Performance of the Model on Intervocalic Consonants

To specifically examine the performance of the model on intervocalic consonants, all disyllabic words that the model was trained on that had 6 or less intervocalic consonants and the percentage of errors that the model made on them was examined. Only disyllabic words were used so that the performance could be compared with Perry, Ziegler, and Zorzi (2010). Consonants that were part of vowel graphemes (e.g., the –r in the –or in *fort*) were not considered intervocalic consonants. The results appear in the table below. Unlike Perry et al., any error on the words was considered an error, including the incorrect selection of graphemes that were not intervocalic consonants. This was done because the selection of intervocalic consonants is contextually dependent on other letters, and thus context may be involved in the incorrect selection of intervocalic graphemes. The results below therefore penalize the current model in terms of how accurate it is compared to Perry et al., who reported that CDP++ was able to correctly assign graphemes to 82.16% of all disyllables with 1 or more intervocalic consonants.

Parsing error rate on disyllabic words with zero to six intervocalic letters

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *N. Intervocalic Letters* | *N. Errors* | *N. Correct* | *% Error* |  |
| 0 | 154 | 388 | 28.41 |  |
| 1 | 382 | 8001 | 4.56 |  |
| 2 | 617 | 10194 | 5.71 |  |
| 3 | 473 | 3217 | 12.82 |  |
| 4 | 163 | 600 | 21.36 |  |
| 5 | 25 | 53 | 32.05 |  |
| 6 | 2 | 5 | 28.57 |  |
| Overall | 1816 | 22458 | 7.48 |  |

As can be seen, the overall error rate for the segmentation of words was relatively low, at least compared to the results of Perry et al. (2010). The higher error rate on the words with no letters between the vowels was mainly caused by words that were split based on morphological or other types of boundaries that the model would typically assign single graphemes to. Note that in Perry et al., these were not examined because these words always created a different number of orthographic and phonological vowels, and hence were not used in training or testing. For example, *rearm* was given the graphemes *r.ear.m* rather than *r.e/ar.m* and *sour* was given the graphemes *s.our* rather than *s.o/ur*. With the model presented here, both of these graphemes were split into two based on there being two phonological syllables in the lexical phonology and hence used in training and testing.

In terms of the words with 4 and 5 intervocalic consonant letters, the sudden jump in error rates appeared to be caused by compound words that use two morphemes. With these words, the morphemes often caused segmentations that are not based on the most common phoneme sequences (e.g., *gu.ar.d.s/m.e.n* vs. *gu.a.r.d/s.m.e.n*), and the model failed to learn many of these, or the morphemes used an –e grapheme which was incorrectly predicted by the model to be a vowel and not a consonant (e.g., *g.a.t.e/c.r.a.sh* vs. *g.a/t.e/c.r.a.sh*).

Parameters used in CDP++.parser

Lexical Route

Features

Feature to letter excitation: 0.005

Feature to letter inhibition: -1.5

Letters

Letter to letter inhibition: -.3

Letter to orthography excitation: 0.02

Letter to orthography inhibition: -1.5

Orthographic lexicon

Orthography to orthography inhibition: -.1

Orthography to letter inhibition: 0

Orthography to phonology excitation: 1.9

Orthography to letter excitation: 0

Phonological lexicon

Phonology to phonology inhibition: -.12

Phonology to phoneme excitation: 0.13

Phonology to phoneme inhibition: -0.135

Phonology to orthography excitation: 2.5

Phoneme Output Buffer

Phoneme to phoneme inhibition: -0.005

Phoneme to phonology excitation: 0.082

Phoneme to phonology inhibition: -.16

Sublexical Route

Graphemic parsing cycles per letter: 10

Sublexical network to phoneme output buffer/stress output node activation: 0.065

Level of activation which a letter must be over before graphemic parsing begins: .20

Temperature (s) in the sublexical network: 3

Learning rate (e) in the sublexical network: 0.05

Dead node level: 15.0

Word stress parameters

Stress node naming criterion: 0.1 (unless otherwise stated)

Phonological lexicon to stress output node excitation: .037

Phonological lexicon to stress output node inhibition: -.023

Stress output node to stress output node lateral inhibition: -.11

Overall parameters

Overall activation rate: 0.15

Lexicon frequency scaling: 0.15 \* (log (word frequency + 2)/log (maximum word

frequency + 2))

Phoneme naming activation criterion: 0.5

Cycle-to-cycle stopping criterion: 0.0023

Results from CDP++.parser on other data sets

Length by lexicality and frequency. Weekes ([1997](#_ENREF_52))

Weekes ([1997](#_ENREF_52)) found that people responded to high frequency words more quickly than low frequency ones, and that the length effect with nonwords was bigger than it was with words. The model displayed a similar pattern, with an effect of word frequency when examining the high frequency versus low frequency words separately, *F*(1, 197) = 257.10, *p* < .001, as well as a Word Length by Lexicality interaction, *F*(3, 282) = 6.86, *p* < .001. (Means: High Frequency Words (3, 4, 5, 6 letters): 75.1, 81.5, 84.4, 88.0; Low Frequency Words (3, 4, 5, 6 letters): 87.5, 92.3, 95.7, 100.3; Nonwords (3, 4, 5, 6 letters): 137.4, 134.8, 148.7, 159.2. Variance Explained (*r*2): Words = 4.2%; Nonwords: 20.4%)

Length by lexicality and body neighborhood. Ziegler, Perry, Jacobs, and Braun ([2001](#_ENREF_56))

Ziegler et al. ([2001](#_ENREF_56)) examined the lexicality by length interaction, and found similar results as Weekes. They also examined and found an effect of body neighborhood (see Ziegler & Perry, [1998](#_ENREF_55)). The results of the model also showed this pattern, with main effects of Length (*F*(3, 144) = 21.33, *p* < .001), Lexicality (*F*(1, 144) = 810.00, p < .001), Body Neighborhood (F(1, 144) = 4.82, p < .05), and an interaction between Length and Lexicality (*F*(3, 144) = 10.66, *p* < .001). (Means, Low Body Neighborhood Words (3, 4, 5, 6 letters): 78.7, 84.5, 85.8, 87.6; Low Body Neighborhood Nonwords (3, 4, 5, 6 letters): 136.6, 133.6, 151.3, 172.3; High Body Neighborhood Words (3, 4, 5, 6 letters): 79.5, 80.8, 83.1, 86.2; High Body Neighborhood Nonwords (3, 4, 5, 6 letters): 122.8, 128.7, 155.7, 157.6. Variance Explained (*r*2), Words = 10.2%; Nonwords = 13.9%)

Consistency, irregularity and neighborhood. Jared (2002)

Jared (2002) examined the effect of summed neighborhood frequency (calculated at the level of the orthographic body) on the responses of words with irregular and inconsistent correspondences. The results of the model were similar to those found in Jared, with words with higher frequency enemies than friends showing significant effects of regularity and consistency but words with higher frequency friends than enemies not showing significant effects of regularity and consistency. (Low Frequency Exception (Enemies > Friends)/Control: 100.5/92.8.0, t(38) = 4.18, p < .001; Low Frequency Exception (Friends > Enemies)/Control: 94.6/92.3, *t*(37) = 1.46, *p* = .16; Low Frequency Inconsistent (Enemies > Friends)/Control, 98.4/92.5, *t*(38) = 3.22, *p* < .005; Low Frequency Inconsistent (Friends > Enemies)/Control: 93.3/91.3, *t*(38) = 1.51, p = .14 ; High Frequency Exception (Enemies > Friends)/Control, 84.2/78.2 *t*(37) = 3.23, p < .005; High Frequency Exception (Friends > Enemies)/Control: 82.4/80.4, *t*(36) = 1.44, *p* = .16; High Frequency Regular Inconsistent/Control: 82.8/ 80.5, *t*(38) = 1.11, *p* =.27. Variance Explained (*r*2), Experiment 1 = 28.2%; Experiment 2 = 36.0%; Experiment 3 = 48.8%; Experiment 4 = 42.6%).

Position of irregularity. Rastle and Coltheart ([1999](#_ENREF_37))

Rastle and Coltheart ([1999](#_ENREF_37)) examined the regularity effect in the first, second, and third phonemic positions of words, and found that words with an irregular correspondence early in their sequence showed a bigger effect than those with a correspondence later in their sequence. The model showed a similar pattern. (Means (Irregular/Control), 1st Position: 101.2/93.5, *t*(38) = 2.78, *p* < .01; 2nd Position: 98.0/ 93.8, *t*(75) = 2.71, p < .01; 3rd Position: 99.1/96.2, *t*(56) = 1.64, *p* = .11. Variance Explained (*r*2) = 26.2%)

Position of irregularity. Roberts and Besner ([2003](#_ENREF_41))

Roberts and Besner ([2003](#_ENREF_41)) also examined the regularity by position interaction. Like CDP++, but unlike the data, the model showed a regularity effect in both groups, and not just the 2nd position irregular group. (Means (Irregular/Control): 2nd Position: 103.7/95.7, *t*(66) = 5.15, *p* < .001; 3rd Position: 101.4/96.3, *t*(35) = 2.43, *p* < .05).

Stress Assignment in Disyllables. Rastle and Coltheart ([2000](#_ENREF_38))

Words. Rastle and Coltheart ([2000](#_ENREF_38)) reported that low frequency stress irregular words (using their definition of stress regularity) displayed a stress regularity effect, but high frequency words did not. To simulate this, we increased the stress response criterion to 0.35 in the model. The results showed a significant stress regularity effect with low frequency words, 126.3 vs. 112.3 cycles (*t*(85) = 3.25, *p* < .005; 18% stress error rate), but not high frequency words, 107.6 vs. 99.2 cycles, (*t*(16) = 1.56, *p* = .14; 4% stress error rate). (Variance Explained (*r*2) = 29.2%). Alternatively, they noted that using only an iambic vs. trochaic stress distinction, no stress effects could be found in either low or high frequency words. This is what the model produced on their stimuli also with the normal parameter set, with the model making only three errors on their stimuli and displaying no significant interaction between frequency and stress regularity (F<1), and displaying no main effect of stress regularity (F < 1). (Variance Explained (*r*2) = 18.6%).

Nonwords. Rastle and Coltheart’s ([2000](#_ENREF_38)) experiments also had a set of disyllabic nonwords that the model can be tested on. The model was run on all of their items, although four nonwords were removed because they were coded graphemically as three syllables. The error rate of the model was 5.2%, and the model displayed a similar level of accuracy as CDP++ with stress assignment in terms of choosing the stress category that the majority of participants gave to each nonword (92.3% with trochaic nonwords and 55.7% with iambic nonwords). Of the nonwords the model made errors on, a number triggered dead-nodes, which would have signalled to the model to use a backup read-out strategy. If these are removed, then the error rate is 4.3%.

Stress assignment and onset cluster complexity. Kelly ([2004](#_ENREF_24)).

Kelly (2004) examined the effect of onset complexity on stress judgements with disyllabic words and found that complex onsets attracted stress to the syllable more often than simple ones. The results of the model were very similar to the data, with the model correctly predicting that nonwords with complex onsets should attract stress more often than those with simple onsets (90.3% vs. 76.6%). The model also did a reasonable job of predicting which nonwords participants gave trochaic or iambic responses to the majority of the time (92.0% of trochaic and 64.3% of iambic nonwords). Note that the stress pattern of what would have been 3 naming errors was entered into the data because stress judgement tasks with disyllables involve a forced choice rather than an overt pronunciation that may have errors in it.

Effect of consonants on vowel pronunciation. Waese and Jared ([2006](#_ENREF_51))

Waese and Jared ([2006](#_ENREF_51)) examined the extent that people use short vowel responses when either a single consonant, two consonants that do not form a legal onset, or two consonants that do form a legal onset occur after a single letter vowel. The model produced similar results to CDP++ (46.8%, 73.7%, 85.7%). Note that this model was trained on words which were removed to fix the distribution of long and short vowels in CDP++ (i.e., words with a middle –e, which generally use long vowel pronunciations). These numbers are therefore likely to represent an underestimate of the number of short vowels that the model would produce if these words were removed. The error rate of the model was 5.0%. However, if words that triggered dead-nodes are removed, then the error rate is 2.8%.

Nonword Consistency. Andrews and Scarratt ([1998](#_ENREF_1))

Andrews and Scarratt ([1998](#_ENREF_1)) examined the extent that body consistency influences people’s pronunciations of nonwords. Of particular interest were the no regular analogy words they used, which are words that share an orthographic body with other words, but where the shared words have bodies that do not have regular pronunciations. With these nonwords, the percentage of regular responses the model made was 25.0% (Experiment 1), 0% (Experiment 2, nonwords that share many bodies) and 39.1% (Experiment 2, nonwords that share few bodies), which is essentially the same as CDP++.

Pseudohomophones. Reynolds and Besner ([2005](#_ENREF_40))

It has been shown that pseudohomophonic nonwords are named faster than non-pseudohomophonic nonwords. There is also a negative correlation between the reaction times of pseudohomophones and the frequency of the words which the pseudohomophones share their pronunciations with (i.e., the base words). To examine this with the new model, the pseudohomophones and their matched controls that appear in Appendix C of Reynolds and Besner ([2005](#_ENREF_40)) were used – unmatched pseudohomophones were discarded, meaning that there were 133 items per group. The model was faster at naming the pseudohomophones than the controls (137.2 vs. 145.1 cycles, *t*(241) = 2.73, *p* <.01) and there was a negative correlation between the naming times of the pseudohomophones and log base word frequency (*r* = -.26, *p* < .005).

Surface Dyslexia. Behrmann & Bub ([1992](#_ENREF_8)); Patterson and Behrmann ([1997](#_ENREF_29))

MP is the perhaps the most studied surface dyslexic. MP showed effects of both consistency and frequency. To simulate this, the parameters of the model were changed in the same way as CDP++, with the frequency scaling being increased to 0.3, the phonological lexicon to phoneme output buffer excitation parameter being decreased to .032, and the phonological lexicon to phoneme output buffer inhibition parameter being set to zero. The results showed that the model displayed a consistency effect similar to MP (Patterson & Behrman, 1997: Percentage accuracy for the Inconsistent/Control words across the three levels of consistency: Model: 33/88, 53/83, 67/100; MP: 38/100, 57/83, 75/83). In terms of Frequency, on the Behrmann and Bub (1992) stimuli, the model also showed a very similar pattern (results of the 6 frequency bands (Inconsistent/Control): Model: 37/73, 24/83, 66/95, 63/77, 72/100, 71/89; MP: 18/72, 28/87, 63/87, 38/65, 61/100, 75/93).

Phonological Dyslexia: LB. Derousné and Beauvois ([1985](#_ENREF_14))

LB is a classic phonological dyslexic and hence reads words much better than nonwords. An interesting aspect of LB’s performance is that when tested on pseudohomophones, a higher performance was found than with non-pseudohomophonic nonwords. In addition, when the pseudohomophones were orthographically close to their base words, LB showed a better performance compared to when they were not. This was simulated here using the same parameter manipulation as CDP++ where all of the inhibitory connections in the model were divided by 3.9, and the output parameter from the sublexical network was reduced (in this case to .038). The results of the new model were very similar to LB (LB/Model, Close-Pseudohomophone: 85/75, Far-Pseudohomophone: 52/55, Control group for Close-Pseudohomophone group: 35/35, Control group for Far-Pseudohomophone group: 27/47.5).

Masked Priming. Forster & Davis ([1991](#_ENREF_15))

Forster and Davis (1991) showed that if a prime was presented for a very small amount of time before a target word, then the target word would be named faster if it shared the same onset. The same was not true of the orthographic body, however -- if the target and prime shared the same body but not the onset, this did not appear to speed responses. This was simulated in the same way as CDP++, using a 20 cycle prime exposure. The results showed that using primes with the same onsets as the target words caused a faster response than using primes that did not share the same onsets (74.7 vs. 79.2 cycles, *t*(25) = 10.81, *p* < .001). However, there was also a weak effect when the orthographic bodies of the prime and target words were the same (78.5 vs. 79.2 cycles, *t*(25) = 3.15, p < .005).

Testing whether schwas can be inferred from null-vowels

A final test of the model we performed was with the consonant-schwa-l pattern that was encoded with no vowel (e.g., *cuddle*), as described in Appendix B. To do this, the 494 disyllabic words with this pattern were extracted and were run through the model as if they were nonwords. The results showed that the model produced a plausible pronunciation that included the shwa on 470 of these words (94.9%). A further 6 (1.21%) of the nonwords had pronunciations that appeared reasonable, but did not include a schwa (e.g., *bristles* read as /brɪst.ləz/). The remaining 19 words (3.8%) were errors, with the majority being parsed incorrectly, where all of the graphemes after the vowel were placed in the consonant slots of the first syllable. This led to strange nonword responses. However, with all of these nonwords, a grapheme was placed in a position where a dead-node occurred (i.e., a slot in the network where nothing was learnt), and thus it would have been simple to identify these words and use a back-up strategy to read them out correctly.

References

Andrews, S., & Scarratt, D. R. (1998). Rule and analogy mechanisms in reading nonwords: Hough Dou Peapel Rede Gnew Wirds? *Journal of Experimental Psychology:Human Perception and Performance, 53*, 567-593.

Behrmann, M., & Bub, D. (1992). Surface dyslexia and dysgraphia: Dual routes, single lexicon. *Cognition Neuropsychology, 9*, 209-251.

Derousené, J., & Beauvois, M.-F. (1985). *The ‘‘phonemic’’ state in the nonlexical reading process: Evidence from a case of phonological alexia*. In K. E. Patterson, J. C. Marshall, & M. Coltehart (Eds.) Surface dyslexia: Neuropsychological and cognitive studies of phonological reading (pp. 399-458). London: Erlbaum.

Forster, K. I., & Davis, C. (1991). The density constraint of form-priming in the naming task: Interference effects from a masked prime. *Journal of Memory and Language, 30*, 1-25.

Kelly, M. H. (2004). Word onset patterns and lexical stress in English. *Journal of Memory and Language, 50*, 231-244.

Patterson, K., & Behrmann, M. (1997). Frequency and consistency effects in a pure surface dyslexic patient. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 1217-1231.

Rastle, K., & Coltheart, M. (2000). Lexical and nonlexical print-to-sound translation of disyllabic words and nonwords. *Journal of Memory and Language, 42*, 342-364.

Rastle, K., & Coltheart, M. (1999). Serial and strategic effects in read reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 482-503.

Reynolds, M., & Besner, D. (2005). Basic processes in reading: A critical review of pseudohomophone ffects in reading aloud and a new computational account. *Psychonomic Bulletin and Review, 12*, 622-646.

Roberts, M. A., Rastle, K., Coltheart, M., & Besner, D. (2003). When parallel processing in visual word recognition is not enough : new evidence from naming. *Psychonomic Bulletin and Review, 10*, 405-414.

Waese, M., & Jared, D. (2006). *The role of intervocalic consonants in disyllabic word naming*. Paper presented at the 47th Annual Meeting of the Psychonomic Society.

Weekes, B. S. (1997). Differential effects of number of letters on words and nonword naming latency. *The Quarterley Journal of Experimental Psychology, 50(A)*, 439-456.

Ziegler, J. C., & Perry, C. (1998). No more problems in Coltheart's neighborhood: resolving neighborhood conflicts in the lexical decision task. *Cognition, 68*, 53-62.

Ziegler, J. C., Perry, C., Jacobs, A. M., & Braun, M. (2001). Identical words are read differently in different languages. *Psychological Science, 12*, 379-384.