In: *Proceedings of CoNLL-2000 and LLL-2000*, pages 13–18, Lisbon, Portugal, 2000.

Pronunciation by Analogy in Normal and Impaired Readers

R.I. Damper
Image, Speech and Intelligent
Systems Research Group,
Department of Electronics and
Computer Science,
University of Southampton,
Southampton SO17 1BJ, UK

Y. Marchand
Cognitive/Clinical Neuroscience Unit,
Department of Psychology,
Dalhousie University,
Halifax, Nova Scotia,
Canada B3H 4J1

Abstract

The prevailing dual-route model of oral reading claims that a lexical route is used for the pronunciation of words and a non-lexical route processes nonwords. Neurological data from patients with acquired dyslexias have been highlighted to support this claim. Models using a lexicon alone are generally held to be incapable of explaining these data. However, by selectively impairing its component parts, it is easily possible to account for phonological and surface dyslexias using a single-route model based upon pronunciation by analogy.

1 Introduction

We have previously developed pronunciation by analogy (PbA) as a model of reading aloud and as a method for automatic phonemisation in text-to-speech synthesis (Sullivan and Damper, 1993; Damper and Eastmond, 1997; Marchand and Damper, 2000). We have also demonstrated (Damper et al., 1999) that the performance of PbA in producing correct pronunciations is vastly superior to manually-written rules and significantly better than the competitor data-driven techniques of back-propagation (Sejnowski and Rosenberg, 1987; McCulloch et al., 1987) and the IB1-IG method based on information gain weighting (Daelemans et al., 1997). Although we cannot claim that PbA is absolutely the best method for pronunciation generation, it must be taken seriously. This view is clearly shared by other workers who are actively developing analogical methods for natural language processing tasks (Pirrelli and Federici, 1995; Jones, 1996; Yvon, 1996; 1997; Bagshaw, 1998; Pirrelli and Yvon, 1999).

Explicit analogy (e.g., Dedina and Nusbaum, 1991; Damper and Eastmond, 1997) retains the lexicon in its entirety, typically as a list of words and their spellings. PbA requires a dictionary in which text and phonemics have been aligned, so that pronunciations corresponding to matching orthographic substrings can be identified. However, many of the necessary computational steps to assemble a pronunciation can be carried out in advance. Thus, in implicit analogy (e.g., Sullivan and Damper, 1993), the lexical database is precompiled to yield a generalised phonological knowledge base which is consulted during pronunciation generation. This done, the (explicit) dictionary can be discarded. Implicit analogy may also attempt to compress the training data, so that some proportion is discarded.

Here, we extend earlier work on modelling pronunciation by normal readers to impaired readers with acquired dyslexias. There are several forms of this: two of the most important are *phonological* and *surface* dyslexia. Cases of phonological dyslexia display good ability to read words (both regular and irregularly-spelled) aloud but poor nonword reading ability (Beauvois and Dérouesné, 1979). In surface dyslexia, however, patients misread irregularly spelled words, which tend to be regularised in their pronunciation (Coltheart et al., 1983). To simulate these dyslexias, we use explicit PbA without compression. The approach is to damage the model and then to observe its ability to replicate the neuropsychological data.

2 Dual and Single Routes to Sound

The nature of the cognitive processes underlying the act of reading aloud has spawned an important and controversial debate in psychology (Humphreys and Evett, 1985; Seidenberg and McClelland, 1989; Coltheart et al., 1993; Plaut et al., 1996). One popular view is that there are two routes from print to sound: a lexical and a nonlexical route (Coltheart, 1978). The former involves access to lexical knowledge for familiar words. The second route concerns the pronunciation of unfamiliar words or pronounceable nonwords and is thought to operate on the basis of a set of abstract spelling-to-sound rules. The strong version of this *dual-route theory* claims that nonwords are segmented at the level of the grapheme and that the pronunciation of nonwords is not influenced by lexical information. A line of evidence generally held to support the model comes from neuropsychological studies of acquired dyslexia. For instance, the patient WB studied by Funnell (1983) is considered a particularly pure case of phonological dyslexia with good reading of words and poor reading of nonwords. This case appears to conform to one of the main predictions of dual-route theory: namely, that neurological damage could selectively impair either processing route, so that a patient may have impaired processing in one system but intact processing in the other.

Nonetheless, the dual-route model has been criticised by different authors (Marcel, 1980; Kay and Marcel, 1981; Glushko, 1981; Shallice et al., 1983; Humphreys and Evett, 1985; McCarthy and Warrington, 1986) who emphasise that nonword pronunciation can be subject to lexical influences and/or argue for "multiple levels" of processing. Two main alternatives have been proposed to counter these objections: a *single-route* framework and a modified dual-route model. The first claims that all print-to-sound conversion is realised through a lexical route. That is, oral reading involves processes that all operate on a lexical database so that words and nonwords can be produced by the same mechanism. However, there has sometimes been a lack of clarity in defining such a single-route mechanism. Often, some kind of analogy process is posited, but its precise form has rarely been specified. Hence, informed commentators have most often been inclined to reform and repair the dual-route theory by relaxing its strong assumptions, either to allow an interaction between routes (Reggia et al., 1988) or to extend the notion of grapheme-phoneme correspondence (Patterson and Morton, 1985) by introducing the notion of *body*—the vowel-plus-terminal-consonant segment of monosyllabic words.

The dual-route model has been more recently questioned by a plethora of single-route computational models based on connectionist principles (Sejnowski and Rosenberg, 1987; Seidenberg and McClelland, 1989; Hinton and Shallice, 1991; Plaut et al., 1996; Bullinaria, 1997; Ans et al., 1998; Zorzi et al., 1998). Less often has analogy been used as the basis of a single-route model. The idea that pseudowords can be pronounced by analogy with lexical words that they resemble has a long history (Baron, 1977; Brooks, 1977; Glushko, 1979). In place of abstract letter-to-sound rules in dual-route models we have specific patterns of correspondence in single-route analogy models.

3 Implementing PbA

In PbA, an unknown word is pronounced by matching substrings of the input to substrings of known, lexical words, hypothesizing a partial pronunciation for each matched substring from the phonological knowledge, and assembling the partial pronunciations. Here, we use an extended and improved version of the system described by Dedina and Nusbaum (1991), which consists of four components: the (uncompressed and previously aligned) lexical database, the matcher which compares the target input to all the words in the database, the pronunciation lattice (a data structure representing possible pronunciations), and the decision function, which selects the 'best' pronunciation among the set of possible ones. The lexicon used is *Webster's Pocket Dictionary*, containing 20,009 words manually aligned by Sejnowski and Rosenberg (1987) for training their NETtalk neural network.

Pattern Matching: An incoming word is matched in turn against all orthographic entries in the lexicon. For a given entry, assume the process starts with the input string and the dictionary entry left-aligned. Substrings sharing contiguous, common letters in matching positions are then found. Information about these matching letter substrings and their corresponding, aligned phoneme substrings in the dictionary entry under consideration is entered into a pronunciation lattice—see below. One of the two strings is then shifted right by one letter and the matching process repeated, until some termination condition is met. This process can be alternatively seen as a matching between substrings of the incoming word, segmented in all possible ways, and the dictionary entries.

Pronunciation Lattice: A node of the lattice represents a matched letter, L_i, at some position, i, in the input. The node is labelled with its position index i and with the phoneme which corresponds to L_i in the matched substring, P_{im} say, for the mth matched substring. An arc is placed from node i to node j if there is a matched substring starting with L_i and ending with L_j. The arc is labelled with the phonemes intermediate between P_{im} and P_{jm} in the phoneme part of the matched substring. Additionally, arcs are labelled with a 'frequency' count which is incremented each time that substring (with that pronunciation) is matched during the pass through the lexicon.

Decision Function: A possible pronunciation for the input corresponds to a complete path through its lattice, from *Start* to *End* nodes, with the output string assembled by concatenating the phoneme labels on the nodes/arcs in the order that they are traversed. (Different paths can, of course, correspond to the same pronunciation.) Scoring of candidate pronunciation uses two heuristics. If there is a unique shortest path, then the corresponding pronunciation is taken as the output. If there are tied shortest paths, then the pronunciation corresponding to the best scoring of these is taken as the output.

This also offers a way of simulating the 'word segmentation' test of Funnell (1983), in which patients have to find words 'hidden' in letter strings. First, there is an initial segmentation in which the input string is segmented in all possible ways, as in 'regular' PbA. Then, if any of these substrings produces a lattice with a length-1 arc, this identifies a lexical word. A single-route connectionist model or abstract rules (or, for that matter, implicit PbA) can not do this without some extension to maintain explicit knowledge of lexical status. Of course, it is possible that a patient can perform the first of these steps, but not the second. This is the difference between our 'unconscious' and 'conscious' segmentations (see below) so-called because, in the latter, the patient is aware that he/she has to find a hidden word.

This particular implementation of PbA does not guarantee an output pronunciation. A complete path through the lattice requires that all nodes on that path (except the first and last) are linked by at least one arc. Clearly, each arc must have a node at either end. Although an arc may have an empty label, a node cannot. Hence, the minimum matching segment length corresponds to a letter bigram. It may be that no matching bigram exists in some cases. So there with be no complete path through the lattice and no pronunciation can be inferred—the 'silence problem'.

Recent Improvements: The implementation used here features several enhancements over the original Dedina and Nusbaum (D&N) system (Marchand and Damper, 2000). First, we use 'full' pattern matching between input letter string and dictionary entries, as opposed to the 'partial' matching of D&N. That is, rather than starting with the two strings left-aligned, we start with the initial letter of the input string $\mathcal{I}$ aligned with the last letter of the dictionary entry $\mathcal{W}$. The matching process terminates not when the two strings are right-aligned, but when the last letter of $\mathcal{I}$ aligns with initial letter of $\mathcal{W}$. Second, multiple (five) heuristics are used to score the candidate pronunciations. Individual scores are then multiplied together to produce a final overall score. The best-scoring pronunciation is then selected as output. Marchand and Damper show that this 'multi-strategy' approach gives statistically significant performance improvements over simpler versions of PbA.

4 Modelling Phonological Dyslexia

By selective impairment of component parts of the PbA model, we have simulated reading data from the two phonological dyslexic patients (WB and FL) studied by Funnell (1983). (The reader is referred to this original source for specifications of the tests and materials.) While the first of these patients has often been cited as a key individual strongly supporting dual-route theory, we believe that FL (who has been largely ignored) is actually a counter-example. FL was unable to supply a sound for single letters (which argues that the abstract rule-based route is impaired) although she could read non-words normally (which contradicts the

Table 1: Reading performance of patient WB and versions of faulty and non-faulty PbA. 'Words (712)' refers to a random sampling of words from the dictionary.

	Tests		Patient	Faulty PbA		Non-faulty
			WB	Version 1	Version 2	PbA
Lexicon	Words (712)		85%	85%	79%	100%
Nonwords	Single letters		0/12	0/12	0/12	10/12
	Nonsense words		0/20	0/20	0/20	17/20
	Pseudo-homophones		1/10	0/10	0/10	7/10
	Isolated suffixes		1/10	1/10	1/10	7/10
	Parkin's test		0/10	0/10	0/10	10/10
Segmentation	Test 1	Parent words	15/15	12/15	7/15	13/15
		Segmented words	30/30	30/30	26/30	30/30
	Test 2	Parent words	14/15	10/15	6/15	15/15
		Segmented words	22/30	24/30	21/30	28/30
	Test 3: Hidden words		15/15	15/15	14/15	15/15

Table 2: Reading performances of patient FL and of faulty and non-faulty PbA.

Tests	Patient FL	Faulty PbA	Non-faulty PbA
Single letters	0/15	0/15	12/15
'Easy' Nonwords	25/34	26/34	31/34
'Difficult' Nonwords	4/6	1/6	3/6

presumption of impaired rules).

For patient WB, two different versions of impaired PbA have been studied. Version 1 supposes that brain damage has induced a partial loss of words from his mental lexicon (the 15% that he can not read aloud) and a total breakdown of his concatenation mechanism. Version 2 supposes that WB's impairment results from injury to one component only; namely, the process of segmentation into all possible substrings is partially damaged. In Version 2, we stress the distinction made earlier between this basic (*unconscious*) segmentation process and Funnell's (*conscious*) segmentation. The unconscious segmentation is that embodied in the PbA pattern matching when WB is asked to read some string. For this specific patient, we postulate damage to the segmentation component such that it can only process substrings of length between 5 and 7. The conscious segmentation is that used when WB is asked to find words within strings and to read them aloud. This process is assumed to be fully operational. For patient FL, a single 'faulty' version of PbA has been developed which postulates a deficiency of (unconscious) segmentation such that substrings of length less than three cannot be used in pattern matching.

Table 1 shows reading accuracy for patient WB for the various tests performed by Funnell together with the corresponding results of simulations of impaired and non-faulty PbA. Table 2 shows the results for patient FL reading aloud and the corresponding simulation of faulty and non-faulty PbA. Evidently, it is possible to reproduce quite well both patients' symptoms. Indeed, with Version 1, we can interpret WB's condition very directly: The concatenation process involved in nonword reading is completely destroyed but the mental lexicon is relatively spared. Because of the absence of some compound words (e.g., *gentle|man*) from the dictionary, the simulations concerning "parent words" (e.g., *father* is the parent of *fat* and *her*) for both Test 1 and Test 2 are not perfect. Version 2 is slightly poorer but still close to the neuropsychological data. For patient FL, the faulty version reproduces her impaired reading of single letters and 'easy' nonwords very well, but does so less well for 'difficult' nonwords.

The simulations also handle the fact that these patients were completely unable to read single letters: the silence problem (see above) can occur for single letters by virtue of the form of the pronunciation lattice used, which requires matching bigrams (at least) at all positions to produce a pronunciation.

5 Modelling Surface Dyslexia

We have also modelled data from patient KT described by McCarthy and Warrington (1986). KT was able to pronounce regular words and nonwords very well but had serious difficulty in reading irregular words, tending to produce regularisation errors. (Again, limitations of space mean we must refer the reader to the original source for details of the reading tests and materials.) Together with WB, these patients have been taken as almost an existence proof of dual routes which can be differentially damaged.

We suppose that KT's impairment results from injury to two components of the PbA model. First, as in phonological dyslexia, we assume that the process of segmentation into all possible substrings is partially damaged. More specifically, we postulate a deficiency concerning the size of the window involved in the pattern matching. Second, it is assumed that one or several (of the total of five) multi-strategies may be degraded.

The simulations were obtained for a model with damage in the third and fourth multi-strategies (see Marchand and Damper, 2000, for detailed specification) and only substrings of length between 2 and 4 can be segmented in pattern matching. Table 3 shows KT's mean reading accuracy over the various tests performed by McCarthy and Warrington together with our corresponding simulation results for impaired and non-faulty PbA. Clearly, it is possible to reproduce quite well the patient's cardinal symptoms: his ability to pronounce regular words much better than irregular ones. The incorrect pronunciations show a clear regularisation effect (not detailed here).

6 Conclusion

Contrary to the claims of dual-route theorists, a single-route PbA model of reading is indeed able to explain both phonological and surface dyslexia, on the basis of selective impairment of its component parts.

References

B. Ans, S. Carbonnel, and S. Valdois. 1998. A connectionist multiple-trace memory model for polysyllabic word reading. *Psychological Review*, 105(4):678–723.

P. C. Bagshaw. 1998. Phonemic transcription by analogy in text-to-speech synthesis: Novel word pronunciation and lexicon compression. *Computer Speech and Language*, 12:119–142.

J. Baron. 1977. Mechanisms for pronouncing printed words: Use and acquisition. In D. LaBerge and S. Samuels, editors, *Basic Processes in Reading: Perception and Comprehension*, pages 175–216. Lawrence Erlbaum Associates, Hillsdale, NJ.

M. F. Beauvois and J. Dérouesné. 1979. Phonological alexia: Three dissociations. *Journal of Neurology, Neurosurgery and Psychiatry*, 42:1115–1124.

L. Brooks. 1977. Non-analytic correspondences and pattern in word pronunciation. In J. Renquin, editor, *Attention and Performance VII*, pages 163–177. Lawrence Erlbaum Associates, Hillsdale, NJ.

J. A. Bullinaria. 1997. Modeling reading, spelling, and past tense learning with artificial neural networks. *Brain and Language*, 59:236–266.

M. Coltheart, J. Masterson, S. Byng, M. Pryor, and J. Riddoch. 1983. Surface dyslexia. *Quarterly Journal of Experimental Psychology*, 35A:469–495.

M. Coltheart, B. Curtis, P. Atkins, and M. Haller. 1993. Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, 100(4):589–608.

M. Coltheart. 1978. Lexical access in simple reading tasks. In G. Underwood, editor, *Strategies of Information Processing*, pages 151–216. Academic Press, New York.

W. Daelemans, A. van den Bosch, and T. Weijters. 1997. IGTree: Using trees for compression and classification in lazy learning algorithms. *Artificial Intelligence Review*, 11(1-5):407–423.

R. I. Damper and J. F. G. Eastmond. 1997. Pronunciation by analogy: Impact of implementational choices on performance. *Language and Speech*, 40(1):1–23.

R. I. Damper, Y. Marchand, M. J. Adamson, and K. Gustafson. 1999. Evaluating the pronunciation component of text-to-speech systems for English: A performance comparison of different approaches. *Computer Speech and Language*, 13(2):155–176.

M. J. Dedina and H. C. Nusbaum. 1991. PRONOUNCE: A program for pronunciation by analogy. *Computer Speech and Language*, 5:55–64.

Table 3: Mean percentage performances of patient KT and versions of faulty/non-faulty PbA for 161 regular words (RW) and 161 irregular words (IW).

Patient KT		Faulty PbA		Non-Faulty PbA	
RW	IW	RW	IW	RW	IW
91	28	83	30	99	100

E. Funnell. 1983. Phonological processes in reading: New evidence from acquired dyslexia. *British Journal of Psychology*, 74:159–180.

R. J. Glushko. 1979. The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, 5:674–691.

R. J. Glushko. 1981. Principles for pronouncing print: The psychology of phonography. In A. M. Lesgold and C. A. Perfetti, editors, *Interactive Processes in Reading*, pages 61–84. Lawrence Erlbaum Associates, Hillsdale, NJ.

G. E. Hinton and T. Shallice. 1991. Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review*, 98:74–95.

G. W. Humphreys and L. J. Evett. 1985. Are there independent lexical and non-lexical routes in word processing? An evaluation of the dual route theory of reading. *Behavioral and Brain Sciences*, 8:689–739.

D. Jones. 1996. *Analogical Natural Language Processing*. UCL Press, London, UK.

J. Kay and A. Marcel. 1981. One process, not two, in reading aloud: Lexical analogies do the work of non-lexical rules. *Quarterly Journal of Experimental Psychology*, 33A:397–413.

A. J. Marcel. 1980. Surface dyslexia and beginning reading: A revised hypothesis of the pronunciation of print and its impairments. In M. Coltheart, K. E. Patterson, and J. C. Marshall, editors, *Deep Dyslexia*, pages 227–258. Routledge and Kegan Paul, London, UK.

Y. Marchand and R. I. Damper. 2000. A multistrategy approach to improving pronunciation by analogy. *Computational Linguistics*, 26:195–219.

R. McCarthy and K. Warrington. 1986. Phonological reading: Phenomena and paradoxes. *Cortex*, 22:359–380.

N. McCulloch, M. Bedworth, and J. Bridle. 1987. NETspeak – a re-implementation of NETtalk. *Computer Speech and Language*, 2:289–301.

K. E. Patterson and J. Morton. 1985. From orthography to phonology: An attempt at an old interpretation. In K. E. Patterson, J. C. Marshall, and M. Coltheart, editors, *Surface Dyslexia: Neuropsychological and Cognitive Studies of Phonological Reading*, pages 335–359. Lawrence Erlbaum Associates, London, UK.

V. Pirrelli and S. Federici. 1995. You'd better say nothing than something wrong: Analogy, accuracy and text-to-speech applications. In *Proceedings of 4th European Conference on Speech Communication and Technology, Eurospeech'95*, volume 1, pages 855–858, Madrid, Spain.

V. Pirrelli and F. Yvon. 1999. The hidden dimension: A paradigmatic view of data-driven NLP. *Journal of Experimental and Theoretical Artificial Intelligence*, 11(3):391–408.

D. C. Plaut, J. L. McClelland, M. S. Seidenberg, and K. E. Patterson. 1996. Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, 103(1):56–115.

J. A. Reggia, P. M. Marsland, and R. S. Berndt. 1988. Competitive dynamics in a dual-route connectionist model of print-to-sound transformation. *Complex Systems*, 2:509–547.

M. S. Seidenberg and J. L. McClelland. 1989. A distributed, developmental model of word recognition and naming. *Psychological Review*, 96(4):523–568.

T. J. Sejnowski and C. R. Rosenberg. 1987. Parallel networks that learn to pronounce English text. *Complex Systems*, 1:145–168.

T. Shallice, E. K. Warrington, and R. McCarthy. 1983. Reading without semantics. *Quarterly Journal of Experimental Psychology*, 35A:111–138.

K. P. H. Sullivan and R. I. Damper. 1993. Novel-word pronunciation: A cross-language study. *Speech Communication*, 13:441–452.

F. Yvon. 1996. Grapheme-to-phoneme conversion using multiple unbounded overlapping chunks. In *Proceedings of Conference on New Methods in Natural Language Processing (NeMLaP-2'96)*, pages 218–228, Ankara, Turkey.

F. Yvon. 1997. Paradigmatic cascades: A linguistically sound model of pronunciation by analogy. In *Proceedings of 35th Annual Meeting of the Association for Computational Linguistics*, pages 429–435, Madrid, Spain.

M. Zorzi, G. Houghton, and B. Butterworth. 1998. The development of spelling-sound relationships in a model of phonological reading. *Language and Cognitive Processes*, 13:337–371.