence of a pronunciation rule system. This rule system *defines* whether a word is regular or not. Thus, all irregular stimuli chosen for experiments on the DRC model are chosen according to the GPC rules. Experiments which attempt to refute the DRC model at any level must also take these rules into consideration when choosing stimuli.

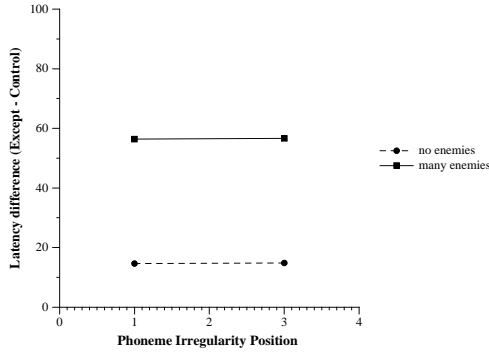Ideally, the same words that the DRC system uses should

Figure 3: Hypothetical Enemy-Ratio-Only Effect

| | **Position** | |
Enemy Ratio	Front	Back
High	aunt	plaid
Low	earl	fluke

Table 2: Sample Experiment Words

vowel for coda consonants, and the following vowel for onset consonants. As expected, consonant irregularities are by far the minority, and are limited to 'CH', 'TH', 'G', 'C', 'Q', and the silent instantiations such as 'T' and 'H'.

Methods

Subjects
Subjects were 23 undergraduate psychology students from University of California San Diego. All subjects had normal or corrected-to-normal vision, and were native North-American-English speakers. They were given course credit for their participation.

Materials
Sixty-four words with irregular grapheme-to-phoneme correspondences (according to the GPC rules of the DRC model) were chosen. Each target was uninflected and monosyllabic, and had between 3 to 6 letters with Kuçera-Francis frequency between zero and twenty-two.

The chosen words had an irregular grapheme-to-phoneme correspondence in either the first ("front") or third ("back") phoneme position, and were divided into 2 lists on that basis. Each list was further divided into two sublists, based on whether the word had only friends in the neighborhood based on the regularity (Enemy Ratio $En = 0.00$) or mostly enemies at that location (Enemy Ratio $0.6 <= En < 1.0$). Since a word's neighborhood by our measure includes itself, words with a neighborhood size of one ("loners") were excluded from consideration. These words correspond to Colheart's categorization of "irregular consistent".

Of the eligible words, the front-enemy condition had only 16 candidate words. Each of the other three conditions were randomly pruned down to size 16 in order to balance the conditions. The resulting average word frequency did not differ significantly between conditions ($M = 4.8281, F(3, 60) = 0.476, p = 0.700$). Each irregular word was then matched with a regular control word. Control words were matched to their irregular partners based on initial phoneme (since different phonemes take longer to trigger the microphone) and number of letters. The controls were also in the zero to 22 Kuçera-Francis frequency range.

An example test word from each of the four conditions is shown in Table 2.

Results
Of the original 25 subjects, data from 2 were unusable (in one case the latency data were accidentally deleted; in the other case the audio recording did not function so errors could not be scored). For the remaining 23 data sets, latencies associated with voice key failures were discarded; if the stimulus was either a test word or a control the associated (control or test) word was similarly disregarded (13 pairs total

be addressed. Thus, irregular words for this experiment were identified using the Australian GPC used in the DRC system, and the Australian pronunciations from the MRC database, from which the GPC rules were initially derived. Neighborhood Enemy Ratios were also calculated using the MRC pronunciations. American English would then be used only for identifying errors in subjects' pronunciations.

A program was written implementing the GPC rules of the DRC system as listed in (Rastle and Coltheart, 1999). A word was considered irregular if the pronunciation generated by the rules did not match the pronunciation provided by the MRC database. From the list of identified exception words, homographs where one generated pronunciation was correct were excluded (/wind/ vs. /wInd/), as were Australian words with spellings not commonly used in the United States (gaol). Also excluded were words marked as irregular by the GPC rule which states that word-final /s/ is *always* converted to /z/. This rule causes all words ending in -ace or -ice (face, mice) to be considered irregular.

Overall, the words which were identified by this procedure using the Australian English were also found to be irregular through a similar procedure using the rules of American English pronunciation from (Venezky, 1970). The details of that investigation are reported with a companion study in (Milostan et al., 2000).

Neighborhoods
Neighborhoods are defined using an extension of the (Taraban and McClelland, 1987) neighborhood rules described in (Milostan and Cottrell, 1998), summarized for single syllable words here:

- Consonant neighborhoods consist of orthographic clusters which correspond to the same location in the word. For one-syllable words, this results in 2 consonant cluster locations: onset and coda.

- Each vowel group is considered within the context of its coda. In order for a word to be in the neighborhood of a test word, it must have the same vowel group ('E' is considered separately from 'EE') and be followed by the same consonant cluster ending that syllable. As an example, the 'OO' neighborhood in 'BOOK' are all those words ending in 'OOK', with the first syllable coda containing only 'K'.

- Consonant cluster neighborhoods include the preceding
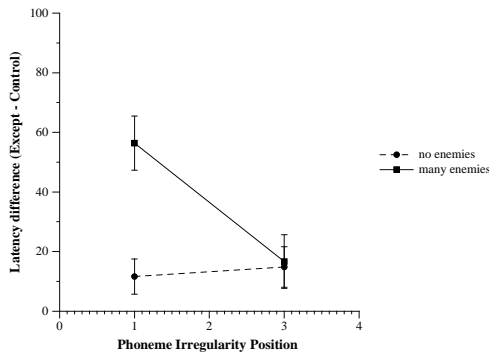
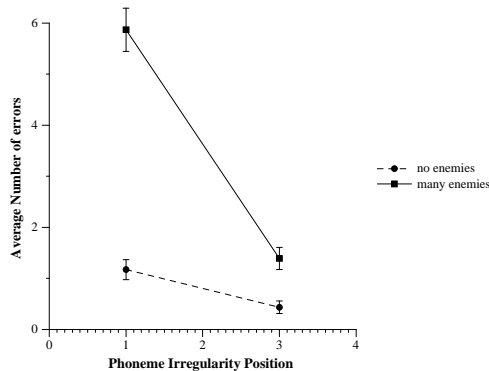Figure 4: Naming Latency Results overall

Figure 5: Error Count Results overall

over all subjects). Latencies for all nonword fillers were also discarded. Words which were pronounced incorrectly, along with the associated match, were removed for separate error analysis.

Naming latency differences were then calculated by subtracting the control word latency from the associated test word latency. Analysis of variance (ANOVA) was then performed on these values. Words in the high enemy ratio condition had significantly greater latency differences than the words in the friend condition ($F(1,22) = 18.16, p < 0.0005$), and there was a significant interaction between enemy ratio and position of irregularity ($F(1,22) = 8.419, p = 0.008$). Latency differences for first and third position irregularities, combining both enemy ratio conditions, approached but did not reach significance ($F(1,22) = 3.766, p = 0.065$). The latency data is shown in Figure 4.

Subjects made a total of 22 errors on control words, and 248 errors on irregular test words. Control words are not considered in the error analysis. Subjects made significantly more errors for front position irregulars than for back position irregulars ($F(1,22) = 32.922, p < 0.0005$), and more errors for high-enemy words than for low-enemy words ($F(1,22) = 326.549, p < 0.0005$). Position and enemy ratio also had a significant interaction in number of errors made ($F(1,22) = 12.415, p = 0.002$). These error data are shown in Figure 5.

Discussion

From the data collected in this experiment, there is a slight effect of irregularity position, but this appears to be completely driven by the words with high enemy ratios (see Figure 4). First-position-irregular words with high numbers of enemies in their neighborhood take longer to name than similar words with friends only. This effect has mostly disappeared for those words with third position grapheme-phoneme irregularities.

This makes sense from a cascaded information processing point of view (McClelland, 1979), since it is possible that any (potential) errors late in a word can be resolved by the time the third phoneme is ready to be produced. This difference in time delays can be considered an effect of the temporal nature of the speech process, and the time available to make online corrections. Words with later irregularities have, by definition, *regular* grapheme-phoneme correspondences at the beginning. The subject can begin pronouncing those phonemes immediately, even if she must then make accommodations later. Thus, the initial phoneme in an irregular (high enemy ratio) word may be produced with the same latency as a completely regular word, while the phoneme at the irregular mapping itself may actually be delayed internal to the word. However, there is currently no way of measuring the latencies of each internal phoneme using only the voice key.

Feed-Forward Network Performance

The feed-forward network of PMSP does not contain a temporal component. Since all phonemes are calculated simultaneously, the irregularity position may not play a part in the latencies calculated from the network as these are actually a measure of the difference between the correct target pronunciation and the network's actual output across the word. Thus, the feed-forward pronunciation network should be affected by enemy ratio alone, as those words with many contradictory spelling-sound mappings will receive less total reinforcement for the correct mapping.

Five feed-forward connectionist networks were trained on 3015 single syllable words as described in (Plaut et al., 1996; Milostan and Cottrell, 1998). This data set is the 2998 words used in PMSP plus 17 additional words used in the current subject experiment. These words were not included in the PMSP data set as they are of frequency rating zero.

Naming latency was then calculated for each test word by using the sum squared error at the output layer, producing the results shown in Figure 6. Unexpectedly, it appears that the back position irregulars take *longer* to name than the front irregulars, regardless of enemy ratio. Remember, however, that naming latency in these feed-forward networks is a measure of error, not of time directly. The representation used for the output layer is a sparse coding of the output phonemes. Of the 62 units, only a small number will be on for any particular word. Thus, the network is exposed to a training set where the majority of the output units are off most of the time. These networks learn how to turn off units very well, and thus there will be less discrepancy between the target and actual output which should be off and is, than for a target which should be on and the related output unit which is actually on. This means that, everything else being equal, training pairs with more units on in the target will inherently produce more error than for those with fewer on targets.

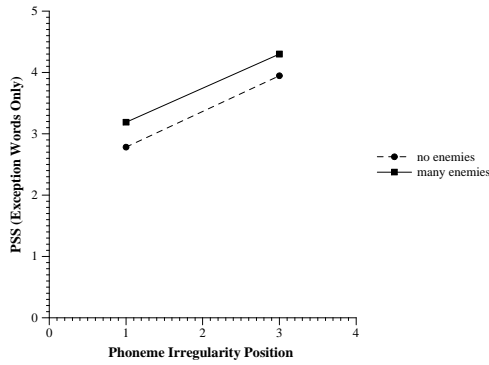As an example, consider a hypothetical network with 10 output units, and compare the results of two targets, one of
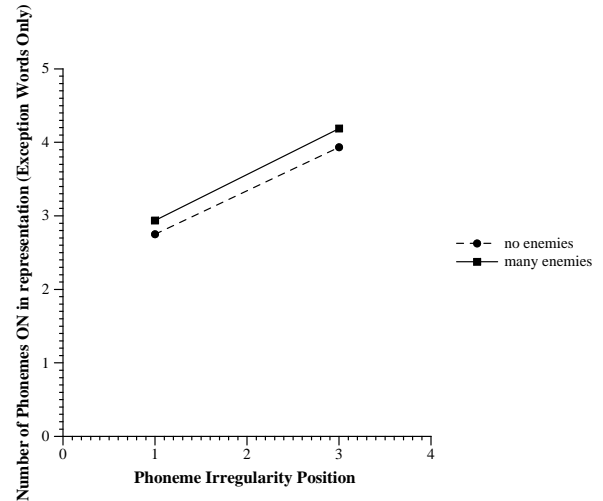
Figure 6: Network Sum Squared Error
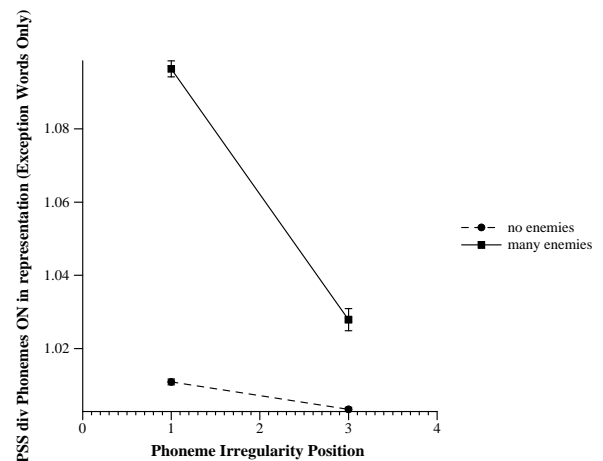

Figure 7: Number of Units ON in Output Representation

Figure 8: Network Error / Number of Phonemes ON

which has one unit on and the other of which has two units on. Since the "off" units receive more training, assume that any units off in the target have an activation of 0.1 in the actual network during testing, while the on-units are activated at 0.8. Both of the hypothetical training sets accurately produce the intended output, but in the case where the target has one unit on the network shows an error of $(0.1^2 * 9) + (0.2^2 * 1) = .13$, while the target with 2 units on has an error of $(0.1^2 * 8) + (0.2^2 * 2) = .16$. This discrepancy becomes more exaggerated as the number of off-units increases. Thus, if there is a systematic difference in the expected number of on-units among the conditions, those targets with more on-units may be unduly penalized. Examination of the output targets for the various test categories reveals that indeed, those in the back position conditions have more on-units than the front position targets, as shown in Figure 7. This means that the words in the back position systematically have one more phoneme than the front-position test words.

In parallel connectionist models, output error is associated with naming latency under the assumption that the more error the output shows, the longer it takes for the system to then converge on a veridical representation for a further stage which will begin the actual production of the speech signal. If the output for each of the ON-bits in the representation can be cleaned up in parallel, then the time required before the next stage may begin is more a measure of the average amount of time required to make the cleanup. Thus, the average ON-bit error provides a more realistic measure of naming latency.

To correct for the bias in number of bits on between first- and third-position words, the total output error for each word was divided by the number of ON-units in that word's output representation. These results are shown in Figure 8. As in the human data, the networks show a significant effect of enemy ratio ($F(1,4) = 478.669, p < 0.0005$) and a significant interaction between enemy ratio and position ($F(1,4) = 621.335, p < 0.0005$). Unlike the human subjects, however, the networks also show a significant effect of position ($F(1,4) = 683.588, p < 0.0005$). Again, this appears to be mostly driven by the high enemy ratio words. This is actually a bit surprising since the networks produce output in parallel, and thus would not have any time to "correct" for later-position errors. The network results instead reflect the finding that English words are more consistent in endings than they are in onsets (Treiman et al., 1995).