Stressing what is important: Orthographic cues and lexical stress assignment

Nada Ševa[1], Padraic Monaghan[2] & Joanne Arciuli[3]

[1]Department of Psychology, University of York, UK

[2]Department of Psychology, Lancaster University, UK

[3]Department of Speech Pathology, University of Sydney, Australia

Accepted for publication in *Journal of Neurolinguistics*

Corresponding author:

Nada Ševa

Department of Psychology

University of York

York

YO10 5DD

UK

E-mail: nadaseva@gmail.com

Tel: +44 1904 433177

Fax: +44 1904 43318

1. Introduction

Computational modelling has enabled links to be forged between neural structure and cognitive processes (see, e.g., Monaghan & Shillcock, in press; Rogers & McClelland, 2004). Computational models have also facilitated insight into the cognitive categories involved in particular tasks. Particularly insightful in this respect have been models of single word reading, where proposals for the precise mechanisms involved in mapping written words onto spoken forms have been tested (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Seidenberg & McClelland, 1989). Yet, these previous computational models of reading have concentrated on determining the mapping from letters, or sets of letters, onto phonemes, or sets of phonemes. In this paper we review the implications for this restriction to phonology in comparing computational models of reading, and show that considering stress assignment in reading is an important distinguishing characteristic between alternative cognitive accounts of word processing.

There are two recent traditions for modelling the cognitive processes involved in mapping letters onto phonemes: the dual-route model, and the connectionist triangle model. The dual-route framework incorporates into the model two systems for forming the mapping between letters and phonemes. The Dual-Route Cascaded (DRC) model (Coltheart, 2000; Coltheart et al., 2001) implemented these two routes in a model of reading, with the lexical route comprising a stored lexicon containing phonological information for all the words known to the hearer, and the second sub-lexical route which applies grapheme-phoneme correspondence rules to convert serially the orthographic input into phonemes. Though the two routes operate simultaneously and in parallel, for word reading, the lexical route is configured to process the written input faster than the sub-lexical route, and so correct naming of irregular words is achieved. For nonwords, there are no entries in the stored lexicon and output from the sub-lexical route determines the pronunciation. A recent development in the dual-route framework is the CDP+ model which provides an impressive fit to item-level naming data (Perry, Ziegler, & Zorzi, 2007). The model is an adaptation of the DRC, except that the grapheme-phoneme correspondence route is implemented as an associative network that is trained on the lexicon to discover the correspondences. In the DRC model, these correspondences are rule-based and provided to the model.

A contrasting tradition in modelling reading is the connectionist triangle model, where the mapping between orthography and phonology is mediated by direct links between these representations and also connections to and from a semantic representation of words. The triangle model has been implemented, to varying degrees of completeness (Harm & Seidenberg, 1999, 2004; Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989), in connectionist models where all connections between representations are learned. So, the model stores statistics about the associations between the representations, and these representations interact in the process of mapping written words onto pronunciation. The two frameworks of modelling reading have shown convergence over many aspects of their architectures, as exemplified by the incremental, nested modelling approach of CDP+, which encompasses both trained, associative networks characteristic of the triangle model tradition, as well as hard-wired localist lexical units inherited from the DRC. However, a key distinction is the nature of nonword reading in these models. In the dual-route model, pronunciation rules are applied to the graphemes of the nonword. In the triangle model, nonwords are read by analogy to similar words and parts of words to which the model has previously been exposed. This distinction proves to be critical for conceptions of how stress is applied to nonword naming, and we will return to this below. Imaging studies of reading and reading impairment support these notions of direct and indirect pathways involved in reading (e.g., Shaywitz, Lyon, & Shaywitz, 2006). Monaghan and Shillcock (in press) have demonstrated how anatomical distinctions in left and right hemisphere processing, and impairments to the specified pathways between left and right hemisphere processing can result in dyslexic behaviour for the reading task. Such links between reading and anatomy are driven by the theoretical advances resulting from implementing computational models of reading.

All the previously described models have focused only on monosyllabic reading. Such a limitation is a reasonable constraint to determine the lower bound of performance for a particular architecture of the reading system – so if a model fails to simulate human performance for reading monosyllables then it is insufficient as a model of the reading process. However, extending models to polysyllabic reading can potentially reveal limitations or over-specifications of models of reading. Take, for instance, the CDP+ model of reading. It requires a pre-mapping of letters into graphemes in the rule-correspondence route of the model. For a word like "graph", this is a relatively trivial task, as the system can recognise that "g", "r", "a", and "ph" are the graphemes to be extracted from the individual letters. Then, these can be mapped with little effort onto the phonemes. But for a word like "hothouse", the mapping becomes more problematic in that the model has to make a decision about whether "th" is one or two graphemes, and to which syllable in the output each grapheme is related. Thus, extending the CDP+ model to polysyllabic words, though possible, highlights a heavy pre-processing requirement in the system for determining the graphemes and the assignment of these graphemes to the different syllables, which were processes irrelevant to monosyllabic reading systems.

So far the only model of polysyllabic reading from a cognitive perspective has been developed for French (Ans, Carbonnel & Valdois, 1998). Although the basic assumptions of the model extend previous connectionist models of reading, it assumes additionally that reading of polysyllabic words is supported by two procedures that operate serially (therefore is in contrast to previous connectionist models of reading): first, a global procedure which uses knowledge about whole words, and second, an analytic procedure applying to word syllable segments which is activated only if the global procedure fails. The model successfully provided an account of basic effects in human reading such as the effect of frequency by consistency interaction and a position-of-irregularity, as well as some effects associated with different subtypes of dyslexia.

An additional challenge for models of polysyllabic reading is the assignment of lexical stress in reading. For some languages, like French, this question is not an issue because stress is always fixed on the last syllable. On the other hand, in languages such as English or Dutch, there is a tendency for words to be stressed on the first syllable (78% and 75% of disyllabic words, respectively, in the CELEX database, Baayen, Piepenbrock, & van Rijn, 1993). However, treating this as a default or regular state still requires that the reading system must determine stress position for the remaining 22-25% of the disyllabic words. How does the reading system accomplish this task? How, for instance does the reading system know that "giraffe" is second syllable stress (iambic pattern of stress), but that "zebra" is stressed on the first syllable (trochaic pattern of stress) in English?

Within the dual-route framework, an obvious solution is to include position of stress in the lexical route. So, we know "giraffe" is second syllable stress because our encoding of the lexical item includes information about both the phonemes within the word and stress position. But lexical-based storage of stress is problematic for nonword reading. Nonwords can be, and generally are, pronounced with stress by readers. The dual-route model simulates nonword reading by using only the GPC rules, as, by definition, there is no lexical entry. Consequently, in current formulations of the DRC model there is no mechanism for stress assignment of nonwords.

This difficulty was addressed by Rastle and Coltheart (2000) who generated a rule-based system for stress assignment for words and nonwords, that included both orthographic and phonological information and that extended to stress assignment in nonword reading. Rastle and Coltheart (2000) based their model on linguistic analysis of stress patterns in English by Fudge (1984) and Garde (1968), which suggested that 54 word beginnings and 101 word endings (most of which were morphemes in English) could influence the placement of stress. In particular, certain morphemes were either reliably stressed or unstressed (such as the suffix –ing, which rarely carries stress. The Rastle and Coltheart model (R&C model) consisted of several steps in an algorithm that used this correspondence between certain morphemes and stress position: 1) identification of first word beginnings and then word endings; 2) translation of the remaining parts of words into a phonological representation by using GPC rules plus a set of additional rules for correction of illegal phoneme combinations; 3) stress assignment based on the stored affix' stress position and the quality of the vowels (presence of schwa); and 4) if no prefix and suffix was identified, application of first syllable stress as the default stress position.

The R&C model was effective in correct stress assignment for 89.7% of English disyllabic words from the CELEX database (Baayen, et al., 1993). Its ability to extend to coverage of nonwords was tested by determining stress placement by participants for a set of disyllabic nonwords. The R&C model agreed with that of the majority of participants' decisions for 84.8% of these stimuli. However, the set of nonwords was somewhat biased to good performance by the model due to the majority of nonwords containing affixes listed in the R&C morpheme list. In particular, second-syllable stress is recognised by the model if the nonword contains a prefix that is habitually unstressed, and 40% of the second syllable nonwords in the stimulus set contained such prefixes. The model also included phonological information in the form of whether vowels were reduced as an additional contribution to stress position assignment which was not available only from the orthography. Thus, the relative role of orthographic and phonological cues for stress assignment remains unclear from this model.

An alternative in studying stress assignment in reading sets is the connectionist tradition, which proposes that the statistical regularities with respect to stress assignment will be learned in the same way as the learning of regularities in the orthography to phonology mapping (Harm & Seidenberg, 1999, 2004; Plaut et al., 1996; Seidenberg McClelland, 1989). Previous attempts in connectionist and other data-driven frameworks have shown that generalisation of lexical stress assignment is possible without using explicit linguistic rules (Arciuli, & Thompson, 2006; Daelemans, Gillis, Durieux, 1994; Gupta & Touretzky, 1994; Zevin & Joanisse, 2000). We extend these previous studies by modelling a large scale, realistic lexicon of disyllabic English words that learns the mapping between orthography and stress position. The model provides an empirical test of the extent to which probabilistic information in orthography can provide information about stress position in a representative lexicon of English. The model also provides a substantial first step in the connectionist modelling of polysyllabic reading, to determine the extent to which stress assignment may be accomplished using the same principles of connectionist models of reading that apply to mapping between orthography and phonology.

Corpus analyses of English polysyllabic words have indicated that there are numerous probabilistic cues available in the phonology and orthography for indicating stress position. In phonology, cues for stress are present both in the rime (reduced vowels are unstressed and consonantal clusters in codas are stressed, Chomsky & Halle, 1968) and in the onset (consonantal clusters tend to be stressed, Kelly, 2004). In orthography, length and complexity of coda and onset, as well as the identity of particular letters (both consonants and vowels) tend to predict stress assignment accurately, albeit, probabilistically (Arciuli & Cupples, 2006, in press; Kelly, Morris & Verrekia, 1998). For instance, there are some word beginnings, like *cu-* (which tends to be stressed) or *be-* (which tends to be unstressed), and endings, like *–um* (typically unstressed) or *–een* (typically stressed) that are not always morphological but still carry reliable information about stress position. Experimental studies have demonstrated that readers are sensitive to such phonological and orthographic cues present in the input (Arciuli & Cupples, 2006, in press; Kelly & Bock, 1988; Kelly et al., 1998). The R&C model of stress assignment incorporates a partial version of the orthographic information potentially available for stress assignment by focusing on frequent, regular patterns of letters that predict certain patterns of stress assignment. Such cues are a part of the orthographic information, but, in the connectionist tradition, they do not have special status beyond other, equally reliable cues that may be present in the letter string to determine stress position. This provides a substantial advantage in that a separate list of morphemes does not have to be incorporated into the connectionist model of stress assignment, and decisions about what does and does not count as a morpheme, and how it can be identified and isolated within the letter string, also do not have to be decided *a priori*.

We report the results of a connectionist model of English stress assignment that tested the extent to which orthographic regularities alone can predict stress assignment in English words and a variety of stimulus sets of nonwords. The first set of simulations aimed to investigate to what extent a connectionist model will be successful in stress assignment in English disyllabic words based only on orthographic representations. The model was then tested on two sets of nonwords, varying in the extent to which they contain affixes, and compared to readers' responses to the stimuli. The same test stimuli were applied also to the R&C model. We hypothesised that the connectionist model would perform well on nonwords that contained and did not contain affixes, whereas the R&C model was predicted to perform well only on nonwords with affixes.

2. Method

2.1. Architecture

We constructed a simple feedforward network that learned to map the orthography of words onto stress position (see Figure 1). The orthographic input layer was composed of 14 letter slots, which was sufficient to encode all the words in the word and nonword corpora. If a letter appeared in the slot then one of 26 units was active to represent the letter. The input layer was fully connected to a layer of 100 hidden units, which in turn was fully connected to one output stress unit. Words in the input were left aligned.

Insert Figure 1 around here

2.2. Training and testing

The language input to the model was all 25017 English disyllabic words extracted from the CELEX lexical database (Baayen et al., 1993), with abbreviations and acronyms excluded. All words were of length 14 letters or less. Words were presented at the input layer left aligned, such that the first letter of each word always occupied the first letter slot. Stress position was taken from CELEX's phonological transcription of the words, which was based on Roach and Harman's dictionary (1997). For first-syllable stress, the stress unit's target activity was 0, for second syllable stress it was 1. 90% of the words were randomly selected for training the models and the remaining 10% were reserved for testing. The learning rate was .005 and the models were trained for 5 million presentations of words, selected according to their log-compressed frequency. Each simulation was run from up to

20 times, with different randomised starting weights and different divisions between the 90% training and 10% testing CELEX corpora. The model was judged to have assigned first syllable stress if activation of the output unit was less than .5, and second syllable stress for activity greater than .5.

The model was also tested on two lists of nonwords. The first was the set of 210 disyllabic nonwords from Rastle and Coltheart (2000), with 115 to which participants responded with a majority decision of first syllable stress, and 95 with majority decision of second syllable stress. The second list was from Kelly's (2004) study of the influence of onset clusters on stress assignment. 94 disyllabic nonwords were generated by varying the complexity of the onset of the word (C (ponveen) vs. CC (plonveen) clusters, see Table 1 for more details). We again used participant-majority judgments on the stress position for the Kelly (2004) nonwords, as a baseline for the evaluation of the performance of both the R&C and the connectionist models.

Insert Table 1 around here

3. Results

The mean percent correct stress assignment for the connectionist model on the 90% of words used for training and the 10% reserved for testing is shown in Figure 2, along with the R&C model's performance on the CELEX disyllabic words. Figure 2 also presents mean d' for the model's discrimination between first and second syllable stress assignment[1].

Insert Figure 2 around here

On the trained word set, the connectionist model learned to categorise most of the disyllabic words, though note that perfect performance on the trained words was not anticipated. This is because there were instances of ambiguously stressed words, which were indistinguishable from the orthography (such as the noun *'record* and the verb *re'cord*, see below for further discussion). In order to compare the connectionist model to the R&C model, however, we tested the model on the 10% of words that had been reserved during the training phase. As the connectionist model is specifically trained on certain words, comparisons between this model and the R&C model for the whole lexicon would be biased in favour of the connectionist model. The connectionist model performed marginally better than DRC on the reserved 10% of the words from the CELEX database. The d' values for classification of each model for each testing set were determined, and compared in a one-way t-test, indicating better classification by the connectionist model, $t(19) = 21.46$, $p < 0.001$. Interestingly, the model exhibited the same pattern as the R&C model of higher accuracy for first syllable words than second syllable words. The effect is mainly due to the higher frequency of first syllable words in the corpus - 83% first syllable versus 17% second syllable words.

Item-based analyses of the connectionist model's performance provides insight into the types of cues that the model is responding to, and revealed that the network was producing three types of errors:

1) approximately 20% of errors were for words which shared orthography but had different grammatical class, for example, *'contrast* as a noun versus *con'trast* as a verb. The orthography of these words was not distinguished in the training set, and so the model is unable to resolve the stress assignment of these homographs.

2) overgeneralization errors appeared with the words which contained the beginning or the ending which were the usual cues for the opposite stress assignment. For example, ab- at the beginning of words usually occurs in second syllable stress words, *a'bout, a'bove, a'broad*. In CELEX, 60 *ab-* words with summed frequency 51897 have second syllable stress, and 21 words with summed frequency 7708 have first syllable stress, and the model incorrectly generalises to assign second syllable stress to *'abject*.

   85% of errors on words beginning with ab- were overgeneralisations of first syllable stress for second syllable target words.

3) A third error type was for word beginnings that were balanced in terms of stress position across the corpus. In this case the connectionist model produced output unit activity close to .5, and so stress assignment varied. For example, in the corpus 101 words (summed frequency 13008) starting with *con-* had first syllable stress and 169 (summed frequency 44292) had second syllable stress. The model made an error in assigning stress to 38 of the first syllable stress and 44 of the second syllable stress words in the corpus. For words beginning *de-*, 156 (summed frequency 11942) had first syllable stress and 408 (summed frequency 31422) had second syllable stress. The model wrongly assigned stress to 38 of the first syllable stress words and 20 of the second syllable stress words.

For the R&C nonwords, the connectionist model's performance is shown alongside that of the R&C model's predictions in Figure 3, as for the reserved word testing set, we computed d' for each run of the model's classifications.

Insert Figure 3 around here

The R&C model performed better than the connectionist model on the R&C nonwords, with significantly higher d' in the one-way t-test, $t(19) = -31.11$, $p < 0.001$. The difference in performance was mainly due to the connectionist model's worse performance on the second syllable stress nonwords, as it may have been over-influenced by the predominance of first syllable stress words in the corpus. In order to test this hypothesis we trained the model again using the same procedure, but now using the set of all polysyllabic words from CELEX that had stress on the first or second syllable and were of length 14 or less (99.9% of the database). This sample contained 51948 words, which represented 89.6% of the polysyllabic word types in the database (the remaining 10.4% had stress outside the first two syllables). The distribution of the stress patterns was now shifted in this set with 68.6% first syllable and 31.4% second syllable words. However, performance on the R&C nonwords did not indicate a significant change in accuracy of classification as measured by d', $t(38) = 0.28$, $p > 0.05$, which indicated that stress assignment is not only based on general characteristics of stress position in the corpus, but is rather due to complex interactions between letter clusters and stress position. A large proportion of the second syllable nonwords that were problematic for the connectionist model contained beginnings that were ambiguous in the corpus with respect to stress position. For example, the nonword *nockate* was assigned second syllable stress by participants and the R&C model. In the CELEX corpus, for the beginning *no-* there were 104 first syllable words (summed frequency 22077) and only 15 (summed frequency 285) second syllable words.

For Kelly's (2004) nonwords, Table 1 reports the proportions of items that contain prefixes or suffixes or both based on Fudge's (1984) list of morphemes. Most of the nonwords did not contain any prefixes or suffixes, which means that the R&C model would assign first syllable stress as a default to the majority of the nonwords.

The one-way t-test showed that the connectionist model resulted in significantly better classification than the R&C model on the Kelly (2004) nonwords, $t(19) = 6.51$, $p <$

0.001. This implies that the connectionist model learned regularities between frequently occurring letter sequences that were regular with regard to stress position, and goes beyond the classifications of stress position based only on affixes. Furthermore, interactions between a range of cues are potentially available to the connectionist model. In Kelly's (2004) experiment, English speakers assigned trochaic stress to disyllabic pseudowords more often when they began with two consonants rather than one. Similarly, for the connectionist model, the mean proportion of trochaic assignments was .94 for two consonant onset pseudowords but only .73 for single consonant onset pseudowords, $t(19) = 9.46$, $p < 0.001$. However, both the connectionist model and the R&C model were less successful in simulating the stress assignment for the Kelly (2004) nonwords than for classifying the R&C nonwords. The results are shown in Figure 4.

Insert Figure 4 around here

4. Discussion

The present study provided a demonstration that stress assignment for words and nonwords can be accomplished with accuracy in a connectionist model that learns to map orthography onto stress position for disyllabic words in English. This implies that stress-assignment does not need to be governed by a set of predefined linguistic rules, but rather it can emerge to a large extent through a combination of different cues present in the orthographic input. The sources of information implemented in the R&C model algorithm, for instance, are potentially available to a connectionist approach where regularities of mapping from single letters, bigrams, trigrams, and so on, onto stress position can be discovered from the lexicon. So, frequently occurring morphemes that either do or do not carry stress are of this type, yet they do not have to be specified or listed in advance. Furthermore, frequently-occurring patterns that are not covered by Fudge's (1984) list can also be used by a connectionist model of stress assignment. This has the consequence that the overlap in information used by the R&C rule-based approach and a statistical approach implemented in our connectionist model of stress assignment may be substantial.

Yet, there are clearly sources of information utilised by the R&C model that are not discovered by the particular implementation of the connectionist model we have presented. In particular, the stress assignment for second syllable nonwords in the R&C nonword set was not as accurate for the connectionist model. This is in part due to the difficulty that the connectionist model had in overcoming the general bias in the corpus to first syllable stress assignment, and also due in part to difficulties in identifying frequently co-occurring beginnings and endings of words, which had an ambiguous relationship to the stress pattern. In addition, it may be that the advantage of the R&C model for the first set of nonwords is due to phonological information which indicates stress position, as in the presence of schwa. Indeed, phonology may have a special role in stress assignment as children have substantial practice at listening and speaking the language prior to learning to read.

On the other hand, there are advantages in the statistical approach to stress assignment over rule-based approaches in that multiple, interacting sources of information are available to the connectionist model, for instance, sensitivity to consonant clusters, as tested in Kelly's (2004) nonwords. The assignment of stress, as indicated by numerous behavioural studies, is influenced by many different sources of information – including grammatical category, orthographic, phonological, and utterance position (Arciuli & Cupples, 2006, in press; Gerken, 1996; Gerken, Jusczyk, & Mandel, 1994; Kelly, 2004). Placing stress in the lexicon, or applying stress via rule-based accounts precipitates against the combination and interaction of these information sources. Such multiple cue accounts have been shown to result in more accurate classification in speech segmentation (Onnis, Monaghan, Chater, & Richmond, 2005) and grammatical categorisation tasks (Monaghan, Christiansen, & Chater, 2007), and as they are statistically related to stress assignment in reading, are likely to be used in concert for this purpose.

A criticism of connectionist models is that determining precisely how they are solving the task is difficult to ascertain. However, the principal aims of the connectionist model was to demonstrate first that probabilistic information can be discovered by a computational system for mapping orthography to stress position for disyllabic words, and to indicate that such a mapping is not problematic for a computational system based on the principles of the connectionist tradition. The current model has accomplished both these goals. Yet, we have also provided several ways in which the model's processing performance *can* be accessed. Certain of the model's errors were shown to relate to ambiguity in the orthographic input – and these errors can be resolved by including information about grammatical category, providing insight into how multiple cues may be coordinated in stress assignment. Other errors of the model were due to the assignment of stress for particular lexical items being contrary to the general stress assignment for words with similar beginnings and endings. In addition, the model's sensitivity to consonant clusters at the beginning of the word indicated that such phonotactic cues to stress assignment are available to the model. Yet, demonstrating the sensitivity of the model to certain interacting cues is just a first step in this research program, and raises several intriguing hypotheses about the types of cues to which readers may be sensitive – the computational models – the R&C model and the connectionist model – indicate different orthographic properties that are contributing to stress assignment for readers, and such predictions will assist in determining which cues are not only useful but also used in the reading process.

A further limitation of the current connectionist model is that it only applies to stress assignment, and does not also provide a mapping between orthography and phonology for polysyllabic words. We have indicated that stress assignment is not problematic for connectionist models extended to polysyllabic reading. Yet, the mapping from orthography to phonology is a substantial additional requirement in such models. Rastle and Coltheart's (2000) model does provide rules for such a mapping, which are included additionally in the process of stress assignment. For a connectionist model, the pronunciation has to be learned from co-occurrences of letters, and learning that certain letters map onto phonemes in particular syllables will be as problematic for connectionist models as for the CDP+ model, for similar reasons described above. Yet, some headway has been achieved in providing connectionist models that learn to map orthography to phonology. Pagliuca and Monaghan (submitted) report a model of reading Italian that learns to read words up to three syllables in length. Italian is more regular in mapping of letters onto phonemes than English, and it also has fewer consonant clusters. Yet, still the model has to solve the task of mapping particular letters onto phonemes in each syllable, which it achieves with a high degree of success. Though a substantially harder task in a language with less letter-phoneme regularity and more complex syllabic structure, similar principles may well go a long way to solving the orthography to phonology mapping for English. Indeed, also including phonology in the architecture may assist the model further in correct stress assignment, as reduced vowels are not stressed – information that proved to be useful for the R&C model.

Thus, an additional benefit of statistical approaches to stress assignment is that, akin with previous computational models of monosyllabic word reading, no particular architectural assumptions about how and when stress is assigned in the reading process are made. In particular, it means that different sources of information can be distinguished in terms of their role in stress assignment, as it is not assumed that phonology precedes stress assignment, as in some branches of the R&C model algorithm. It may be that a phonological code in reading words may proceed in parallel with the generation of stress assignment in reading, and whether better performance is achieved by prior phonological encoding in the connectionist model may assist in answering this question. We do not however lay special claim to the particulars of the architecture of the model we have presented here. Our aim was to indicate the potential value of a system that responds to the statistical properties present in orthography for learning stress assignment during reading. We have indicated that, for a large lexicon of English disyllabic words, the model can learn accurate stress assignment and can also accurately generalise to a subset of words reserved for testing. The performance of the connectionist model on stress assignment for nonwords shows mixed results. The connectionist model does not pick up on all the potential regularities present in the orthography of suffixes and prefixes of English. The model does, however, indicate how multiple cues in orthography (such as consonant clusters) may be exploited by a system that responds to the generalities learned through exposure to the lexicon and without pre-specifying this information in advance. In addition, we anticipate that phonological and grammatical information will interact with orthography as a set of multiple probabilistic cues for stress position in the reading system.

Acknowledgements

This work was supported by the ESRC/ARC Bilateral Research Awards Grant, RES 000-22-1975/LX0775703. We would like to thank Kathy Rastle and an anonymous reviewer for helpful comments on an earlier version of this paper.

References

Ans, B., Carbonnel, S., & Valdois, S. (1998). A connectionist multi-trace memory model of polysyllabic word reading. *Psychological review*, *105*, 678-723.

Arciuli, J., & Cupples, L. (2006). The processing of lexical stress during visual word recognition: Typicality effects and orthographic correlates. *Quarterly Journal of Experimental Psychology*, *59*(5)*,* 920-948.

Arciuli, J., & Cupples, L. (in press). Would you rather 'embert a cudsert' or 'cudsert an embert'? How spelling patterns at the beginning of English disyllables can cue grammatical category. To appear in: Schalley, A., & Khlentzos, D. (Eds.), *Mental States: Language and Cognitive Structure*.

Arciuli, J., & Thompson, J. (2006). Improving the Assignment of Lexical Stress in Text-to-Speech Systems. *Proceedings of the 11th Australian International Conference on Speech Science and Technology*, University of Auckland, New Zealand. December 6-8, pp. 296-300.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993) *The CELEX Lexical Database* (CD-ROM). University of Pennsylvania, Philadelphia, PA: Linguistic Data Consortium.

Chomsky, N. & Halle, M. (1968). *The sound pattern of English*. MIT Press, Cambridge, Massachusetts.

Coltheart, M. (2000). Dual routes from print to speech and dual routes from print to meaning: Some theoretical issues. In Kennedy, A., Radach, R., Pynte, J. & Heller D. (Eds.), *Reading as a Perceptual Process*. Oxford: Elsevier, 2000.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A Dual Route Cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204-256.

Daelemans, W., Gillis, S., & Durieux, G. (1994). The Acquisition of Stress: A Data-oriented Approach. *Computational Linguistics*, *20*, 421-451.

Fudge, E. (1984). *English word-stress*. London: George Allen & Unwin.

Garde, P. (1968). *L'Accent.* Paris: Presses Univenitaires de France.

Gerken, L.A. (1996). Prosody's role in language acquisition and adult parsing. *Journal of Psycholinguistic Research*, *25*, 345-356.

Gerken, L.A., Jusczyk, P.W. & Mandel, D.R. (1994). When prosody fails to cue syntactic structure: Nine-month-olds' sensitivity to phonological vs. syntactic phrases. *Cognition*, *51*, 237-265.

Gupta, P. & Touretzky, D. (1994). Connectionist models and linguistic theory: Investigations of stress systems in language. *Cognitive Science*, *18*(1), 1-50.

Harm, M. W., & Seidenberg, M. S. (1999). Reading acquisition, phonology, and dyslexia: Insights from a connectionist model. *Psychological Review*, *106*, 491-528.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the Meanings of Words in Reading: Cooperative Division of Labor Between Visual and Phonological Processes. *Psychological Review*, *111*, 662-720.

Kelly, M.H. & Bock, J.K. (1988). Stress in time. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 389–403.

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

Kelly, M.H., Morris, J., & Verrekia, L. (1998). Orthographic cues to lexical stress: Effects on naming and lexical decision. *Memory & Cognition*, *26*, 822-832.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A User's Guide (2nd ed.)*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Monaghan, P. & Shillcock, R.C. (in press). Hemispheric dissociation and dyslexia in a computational model of reading. *Brain and Language*.

Monaghan, P., Christiansen, M.H., & Chater, N. (2007). The phonological-distributional coherence hypothesis: Cross-linguistic evidence in language acquisition. *Cognitive Psychology*, *55*, 259-305.

Onnis, L., Monaghan, P., Richmond, K., & Chater, N. (2005). Phonology impacts segmentation in speech processing. *Journal of Memory and Language*, *53*(2), 225-237.

Pagliuca, G. & Monaghan, P. (submitted). DANTE++: A connectionist model for reading polysyllabic words in Italian. Proceedings of the 11th neural computation and psychology workshop. Hackensack, NJ: World Scientific.

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modelling in the development of computational theories: the CDP+ model of reading aloud. *Psychological Review*, *114*, 273-315.

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

Roach, P. & Hartman, J. (1997). *The English pronouncing dictionary*, 15th Edition. Cambridge University Press: Cambridge.

Rogers, T. T., and J. L. McClelland (2004). Semantic Cognition: A Parallel Distributed Processing Approach. Cambridge, MA: MIT Press.

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

Shaywitz, B.A., Lyon, G.R., & Shaywitz, S.E. (2006). The role of functional magnetic resonance imaging in understanding reading and dyslexia. *Developmental Neuropsychology*, *30*, 613-632.

Zevin, J. & Joanisse*,* M. (2000). Stress assignment in nonword reading. *Journal of Cognitive Neuroscience, 41*B, S5.

Footnotes

1 d' (d-prime) is a measure used mainly in signal detection theory as an index of the detectability or discriminability of a signal, given by the difference between the means (or separation between the peaks) of the signal-plus-noise and the noise-only probability distributions, divided by the standard deviation of the noise-only distribution (Macmillan & Creelman, 2005). We used d' measures to avoid problems of high accuracy resulting from models producing default responses to a frequent category, as it takes into account in a single measure the proportion of correct classifications as well as misclassifications.

Figure 1. The architecture of the network.

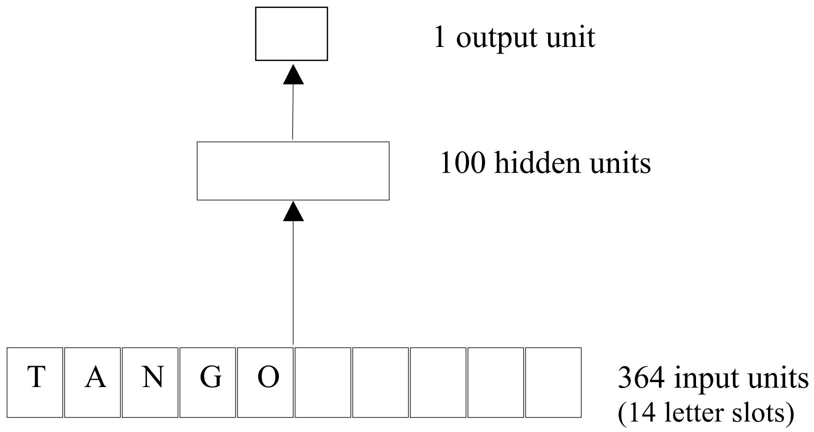

Table 1

Distribution of Different Types of Words in Kelly's (2004) nonwords, Based on Fudge's (1984) Classification of Prefixes and Suffixes, Complexity of the Onset and Stress Placement

		C onset	CC onset
Trochaic	nosufix+noprefix	22.3	40.4
	nosufix+prefix	10.6	0.0
	suffix+noprefix	2.1	6.4
	suffix+prefix	1.1	0.0
Iambic	nosufix+noprefix	8.5	1.1
	nosufix+prefix	1.1	0.0
	suffix+noprefix	4.3	2.1
	suffix+prefix	0.0	0.0

Figure 2. The mean percent correct stress assignment (and 2 S.D.) and d' for the connectionist model and R&C model for testing and training on CELEX set of disyllabic English words.

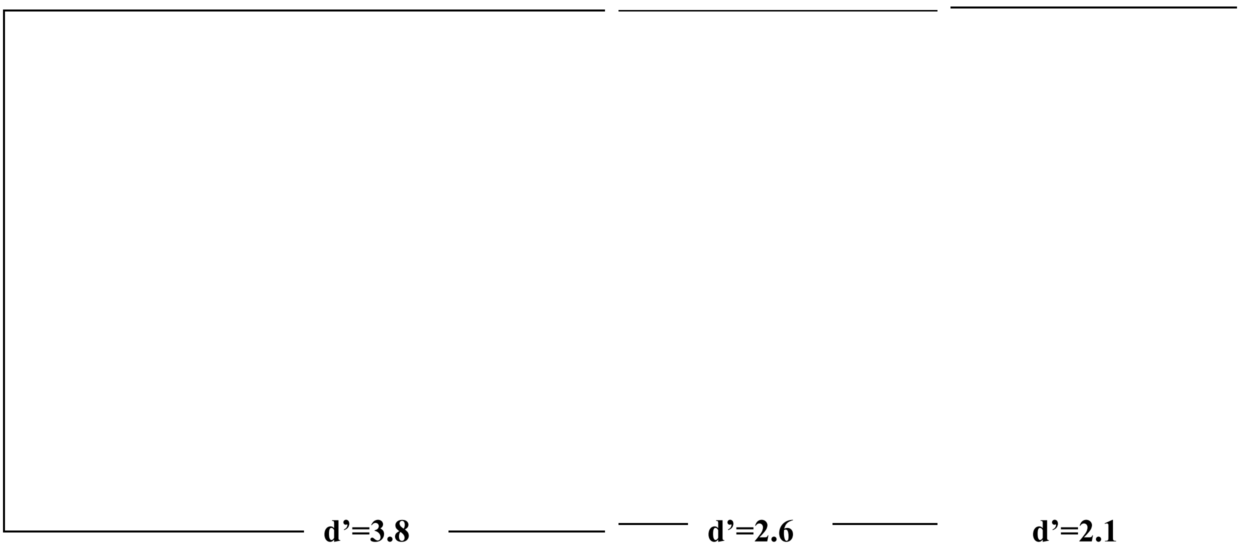

Figure 3. The mean percent correct stress assignment (and 2 S.D.) and d' for the connectionist model and R&C model for testing on R&C nonwords.

d'=1.2 **d'=2.1**

Figure 4. The mean percent correct stress assignment (and 2 S.D.) and d' for the connectionist model and R&C model on Kelly (2004) nonwords.

d'=1.0 **d'=0.6**