

# Connectionist Modelling of Chinese Character Pronunciation Based on Foveal Splitting

Janet Hui-wen Hsiao (h.hsiao@sms.ed.ac.uk)

Richard Shillcock (racs@inf.ed.ac.uk)

School of Informatics, University of Edinburgh  
2 Buccleuch Place, Edinburgh, EH8 9LW, UK

## Abstract

We describe a connectionist model designed to reflect some of the anatomy of the visual pathways, notably the precise division of the human fovea and its subsequent contralateral projection to the cortex. The model was trained on a realistically large-scale problem, mapping between Chinese orthography and phonology. This split-fovea model replicated the interaction between character regularity and frequency that has been found in Chinese phonetic compound naming tasks. It also provided cross-language support for the hemispheric desynchronization account of dyslexia. Finally, the model predicted different regularity effects between characters with different phonetic radical positions.

## Introduction

Cognitive scientists aim to understand language processing universals. Seidenberg and McClelland's "triangle model" of the reading of monosyllabic English words has been substantially developed (e.g. Harm & Seidenberg, 1999). However, there is still little application to languages other than English. The cognitive modelling of the processing of Chinese orthography suffers from an input representativeness problem (cf. Chater & Christiansen, 1999) due to its complexity; there is ongoing debate as to how to represent Chinese characters in a psychologically realistic way. Most of the proposed computational models of Chinese character reading either have not been computationally implemented (e.g. Perfetti & Tan, 1999), or have employed relatively small-scale training data (e.g. Chen & Peng 1994). Cognitive modelling research in Chinese reading thus has lagged behind research in the reading of English.

Chinese has a radically different orthography from any alphabetic language. The basic writing units of Chinese are characters, which usually contain meaningful morphemes, instead of the letter-based representations of speech segments found in alphabetic languages. In general, there are four different ways of composing Chinese characters: *pictographs*, *indicatives*, *ideographs*, and *semantic-phonetic compounds*. A semantic-phonetic compound (or simply a *phonetic compound*) contains both semantic and phonological

information. Such compounds comprise about 81% of the 7,000 most frequent characters in the Chinese dictionary (Li & Kang, 1993). Hence, understanding how Chinese readers recognize these phonetic compounds is an important goal in psycholinguistic cognitive modelling.

A phonetic compound can be decomposed into two major components: a semantic component, which bears information about the meaning of the character, and a phonetic component, which may have partial information about the pronunciation of the character. Most phonetic compounds have their semantic radicals on the left and phonetic radicals on the right (SP characters). For example, the character “沐” means “taking a bath” in English and is pronounced as “mu4” in Pinyin<sup>1</sup>. It consists of a semantic radical on the left, which means “water”, and a phonetic radical on the right, which is pronounced the same as the character. We call these characters *regular phonetic compounds*. Some characters may be pronounced slightly differently from their phonetic radicals, such as “柚”. Its phonetic radical “由” is pronounced as “iou2” in Pinyin. However, “柚” has a different tone – it is pronounced as “iou4”. These characters are referred to as *semi-regular phonetic compounds*. Finally, there are some characters pronounced very differently from their phonetic radicals, such as “洒” (sa3) and “西” (xi1). We call them *irregular phonetic compounds*. The opposite structure to an SP character exists, in which the phonetic radical is on the left and the semantic radical is on the right (PS characters). The ratio of SP characters to PS characters is about 9:1 (Hsiao & Shillcock, *in preparation*).

A *regularity effect* has been found in the processing of Chinese phonetic compounds: Chinese readers name regular characters faster than irregular characters. There is also a frequency by regularity interaction in Chinese, as in English (see, .e.g., Hue, 1992; Liu, Wu & Chou, 1996; Seidenberg, 1985.)

Researchers have also studied Chinese character reading in brain-damaged patients (Yin & Butterworth, 1992) and found similar disorders as those found in

---

<sup>1</sup> The Chinese Pinyin system is a spelling system based on the Latin alphabet. The number at the end indicates the tone type.

English word reading. Chinese deep dyslexics were found to be able to pronounce irregular characters well but had difficulties pronouncing pseudo-characters with real semantic and phonetic radicals. On the other hand, Chinese surface dyslexics tended to regularize irregular characters and were able to pronounce about 50% of pseudo-characters according to their phonetic radicals (Zhou, 1999).

There is clear evidence that the human fovea is split precisely about the vertical midline: the left and right visual hemifields are projected contralaterally to the right and left hemispheres respectively (see, e.g. Fendrich & Gazzaniga, 1989). On the basis of anatomical and other evidence, a “split-fovea model” of English word reading has successfully captured several reading phenomena (see, e.g., Shillcock Ellison & Monaghan, 2000; Shillcock & Monaghan, 2001). Chinese phonetic compounds provide opportunities not available in alphabetic languages for examining the plausibility of this split-fovea model, since phonological information only comes directly from half of a character. In other words, the split fovea architecture seems to correspond fortuitously to the major functional division in the structure of Chinese phonetic compounds; the model “carves the problem at its joints”. Also, when an input character is irregular, the model faces an XOR-like problem, which makes interaction between the two halves necessary. Here we report our results of applying this split-fovea architecture to the modelling of Chinese character pronunciation.

## Simulations

### Phonological Representations

The sound system of Chinese differs from that of English. One of the most salient differences is the four distinct tones in standard Chinese (i.e. Mandarin)<sup>2</sup>. The pronunciation of each character has only one syllable, and every syllable has a nucleus and a tone associated. Characters with the same nucleus but different tones are usually not related in their meanings or orthography. In addition to a nucleus and a tone, there are three optional components associated with a syllable: a consonant at the beginning, a glide in the middle, and a glide or a consonant from a restricted class at the end (Wang, 1973). In total, syllables in Chinese have eight possible forms.

In Chinese syllables, all consonants can appear in the initial consonant position, and all vowels can appear in the nucleus position. On the other hand, there are only three possible vowels in the medial glide position, and five possible consonants and vowels in the ending

<sup>2</sup> Some dialects in China, such as Cantonese or Southern Min, may have more than four different tones.

position. According to the phonetic features of the Chinese Pinyin system (“Mandarin Consonants and Vowels”), there are 14 features for consonants: *bilabial, labiodental, dental, alveolar, palatal, velar, stop-aspirated, stop-unaspirated, nasal, fricative, affricative-aspirated, affricative-unaspirated, glide, and liquid*. Hence, we encoded every consonant in terms of these 14 features. Vowels were encoded with 8 features: *front, central, back, high, mid, low, unround, and round*.

In our phonological representation, the two major parts were the initial consonant, which consisted of 14 nodes for the 14 consonant features, and the nucleus vowel, which consisted of 8 nodes for the 8 vowel features. The glide was represented together with the vowel features in the nucleus vowel section. The same applied to the vowel features in the ending position. After 8 vowel feature nodes, we used 3 nodes to represent the features of the consonant in the ending position (*nasal, dental, and velar*). Notice that there are only two consonants (n and ng) possible in the final position. The last 2 nodes represented high and low tones respectively. 4 different tones in Chinese were represented with different combinations of the high and low tones (Yip, 2002). In total, the distributed phonological representation consisted of 27 nodes (see Figure 1).

Initial consonant features 14 nodes	Nucleus vowel features 8 nodes	Final consonant 3 nodes	Tones 2 nodes

Figure 1: The phonological representation.

### Orthographic Representations

Chinese characters consist of several individual strokes. There are some 20+ distinct strokes in Chinese orthography. Together, a few strokes may comprise a “stroke pattern”, a recurrent orthographic unit of Chinese characters. Some stroke patterns can be characters by themselves. Units can be constructed recursively to form another composite unit. Those units that are integral stroke patterns and cannot be further decomposed into other units have been referred to as *single bodies* (Chen et al, 1996).

Researchers have long believed that Chinese character recognition starts from an analysis of features and the number of individual strokes (e.g., Seidenberg, 1985), in contrast with letters in alphabetic writing systems. In recent years, researchers have found evidence that this recognition by skilled readers is based upon well-defined orthographic constituents, instead of individual strokes (Chen, Allport, and

Marshall, 1996; Zhou & Marslen-Wilson, 1999). Hence, in the orthographic representation, we used the basic stroke patterns defined in Cangjie, a Chinese transcription system developed by Ban-fu Chu in 1978. From a database analysis, there are 179 such stroke patterns comprising the radicals of all left-right structured Chinese phonetic compounds (Hsiao & Shillcock, *in preparation*). Hence, we used these 179 stroke patterns to encode the orthographic representation of the Chinese characters whose pronunciation we modelled.

### Training and Test Corpora

The training corpus contained all left-right structured Chinese phonetic compounds and all their radicals which exist as characters on their own. During training, each character was presented according to its log token frequency, taken from a Chinese lexical database (Hsiao & Shillcock, *in preparation*). The database contains about 3,000 of the most frequent Chinese phonetic compound characters. Among them there are 2,159 left-right structured phonetic compounds and 880 radicals that are also existing characters. The test corpus contained the same phonetic compounds, but not the radicals on their own.

### Network Architecture

Anatomical evidence has shown that the human fovea is precisely split about a vertical midline: when an alphabetic word or a Chinese character is fixated, the parts to the left and right of the fixation point are directly projected contralaterally. In modelling Chinese character recognition, we initially abstracted from real fixation behaviour and assumed that a character consisting of a semantic and a phonetic radical side by side could receive three possible fixations (see Figure 2). Characters were presented in the three fixation positions equally frequently during training. The task for the model, as for the reader, was to coordinate the information across the hemifields/hemispheres (Shillcock *et al.*, 2000).

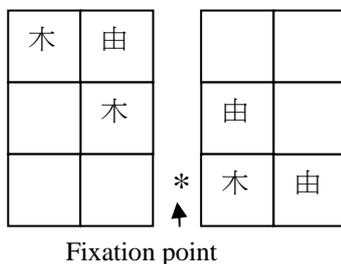


Figure 2: The complete pattern of inputs.

The network consisted of three layers. Adjacent layers were fully connected. Input units were localist

representations of stroke patterns, capturing the claim that stroke patterns are functional units of character recognition. The characters were all represented in each of the three positions necessary to accommodate the input schema shown in Figure 2. Each position represented each of the 179 possible stroke patterns. The input was mapped, via a hidden layer, onto a feature-level phonological output. For characters with more than one pronunciation, only the most frequent pronunciation was employed.

The model is shown in Figure 3. To solve the task, “interhemispheric” communication is necessary, in the form of “callosal” connections between the two sets of hidden units.

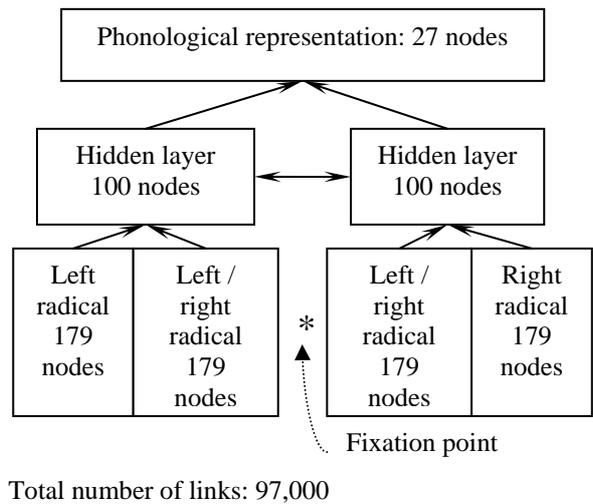


Figure 3: The split-fovea model for mapping Chinese orthography to phonology, with callosal connections.

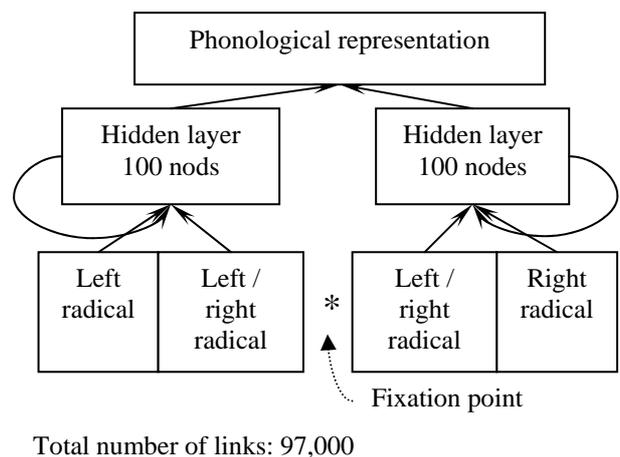


Figure 4: The model with no callosal connections.

Figure 4 shows a comparison model with no callosal connections in the hidden layers, which was trained on

the same task. In order to compare the performance of the two different architectures, we equalized their computational power by putting recurrent links on the hidden layers of the model with no callosal connections. Hence, both models had identical parameters and numbers of weighted connections. Thus, the principal difference between the models was the network architecture. We report elsewhere the more comprehensive comparison with a non-split model. The learning algorithm was discrete back propagation through time (Rumelhart, Hinton, and Williams, 1986).

## Results

We ran each of the two different models three times and analyzed their average performance. Figure 5 shows the performance of the two models on regular and irregular characters, in terms of summed square error (SSE) at different stages during training. Neither of the two models had difficulty learning this task well. The split architecture encouraged the model to discover the formal similarities within the radicals in the two halves of the characters; that is, that most phonological information came from the right half of the characters. The divided visual system fortuitously mirrored this distinction in the orthography.

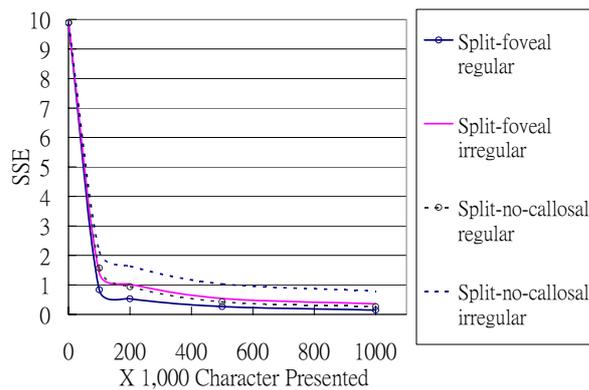


Figure 5: Performance of different models on regular and irregular characters.

The implemented split-fovea model provides an approach to understanding dyslexia in terms of hemispheric desynchronization (Shillcock & Monaghan, 2001). In the current simulations, the split-fovea model with callosal connections outperformed the model with no callosal connections (equivalent to extreme hemispheric desynchronization) on both regular and irregular characters; it especially exhibited more difficulty learning irregular characters, which constitute an XOR-like problem for the model with no callosal connections. Chinese surface dyslexics demonstrate reading impairments similar and analogous to those of dyslexics in alphabetic languages: poorer performance

reading irregular characters (Yin & Butterworth, 1992). Hence, the implemented split-fovea model provides cross-language support for the hemispheric desynchronization account of dyslexia.

The model with no callosal connections made regularization errors on irregular characters, as we might predict from the nature of the problem it faced. Table 1 shows some examples of such regularization errors. As can be seen, most characters were mistakenly pronounced exactly like their phonetic radicals; some were given the same pronunciation but with a different tone. Interestingly, we found some which were pronounced as other irregular characters with the same phonetic radical (e.g., 裕 in Table 1). This shows that the pronunciation of an irregular character was not only affected by its phonetic radical, but also by orthographic “neighbours” which share the same phonetic radical.

Character	Correct pronunciation	Generated pronunciation	Phonetic radical pronunciation
猜	cai1	qing1	qing1 (青)
帖	tie3	zhan4	zhan4 (占)
橫	heng2	huang2	huang2 (黃)
裕	Su2	yu4 (欲, 裕)	gu3 (谷)
沙	sha1	shao2	shao3 (少)
冶	ye3	tai2	tai2 (台)
杯	bei1	bu4	bu4 (不)

Table 1: Examples of regularization errors generated by the split model with no callosal connections.

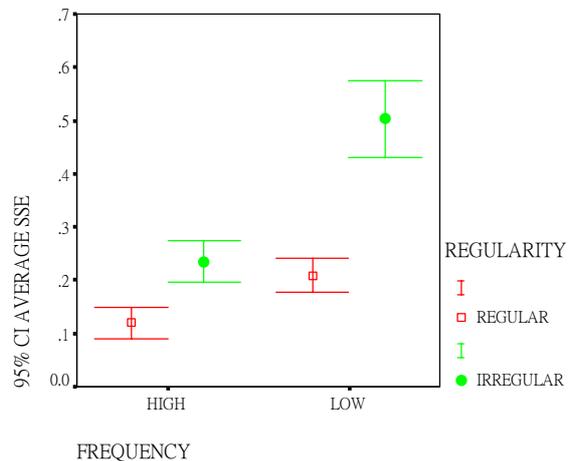


Figure 6: Interaction between frequency and regularity in the model with callosal connections after two million character presentations.

Figure 6 shows the interaction between frequency and regularity effects in the split-fovea model with callosal connections, after two million character presentations. This same interaction has been shown in experiments on Chinese character recognition (see, e.g., Shu et al, 2000; Hue, 1992; Liu, Wu & Chou, 1996; Seidenberg, 1985.). The model also produced this behaviour: the regularity effect was clearer among low frequency characters; there was a significant interaction between regularity and frequency (ANOVA analysis,  $F(1,1075) = 16.296$ ,  $p < 0.001$ ). The same significant interaction was also found in the version of the model with no callosal connections ( $F(1,175) = 6.809$ ,  $p < 0.01$ ).

We also examined the model's behaviour on SP and PS characters. It showed that there was no significant difference in the average SSE between the two groups in both split models with and without callosal connections ( $F(1,2155) = 1.730$ ,  $p > 0.05$ ;  $F(1,2155) = 2.117$ ,  $p > 0.05$ ). A significant interaction between position of the phonetic radical (i.e. SP or PS characters) and regularity was also found in both models ( $F(1,2155) = 4.719$ ,  $p < 0.05$ ;  $F(1,2155) = 5.479$ ,  $p < 0.05$ . See Figure 7 and 8). In the split model with callosal connections, there was a significant regularity effect among SP characters ( $F(1,1940) = 127.486$ ,  $p < 0.001$ ), but not among PS characters ( $F(1,215) = 3.048$ ,  $p > 0.05$ ). This may reflect the fact that only 24% characters are regular in the PS group, compared with 39% in the SP group (Hsiao & Shillcock, *in preparation*). On the other hand, the split model with no callosal connections did not exhibit the same behaviour: there were significant regularity effects among both SP characters ( $F(1,1940) = 140.654$ ,  $p < 0.001$ ) and PS characters ( $F(1,215) = 6.493$ ,  $p < 0.001$ . See Figure 8). Here the modelling makes a testable prediction regarding human behaviour. Elsewhere we verify this prediction (Hsiao & Shillcock, *submitted*).

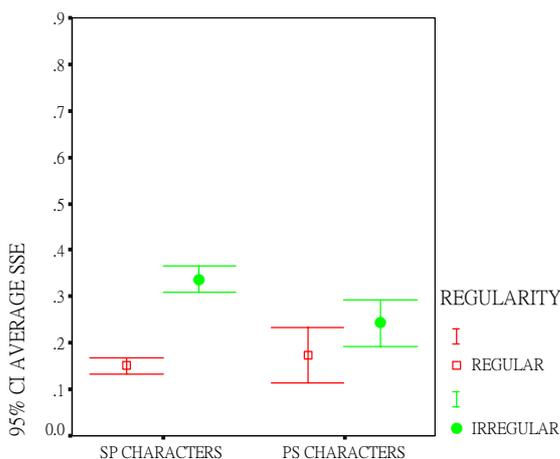


Figure 7: Interaction between position of phonetic

radicals and regularity of characters in the split model with callosal connections.

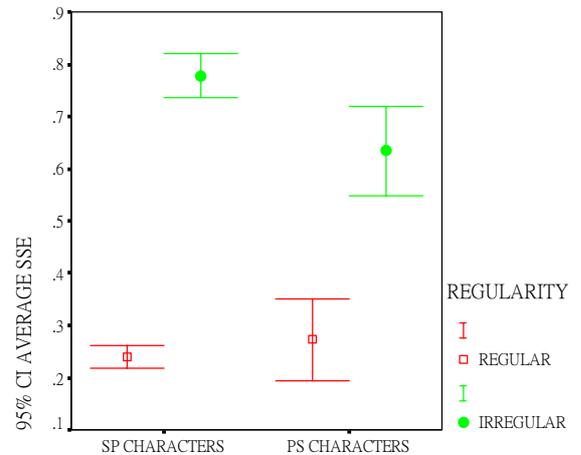


Figure 8: Interaction between position of phonetic radicals and regularity of characters in the split model with no callosal connections.

## Conclusion and Discussion

We have presented a connectionist model of Chinese character recognition, an extension of the anatomically based split-fovea model, and we have compared the processing of Chinese phonetic compounds in architectures with and without callosal connections. We have incorporated several simplifications concerning the nature of the orthographic input and fixation behaviour, but several dimensions of our modelling have been of a psychologically realistic scale and the modelling has succeeded in capturing a number of behaviours and also in making experimentally testable predictions.

On the task of orthography to phonology mapping, the split-fovea architecture facilitates the network's discovery of the relationship between character substructure and pronunciation. The split architecture fortuitously corresponds to the major functional division in the stimuli we have used. This modelling further demonstrates the potential value of incorporating the anatomical constraints of the visual pathways into the computational modelling of reading: the requirement of a staggered input (Figure 2) effectively parses the stimuli (a process that is more apparent in modelling the reading of alphabetic inputs).

Also, we have examined the performance of the model with no callosal connections and found behaviour similar to that of Chinese surface dyslexics. The performance of the "callosally impaired" model is worse than the split-fovea model especially on irregular characters. A further examination showed that most

errors made were regularization errors, which matches the behaviour of surface dyslexics. The modelling hence provides cross-language support for the hemispheric desynchronization account of surface dyslexia.

The model also has made some testable predictions from its performance. It shows that the regularity effect is more salient among characters with their phonetic radicals on the right than on the left. This interaction reflects a statistical fact that there is greater regularity among characters with phonetic radicals on the right. Hence, these phonetic radicals become better cues for pronunciation.

## References

- Chen, Y., and Peng, D. (1994). A connection model of Chinese character recognition and pronunciation. *Advances in the study of Chinese language processing*, pp.211-240.
- Chen, Y. P., Allport, D. A., & Marshall, J. C. (1996). What are the functional orthographic units in Chinese word recognition: The stroke or the stroke pattern? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 49(4), 1024-1043.
- Coltheart, M. (1981). Disorders of reading and their implications for models of normal reading. *Visible Language*, 15, 245-286.
- Christiansen, M. H., & Chater, N. (2001). Connectionist psycholinguistics: Capturing the empirical data. *Trends in Cognitive Sciences*, 5, 82-88.
- Fendrich, R. & Gazzaniga, M.S. (1989). Evidence of foveal splitting in a commissurotomy patient. *Neuropsychologia*, Vol. 27, No. 3, pp. 273-281.
- Goswami, U., & Bryant, P.E. (1990). *Phonological skills and learning to read*. Hillsdale, NJ: Erlbaum.
- Chinese characters: A Genealogy and Dictionary* by Harbaugh
- Harm, M. W. & Seidenberg, M.S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psych. Rev.*, 106, 491-528.
- Hsiao, J. H. & Shillcock, R. (2004). Analysis of A Chinese lexical database (*in preparation*).
- Hsiao, J. H. & Shillcock, R. (2004) Regularity Effect in Naming Chinese Phonetic Compounds with the Phonetic Radical on the Left or Right (*submitted*)
- Hue, C. W. (1992). Recognition processes in character naming. In *Language Processing in Chinese*, H.C. Chen and O.J.L. Tzeng (Eds.) 1992 Elsevier Science Publishers B.V.
- Li, Y. & Kang, J. S. (1993). Analysis of phonetics of the ideophonetic characters in Modern Chinese. In Y. Chen (ed.), *Information analysis of usage of characters in Modern Chinese*, pp.84-98. Shanghai: Shanghai Education Publisher. (In Chinese)
- Liu, I. M., Wu, J. T., & Chou, T. L. (1996). Encoding operation and transcoding as the major loci of the frequency effect. *Cognition*, 59, 149-168.
- Mandarin consonants and vowels. (n.d.). Retrieved May 13, 2004, from <http://personal.cityu.edu.hk/~cttomlai/doc/teach/stuff/mancon.htm>
- Perfetti, C.A. & Tan, L. (1999). The constituency model of Chinese word identification. In J. Wang, A. Inhoff & H. Chen (Eds.) *Reading Chinese Script*, Erlbaum: London, 115-134.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart and J. L. McClelland, editors, *Parallel Distributed Processing: Explorations in the microstructure of cognition; Vol. 1: Foundations*, Cambridge, Massachusetts. The MIT Press.
- Seidenberg, M.S. (1985). The time course of phonological code activation in two writing systems. *Cognition*, 19, 1-30.
- Shillcock, R., Ellison, T. M., and Monaghan, P. (2000). Eye-Fixation Behavior, Lexical Storage, and Visual Word Recognition in a Split Processing Model. *Psychological Review*, 2000, Vol. 107. No. 4, 824-851.
- Shillcock, R. C. & Monaghan, P. (2001). Connectionist modelling of surface dyslexia based on foveal splitting: Impaired pronunciation after only two half pints. *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*. Edinburgh: LEA. pp.916-921.
- Shu, H., Anderson, R.C., and Wu. (2000). Phonetic Awareness: Knowledge of Orthography-Phonology Relationships in the Character Acquisition of Chinese Children. *Journal of Educational Psychology*, Vol. 92, No. 1, 56-62.
- Wang, W. S. Y. (1973). The Chinese language. In Freeman (Ed.), *Readings from Scientific American: Language, Writing and the Computer* (pp. 50-60): Scientific American.
- Yin, W. & Butterworth B. (1992). Deep and Surface Dyslexia in Chinese. In H.C. Chen and O.J.L. Tzeng (Eds.), *Language Processing In Chinese*, pp. 349-366.
- Yip, M.J. (2002). *Tone*. Cambridge University Press.
- Zhou, X. & Marslen-Wilson (1999). Sublexical Processing in Reading Chinese. In J. Wang, A. Inhoff & H. Chen (Eds.) *Reading Chinese Script*, 37-63, Erlbaum: London.