Differences of Split and Non-split Architectures Emerged from Modelling Chinese Character Pronunciation

Janet Hui-wen Hsiao (h.hsiao@sms.ed.ac.uk)
Richard Shillcock (rcs@inf.ed.ac.uk)
School of Informatics, University of Edinburgh
2 Buccleuch Place, Edinburgh, EH8 9LW, UK

Abstract

The split fovea model, which reflects some aspects of the anatomy of the visual pathways, has successfully addressed several phenomena in visual word recognition (e.g., Shillcock, Ellison & Monaghan, 2000). However, it is still unclear what qualitative processing differences exist between a split architecture and a non-split counterpart. In the current study, we compare the performance of split and non-split architectures in modelling Chinese character pronunciation and show that Chinese left-right structured phonetic compounds create a unique opportunity for understanding the qualitative processing differences between the two possible versions of cognitive architectures.

Keywords: Connectionist modelling; Chinese character pronunciation; foveal splitting.

Introduction

In Chinese orthography, characters are the smallest units of the orthography. There exists a dominant type of Chinese characters, *phonetic compounds*, in which a semantic radical signifies the meaning of the character and a phonetic radical potentially informs the pronunciation of the character. Phonetic compounds comprise about 81% of the 7,000 frequent characters in the Chinese dictionary (Li & Kang, 1993).

The phonetic compounds have different relationships with their phonetic radical. For current purposes, *regular characters* are referred to as the characters that have the same pronunciation and tone as their phonetic radical; *semiregular characters* have the same pronunciation as their phonetic radical, but with a different tone; *irregular characters* have different pronunciations from their phonetic radical. In the Chinese lexicon, about half of the phonetic compounds are irregular, the other half are either regular or semiregular (Hsiao & Shillcock, *submitted*(a); unless otherwise stated, all Chinese language statistics cited here are from this paper). A *regularity effect* has been reported in Chinese phonetic compound recognition: regular characters are named faster than irregular characters. There is also an interaction between character frequency and regularity in Chinese, as in English (see, e.g., Hue, 1992; Liu, Wu & Chou, 1996; Seidenberg, 1985). The regularity effect and its interaction with frequency have been commonly used to examine cognitive plausibility of computational models (see, e.g., Plaut, McClelland, Seidenberg, & Patterson, 1996).

About two-thirds of phonetic compounds have a left-right structure. This left-right structure is the most tractable aspect in Chinese orthography, and has been a focus for understanding Chinese character recognition processes. The majority of the left-right structured phonetic compounds have the semantic radical on the left and the phonetic radical on the right – *SP characters*. The opposite arrangement also exists, in which the phonetic radical appears on the left and the semantic radical on the right – *PS characters* (Figure 1). The ratio of SP and PS character types in Chinese lexicon is about nine to one. They have correspondingly different internal information profiles. From an entropy analysis, there is greater entropy on the right side of the SP characters, reflecting greater variability of the phonetic radicals on the right as opposed to the semantic radicals on the left. In contrast, there is greater entropy on the left side of the PS characters, where the phonetic radicals appear. The distinction between the SP and PS characters provides important opportunities to examine hemispheric processing in language, given the very different information contained in the two halves of the character, interacting with the split fovea, which we discuss below.

採 彩

Figure 1. An SP and a PS character

The fovea is the part of the retina across which a fixated word is projected. It is responsible for fine-grain, focal visual processing. From anatomical and behavioural studies, it has become increasingly clear that the human fovea is precisely vertically split (see, e.g., Frendrich & Gazzaniga, 1989; Gray, Galetta, Siegal, & Schatz, 1997). This fact has fundamentally important implications for visual word recognition (Shillcock et al., 2000): when a word is fixated, the left part of the word is initially projected to the right hemisphere (RH) and the right part to the left hemisphere (LH). Thus, visual word recognition can be reconceptualised in terms of coordinating the information in the two hemispheres. The splitting is sufficiently precise that a single Chinese character, under normal reading conditions, is split precisely at the fixation point, with the semantic and phonetic radicals projected contralaterally to the two hemispheres (Hsiao, Shillcock, & Lavidor, *submitted*).

The split fovea model of English word reading has successfully captured several reading phenomena (see. e.g., Monaghan & Shillcock, *submitted*; Shillcock et al., 2000; Shillcock & Monaghan, 2001a). The split architecture fortuitously corresponds to the functional division of the phonetic compound structure; it "carves the problem at its joints". Alphabetic languages like English contain uneven distributions of information within words, but do not contain the dramatic difference observed in Chinese orthography. In modelling word recognition phenomena in English, the split architecture significantly accentuates phenomena such as the exterior letters effect (Shillcock & Monaghan, 2001b), but it does not trade on the qualitative differences in representations found in Chinese orthography. In view of this opportunity – not available in alphabetic languages – for examining the plausibility of the split fovea claim, Hsiao and Shillcock (2004) applied the split fovea architecture in modelling Chinese phonetic compound pronunciation. Their model successfully addressed some of the known effects in Chinese character recognition, such as the regularity effect and the regularity by frequency interaction, and provided cross-language support for the hemispheric desynchronization account of surface dyslexia (see Monaghan & Shillcock, *submitted*).

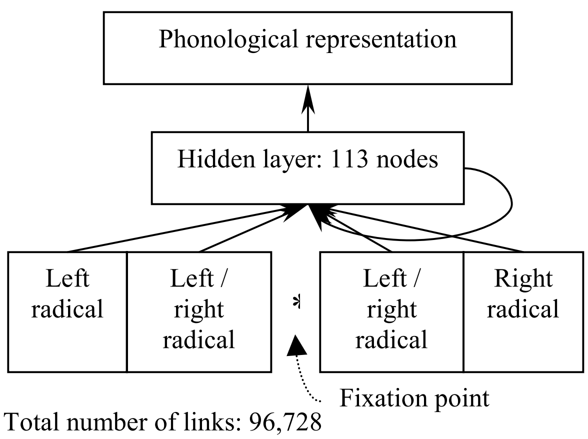

Figure 2. The non-split model for mapping Chinese orthography to phonology.

Previous efforts in connectionist modelling of Chinese character recognition usually adopted a feed-forward network architecture with a single hidden layer (Chen & Peng, 1994. See Figure 2). This architecture can be considered as a non-split architecture as opposed to the split fovea architecture, which has two interconnected hidden layers, left hidden layer (LHL) and right hidden layer (RHL), receiving input from left and right halves of the input layer respectively (Figure 3). These interconnections, or "callosal" connections, between the two hidden layers enable the "interhemispheric" communication between the two halves of the input. Nevertheless, non-split models also have been reported to be able to capture the regularity effect and the regularity by frequency interaction (Chen & Peng, 1994). It is hence unclear whether the split fovea model has a qualitatively different processing style and behaviour from the non-split model. Here we explore this issue by training both the split and non-split models with a realistic large-scale lexicon, which has an imbalanced distribution of SP and PS characters, and an artificial lexicon that has a balanced distribution of SP and PS characters. We show that the difference between the two architectures is revealed in processing the lexicon with an imbalanced SP and PS character distribution. This difference hence has important implications for the hemispheric processing of language.

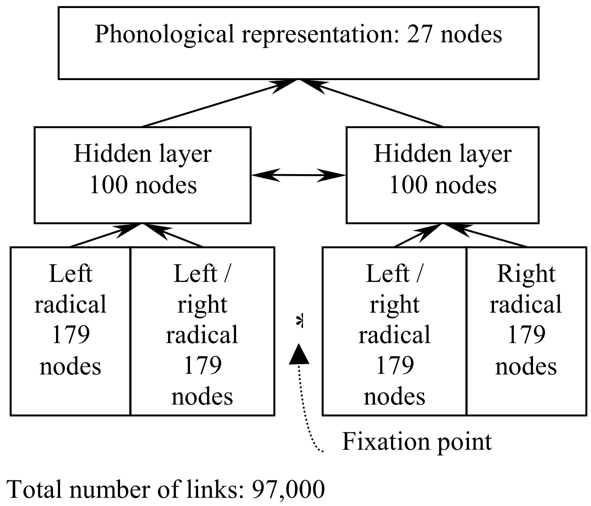

Figure 3. The corresponding split-fovea model.

Simulations

Phonological Representation

In the current modelling, we adopted a distributed, feature-based phonological representation. The pronunciation of each Chinese character has only one syllable, which can be divided into three segments: the initial consonant, the nucleus vowel, and the final consonant. Each character also has a tone associated. We allocated 14 features for the initial consonant, 8 features for the nucleus vowel, 3 features for the final consonant, and 2 features for the tone. In total, the phonological representation consisted of 27 feature nodes (for details, see Hsiao & Shillcock, 2004).

Orthographic Representations

Chinese orthography consists of several individual strokes. Some strokes may comprise a "stroke pattern", which is a recursive constituent of Chinese characters. In the orthographic representation, we used basic stroke patterns defined in Cangjie, a Chinese transcription system (Chu, 1978), to reflect the observation that the recognition by skilled readers is based upon well-defined, integral orthographic units (Chen, Allport, & Marshall, 1996; Zhou & Marslen-Wilson, 1999). There are 179 such basic stroke patterns comprising the radicals of all left-right structured Chinese phonetic compounds. These 179 stroke patterns were used to encode each Chinese character in the current models (see Hsiao & Shillcock, 2004).

Training and Test Corpora

Two sets of training and test corpora were used. In the first set, the training corpus contained all left-right structured Chinese phonetic compounds and their phonetic radicals which exist as characters on their own. During training, each character was presented according to its log token frequency. The database contains 2,159 of the most frequent left-right structured phonetic compound characters and 880 radicals that are also existing characters. The test corpus contained the same phonetic compounds as those in the training corpus, but not the phonetic radicals.

In the second set of corpora, an artificial lexicon was created in which the SP and PS characters had a balanced distribution. This artificial lexicon had 200 SP and 200 PS characters. The character type distribution in both the SP and PS character groups was proportional to the distribution of the SP characters in the real lexicon. Hence, among the 200 characters in either the SP or PS group, there were 74 regular characters (37%), 26 semiregular characters (13%), and 100 irregular characters (50%). Within the 100 irregular characters, 53 characters had the same rime as their phonetic radical, 12 characters had the same onset as their phonetic radical, and 35 characters had a radically different pronunciation from their phonetic radical.

The radicals that comprised the 200 SP characters consisted of 10 semantic radicals that only appeared on the left of the characters, and 40 phonetic radicals. The 200 PS characters consisted of the same 40 phonetic radicals as those in the SP characters, and another set of 10 semantic radicals which only appeared on the right of the characters. The 40 phonetic radicals could appear on either the left or right of a character. The characters in the SP group were randomly generated from different combinations of the left semantic radicals and the phonetic radicals; the semantic radicals of the characters in the PS group had the same combinations with the phonetic radicals as those in the SP group. The training corpus contained all the 400 phonetic compounds and the 40 phonetic radicals. Each character was presented with equal frequency. The test corpus contained the same phonetic compounds but not the phonetic radicals.

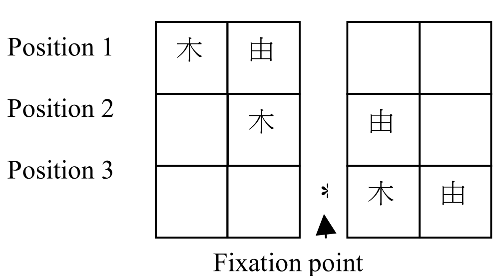

Figure 4: Three fixation positions in the input layer.

Network Architecture

In our split fovea model for Chinese character pronunciation, real fixation behaviour was idealized and a character was presented in each of three fixation positions equally frequently (Hsiao & Shillcock, 2004. see Figure 4). We adopted the same idealization in the current modelling.

Figures 2 and 3 show the non-split and split network architectures respectively for modelling the real lexicon. Adjacent layers were fully connected. The four blocks in the input layer were used to accommodate the input schema shown in Figure 4. Each node in a block represented one of the 179 possible stroke patterns. This orthographic input was mapped onto a feature-based phonological output, where the most frequent pronunciation of the input character was presented. We equalized the computational power of the two networks as much as possible by adding recurrent links to the hidden layer of the non-split model and making the number of weighted connections in the two models as closed as possible. Hence, the principal difference between the two models was the network architecture [1]. The corresponding split and non-split models for modelling the artificial lexicon is shown in Figure 5. The learning algorithm was discrete back propagation through time (Rumelhart, Hinton & Williams, 1986).

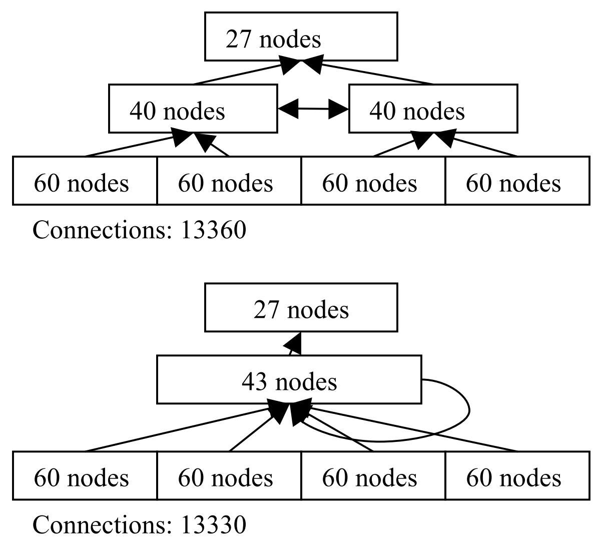

Figure 5. Corresponding split and non-split architectures for modelling the artificial lexicon.

[1] Note that because of the inherent difference in architecture between the two models, it is not possible to equalize the computational power of the two models in terms of both the number of nodes and the number of weighted connections while keeping connections between layers all fully connected. We chose to match the number of weighted connections since the weights on the connections are the only trainable parameters during training according to the learning algorithm, and hence the number of weighted connections is a good indication in the models' learning capacity.

Results and Discussion

We ran each model ten times and analyzed their average performance. The performance of the two models was very similar but in each of the two cases, the artificial lexicon and the real one, the split model had a slight but significant advantage over the non-split model (paired t-test, $t(1, 399) = -3.778$, $p < 0.001$ and paired t-test, $t(1, 2158) = -6.363$, $p < 0.001$, respectively). This slight advantage suggests that the split architecture encouraged the model to discover the functional division between the radicals in the two halves of the characters. The divided visual system fortuitously mirrored the distinction between the phonetic and semantic information in the orthography[2] (Hsiao & Shillcock, 2004).

As for their performance on different types of characters, both models captured the regularity effect and the regularity by frequency interaction in the real lexicon. For the top 10% high frequency and bottom 10% low frequency characters, the regularity by frequency interaction was significant ($F(1, 428) = 8.052$, $p < 0.01$); the interaction between the two models was not significant ($F(1, 428) = 1.456$, *n. s.*).

The two models also had similar performance on SP and PS characters. In the split model, there was a significant interaction between position of the phonetic radical and character regularity ($F(1, 2155) = 4.161$, $p < 0.05$); this interaction was marginal in the non-split model ($F(1, 2155) = 2.938$, $p = 0.087$). The difference between the two models in this interaction was not significant ($F(1, 2155) = 0.646$, *n. s.*). This interaction reflected the fact that only 34% of characters are regular or semiregular among the PS characters, compared with 50% among the SP characters.

When examining the performance difference between the two models, we observed a significant three-way interaction between network architecture, fixation position (see Figure 4), and position of the phonetic radical ($F(2, 4310) = 6.594$, $p = 0.001$). When we examined the models' performance in different fixation positions separately, we found that when characters were centrally presented in fixation position two, there was a significant interaction between network architecture and position of the phonetic radical: compared with the SP characters, the PS characters were relatively more difficult to process in the non-split model, but relatively easier in the split model ($F(1, 2155) = 6.161$, $p = 0.013$. See Figure 6). In contrast, when characters were presented in fixation position one or three, this interaction was not significant. In other words, the split model's behaviour in the fixation position one or three was very similar to the non-split model.

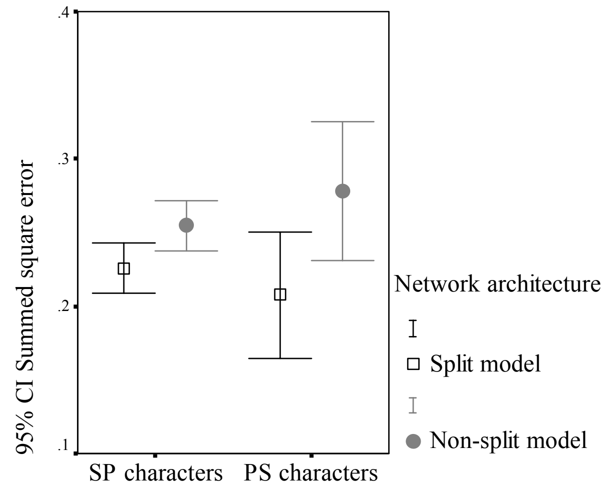

Figure 6. Interaction between position of phonetic radicals and network architecture. Error bars show 95% confident interval for mean.

This interaction can be explained in terms of the information profile in SP and PS characters and the qualitatively different processing style of the two models. As mentioned earlier, there is greater variability on the right of the SP characters, and this distribution is reversed in the PS characters. The overwhelming majority of SP characters leads to a greater variability on the right of the left-right phonetic compounds. When characters were centrally presented in the split model, the left and right radicals were projected to the LHL/RH and RHL/LH respectively, and communicated through the interconnected callosal connections (for the importance of these callosal connections, see Hsiao & Shillcock, 2004). Compared with the LHL, the RHL had a heavier processing load due to the greater variability, or entropy in information theory, on the right of the characters (Figure 7). For a centrally presented SP character, the RHL was where its phonetic radical initially projected. It hence had more processing difficulty than a centrally presented PS character, whose phonetic radical was initially projected to the LHL. Thus, in the split architecture, centrally presented SP characters were relatively more difficult to process than PS characters; this was especially true for irregular SP characters, for which more computational resource was required.

In contrast, in the non-split architecture, both left and right radicals were projected to and processed in the same single hidden layer. Also, as shown in Figure 7, the staggered input scheme (Figure 4) made the input entropy of the two sides of a centrally presented stimulus balanced (block 2 and 3 in Figure 7). Hence, for centrally presented characters, the minority PS characters became more difficult to process than the majority SP characters and this was especially true for regular PS characters, due to their unrepresentative nature: phonological information came from the left of the characters, as opposed to the normal

[2] Another possibility concerning the observed advantage for the split model was the slightly larger number of weighted connections in the split model than in the non-split model. Nevertheless, the number of connections was already matched as closely as possible between the two models, and hence the performance difference observed was less likely to be due to the small difference in the total number of connections.

cases in which phonological information came from the right of the characters. These two qualitatively different processing styles gave rise to the significant interaction observed between network architecture and position of the phonetic radical.

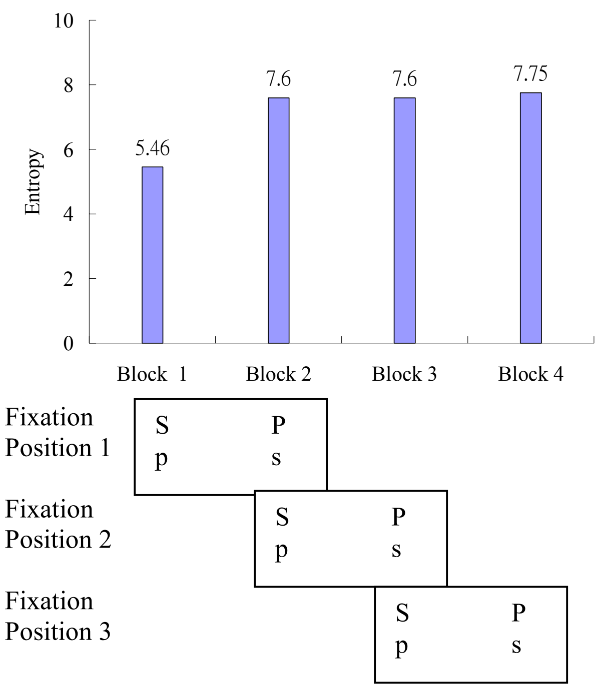

Figure 7. Entropy Analysis of the four blocks in the input layer.

Nevertheless, such interaction was only observed when the distribution of SP and PS characters in the lexicon was imbalanced. Figure 8 compares the performance of the two architectures on the artificial lexicon with a balanced distribution of SP and PS characters. In the split model, the LHL and RHL received the same processing load without any bias toward either PS or SP characters. Neither did any bias exist in the non-split model given an exactly balanced distribution of SP and PS characters. Hence, the interaction between network architecture and position of the phonetic radical was absent ($F(1, 396) = 0.466$, *n. s.*).

What can be inferred from these findings is that, the qualitatively different processing styles between the split and non-split architectures is best observed when there is an imbalanced distribution of two groups of stimuli with opposite internal information distributions. The distinction between Chinese SP and PS characters represents a unique opportunity for this examination. The processing difference between the two architectures hence has important implications for understanding hemispheric processing in language and examining the cognitive plausibility of the two architectures.

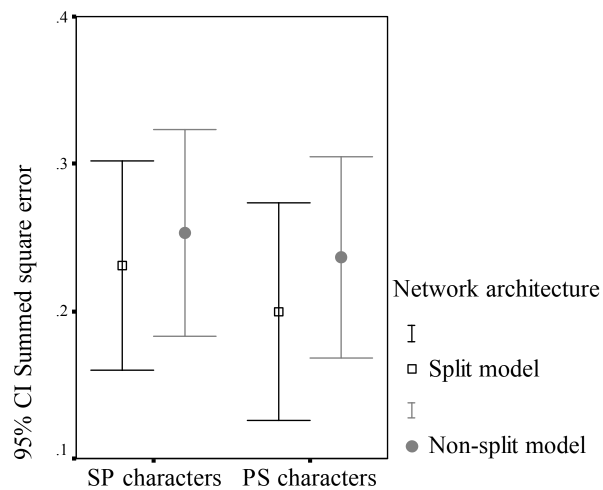

Figure 8. Performance of the two architectures on the artificial lexicon. Error bars show 95% confident interval for mean.

Conclusion

We have compared the performance of two computational architectures, the split fovea model and its non-split counterpart, in modelling Chinese character pronunciation. Both models have successfully addressed the regularity effect and the regularity by frequency interaction found in behavioural studies. When the computational power of the two models is closely matched as much as possible, the split fovea model slightly outperforms the non-split model. The split architecture fortuitously mirrors the functional distinction between the semantic and phonetic radicals, and hence facilitates discerning where the phonological information comes from.

The difference in processing style between the two models emerged when comparing their performance on centrally presented SP and PS characters. Due to the imbalanced distribution of SP and PS characters in the real lexicon, in the split architecture, the LHL/RH typically receives less processing load than the RHL/LH, and consequently facilitates the processing of centrally presented PS characters. In contrast, in the non-split architecture, both SP and PS characters are projected to and processed in the same single hidden layer; the unrepresentative nature of PS characters consequently induces more processing difficulties. Hence, there is a significant interaction between the network architecture and the phonetic radical position. Nevertheless, such interaction is not present when training the networks on an artificial lexicon with a balanced distribution of SP and PS characters. The distinct structures and skewed distribution of Chinese SP and PS characters hence has provided a unique opportunity for examining the difference between the split and non-split architectures.

Do the visual pathway anatomies, reflected in the split fovea architecture, really matter in attempts to model reading behaviour? The different processing styles of the two architectures have made different predictions about Chinese readers' behaviour when naming centrally presented SP and PS characters. The future work hence is to examine the cognitive plausibility of the two computational architectures by verifying these testable predictions through behavioural studies (Elsewhere we compare the modelling results with behavioural data; see Hsiao & Shillcock, *submitted*(b)).

Acknowledgments

We are grateful to the Economic and Social Research Council.

References

Chen, Y., and Peng, D. (1994). A connection model of Chinese character recognition and pronunciation. *Advances in the study of Chinese language processing*, 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*, 1024-1043.

Chu, B. (1979). Laboratory of Chu Bong-Foo. Retrieved Aug. 25, 2004, from http://www.cbflabs.com/.

Fendrich, R. & Gazzaniga, M. S. (1989). Evidence of foveal splitting in a commissurotomy patient. *Neuropsychologia*, *27*, 273-281.

Gray, L. G., Galetta, S. L., Siegal, T., & Schatz, N. J. (1997). The central visual field in homonymous hemianopia – evidence for unilateral foveal representation. *Archives of Neurology*, *54*, 312-317.

Hsiao, J. H. & Shillcock, R. (2004). Connectionist Modelling of Chinese Character Pronunciation Based on Foveal Splitting. *Proceedings of the Twenty Sixth Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum Associates.

Hsiao, J. H. & Shillcock, R. (*submitted*(a)). Analysis of a Chinese Phonetic Compound Database: Implications for Orthographic Processing.

Hsiao, J. H. & Shillcock, R. (*submitted*(b)). Chinese orthography is processed differently in the male and female brain.

Hsiao, J. H., Shillcock, R. & Lavidor, M. (*submitted*). A TMS Examination of Semantic Radical Combinability Effects in Chinese Character Recognition.

Hue, C. W. (1992). Recognition processes in character naming. In H. C. Chen & O. J. L. Tzeng (Eds.), *Language Processing in Chinese* (pp. 93-107). Amsterdam: North-Holland.

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*. 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.

Monaghan, P. & Shillcock, R. (*submitted*). Hemispheric desynchronisation and developmental dyslexia in a computational model of reading.

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.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart & J. L. McClelland (Eds), *Parallel Distributed Processing: Explorations in the microstructure of cognition; Vol. 1: Foundations* (pp. 318-362). 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, 107*, 824-851.

Shillcock, R. C. & Monaghan, P. (2001a). Connectionist modelling of surface dyslexia based on foveal splitting: Impaired pronunciation after only two half *pint*s. *Proceedings of the 23rd Annual Conference of the Cognitive Science Society.* Edinburgh: LEA.

Shillcock, R. C. & Monaghan, P. (2001b). The computational exploration of visual word recognition in a split model. *Neural Computation*, *13*, 1171-1198.

Zhou, X. & Marslen-Wilson (1999). Sublexical Processing in Reading Chinese. In J. Wang, A. Inhoff & H. Chen (Eds.), *Reading Chinese Script* (pp. 37-63). Erlbaum: London.