Proceedings of the Annual Meeting of the Cognitive Science Society

A Model of the Sound-Spelling Mapping in English and its Role in Word and Nonword Spelling

https://escholarship.org/uc/item/5r371758

Proceedings of the Annual Meeting of the Cognitive Science Society, 20(0)

Houghton, George
Zorzi, Marco

1998

A Model of the Sound-Spelling Mapping in English and its Role in Word and Nonword Spelling

George Houghton (g.houghton@ucl.ac.uk)
Institute of Cognitive Neuroscience, University College London
Gower St., London WC1E 6BT (UK)

Marco Zorzi (zorzi@univ.trieste.it)
Department of Psychology, University of Trieste
via dell'Universita` 7, 34123 Trieste (Italy)

Abstract

A model of the productive sound-spelling mapping in English is described, based on previous work on the analogous problem for reading (Zorzi, Houghton & Butterworth, 1998a, 1998b). It is found that a two-layer network can robustly extract this mapping from a representative corpus of English monosyllabic sound-spelling pairs, but that good performance requires the use of graphemic representations. Performance of the model is discussed for both words and nonwords, direct comparison being made with the spelling of surface dysgraphic MP (Behrmann & Bub, 1992). The model shows appropriate contextual effects on spelling and exactly reproduces many of the subject's spellings. Effects of sound-spelling consistency are examined, and results arising from the interaction of this system with a lexical spelling system are compared with normal subject data.

Introduction

The study of reading has recently been much enlivened by the development of explicit computational models of reading aloud and word recognition (Coltheart, Curtis, Atkins & Haller, 1993; Grainger & Jacobs, 1996; Plaut, McClelland, Seidenberg & Patterson, 1996; Seidenberg & McClelland, 1989; Zorzi, Houghton & Butterworth, 1998a). At the same time, progress has been made in the related problem of spelling, both from memory (Glasspool & Houghton, submitted; Houghton, Glasspool & Shallice, 1994) and by sound-spelling conversion (Glasspool, Houghton & Shallice, 1995; Olson & Caramazza, 1994; Shallice, Glasspool & Houghton, 1995).

In this paper, we look at results produced by a particular model of the sound-spelling mapping for English monosyllables. The basic architecture of the model is based on the reading model of Zorzi, Houghton and Butterworth (1998a, 1998b). The particular aspects of spelling which are examined are (1) the nature of orthographic output representations and their role in facilitating the learning of the productive sound-spelling mapping for English, (2) the minimal computational requirements for an adequate learning of this mapping, with the emphasis on generalization to nonword spelling, (3) direct comparison of model output with nonword spelling in surface dysgraphia, and (4) the model's treatment of words and its interaction with the lexical spelling mechanisms.

Division of labor in reading and spelling

Skilled problem solving generally involves the interaction of at least two kinds of knowledge, which we will refer to as "generative" and "case-specific" knowledge. The first type, generative knowledge, exploits regularities found to apply across specific examples of the problem class, and, in principle, may be applied to any problem instance, in particular novel examples. Case-specific knowledge only applies to particular instances and is based on memory for previously experienced examples. This latter form of recognition-based problem solving may have the advantage of relative efficiency, and in particular can be applied to non-regular or exceptional cases when the problem domain is only "quasi-regular".

With respect to spelling (and reading) this general perspective translates into the idea that there will be a distinction between "rote" performance, based on memory for the spelling of a word, and a more productive capacity for mapping from the constituent sounds of a word to constituent letters (and vice versa for reading). The problems of reading and spelling English are good examples of a quasi-regularity. Thus English spelling to sound relationships are basically systematic but with many inconsistencies and specific exception words. However, skilled spellers can also easily spell new or non- words, and generally do so in a way which conforms to the dominant regularities which exist amongst the words they know. The idea that the regularities employed in spelling and reading nonwords (the generative route) may be separately represented from knowledge of individual words (case-specific route) has a long history (see e.g., Baron & Strawson, 1976; Coltheart, 1978). With respect to spelling, specific evidence for distinct lexical and generative routes has come from neuropsychological studies of dissociations (Beauvois & Derouesne, 1981; Shallice, 1981).

However, beginning with the work of Rumelhart and McClelland (1986) on learning the past tense of English verbs, studies with neural networks have led to the psychological validity of this distinction being questioned, at least for some domains. It is argued that a neural network with a single, homogenous route from input to output can handle both the regular and irregular cases and still be able to generalize the regularities to novel cases (e.g., Plunkett & Marchman, 1993; Plaut et al., 1996; Rumelhart & McClelland, 1986; Seidenberg & McClelland, 1989). In the case of reading, however, it has become clear that no single route model can account for both experimental and neuropsychological data (Plaut et al., 1996; Zorzi et al., 1998a).

Connectionist modeling must not necessarily commit to a single route approach (see, e.g., Jacobs, Jordan & Barto, 1991). In fact, Zorzi et al. (1998a, 1998b) describe a connectionist model of reading where a dual-route processing system emerges from the interaction of task demands and initial network architecture in the course of reading acquisition (for brevity we will henceforth refer to this as the ZHB model). In this model, the distinction between generative (phonological assembly) and case-specific (lexical retrieval) knowledge is realized in the form of the connectivity (either direct or mediated) between orthographic input and phonological output patterns. In particular, it was found that, when spelling and sound are directly connected in a two-layer network (i.e., interactions are <u>not</u> mediated by hidden units), then this direct pathway learns the most reliable spelling-sound correspondences, without forming representations of individual training items (i.e., lexical representations these will form in the mediated pathway, i.e., one containing a layer of hidden units). This is so even though the training set contains numerous inconsistent and exception pairs. This two layer model, trained with the Delta rule, was thus found to implement a phonological assembly mechanism. It generates correct pronunciations of regular words, and tends to regularize irregular words (with a pattern similar to that of a surface dyslexic patient). On nonword reading, the model closely matches data from normal subjects. Further studies (Zorzi et al., 1998b) showed that a multilayer network with no direct connectivity between input and output, when given the same training, performs much less well as a phonological assembly mechanism, tending to simply memorize its training set (i.e., act as a lexical route). Zorzi et al. conclude that it is not necessary to specify *a priori* the functional roles of the different routes in reading but that they can emerge as the optimal use of specific computational resources, viz. direct vs. mediated connectivity.

Mapping sound onto spelling

Following the ZHB model results, Glasspool and colleagues (Glasspool et al., 1995; Shallice et al., 1995) looked at the capabilities of two-layer networks on the sound to spelling mapping. In this model input phonemes arrive serially using a "sliding window" technique (cf. Olson & Caramazza, 1994; Sejnowski & Rosenberg, 1988). During learning, the spelling of the word is also presented serially as a time-varying pattern of activation over a *single set* of nodes representing graphemes (letter groups corresponding to phonemes, e.g., TH, EA, SH, etc.). Grapheme position was not explicitly represented. After 60 epochs of training on a set of 3957 spelling-sound pairs the model achieved about 80% acceptable (i.e., phonologically plausible) spellings. The results of Glasspool et al. are encouraging; however a lower rate of implausible spellings is desirable. In the model reported here, we dispense with the serial aspects of the Glasspool et al. model and describe results from an "inverse" version of the ZHB model.

Description of the model

The ZHB model of phonological assembly uses a slot based representation for input and output, with a distinction between onset and rime. Input units represent the presence of individual letters in particular positions in the orthographic onset and rime. Only representations of individual letters were used (i.e., no complex graphemes). Letters occurring before the vowel letter fill successive onset slots (up to a maximum of 3). Letters following the first vowel letter fill successive rime slots. Output phonology is represented in the same way. In the first version of this model we simply inverted the ZHB assembly model and trained it from sound to spelling. We found for this model that the level of acceptable spelling produced was rather low (around 70%) and much of the problem was due to the model producing blends of alternative "graphemes" for a given sound, e.g., for the sound */f/*, blending F and PH, to produce FH (note that PH is not represented as a single unit). This suggested that grapheme units should be used in the output representation. Interestingly, Barry and Seymour (1988) present data from priming studies which support the use of graphemes in spelling. We therefore developed a second model with the same onset-rime structure in the orthography but with the addition of nodes for complex graphemes.

In this model the output representation requires 7 groups of grapheme units, consisting of three successive onset groups, and 4 rime groups starting with the orthographic vowel. The graphemes used (in addition to the standard single letters) are shown in table 1.

Table 1 - Graphemes used in orthographic output representation.

Onset	CH KN PH KN QU SH TH WH WR
Vowel	AI AR AU AW AY EA EE EI ER EW EY IE IR OA OE OI OO OU OR OW OY UA UE UI UR UY YE
Coda	CH CK DD DGE FF GH GHT GUE LL MB NG PH QUE SH SS TCH TH TT ZZ

Most onset graphemes are initial and hence only occur in the first orthographic position. This scheme produces output representations such as: *the* = TH - - E - - -; *church* = CH UR CH - -; *strength* = S T R E NG TH -; where a hyphen indicates an empty slot (no graphemes activated). The architecture of the model is depicted in Figure 1.

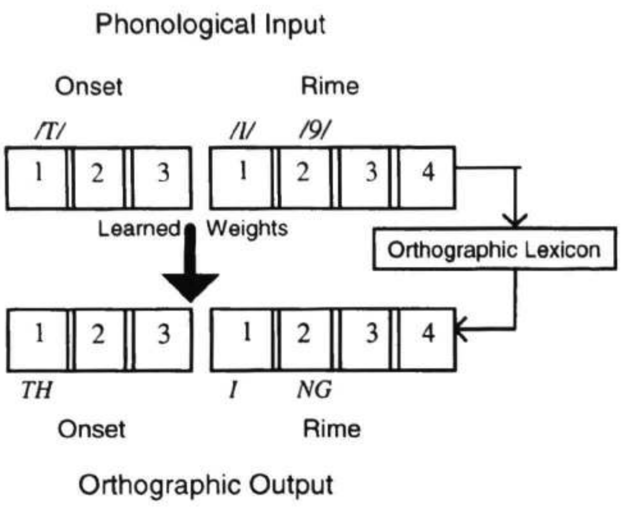

Figure 1. Structure of the model. The direct sound-spelling network maps from a (slot-based) onset/rime representation of phonology to an onset/rime representation of spelling via a single set of learned weights. The output representation uses graphemes and the figure illustrates the representations in the mapping from */T I 9/* to THING (see Table 1 for grapheme set, Table 2 for phoneme symbols). Input and output may also be linked by a lexical route.

The network was trained on a set of 2500 monosyllabic sound-spelling pairs using the Delta rule error correcting algorithm (Widrow & Hoff, 1960). For each spelling-sound pair in the training set, an appropriate phonological input is established, setting each activated phoneme node to a value of 1. Activations propagate to the output layer, using the dot product net input rule to calculate the inputs to each grapheme unit. Connection weights are all initialized to zero, and units have no bias term. Phonemic activations are a sigmoidal function f of their net input, bounding phoneme activations in the range [0,1], and with $f(0) = 0$ (no input, no output). This output activation is compared with the target activation (nodes that should be on have a target activation of 1, nodes that should be off a target of 0). The error for each grapheme unit is the difference between the target and actual activations. Where errors occur, weights to the offending units are changed according to the delta rule, i.e.,

$$\Delta w_{ij} = \lambda a_i \left(t_j - a_j\right)$$

where w_{ij} is the weight from input unit u_i to output unit u_j, a_i and a_j are activations of the input and output units respectively, t_j is the target activation for output unit j, and λ is a learning rate parameter. Training stopped when the mean error reached asymptote. This was typically after around 20-30 epochs of training. Following this the model's performance on words and nonwords was tested. Testing consists of presenting the phonemic representation at the input and then reading out the spelling by selecting the most active grapheme above a threshold at each output grapheme position (this selection can be achieved using lateral inhibition between nodes within a "slot", but we ignore this complication here; see Zorzi et al., 1998a, for the description of a similar competitive output system). The threshold was set at 0.2.

Results

Basic performance After learning reached asymptote the model's spelling was tested on 500 randomly selected words from the training set. Of these, about 96% were judged to be phonologically plausible spellings, in the sense that the pronunciation could be reliably retrieved from the model's spelling. This was clearly a great improvement over the inverse ZHB model, showing that the use of complex grapheme representations is highly functional for this basic architecture. This supports the Barry and Seymour (1988) priming results showing the use of graphemes by adults spelling nonwords to dictation. The model also performs better than the Glasspool et al. (1995) model in terms of avoiding implausible spellings. We take this improvement to be due to the explicit provision of orthographic syllable structure in the output representation.

For regular words (*plank, match, bill, etc.*) the model's spelling was correct, while irregular and inconsistent words were regularized, e.g. *soap - sope; sauce sorse; spear - speer.* This aspect of the model's performance is investigated more systematically below.

Many of the errors involved adding an unlikely final "E" on words with short vowels and a final */s/* or */z/*, so for instance, */s w I s/* was spelt *swisse*. Although these were scored as errors, we suspect many would be pronounced correctly.

Nonword spelling Of particular interest is the model's nonword spelling (generalization). As a simple test of the model's nonword spelling, Table 2 provides a comparison of the model's spelling of a set of 58 nonwords with those produced by MP (Behrmann & Bub, 1992), a surface dyslexic subject who showed near perfect performance on nonword spelling and reading.

The model makes one clear error, */g e d/*→ *gd*. In this case no vowel grapheme has been activated above the 0.2 threshold. However, the most active subthreshold vowel letter is *e*. Apart from this the model shows the tendency mentioned above to overuse a final *e*, e.g., in *noode, degse, draite*, although this does not generally lead to unacceptable spellings. In general this is due to the fact that "silent e" spellings, as in *late, pole, mule* etc., are not represented in the model as distinct (orthogonal) output options and hence cannot compete with alternative spelling as a single coherent "group". So for instance, in the case of *draite*, there is competition between the alternative spellings *drait* and *drate*. However, the final *e* does not form a group (or discontinuous grapheme) with the *a* of *drate* with the result that when the *a* loses the competition in the vowel position to the grapheme *ai*, the final *e* is not affected and is still produced. To the extent that people do not produce such spellings, it suggests that the *vowel + e* option should be represented as a single unit which can be selected or rejected.

Table 2: Comparison of model's nonword spelling with that of surface dyslexic subject MP (Behrmann & Bub, 1992, appendix 2).
Key. Vowels: /&/ as in cAt, /e/ as in bEt, /I/ as in hIt, /0/ as in hOt, /V/ as in hUt, /I/ as in bEAt, /u/ as in bOOt, /U/ as in pUt, /O/ as in dOOR, /@U/ as in grOve, /aI/ as in fIle, /eI/ as in wAIt, /3/ as in bURn, /I@/ as in chEER.
Consonants: most have standard values, e.g., /d/ as in Door. Also /S/ as in Shed, /9/ as in siNG, /T/ as in Thin. (.) = subthreshold response; * denotes response marked as an error by Behrmann and Bub (1992).

Word	MP	Model	Word	MP	Model
/b I l/	beel	beel	/b l @U m/	blome	bloame
/s e k/	sek	seck	/k O m/	caum	corm
/m e l/	mell	mell	/tS & m/	cham	cham
/s p & l/	spal	spall	/h i n/	heen	heen
/k e d/	ked	ked	/d i tS/	deetch	deach
/f l eI t/	flate	flate	/d r eI t/	drate	draite
/n e l d/	neld	neld	/d 0 l d/	dold	dold
/k @U b/	cobe	cobe	/f l & m/	flam	flam
/b O s/	bauce	borse	/l & t s/	lats	latse
/d I m p/	dimp	dimp	/s V f/	suff	suff
/w @U l/	wole	wole	/w V S/	wush	wush
/l i m/	leem	leam	/d r i s/	dreece	driese
/s eI t/	sate	sate	/s k 3 l/	skirl	skirl
/k 0 dZ/	cudge*	codge	/n i tS/	neech	neach
/f r I m/	frim	frim	/s w i m/	sweam	sweam
/n I l t/	nilt	nilt	/m u p/	moop	moope
/t r I s t/	trist	trist	/t eI z/	taze	tase
/f l i p/	fleep	fleep	/h oI l/	hoil	hoil
/b l i m/	bleam	bleam	/b r i s/	brease	brease
/g e d/	ged	g(e)d	/w i tS/	weech	weach
/r I l t/	rilt	rilt	/p l i k/	pleak	pleak
/r i s/	reese	rease	/h oI s/	hois	hoice
/m aI s/	mise	mice	/k eI l/	kail	cale
/f e n t/	fent	fent	/d e g s/	degs	degse
/r e l/	rell	rell	/s t 0 l d/	stold	stold
/p l 0 k/	pluck*	plock	/s k I m/	skim	skim
/r 0 g/	rog	rog	/p e tS/	petch	petch
/b l I k/	blick	blick	/l 0 t/	lot	lot
/f r i tS/	freech	freach	/n u d/	nood	noode

Apart from this, the model clearly performs well, and in particular it produces identical spellings to MP in those cases where the sound to spelling mapping is highly consistent, e.g., *dimp, wush, fent, cham,* etc. Where there is more inconsistency, the model may produce an alternative acceptable spelling, e.g., *freech* vs. *freach, caum* vs. *corm, kail* vs. *cale.* In many of these cases, inspection of the activations of the grapheme layer prior to output competition show that the spellings produced by MP have been activated but have not won the output competition. For instance, in spelling */f r i tS/* the grapheme *ee* has an initial activation of 0.3, but is suppressed by the more active *ea* grapheme. Note also that the model exhibits an appropriate degree of context sensitivity in performing the mapping. For instance MP spells an initial /k/ as *k* in *ked*, but as *c* in *cobe, caum, cudge*; she spells final /k/ as *k* in *pleak*, but as *ck* in *blick*. The model produces exactly the same results.

Consistency effects in word spelling Effects of the spelling-sound consistency of particular words on such measures as RTs have been important for models of reading aloud. Unfortunately, equivalent data regarding possible effects of sound-spelling consistency do not appear to exist for writing (though see Glover & Brown, 1994, for some results on oral spelling). However, we believe this will prove important in further studies of spelling, for instance of errors and problems in learning to spell. In addition, it may turn out to prove important in modeling subtle aspects of reading, as Stone, Vanhoy and Van Orden (1997) have produced evidence that *sound-spelling* (as opposed to *spelling-sound*) consistency can effect visual lexical decision.

To make a more systematic comparison between the model's spelling of consistent and inconsistent words (with consistency defined with respect to the sound-spelling, rather than spelling-sound mapping) we used word lists from the study by Stone et al. (1997)

mentioned above. Stone et al. produced list of words based specifically on their sound-spelling consistency (also see Ziegler, Stone & Jacobs, 1997). Some words which have a "consistent" pronunciation, given their spelling (e.g., "bean"), may be inconsistent in terms of the spelling, given the pronunciation (i.e., the syllable /b i n/ might be spelt *bean*, *been*, or even *bene*). Words such as *plank*, *march*, *loft*, however, are considered sound-spelling consistent, as there is (putatively) no other plausible way to spell them. If the model has extracted the basic regularities of the sound-spelling mapping for English, it should show a clear difference between these sets of words, spelling the consistent words the way they are actually spelt, but producing alternative spellings for more of the inconsistent words. The results are shown in Figure 2.

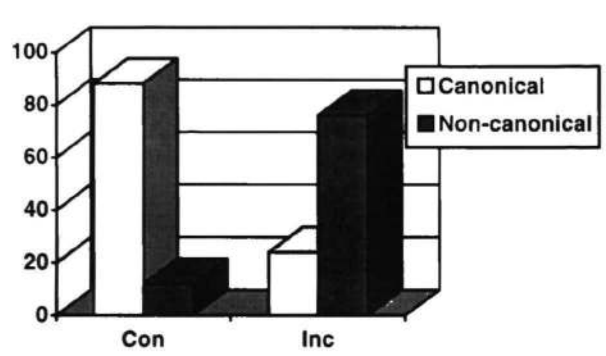

Figure 2: Percentages of canonical (i.e., correct) and non-canonical (i.e., phonologically plausible) spellings produced by the model on sound-spelling consistent and inconsistent words (from Stone et al., 1997, experiment 1).

Of the 25 Stone et al. (1997) consistent words the model spelled 22 of them canonically (88%). The three "errors" were *belch* spelt *bellch*, *fluke* spelled *flooke*, and *loaf* spelled *loafe*. The last 2 show the final "e" problem alluded to earlier, while the first produces a doubled *l* in an (orthographically) inappropriate position. However the spellings remain phonologically acceptable. Of the 25 inconsistent words, the model only spelt 6 canonically (24%) while all 25 spellings were phonologically acceptable. The model thus shows clear sensitivity to the dominant sound-spelling regularities of English when spelling words on which it has been trained. If the word is an inconsistent or exception word this could in principle act as a source of interference with respect to lexical (case-specific) spelling. This raises the issue of the interaction with lexical spelling.

Interaction with lexical spelling The two-layer model (which will refer to here as the "assembly route") regularizes exception and inconsistent words, even if it has been repeatedly trained on them, e.g., *soap* is spelled *sope*. A more general model must therefore be able to represent learned spelling of specific words, which we argue (Zorzi et al., 1998a, 1998b) is best done by a separate set of units (e.g., a lexicon) laying between input and output. We have implemented such a route as a localist, lexical network, individual units being accessed from phonology and sending activation to their constituent graphemes in the same competitive output layer that the assembly model uses. The model can generate a spelling when input arrives from both routes in parallel. For regular words, the inputs are generally the same and simply reinforce each other (though see the example of *belch* above) . For inconsistent and exception words there can be varying degrees of "interference" due to the assembly route trying to regularize some part of the spelling (most often the vowel). However, the lexical spelling will dominate if its input is sufficiently strong. This is because, when the assembly route is "uncertain" regarding a spelling (due to inconsistency in the mapping in the training set), it produces a set of competing candidates at the point of uncertainty. Each of these candidates achieves a lower activation than letters for which there is no uncertainty. In addition, one of the candidates will often be the lexically correct spelling, which supports the lexical input. This permits the lexical input to inhibit the regularizing effect of the assembly procedure. However, the conflicting input still causes interference in the competition process, and time taken for the network to produce an unambiguous spelling can vary as a result. The model thus predicts effects of regularity in spelling production if, for instance, time to initiate a spelling depends on having a "clean" orthographic spelling plan activated before spelling begins (as is required by models of the serial operation of the graphemic output buffer, e.g., Houghton et al., 1994).

Glover and Brown (1994) measured both response initiation time and response duration in oral spelling of known words, and found significant effects due to the sound-spelling consistency of individual words, inconsistency tending to increase both response initiation and duration times, in line with the model. Glover and Brown concluded that adult spelling of known words does make use of sound-to-spelling constraints, suggesting that access to the assembly procedure may be automatic.

Conclusions

The above results show that a two layer network with structured input-output representations, including grapheme nodes, will extract the basic regularities of the sound-spelling mapping for English when trained on a representative sample of monosyllabic sound-spelling pairs, including numerous exceptions and inconsistencies. We suggest that this is probably the "minimal architecture" necessary for acceptable performance on this task. Due to its absence of hidden units, the model cannot learn the training set, and in trying to minimize its error on the words it sees, it extracts the most reliable sound-spelling relationships. These generalize well to nonwords, and cause regularization errors on inconsistent and exception words. In short, the network, though trained on words, end ups behaving as a model for assembled spelling, mirroring the results of Zorzi et al. (1998a) for reading. When combined with a lexical spelling route the model produces interference effects in the time taken to generate an unambiguous spelling "plan", when the two routes are in disagreement. This fits with what data is currently available. We conclude that the distinction outlined at the beginning, between generative and case-specific knowledge, is applicable to spelling, and that the generative knowledge can arise as the result of training on actual words in a network not capable of learning them individually.

References

Baron, J., & Strawson, C. (1976). Use of orthographic and word-specific knowledge in reading words aloud. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 386-392.

Barry, C., & Seymour, P. H. K. (1988). Lexical priming and sound-to-spelling contingency effects in nonword spelling. *Quarterly Journal of Experimental Psychology, 40A,* 5-40.

Beauviois, M. F., & Derouesne, J. (1981). Lexical or orthographic agraphia. *Brain, 104,* 21-29.

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

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

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993), Models of reading aloud: Dual-route and parallel-distributed processing approaches. *Psychological Review, 100,* 589-603.

Glasspool, D. W., & Houghton, G. (submitted). Serial order and consonant-vowel structure in a model of disordered spelling.

Glasspool, D. W., Houghton, G., & Shallice, T. (1995). Interactions between knowledge sources in a dual-route connectionist model of spelling. In L. S. Smith & P. J. B. Hancock (Eds.), *Neural Computation and Psychology*. London: Springer-Verlag.

Glover, P., & Brown, G. D. A. (1994). Measuring spelling production times: Methodology and tests of a model. In G. D. A. Brown & N. C. Ellis (Eds.), *Handbook of spelling: Theory, process and intervention*. Wiley: Chichester.

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

Houghton, G., Glasspool, D. W., & Shallice, T. (1994). Spelling and serial recall: Insights from a competitive queueing model. In G. D. A. Brown & N. C. Ellis (Eds.), *Handbook of spelling: Theory, process and intervention*. Wiley: Chichester.

Jacobs, R. A., Jordan, M. I., & Barto, A. G. (1991). Task decomposition through competition in a modular connectionist architecture: The What and Where vision tasks. *Cognitive Science, 15,* 219-250.

Olson, R. K., & Caramazza, A. (1994). Representation and connectionist models: The NETspell experience. In G. D. A. Brown & N. C. Ellis (Eds.), *Handbook of spelling: Theory, process and intervention*. Wiley: Chichester.

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

Plunkett, K., & Marchman, V. A. (1993). From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition, 48,* 21-69.

Rumelhart, D. E., & McClelland, J. L., (1986). On learning the past tenses of English verbs. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 1: Foundations.* 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.

Sejnowski, T. J., & Rosenberg, C. R. (1987). Parallel networks that learn to pronounce English text. *Complex Systems, 1,* 145-168.

Shallice, T. (1981). Phonological agraphia and the lexical route in writing. *Brain, 104,* 413-429.

Shallice, T., Glasspool, D. W., & Houghton, G. (1995). Can neuropsychological evidence inform connectionist modelling? Analyses from spelling. *Language and Cognitive Processes, 10,* 195-225.

Stone, G. O., Vanhoy, M., & Van Orden, G. C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory and Language, 36,* 337-359.

Widrow, G., & Hoff, M.E. (1960). Adaptive switching circuits. *Institute of Radio Engineers, Western Electronic Show and Convention Record, Part 4* (pp. 96-104).

Ziegler, J., 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. *Behavioral Research Methods, Instruments, & Computers, 29,* 600-618.

Zorzi, M., Houghton, G., & Butterworth, B. (1998a). Two routes or one in reading aloud? A connectionist dual-process model. *Journal of Experimental Psychology: Human Perception and Performance, 24.*

Zorzi, M., Houghton, G., & Butterworth, B. (1998b). The development of spelling-sound relationships in a model of phonological reading. *Language and Cognitive Processes, 13*.