

Running Headers: Writing to read in Chinese

Exploring the effects of knowledge of writing on reading Chinese characters in skilled readers

Mingjun Zhai* & Simon Fischer-Baum

Department of Psychology
Rice University, Houston TX

*Corresponding author

Department of Psychology, MS-25

Rice University

P.O. Box 1892

Houston, TX 77251

Word Count: 5273

mz21@rice.edu

Abstract

Knowledge about how characters are written has been argued to play a particularly important role in how children learn to read Chinese. In the current study, we investigate the role that knowledge about writing characters plays in visual word processing in skilled adult readers. While there is clear neuropsychological evidence against the strong version of the hypothesis that reading depends on writing in Chinese even once literacy is acquired, it is still possible that writing could have a modulatory influence on how visually presented Chinese characters are processed in literate readers. The present study addressed this hypothesis using a visual same/different judgement task on pairs of characters that vary in how similar the two characters are visually and how similar they are in terms of motor plan, using 24 expert readers and writers of Chinese and 24 naïve participants with no prior experience with written Chinese as subjects. The results of linear mixed-effects modeling indicate that the speed of same/different judgements is influenced by visual similarity, but not by how similar they are written, even in the group of skilled readers. These results suggest that knowledge of how Chinese characters are written does not influence visual character processing in skilled readers.

Keywords: character perception, reading, writing, Chinese characters, expertise

Understanding the representations and processes involved in recognizing written language depends, at least in part, in understanding variation in writing systems (Frost, 2012; Share, 2008, 2014). In alphabetic languages, like English, words are composed of a linear sequence of letters. In contrast, the Chinese language has a logographic script. Each character is a box-like symbol occupying equal sized square space in which sub-character components, like strokes and radicals, are assembled in a nonlinear way, resulting in a complicated visual configuration. In English, letters or letter clusters have a relatively transparent mapping to phonemes in the spoken language. In contrast, although some Chinese characters contain phonetic radicals that indicate phonological information of characters, links between orthography and phonology are much more opaque. These differences between Chinese and alphabetic languages may have an impact on the cognitive processes involved in reading words. For example, in English and other alphabetic languages, because phonological decoding from print is relatively straightforward, learning the mapping from print to meaning may rely more on the existing connection between phonology and semantics in spoken language. As a consequence, phonological awareness has been argued to be a critical part of learning to read in alphabetic languages (Ziegler & Goswami, 2005). In contrast, Tan and colleagues (2005) found that these phonological skills had only a weak relationship with Chinese reading ability. Instead, a much stronger relationship was found with tasks that tap into motor processes similar to those used in writing characters. They reasoned that since the visual-spatial configuration of Chinese characters is complicated, when students learn Chinese, there is a greater emphasis on repeated copying the strokes, radicals and characters, than is seen in alphabetic languages. Therefore, Tan and colleagues (2005) argue that learning how to write, and knowledge of the stroke-motor processing involved in producing

characters, may be an integral part of learning to read specifically for logographic writing systems like Chinese, what we will call the writing-to-read hypothesis.

Various studies have shown that repeated writing has a positive effect on reading or memorizing Chinese characters, with both children (Chan et al., 2006; McBride-Chang et al., 2011; Naka & Naoi, 1995) and adult learners (Cao et al., 2013; Guan et al., 2011; Guan, Perfetti & Meng, 2015; Naka, 1998). Less well understood is whether this knowledge about how characters are written continues to have an influence on how literate adults read Chinese. Bi and colleagues (2009) reported a case study of a patient with acquired dysgraphia following brain damage, whose performance suggested that he had little knowledge of stroke motor programming involved in writing Chinese but not a general fine-motor impairment. Despite this writing deficit, his ability to read aloud and understand written language was unimpaired. This clear dissociation between writing and reading in Chinese rules out the strongest version of a writing-to-read hypothesis in literate readers, that if knowledge of how characters are written is lost then the ability to read those characters is also necessarily lost. However, weaker versions of the writing-to-read hypothesis are possible; knowledge of how characters are written is not necessary for character recognition but may still influence how skilled readers read.

Neuroimaging studies have also supported this connection between perceiving and producing Chinese characters in literate adults, showing activation of motor areas associated with writing during a reading task (Siok et al., 2004; Tan et al., 2003). However, while knowledge about how a character is written might be activated during visual word processing tasks, these studies do not demonstrate that this knowledge has any influence on how visual characters are actually processed. Behavioral evidence for this writing-to-read effect in skilled readers has largely come from studies that investigate whether knowledge about the order in

which the strokes are produced when writing a character influences how it is recognized. Several studies have shown stroke order effects on reading Chinese. Flores d'Arcias (1994) reported that character naming was faster when characters were primed by the earliest strokes of the character compared to when characters were primed by the last strokes of the character. Furthermore, he showed that participants were more likely to incorrectly judge two characters as being identical if they shared early strokes than if they shared late strokes. Yan and colleagues (2012) showed that, when removing a fixed proportion of strokes from each character in a sentence reading task, removing initial strokes was more disruptive to sentence reading than removing final strokes (see also Tseng, Chang & Wang, 1965).

The fact that information about which stroke is produced first during writing a character seems particularly important for reading that character has been taken as evidence for the writing-to-read hypothesis in skilled readers. However, an alternative interpretation of these studies is that these effects can be explained by low-level visual differences between the conditions, rather than by differences in stroke order information. **As shown in Yan et al. (2012), the initial strokes were mostly located in the top-left part of the characters. Using an algorithm which has no information about stroke writing order, Wang and colleagues (2013) examined the same set of characters and identified the left part as the most important region for recognizing a character¹. Their analyses revealed that the effect of removing beginning strokes in Yan et al. (2012) could be due to the degraded character contours. Therefore, stroke order effects may be understood as arising from the visual, rather than stroke-motor, properties of the stimuli.** Methods that more clearly distinguish the

¹ Wang et al. (2013) compared the distribution of removed elements from the conditions in their study with Yan et al. (2012) and analyzed character configuration of original and degraded characters by measuring contour of character. The implication of the discrepancies between these two studies was discussed in Wang et al. (2013).

contribution of visual and stroke-motor features in normal character reading are needed to test the writing-to-read-hypothesis.

In the present study, we test the separable contributions of both visual and stroke-motor features to character processing in skilled readers of Chinese, using linear mixed-effect modeling (LMM) to predict reaction times in a same/different judgement task. Wiley, Wilson and Rapp (2016) recently applied this method to discriminating pairs of Arabic letters. Predictors, based on how similar the two letters were in terms of visual features or on knowledge-based factors, like stroke-motor knowledge and letter-sound correspondence knowledge, were used to predict reaction time. For naïve readers with no previous exposure to Arabic, only visual similarity contributed to explaining variation in reaction time. For skilled readers of Arabic, models that included knowledge-based factors in addition to visual similarity fit the data better than models that only included visual similarity. These results indicated that for skilled readers these knowledge-based factors contributed to the processing of Arabic letters.

We apply this same logic to processing Chinese characters, with a specific focus on one knowledge-based factor, knowledge of the stroke-motor program required for writing that character. Using a visual same/different judgement task on pairs of characters, we tested the hypothesis that literate Chinese readers use knowledge of writing characters to recognize them. Characters used in the study were selected to distinguish the role of visual and stroke-motor similarity during character recognition and comparisons were made between expert readers and writers of Chinese and naïve participants. If the knowledge of writing Chinese characters facilitates character perception in skilled readers, we expect to find that models that include stroke-motor similarity should fit the data better than models with only visual similarity. Furthermore, this improvement should be only observed with the expert group but not the naïve

group. If knowing how to write does not contribute to character processing, then stroke-motor similarity should not improve model fit, though models that include other knowledge-based factors should improve the fit of the model, but only for the expert readers.

Methods

Participants

An expert group (11 males and 13 females, average age = 20.2) was recruited consisting of 24 native Mandarin speaking adults who are literate in simplified Chinese characters and have at least a high school education in Mainland China. In addition, 24 native English speakers with no prior experience reading or writing Chinese characters were recruited to form a naïve group (4 males and 20 females, average age = 19.7). The students from both participant groups were undergraduate students taking courses offered by Psychology Department of Rice University. All participants were compensated with course credit. This study was approved by institutional review board at Rice University.

Stimuli

All Chinese characters can be described by a fixed sequence of strokes. A standardized set of 28 strokes taught in the official first-grade Chinese textbook in mainland China. However, some of these strokes can be further decomposed into simpler stroke features. For example, as shown in Figure 1a, the stroke “shugou” is composed of the two simpler stroke features “shu” and “gou”. In total, the 28 strokes can be described by a set of 11 stroke features (see Supplemental Material for a full characterization of this decomposition), and all characters can be described by a fixed sequence of these stroke features. We assume this sequence of stroke features represents the stroke-motor plan for writing each character. For example, the stroke-motor plan for the characters 水 and 永 are shown in Figure 1b. For the character 水, the sequence of stroke features

is “shu” “gou” “heng” “pie” “pie” “na”, while for the character 氵, the sequence of stroke features is “dian” “heng” “zhe” “gou” “heng” “pie” “pie” “na”.

Characters that have similar stroke-motor representations frequently also are visually similar, particularly characters with multiple radicals. Given that our goal was to assess independent contributions of stroke-motor and visual representations to character processing, we selected our stimuli from the set of 241 simple characters that only contain one radical and consist of no more than six strokes, identifying a subset of these characters that could best distinguish stroke-motor and visual representations. To do this, we first calculated stroke-motor and visual similarity between all pairs of these 241 characters. Stroke-motor similarity was defined by the Levenshtein distance, or the minimum number of single-stroke feature edits (insertion, deletion, substitution) required to change one string to another, normalized by dividing by the length of the longer string, in this case, the greater number of stroke features, because it denotes the maximum number of possible edits. For example, the Levenshtein distance between 水(shu gou heng pie pie na) and 氵 (dian heng zhe gou heng pie pie na) is 3 (Substitute shu for dian, Add heng, Add zhe) divided by 8, the greater number of stroke features, that is, the number of stroke features in 氵, for a distance measure of .375. Similarity was calculate as $1 - \text{distance}$, so the stroke motoric similarity between 水 and 氵 was calculated to be 0.625. Visual similarity was calculated by calculating the number of overlapping pixels divided by the number of intersected pixels. To calculate the pairwise visual similarity between two characters, we tried multiple alignments, shifting the relative position of the characters in KaiTi font and recalculating the voxel-wise overlap, with visual similarity reflecting the value at the maximum overlap.

An experimental set of 31 characters was selected from the original 241-character search space, which resulted in 465 pairwise similarity calculations between characters. A brute force, randomization procedure was used in which we randomly selected 31 character a billion times and analyzed the 465 pairwise stroke-motor similarity measures and the 465 pairwise visual similarity measures, looking at the standard deviation of each measure as well as the correlation between the two measures. From these 1,000,000,000 possible stimuli sets, we selected the set of characters that minimized the correlation between stroke motoric similarity and visual similarity matrices of characters (Pearson's $r = .146$) while maintaining a good range on both the visual similarity and the stroke-motor similarity measures (see Supplemental Materials). Some character pairs had high visual similarity but low stroke-motor similarity (e.g., 立 vs. 互, visual similarity: 0.43; stroke-motor similarity: 0.17) while other had low visual similarity but high stroke-motor similarity (e.g., 风 vs. 丸, visual similarity: 0.26; stroke-motor similarity: 0.86).

In addition to the stroke-motor and visual similarity measures, we also considered the fact that, for expert readers, how similar the characters are in meaning and pronunciation may affect their performance in same/different judgment task. Thus, we included the phonological and semantic similarity as knowledge-based factors. The phonological similarity was measured by counting the number of identical onsets or rimes in each pair of characters (see Shu et al., 2005 for motivation). For semantic similarity, we recruited 12 native Chinese speakers who did not participate our experiment to rate how similar each pair of characters are in meaning on a 7-point scale. Correlations between the four predictors are reported in Table 1.

The task involves showing participants two character tokens and asking them to make judgements about whether the characters are the same or different. Each of our 31 characters was

paired with themselves 15 times for a total of 465 same trials and with each of the other 30 characters a single time for a total of 465 different trials, or 930 trials in total in the experiment. Two stimulus lists were created to counterbalance the screen position of each character in different pairs across participants (left vs. right). The order of presenting stimulus was pseudorandomized with no more than three consecutive same or different pairs. Font KaiTi and size 96 were used (each character occupied a visual angle of 1.5°). Pilot data with naïve participants showed good discrimination with these stimuli of this size. Another five pairs of characters were used to construct a practice session to familiarize participants with the procedures of the experiment.

Procedure

Participants were tested individually. They were asked to perform same/different judgment towards pairs of Chinese characters as quickly and accurately as possible. In each trial, they were first presented with a central fixation cross for 500 ms. Then two horizontally aligned characters (separated by a visual angle of 1°) were presented for 2500ms or until a response was given by participants. Participants were required to judge if the two characters were physically identical pressing a yes or no key. The inter-trial interval blank screen lasted for 500 ms. The 930 trials were separated into three blocks each containing 310 trials. Participants were instructed to take breaks between blocks.

After they finished the main experiment, participants did a simple visual detection task in which they were required to press one of the response keys as soon as they saw a visual stimulus (a pair of tilde symbols) on the screen, with 25 trials for each response key. Trial structure was identical in the detection and discrimination tasks, except that fixation duration varied from 600 to 1000 ms in the detection task. In total, the experiment lasted approximately 45 minutes.

Data Analyses

The simple detection task was used to calculate discrimination time from raw reaction time, by subtracting out detection and response time, or the time it took participants to see that a character was present and response key. Detection and response time was calculated by averaging the 20 fastest reaction times in the simple detection task for the response key that corresponded to different trials. Discrimination time was then calculated by subtracting this value from each raw reaction time (Courrieu et al., 2004).

One naïve reader was excluded due to low accuracy (65%) and two experts were excluded for abnormally long reaction times (twice the average reaction time of rest of subjects). We replaced them with three new participants with normal accuracy and reaction times. Within subjects, individual trial outliers were defined as being either outside 3 IQRs or shorter than 200 ms. Trials with discrimination times less than zero were also removed. In total, 4% of the data was removed. Discrimination times were then log-transformed to address positive skew and normalized for model-based analysis.

A nested model comparison approach with mixed effect models was taken using visual, stroke-motor, phonological and semantic similarity between the two different characters as predictors of log-transformed normalized discrimination times. Random intercepts for subject and item were included as random effects for this experimental design. Chi-square tests were used to compare nested models to determine if additional factors contributed significantly to model fit. This procedure was carried out separately for the naïve and expert groups. We also carried out an analysis that combined the data of two groups and include all predictors and the interaction between each of these factors and group. The beta weights of four predictors and the

interaction terms were examined to assess the difference between groups in the contribution of these factors.

Results

The mean and 95% confidence interval of raw reaction time, discrimination time and accuracies by participant group are reported in Table 2. The expert group discriminates character pairs faster (560 ms) than the naïve group (654 ms, $t(46) = 3.19, p < .05$), as would be expected, given differences in experience with Chinese. Both groups were very accurate on this task ($> 97\%$ correct).

Model-based analysis

The four similarity measures were added as predictors to a null model with only the random effects one by one in following sequence: visual, phonological, semantic and stroke-motor similarity. By comparing each model with the previous one, we evaluate the contribution of each predictor above and beyond the previous ones.

Naïve group. For the naïve group, the model with visual similarity added as a predictor had a significantly better fit than the null model with only the random effects ($\chi^2(1) = 44.76, p < .001$), indicating that visual similarity makes a significant contribution to predicting discrimination time. For example, the discrimination time for 水 and 永 (visual similarity: 0.49) was on average 478 ms while the discrimination time for the 兮 and 各 (visual similarity: 0.26) was on average 331 ms. Model fit was not improved significantly by adding in phonological similarity ($\chi^2(1) = 0.45, p = .50$), semantic similarity ($\chi^2(1) = 1.13, p = .29$), or stroke-motor similarity as a predictor ($\chi^2(1) = 0.44, p = .51$). The failure to find improvements in model fit by knowledge-based factors is consistent with the fact that participants in the naïve group had no previous experience with Chinese characters and therefore did not know how they were pronounced, what they meant

or how they were written. The beta weights and *t*-values of the full models with all predictors suggest only visual similarity has significant effect (at a threshold of *t* greater than 2, equivalent to an alpha at .05) on discrimination time ($\beta = 0.032, t = 6.42$), which is consistent with the results of the chi-square tests.

Expert group. For the expert group, the model with visual similarity added as a predictor also had a significantly better fit than the null model with only the random effects ($\chi^2(1) = 19.58, p < .001$). The model fit was not significantly improved when phonological similarity was added above and beyond visual similarity ($\chi^2(1) = 3.35, p = .07$). However, further adding semantic similarity did improve the fit of the model ($\chi^2(1) = 5.08, p < .05$). Finally, and most critically for the current study, adding stroke-motor similarity to the previous model did not improve the fit of the model above and beyond the model that contained visual, phonological and semantic similarity as predictors ($\chi^2(1) = 0.08, p = .77$). Again, the beta weights and *t*-values of the full model are in line with the results of the chi-square tests: only visual ($\beta = 0.018, t = 3.87$) and semantic similarity ($\beta = 0.010, t = 2.19$) are significant predictors.

Given that the contribution of stroke-motor similarity to predicting these different discrimination times is the critical test for the central question in this experiment, we took an exploratory approach to further test whether stroke-motor similarity has any impact on character perception. However, when we compared the model that only include stroke-motor similarity with the null model, we still did not observe any significant change in model fit ($\chi^2(1) = 1.37, p = .24$), indicating that, while other knowledge-based factors, specifically semantics, make some contribution to predicting discrimination time in the expert group, knowledge of how the character is written makes no contribution.

Combining two groups. We examine the beta weights and *t*-values of the full models with all predictors and the interaction between each predictor and group (Table 3). The main effect of group confirmed that expert readers discriminate character pairs faster than naïve readers ($\beta = 0.22, t = 3.30$). The interaction between visual similarity and group was significant in predicting discrimination time ($\beta = 0.014, t = 2.88$), indicating that although higher visual similarity predicts slower responses for both the expert group and naïve group, the effect is significantly larger for the naïve group compared with expert group. This finding is reasonable as visual similarity is the only information naïve readers can use to discriminate pairs of characters. None of the interactions between each of three knowledge-based factors and group are significant (phonological similarity by group: $\beta = -0.0006, t = -0.21$; semantic similarity by group: $\beta = -0.004, t = -1.51$; stroke-motor similarity by group: $\beta = 0.002, t = 0.85$), indicating that the effect was similar in these two subjects group, though there is a tendency for the semantic similarity effect to be larger in the expert group.

Of particular importance is the lack of an interaction between the two groups with respect to stroke-motor similarity. Discrimination times were no more influenced by stroke-motor knowledge for the expert group, who had years of experience writing Chinese, than for the naïve group, who had no experience in writing Chinese. Indeed, stroke-motor similarity had no effect on discrimination times for either group. This finding goes against the predictions of the writing-to-read hypothesis which states that knowledge of how characters are written influences how they are recognized. One important possibility to consider however, is that stroke-motor similarity does play a role in processing visually presented Chinese characters, but that our measure of stroke-motor similarity does not actually correspond to the stroke-motor programs used when writing Chinese characters. For example, we assumed that the units of these motor

programs were stroke features. Alternatively, each stroke might be a single unit that is not decomposable into stroke features. Therefore, in the next section, we consider several alternative methods of computing stroke-motor similarity to see if these alternative measures improve model fit.

Other theories of stroke motoric similarity. One of the advantages of linear mixed effect modeling approach taken in the current study is that it is easy to test other theories of similarity measures with the current data set. Thus, we introduce three additional measures of stroke-motor similarity and test them in the same way as the stroke-motor similarity measure in previous section, that is, ask whether these measures improve model fit above and beyond the model that includes visual, semantic and phonological similarity.

Stroke string similarity. Each Chinese character was coded into a sequence of its component strokes, instead of stroke features. The stroke string similarity of a pair of characters was calculated by subtracting the normalized Levenshtein distance between two sequence of strokes from one. When stroke string similarity was added to the model as a predictor, the model fit was not improved significantly for either the expert ($\chi^2(1) = 0.43, p = .51$) or naïve group ($\chi^2(1) = 0.31, p = .58$).

Shared first stroke. Yan et al. (2012) and Flores d'Arcias (1994) both argue that initial strokes are especially crucial for character identification and discrimination. Therefore, we added a predictor indicating simply whether two characters shared the first stroke feature into model (1 = two characters share first stroke feature, 0 = two characters do not share the first stroke feature). Adding this predictor into the model did not significantly improve model fit for either the expert ($\chi^2(1) = 0.79, p = .37$) or naïve group ($\chi^2(1) = 2.92, p = .09$).

Stroke bigram similarity. Wiley et al. (2016) used stroke bigrams – or pairs of consecutive strokes – as the relevant units of stroke-motor representation to measure motor stroke similarity between Arabic letter pairs. Here we adopted the same method to stroke features of Chinese characters. We described each character by its set of stroke feature bigrams and calculated the proportion of shared stroke bigrams for each character pair. Again, the model fit was not improved significantly by adding in this stroke-motor similarity measure for either the expert group ($\chi^2(1) = 1.04, p = .31$) or the naïve group ($\chi^2(1) = 1.00, p = .32$).

Bayes factor analysis. None of the methods for computing the stroke-motor similarity improved model fit. To test whether there is evidence in favor of a null effect of stroke motoric similarity, we computed Bayes Factor between the two nested models, giving the strength of evidence for the model without the stroke-motor predictor for four different theories of stroke motoric similarity metrics. For the stroke-motor similarity metrics, the Bayes Factor (null/alternative) is 13.8, suggesting it is 13.8 times more likely that the null hypothesis is true than alternative hypothesis. Similarly, the Bayes Factors (null/alternative) for Stroke string similarity, Shared first stroke, and Stroke bigram similarity imply that null hypothesis is favored by a factor of 11.4, 10.2, and 9.5 respectively. These Bayes Factors provide strong evidence for null hypothesis, that is, a model without stroke-motor predictor (Jeffreys, 1961, Appendix B).

It is possible that we have not considered the correct theory of stroke motor similarity in the previous section. However, we believe that the theories describe above likely capture some aspects of the stroke motor plan of Chinese characters and therefore should at least be correlated with the correct theory of the stroke motoric features. The fact that all four theories only has minimal influence on discrimination time increases our confidence that representations of how to

write a character do not influence the speed with which it takes to distinguish characters in this task, for skilled readers.

Discussion

Across a range of analyses, we failed to find any evidence that knowledge of the stroke-motor plans for how characters are written influence the speed with which skilled readers of Chinese discriminate characters. There is a negligible effect of stroke-motor similarity on predicting the performance of the expert readers in this task and there is no difference between skilled readers who have experience writing these characters and the naïve group with no knowledge of how these characters are actually written. Knowledge about how characters were written does not help in this discrimination task.

The experiment had several positive manipulation checks built in. The fact that visual similarity predicted discrimination time for both groups supports the validity of our methods. The fact that semantic similarity played a role for expert readers confirmed that knowledge of Chinese characters, beyond just their visual form, influences their perception once literacy is acquired (Wiley et al., 2016). The Bayes Factors for varies theories of stroke-motor predictors make the null effect of stroke-motor similarity a compelling finding. Specifically, the results serve as clear evidence that, in an adult literate population, knowing how to write a Chinese character does not affect the process of recognizing that character.

The case study reported by Bi et al. (2009) already provided evidence against the strongest version of the writing-to-read hypothesis in skilled readers, as their patient convincingly lost knowledge of how to write characters without losing his ability to read. Our study provides evidence against a weaker version of this hypothesis, that is, knowledge of how a character is written has a modulatory influence on reading in literate readers. While several

previous lines of evidence have supported this modulatory effect, these studies all have clear methodological issues. In the present study, we adopted a method that can evaluate the role of visual and stroke-motor features of characters separately and found that visual features play a crucial factor in character discrimination but stroke-motor features do not.

Although our finding argues against the hypothesis that literate adults use knowledge about how Chinese characters are written when they read, it leaves open the possibility that knowledge of writing may be an integral component of learning to read (Tan et al., 2005). When Chinese children learn to read, they spend a lot of time copying the visually complex characters. This pairing of character recognition and character production during learning might result in a clear link in how these two skills are acquired, and early on, motoric information might aid in character recognition. This explains why strong association between writing and reading has been observed in children (Chan et al., 2006; McBride-Chang et al., 2011; Naka, 1998) as well as in adult learners (Cao et al., 2013; Guan et al., 2011; Guan, Perfetti & Meng, 2015; Naka & Naoi, 1995). However, our finding implies that once readers become experts at recognizing characters, our results suggest that stroke-motor information no longer influences character recognition.

It is worth noting that all participants in our expert reader group were Mandarin-English bilinguals who were currently in an English dominant environment. Many participants in our naïve group were also bilingual, though for the most part, if they spoke two languages both were written with an alphabetic script. While it is possible that experience with an alphabetic script has altered the way that our expert participants read Chinese such that they depend less on stroke motor processing, we think it is unlikely. All of our expert participants were educated in mainland China at least through high school. Even if we conducted the experiment in China, it

would be very hard to find pure monolingual Mandarin speakers with comparable age and education level since English is a mandatory course in high school. Therefore, we think that our sample provides insights into how literate adult readers process Chinese characters.

It is also worth noting that the task used in the present study, the visual judgment task, only requires superficial processing of the characters. The cognitive processes engaged in this task are likely different from those involved in normal reading, and these task effects might potentially modulate the effect of visual similarity and/or stroke motoric similarity on the discrimination time. The fact that semantic similarity influenced expert readers' performance demonstrates that this task relies on more than simply visual perception. However, converging evidence from a task that more closely resembles natural reading process would bolster the finding that knowledge of writing does not influence reading in skilled readers.

A strong connection between reading and writing has been argued play a more significant role in reading processes in Chinese compared to other languages (Tan et al., 2005). However, our findings suggest that the significant contribution of writing to reading, which has been repeatedly demonstrated in research on learning to read, is no longer present after literacy is obtained. This finding contributes to our understanding of what is universal and what is distinctive between the writing systems of the world; while the path to literacy might differ between languages, the fully mature, intact system is similar across different writing systems (Rueckl et al., 2015).

These results also speak to the link between reading and writing in general. The idea that recognizing written words and letters relates to the ability to produce written languages has been proposed in alphabetic languages as well (Longcamp et al., 2003; James & Gauthier, 2006; Cerni et al., 2016), and is in line with a broader debate within the cognitive sciences about the link

between action and perception. In speech perception, simulationist theories assume that to perceive a phoneme one needs to internally simulate the process of producing it (e.g. Liberman et al., 1967). The same type of theory applied to written language would predict that written word processing should be influenced by simulating the hand gesture required to write the words. However, the present study, along with the compelling case study reported by Bi and colleagues (2009), provides evidence against this theory. Even in the case of Chinese, the language that has been argued to be mostly likely to employ the writing-to-read strategy, knowledge of writing is not necessary for being able to read and stroke-motor knowledge does not even appear to have a modulatory impact on visual word processing in skilled readers.

References

- Bi, Y., Han, Z., & Zhang, Y. (2009). Reading does not depend on writing, even in Chinese. *Neuropsychologia*, 47(4), 1193-1199.
- Cerni, T., Velay, J. L., Alario, F. X., Vaugoyeau, M., & Longcamp, M. (2016). Motor expertise for typing impacts lexical decision performance. *Trends in Neuroscience and Education*, 5(3), 130-138.
- Chan, D. W., Ho, C. S. H., Tsang, S. M., Lee, S. H., & Chung, K. K. (2006). Exploring the reading-writing connection in Chinese children with dyslexia in Hong Kong. *Reading and Writing*, 19(6), 543-561.
- Cao, F., Vu, M., Chan, L., Ho, D., Lawrence, J. M., Harris, L. N., ... & Perfetti, C. A. (2013). Writing affects the brain network of reading in Chinese: A functional magnetic resonance imaging study. *Human Brain Mapping*, 34(7), 1670-1684.
- Flores d'Arcais, G. (1994) Order of strokes writing as a cue for retrieval in reading Chinese characters. *Eur. J. Cogn. Psychol.* 6, 337–355
- Frost, R. (2012). Towards a universal model of reading. *Behavioral and Brain Sciences*, 35(5), 263-279.
- Guan, C. Q., Liu, Y., Chan, D. H. L., Ye, F., & Perfetti, C. A. (2011). Writing strengthens orthography and alphabetic-coding strengthens phonology in learning to read Chinese. *Journal of Educational Psychology*, 103(3), 509.
- Guan, C. Q., Perfetti, C. A., & Meng, W. (2015). Writing quality predicts Chinese learning. *Reading and Writing*, 28(6), 763-795.
- James, K. H., & Gauthier, I. (2006). Letter processing automatically recruits a sensory-motor brain network. *Neuropsychologia*, 44(14), 2937-2949.

- Jeffreys, H. (1961). Theory of probability (3rd Edition). Oxford: Oxford University Press
- Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74(6), 431-461.
- Longcamp, M., Anton, J. L., Roth, M., & Velay, J. L. (2003). Visual presentation of single letters activates a premotor area involved in writing. *Neuroimage*, 19(4), 1492-1500.
- McBride-Chang, C., Lam, F., Lam, C., Chan, B., Fong, C. Y. C., Wong, T. T. Y., & Wong, S. W. L. (2011). Early predictors of dyslexia in Chinese children: Familial history of dyslexia, language delay, and cognitive profiles. *Journal of Child Psychology and Psychiatry*, 52(2), 204-211.
- Naka, M. (1998). Repeated writing facilitates children's memory for pseudocharacters and foreign letters. *Memory & cognition*, 26(4), 804-809.
- Naka, M., & Naoi, H. (1995). The effect of repeated writing on memory. *Memory & cognition*, 23(2), 201-212.
- Rueckl, J. G., Paz-Alonso, P. M., Molfese, P. J., Kuo, W. J., Bick, A., Frost, S. J., ... & Lee, J. R. (2015). Universal brain signature of proficient reading: Evidence from four contrasting languages. *Proceedings of the National Academy of Sciences*, 112(50), 15510-1551.
- Share, D. L. (2008). On the Anglocentricities of current reading research and practice: the perils of overreliance on an "outlier" orthography. *Psychological Bulletin*, 134(4), 584.
- Share, D. L. (2014). Alphabetism in reading science. *Frontiers in psychology*, 5.
- Shu, H., Xiong, H., Han, Z., Bi, Y., (2005). The selective impairment of the phonological output buffer: Evidence from a Chinese patient. *Behavioral Neurology*, 16, 179-189.
- Siok, W. T., Perfetti, C. A., Jin, Z., & Tan, L. H. (2004). Biological abnormality of impaired reading is constrained by culture. *Nature*, 431(7004), 71-76.

- Tan, L. H., Spinks, J. A., Perfetti, C. A., Fox, P. T., & Gao, J. H. (2003). Neural systems of second language learning are shaped by native language. *Neuroimage*, 13(6), 612.
- Tan, L. H., Spinks, J. A., Eden, G. F., Perfetti, C. A., & Siok, W. T. (2005). Reading depends on writing, in Chinese. *Proceedings of the National Academy of Sciences of the United States of America*, 102(24), 8781-8785.
- Tseng, S., Chang, L., & Wang, C. (1965). An informational analysis of the Chinese language: I. The reconstruction of the removed strokes of the ideograms in printed sentence-texts. *Acta Psychologica Sinica*, 10, 299-306.
- Wang, H. C., Schotter, E. R., Angele, B., Yang, J., Simovici, D., Pomplun, M., & Rayner, K. (2013). Using singular value decomposition to investigate degraded Chinese character recognition: evidence from eye movements during reading. *Journal of research in reading*, 36(S1), S35-S50.
- Wiley, R. W., Wilson, C., & Rapp, B. (2016). The Effects of Alphabet and Expertise on Letter Perception.
- Yan, G., Bai, X., Zang, C., Bian, Q., Cui, L., Qi, W., ... & Liversedge, S. P. (2012). Using stroke removal to investigate Chinese character identification during reading: Evidence from eye movements. *Reading and Writing*, 25(5), 951-979.
- Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychological Bulletin*, 131(1), 3-29.