

Article

Not peer-reviewed version

Learning Chinese Characters of Visual Similarity: The Effects of Presentation Style and Color Coding

[Junmin Li](#) , [Mengya Shi](#) , [Xin Wang](#) *

Posted Date: 1 September 2025

doi: 10.20944/preprints202508.2217.v1

Keywords: stroke count; Chinese character learning; pattern complexity; presentation style; color cue



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Learning Chinese Characters of Visual Similarity: The Effects of Presentation Style and Color Coding

Junmin Li ¹, Mengya Shi ² and Xin Wang ^{3,*}

¹ Hangzhou City University

² Beihang University

³ Macquarie University

* Correspondence: x.wang1@mq.edu.au

Abstract

This study examined how beginners learn two types of visually similar Chinese characters: those with identical strokes (e.g., 人 and 入) and those differing by an additional stroke (e.g., 日 and 白), while examining color coding and presentation effects. A total of 183 non-tonal native speakers with no prior experience of Chinese characters participated in the study. Using a 2x2x2 experimental design, the study assessed the influence of color coding (with vs. without), presentation style (single vs. paired characters), and stroke similarity (identical vs different). Results showed (1) Characters with stroke differences were learned more easily than identical-stroke characters; (2) Simultaneous character presentation enhanced discrimination of subtle stroke differences; but (3) Color coding slowed reaction times, suggesting visual overload. These findings demonstrate that perceptual similarity - not just complexity - impacts character learning difficulty. Pedagogically, the results support using paired character presentation while cautioning against excessive visual enhancements. The study provides empirical evidence for optimizing Chinese character instruction by balancing discriminability and cognitive load in beginning learners.

Keywords: stroke count; Chinese character learning; pattern complexity; presentation style; color cue

1. Introduction

Learning Chinese characters poses significant challenges for beginners, non-native learners, particularly those from alphabetic language backgrounds at the beginning level (Hu, 2010; Wu, 2014; Yang, 2018; Zhang & Li, 2010). Chinese characters use a logographic system where each character conveys meaning through visual elements rather than phonetics (Shen, 2013). Strong visual-spatial skills are vital for learning Chinese characters, as the ability to distinguish subtle visual differences is essential due to the large number of similar-looking radicals and characters in Chinese orthography (e.g., Chen, et al, 1996; McBride-Chang, et al., 2005, etc.). In this study, we aim to explore whether visual factors such as color coding, stroke structure and presentation style can assist non-native learners in acquiring visually similar Chinese characters by means of reducing cognitive load and thus enhancing memory retention.

The orthographic structure of Chinese characters can be analyzed hierarchically to its components: radicals and strokes. Radicals often serve as semantic or phonetic indicators. For example, the semantic radical 氵 (three dots/strokes to indicate *water*) is found in characters related to water, like 河 (river) and 海 (sea). Likewise, the phonetic radical 青 (pronounced as qing) in 清 (qing1), 情(qing2), 请(qing3), etc., produces similar pronunciation. They are essential for organizing and understanding Chinese writing, as they help group characters with similar meanings or pronunciations. Strokes are the basic components of Chinese radicals and characters. A stroke in Chinese writing is created with a single, continuous motion, often involving changes in direction (Anderson et al., 2013). Shu, et al. (2003) analyzed 2,570 characters taught in Chinese elementary schools and identified eight types of strokes that vary in length and spatial orientations. Strokes can

take various forms—vertical, horizontal, or diagonal; straight or slightly curved; and some may include a “hook” (Taylor & Taylor, 1995). For instance, the simple character 木 (wood) contains four types of strokes: vertical (丨), horizontal (一), left-slanting (丿), and right-slanting (㇏). Additionally, shorter stroke segments, such as dots, can vary in direction, as seen in the character 兴 (prosperous), as compared to the character 河 (river).

Characters are categorized into two types based on their structures: integral and compound characters. Integral characters contain only one radical, such as the character 人, which includes a single radical that conveys the meaning “person” (Zhang 1992: 31–32). Integral characters can be formed with as simple as one stroke (一) or as complex as multiple strokes (東). One dimension of visual complexity of Chinese characters is the relative positions of different strokes. For example, the same two strokes can create different characters by varying the stroke positions. For instance, both 人 (person) and 入 (enter) are composed of two strokes: 丿 and ㇏, but their differing arrangements result in distinct characters.

Similarly, in compound characters, the position of sub-lexical orthographic units (radicals) can change to form entirely different characters (Zhang 1992: 60–61). For example, 杏 (apricot) and 呆 (dumbness) both include the components 木 (wood) and 口 (mouth) arranged in a top-down pattern. When the positions of these components are reversed, a completely different character is formed. In another example, 晾 (air/sundry) and 景 (scenery) both contain the radical 日 (sun) and 京 (capital). However, the structural arrangement differs: 晾 follows a left-right pattern, while 景 adopts a top-down structure.

The phenomenon where Chinese characters share the same strokes but differ in stroke structure is intriguing, especially when considering how L2 beginners of alphabetic language backgrounds learn and identify these visual complexities to map to the meanings eventually. Previous studies on visual complexity typically focus on stroke counts, assuming that more strokes equate to greater visual complexity (e.g., Shu et al., 2013; Yu et al., 2018) when considering L2 learning of Chinese characters. Yet, learning characters not only involves learning the strokes composed of a character, but also identifying characters that look similar but differ in stroke structures. The latter is probably harder to learn, in particular, at the early stage of learning, given the subtlety. To our knowledge, no study has yet explored whether characters of the same stroke counts but presenting different stroke structures pose additional difficulties for Chinese L2 learners. Thus, the aim of the study is to examine how visual similarity (i.e., stroke structure) influences the early stage of L2 Chinese character learning. We focused on integral characters, because as compared to compound characters, they contain fewer strokes and simpler stroke patterns, normally defined as “spatially contiguous patterns of strokes” (Anderson et al., 2013, p. 45).

1.1. Visually Similar Words in Chinese Orthography

A basic examination of the Chinese writing system quickly highlights its stronger emphasis on visual-spatial processing compared to alphabetic orthographies that bear certain phonological relationships with words. That is, the alphabet usually indicates how words sound to a large or small extent, depending on the orthographic depth of the writing system. This is usually referred to Grapheme-Phoneme-Correspondence (GPC) rule. Thus, unlike alphabetic systems, which rely on a segmental and “atomistic” structure (Ho & Bryant, 1999), Chinese lacks this rule. In other words, unlike alphabetic languages, Chinese characters bear little relation to the way a character sounds. As a result, learning to read Chinese may be primarily driven by identifying visual/stroke structures, rather than relying heavily on phonological skills as Jin (2006) found that learners demonstrated the best retention of new characters when the character's configuration/radical was highlighted, rather than highlighting the stroke order or pinyin alone. The nonlinear spatial arrangements of strokes within Chinese characters, patterned as square-shaped structures, —and their high spatial frequency (visual characteristics of Chinese characters that involve fine details and rapid changes within a small spatial area. For example, characters with many small strokes or detailed radicals would be considered to have high spatial frequency) may also place greater demands on spatial memory

compared to the linear layout of alphabetic scripts (e.g., Wang, Koda & Perfetti, 2003). The development of sensitivity to the internal structure of Chinese characters is a gradual process. Gao and Meng (2000) found that beginning learners were significantly less accurate than advanced learners at distinguishing between visually similar characters (e.g., 会 vs. 公, 休 vs. 体) in a reading context. Beginners often struggle to notice subtle differences between similarly shaped characters and are not yet familiar with perceiving the details of stroke-based writing.

The literature contains various measures of visual complexity for Chinese characters, which aim to accurately reflect their complexity. For example, these measures include stroke frequency (Majaj et al., 2002; Zhang et al., 2007) and the “ink” measure (Pelli et al., 2006). Among these, Visual complexity is typically measured by the number of strokes (Liversedge et al., 2014; Yu et al., 2018). However, Chinese characters of the same number of strokes can represent entirely different words, such as 犬 (dog) and 太 (too). Although both characters consist of three strokes, their stroke patterns and spatial arrangements vary significantly. This complexity extends beyond stroke count to include differences in radicals, spatial arrangements, and even subtle variations in individual strokes. In a very similar way, characters like 大 (big) and 犬 (dog), where the latter is formed by adding a single stroke to the former, are normally considered as visually similar characters. These conditions highlight the intricate nature of Chinese orthography and thus raise an intriguing question: when comparing characters like 大 (big) vs. 犬 (dog), and 犬 (dog) vs. 太 (too), where the former pair differs by the addition of a stroke to the first character and the latter pair consists of two characters with the identical number of strokes but different patterns, which pair can be considered more visually complex or more difficult to learn?

Research on Chinese visual word recognition, particularly through the stroke number effect, remains limited. Jiang et al. (2020) conducted a study comparing native Chinese speakers (NS) and Chinese as a second language (CSL) learners using a lexical decision task involving disyllabic compound words (each containing two characters) and an equal number of nonwords. The results revealed that while both groups were influenced by word frequency and familiarity, only the CSL group demonstrated a stroke number effect. In a subsequent study, Jiang and Feng (2021) used single Chinese characters as stimuli and found that CSL learners again exhibited a stroke number effect in a lexical decision task. They argued that adult L2 Chinese learners process Chinese characters analytically, much like beginning readers, focusing on individual strokes or components rather than recognizing the overall character structure (Yeh et al., 2003).

If the number of strokes plays a key role in L2 character learning, we hypothesize that visually similar word pairs with an additional stroke will be more challenging to learn. However, there is currently no clear evidence indicating whether integral characters with the same number of strokes are equally difficult or harder to process compared to those with an additional stroke. Could English beginning readers of Chinese exhibit less-developed visual processing skills, due to lack of exposure, when distinguishing between integral characters with the same number of strokes and those with an additional stroke, such that this stroke number effect would not be as obvious as native readers?

1.2. Color Coding in Word Learning

Color coding has been shown to facilitate word learning by enhancing morphological awareness and supporting the decoding of linguistic structures (e.g., Tortorelli et al., 2021; Songsangkaew et al., 2023). For example, Tortorelli et al. (2021) reviewed instructional approaches to teaching code-related skills in English reading and highlighted the effectiveness of strategies such as separating words into phonemes, syllables, and morphemes, as well as teaching affixes explicitly. These methods of integrating colors into linguistic structures are instrumental in helping learners navigate the complexities of an alphabetic language. Supporting this, Songsangkaew et al. (2023) demonstrated that color-coding derivational and inflectional morphemes significantly improved morphological awareness, lexical inferencing ability, and reading comprehension among non-English major undergraduates. By visually distinguishing morphemes, color coding may provide

learners with a clearer understanding of word structures, aiding the acquisition of both vocabulary and reading skills.

If visual support, like color coding, can benefit learning alphabetical languages as evidenced in various studies, this strategy might be also helpful in learning Chinese characters which rely more on visual processing given their complex stroke structures. Though relevant research is limited and far from being conclusive. To start, some researchers consider lower-level structures of characters as a unit to learn, i.e., radicals, which facilitate L2 character learning by providing integrated structural features, unlike the complexity and variability of individual strokes (e.g., Huang & Chen, 1988). Radicals help to organize stroke-level information (Taft & Chung, 1999) and draw learners' attention to the internal structure of characters to some extent (Cao et al., 2013a). For example, Taft and Chung (1999) conducted a study in which participants learned 24 Chinese characters composed of two radicals paired with meaning (e.g., 咀-CHEW) under four different radical presentation conditions to see whether understanding the internal structure of Chinese characters aids novice learners in memorization. Participants were exposed to the character-meaning pairs three times and then presented with each character alone to recall its associated meaning. In a between-subject design, four groups received different types of radical instructions: **Radicals Before** (radicals pre-learned before exposure), **Radicals Early** (radicals informed during the initial exposure), **Radicals Late** (radicals identified during the third exposure), and **No Radical** (no instruction on radicals). The **Radicals Early** group performed best, suggesting that introducing radicals at the first exposure enhances learning.

If radical learning serves as the foundation of character learning as indicated by previous studies (Taft & Chung, 1999), strategies like color-coding to mark the boundaries of radicals within a character might be facilitatory in L2 character learning as well. However, existing empirical evidence does not seem to support this rationale. Among the limited studies investigating whether marking radicals with different colors—a common teaching technique—supports L2 orthographic and semantic learning, Hou and Jiang (2022) explored its impact on Chinese L2 learners of alphabetic language backgrounds. In a Latin square design, participants were required to learn characters presented across four distinct conditions: (a) radical markings with stroke animations, (b) no radical markings with stroke animations, (c) radical markings without stroke animations, and (d) neither radical markings nor stroke animations. They were shown 40 high-frequency fillers and 120 low-frequency target words, with each character displayed for 1,000 milliseconds. In the study, compound characters were divided into two radicals, either left and right or top and bottom, with the radicals marked in red and blue, respectively. For instance, in the character “苛,” the radical “艹” on top was marked in red, while the radical “可” on the bottom was marked in blue. After learning, participants completed character recognition and meaning-matching tests. The results showed that marking radicals increased participants' reaction times and reduced their accuracy in character recognition tests. Similarly, stroke-order animations also negatively affected recognition accuracy. These findings suggest that providing radical and stroke information may interfere, rather than facilitate character learning, as excessive visual information can increase cognitive load for L2 learners. Although strokes and radicals are key sub-character units in native Chinese speakers' character processing (e.g., Peng & Wang, 1997; Li et al., 2005), their utility for L2 learners in developing Chinese orthographic representations and orthography-semantics connections remains controversial. It is empirical to test whether color-coding a single critical stroke in a character would change the learning outcome.

In the present study, instead of applying color to the entire character (radical), we focused on highlighting one or two specific strokes that distinguish paired words from one another. Su and Samuels (2010) suggest that “beginning Chinese readers process characters in an analytic way”. To emphasize the critical differences between paired words, we applied color coding to the distinguishing strokes. For example, in word pairs with the same number of strokes, such as 犬 and 太, the differentiating dot was highlighted in color to provide critical input. Similarly, for words with a different number of strokes, such as 日 and 旦, the distinguishing stroke 一 was color-coded.

1.3. Presentation Style of characters

Given the pivotal role of generalization in cognition, there has been substantial research exploring the factors that facilitate it. Among them, the timing of instance presentation is particularly important. Research findings reveal a paradox: both simultaneous presentation (presenting instances at the same time), which allows for direct comparison (e.g., Gentner, Loewenstein, Thompson, & Forbus, 2009; Oakes & Ribar, 2005), and spaced presentation (presenting instances apart in time), which separates instances over time (e.g., Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006), have been shown to enhance generalization.

According to structure mapping theory (Gentner, 1983; Namy & Gentner, 2002; Thompson & Opfer, 2010), presenting exemplars at the same time draws attention to shared features, encouraging mental comparison, and the identification of commonalities within a category. Researchers suggest that simultaneous presentations help learners focus on the similarities between exemplars while reducing the short-term memory load required to compare them. As all exemplars are visible during this process, learners do not need to rely on memory to recall previously seen examples, which reduces cognitive demands. In addition, identifying these shared features helps learners generalize to new exemplars within the category. In short, simultaneous presentations support generalization by directing visual attention to key similarities and easing memory requirements.

On the other hand, other researchers suggest that both simultaneous and spaced learning facilitate generalization. For example, Vlach et al. (2012) investigated the effects of simultaneous, massed, and spaced presentations on 2-year-old children's performance in a novel noun generalization task. In this task, four new objects were introduced and labeled (e.g., "fep") using simultaneous, massed, or spaced presentation methods in the learning phase. In the simultaneous condition, all exemplars were presented at once, allowing children to visually compare them. In the massed condition, exemplars were presented one at a time in immediate succession, with less than 1 s between presentations. In a spaced condition, 30 s elapsed between each presentation. The massed and spaced presentation belonged to sequential presentation. In the test phase, four objects were shown, and children were asked to identify the target object (e.g., "Can you hand me the fep?"). For children in the immediate testing condition, the test was conducted right after the learning phase. For those in the delayed testing condition, the test took place 15 minutes later. Results showed that children in the simultaneous condition outperformed those in the other two conditions on immediate tests, while children in the spaced condition performed best after a delay (15 minutes). Overall, research suggests that visually comparing multiple exemplars presented simultaneously promotes better skills of abstraction, retention, and generalization compared to sequential (massed) presentations.

In their ongoing research program, Vlach et al. (2022) investigated how preschool-aged children learn science category exemplars under three schedules: simultaneous, massed, and spaced presentations. While no differences were observed between conditions during the immediate test, their first experiment showed that children in the spaced condition demonstrated the strongest generalization performance at a delayed post-test. However, further experiments revealed that spaced learning led to reduced visual attention and increased forgetting during the learning process.

Previous studies have investigated the effects of simultaneous, massed, and spaced learning on students' performance in science education. The presentation style of literacy development has not been extensively studied. While spaced presentation has been found to reduce visual attention and increase forgetting (Vlach et al., 2022), in the present study, we aim to adopt simultaneous and massed presentation methods to explore their impact on word learning. This study aims to address this gap by investigating how visual stimuli, specifically presenting characters in sequence or in pairs, impact learners' ability to acquire and retain Chinese characters.

In summary, this study aims to address the following questions:

1. Do L2 beginning Chinese learners learn visually similar characters of identical number of strokes more effectively than those of different stroke numbers (word pairs differ by one stroke), or vice versa?

2. Does the use of color coding to highlight the critical strokes that differentiate visually similar characters facilitate L2 character learning?
3. Does simultaneous presentation (2 characters presented together) enhance L2 character learning more effectively than massed presentation (1 character presented individually)?

To address the above questions, we designed a study of four different experimental conditions, namely, "One Character, Color-Coded", "One Character, No Color", "Two Characters, Color-Coded", "Two Characters, No Color", such that we can test how L2 learners learn Chinese characters in different contexts. L2 learners were expected to learn both the stroke structures and the meanings of characters. To investigate this, we employed a $2 \times 2 \times 2$ experimental design of three dependent variables: (1) Critical strokes cued with vs. without color, (2) Presentation with 1 vs. 2 characters, and (3) Characters with identical vs. different stroke numbers. Color-coding and Presentation Style were between-subjects factors, while Stroke Number was a within-subjects factor, as all learners were exposed to characters with both the same and different stroke numbers. Participants received the training first, followed by an immediate test to see how well they could recognize the meaning of the words. The next day, a delayed test was conducted to measure how well they retain the memory of the words.

2. Materials and Methods

2.1. Participants

The study recruited 202 undergraduate students aged from 19 to 24 (Mean =21.47, SD =1.39) at XXX University. 19 participants, who reported having some or extensive knowledge of Chinese characters, were excluded from the final analysis that included a total of 183 participants. They were native speakers of English, with no prior knowledge or minimal prior exposure to Chinese. They were randomly and equally assigned to one of the four experimental conditions.

2.2. Materials

The stimuli consisted of 48 Chinese characters selected for their visual simplicity and structural variability. These characters were grouped into 24 pairs: 12 pairs of identical strokes but different stroke structures (e.g., 人 and 入) and 12 pairs only differing by one stroke (e.g., 日 and 旦). Specifically, in the identical-stroke condition, characters like 人 and 入 share identical strokes but differ in their orientations, forming distinct characters. In the different-stroke condition, characters like 日 and 旦 differ by one single stroke. Each pair was presented alongside an English translation to facilitate learning and recognition. The average stroke number for the identical-stroke condition was 3.5 (SD = 0.9) while for the different-stroke condition, the average stroke number was 3.92 (SD = 0.79) and 4.92(SD = 0.79) respectively. There was no significant difference in stroke number between these two conditions ($t = 0.001$, $p = .999$).

To measure visual complexity of Chinese characters, we compared the four metrics Wang et al. (2014) proposed, namely stroke count, ink density, stroke frequency, and perimetric complexity, among which, the perimetric complexity was the one that could best be applied to both alphabets and Chinese characters. This metric was defined as the perimeter square of a symbol, divided by the 'ink' area (Pelli, Burns, Farell, & Moore-Page, 2006). The width and height of the characters were the same and they were treated as an image with fixed size. Each character was stored as a binary image, with the strokes in black background in white. The ink density is defined by the ratio of the number of black and the total number of pixels in a character image. An independent samples t-test was conducted to compare perimetric complexity between the identical-stroke and different-stroke conditions. There was no significant difference ($t(46)=0.39$, $p = 0.969$) between the identical-stroke condition ($M = 0.51$, $SD = 0.16$, $n = 24$) and the different-stroke condition ($M = 0.59$, $SD = 0.11$, $n = 24$).

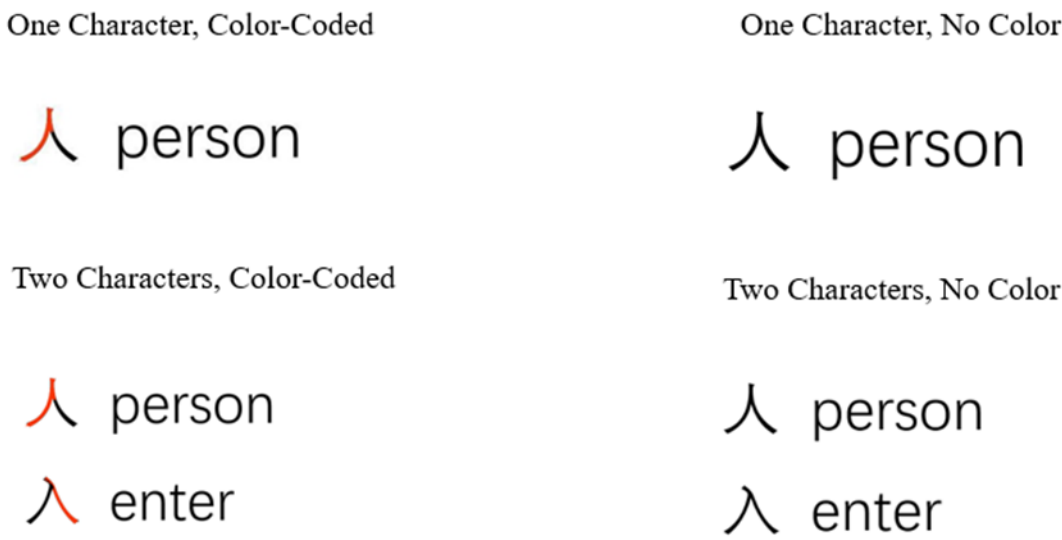
2.3. Design and Procedure

At the initial stage of learning Chinese characters, a common approach is to introduce new characters alongside their *pinyin* (a pronunciation cue) and English equivalent. Research by Lee and Kalyuga (2011a) found that presenting *pinyin* with new words improved word retention for learners who had achieved automatic *pinyin* reading but provided little benefit for those who had not yet mastered the *pinyin* system. Additionally, the widely used horizontal layout for presenting *pinyin* can impose a high cognitive load and hinder character learning, particularly for beginners (Lee & Kalyuga, 2011b). To avoid potential confounding factors, the present study excluded *pinyin* and phonological information during the training phase.

This study employed a 2x2x2 design with two between-subject variables: presentation style (massive presentation vs simultaneous presentation, i.e., one character vs. two characters) and color coding (with vs. without color cues). Stroke number (identical-stroke vs. different-stroke) was treated as a within-subject variable. Thus, in total, there were 4 training conditions.

Participants were randomly assigned to one of the four training conditions presented in **Figure 1**:

Figure 1. Four Training conditions



This study was conducted online using the Gorilla Experiment Builder (www.gorilla.sc; Anwyl-Irvine et al., 2021). First, a questionnaire was designed to gather participants’ Chinese learning background. The experiment comprised of two phases: a training phase and a testing phase. During the training phase, participants were exposed to Chinese characters based on their assigned conditions. They were instructed to observe each character, count the number of strokes, and enter the stroke number before advancing to the next trial. No feedback was given during this phase. Each character was displayed for a fixed duration, and participants pressed a button to proceed to the next trial.

The testing phase involved two recognition tests: an immediate test on Day 1 and a delayed test 24 hours later Day 2. Both tests used a two-alternative forced-choice task, where participants were shown an English translation and asked to identify the corresponding Chinese character from two options. Each test included 48 trials including all the new characters learned during the training phase. The study followed a between-subjects design, ensuring that each participant experienced only one training condition. Figure 2 illustrates the training phase of the “Two Characters, Color-Coded” condition, while the testing phase was the same for all conditions.

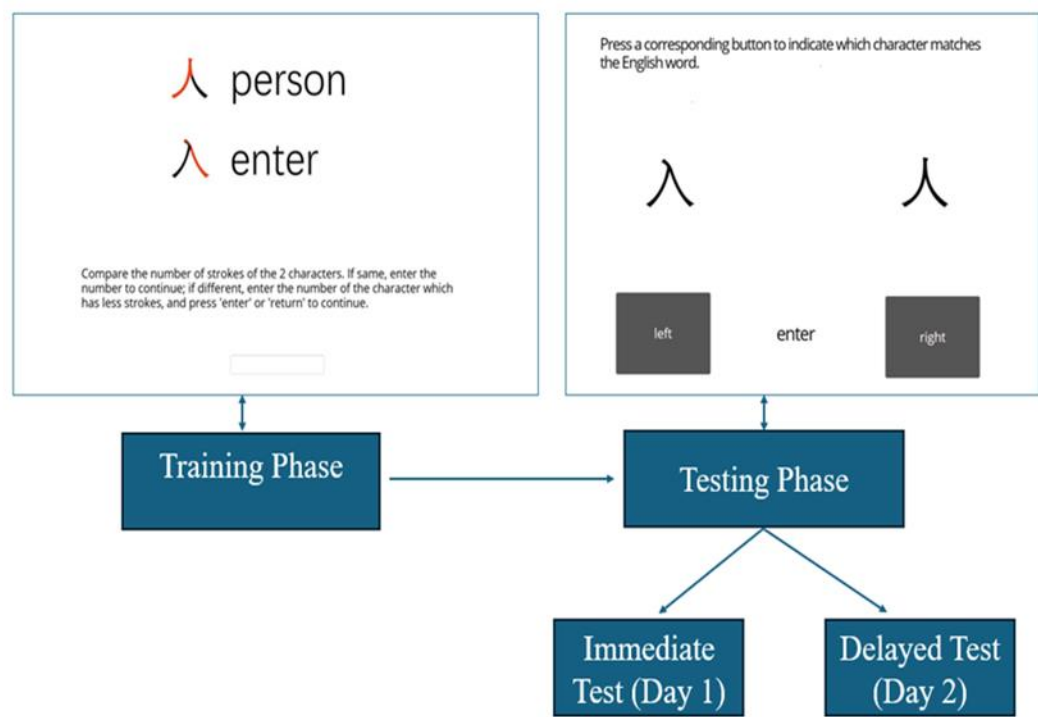


Figure 2. The design schemata of the experiment.

2.4. Statistical Analysis

The statistical analyses were performed using the lme4 package (Bates et al., 2015) within the R statistical software environment, maintained by the R Core Team (2020). In these analyses, accuracy and reaction times (RT) for each trial were treated as dependent variables. A linear mixed-effects model was applied, with stroke number, color cue, and presentation type as fixed-effect predictors. The coding for these predictors was as follows: different- stroke (0.5) vs. identical-stroke (-0.5), presence of a color cue (C) (0.5) vs. absence of a color cue (NC) (-0.5), and massed presentation (one character) (0.5) vs. simultaneous presentation (two characters) (-0.5). The random effect structure included varying intercepts and slopes for both participants and items, with no initial correlation assumed between these intercepts and slopes. In the event of convergence issues, the models were adjusted by removing random terms sequentially, starting with the term having the smallest value (Barr et al, 2013).

Before conducting the analysis, we used a density plot to visualize response times (RTs) and identify potential outliers. RTs below 300 ms or above 1,0000 ms were considered outliers and were excluded from the dataset. Based on the results of Box-Cox tests (Box & Cox, 1964), a logarithmic transformation was applied to the RTs, and the transformed values were used as the dependent variable. The initial model was fitted to the complete dataset, and subsequently, following the residual trimming method proposed by Baayen et al. (2008) and Baayen & Milin (2010), data points with residuals exceeding 2.5 standard deviations were removed. The final results are based on the refined model. For response accuracy, generalized linear mixed-effects models were used, with accuracy coded as 1 (correct) or 0 (error) as the dependent variable. These models were structured similarly to those used in the RT analysis. (Do we need to specify the data availability statement here?)

3. Results

In this analysis, we examine the main effects of three variables—visual complexity (stroke number), presentation style (simultaneous or massed), and color coding—as well as their interactions

in both ER (error rate) and RT (reaction time). To address the research questions, participants’ performance on both the immediate and delayed tests were evaluated using linear mixed models.

3.1. Immediate test analysis

3.1.1. Accuracy analysis

The accuracy rates of the four training conditions, as well as their SD, were presented in **Table 1**.

Table 1. Mean Accuracy in % (SD) of the Four Training Conditions in Immediate Tests Analysis.

	Identical Stroke	Different Stroke
One Character, Color-Coded	0.594 (0.491)	0.683 (0.465)
One Character, No Color	0.586 (0.493)	0.666 (0.472)
Two Characters, Color-Coded	0.586 (0.493)	0.627 (0.483)
Two Characters, No Color	0.608 (0.488)	0.617 (0.486)

A final mixed-effects logistic regression model was adopted to analyze the accuracy data, including the correct response as a dependent variable, with a random intercept, as well as a random slope for stroke number by subject and the interaction between colour cue and presentation style by item, as in (1)

Mixed(correct ~ color * presentation * stroke_num + (stroke_num | subj) + (color * presentation | itemN)).
(1)

The model output indicated a significant main effect of stroke number ($\beta = 0.14$, $SE = 0.05$, $z = 2.64$, $p = .008$), which suggested that the retention of the different-stroke pairs was much better than that of the identical- stroke pairs. The main effect of the color coding was not significant ($\beta = 0.01$, $SE = 0.05$, $z = 0.28$, $p = .78$). Neither was the main effect of presentation style ($\beta = 0.05$, $SE = 0.06$, $z = 0.82$, $p = .41$). There was a significant interaction between presentation style and stoke number ($\beta = 0.07$, $SE = 0.03$, $z = 2.11$, $p = .035$).

Further pairwise analyses using emmeans revealed that the accuracy rate for characters with different strokes was significantly higher than that for characters with identical strokes in both the One Character, Color-Coded condition ($\beta = 1.56$, $SE = 0.22$, $z = 3.08$, $p = .002$) and the One Character, No Color condition ($\beta = 1.49$, $SE = 0.22$, $z = 2.71$, $p = .007$). However, no significant differences were observed between the accuracy rates for characters with different strokes and those with identical strokes in either the Two Characters, Color-Coded condition ($\beta = 1.25$, $SE = 0.19$, $z = 1.49$, $p = .14$) or the Two Characters, No Color condition ($\beta = 1.07$, $SE = 0.15$, $z = 0.45$, $p = .65$).

The post-hoc pair-wise comparison analysis showed that the massed presentation condition, where the characters were presented one by one, produced more accurate results in learning different-stroke pairs than the identical-stroke pairs ($\beta = 1.52$, $SE = 0.19$, $z = 3.34$, $p < .001$). In contrast, the simultaneous presentation condition showed no difference between the identical-stroke and different- stroke pairs ($\beta = 1.15$, $SE = 0.14$, $z = 1.15$, $p = .250$).

3.1.2. RT Analysis

Reaction times (RTs) for correct responses were analyzed, comprising of a total of 5,327 trials. Outliers of reaction times greater than 10,000 ms or less than 300 ms were excluded, accounting for 3.17% of the data, resulting in 5,158 trials in final analyses. An different 1.9% trials were removed as these data points were standardized residuals beyond ± 2.5 standard deviations. Mean RT of the four conditions were presented in **Table 2**.

Table 2. Mean RT in ms (SD) of the Four Training Conditions in Immediate Tests Analysis.

	Identical Stroke	Different Stroke
One Character, Color-Coded	3283 (1782)	2935 (1590)
One Character, No Color	2647 (1645)	2483 (1507)
Two Characters, Color-Coded	2671 (1605)	2611 (1496)
Two Characters, No Color	2791 (1800)	2586 (1587)

A liner mixed-effect model was fitted with RT data, including as a dependent variable and a random intercept and slope for stroke number by subject and a random intercept for items as in (2).

Mixed(logrt ~ color * presentation * stroke_num + (1 + stroke_num1 | subj) + (1 | itemN)) (2)

The final model showed a significant colour cue effect ($\beta=0.09$, $SE=0.04$, $z=2.41$, $p=.017$), which indicated that colour coded characters were processed slower than those of no colour cues. We also observed a significant main effect of stroke number ($\beta=-0.03$, $SE=0.01$, $z=-2.48$, $p=.017$), which showed that different stroke pairs were quicker to response than those of identical stroke, consist with the accuracy data. There was no main effect of presentation style ($\beta=0.04$, $SE=0.04$, $z=1.19$, $p=.23$). Neither was there any interaction between any two variables among the three (all p 's>0.12).

Pairwise comparisons revealed that RTs for characters with different strokes were significantly faster than those with identical strokes in the One Character, Color-Coded condition ($\beta=-0.12$, $SE=0.03$, $z=-3.48$, $p<.001$). In contrast, no significant differences in RTs were observed between characters with different strokes and those with identical strokes in the other three conditions (all p 's > 0.13).

In the massed presentation format, the colored-cued condition significantly hindered responses compared to the no-color condition ($\beta=0.26$, $SE=0.11$, $z=2.50$, $p=.01$). Conversely, in the simultaneous presentation format, there was no significant difference in responses between the colored and no-color conditions ($\beta=0.10$, $SE=0.11$, $z=0.92$, $p=.36$).

3.2. Delayed Test Analysis

Among those who participated in the Day 1 training sessions, a total of 191 participants took part in Day 2, delayed tests. Again, participants who reported knowledge of Chinese characters were excluded, resulting in 172 participants included in the final analysis. The analysis procedure followed the same as that in the immediate tests (Day 1) analysis.

3.2.1. Accuracy Analysis

The accuracy rates of the four conditions were shown in **Table 3**.

TABLE 3. Mean Accuracy in % (SD) of the Four Training Conditions in Delayed Tests Analysis.

	Identical Stroke	Different Stroke
One Character, Color-Coded	0.600 (0.490)	0.688 (0.463)
One Character, No Color	0.573 (0.495)	0.611 (0.488)
Two Characters, Color-Coded	0.588 (0.492)	0.614 (0.487)
Two Characters, No Color	0.619 (0.486)	0.624 (0.485)

A logistic mixed-effect model was fitted with a random intercept and slope for stroke number by subject and presentation by item in (3).

$$\begin{aligned} \text{mixed}(\text{correct} \sim & \text{color} * \text{presentation} * \text{stroke_num} + (\text{stroke_num} | \text{subj}) \\ & + \text{color} * \text{presentation} | \text{itemN}) \end{aligned} \tag{3}$$

The model output showed a marginal main effect of stroke number ($\beta = 0.10$, $SE = 0.05$, $z = 1.88$, $p = .06$), but neither main effect of color ($\beta = 0.04$, $SE = 0.05$, $z = 0.82$, $p = .42$) or effect of presentation style ($\beta = 0.02$, $SE = 0.06$, $z = 0.30$, $p = .76$). In addition, a marginal significant interaction between stroke number and presentation style was observed ($\beta = 0.06$, $SE = 0.03$, $z = 1.86$, $p = .062$).

Pairwise analysis indicated that the accuracy rate for characters with different strokes was significantly higher than that for characters with identical strokes in the One Character, Color-Coded condition ($\beta = 1.57$, $SE = 0.24$, $z = 2.91$, $p = .004$). However, no significant differences in accuracy rates were found between characters with different strokes and those with identical strokes in the other three conditions (all p 's > .20).

Post hoc analyses showed that the massed presentation condition produced better results in learning different-stroke pairs than identical stroke pairs ($\beta = 1.37$, $SE = 0.18$, $z = 2.47$, $p = .014$). However, the simultaneous presentation condition showed no difference between identical and different stroke pairs ($\beta = 1.09$, $SE = 0.13$, $z = 0.71$, $p = .476$). These patterns were consistent with the results from the immediate tests.

3.2.2. RT analysis

Using the same data trimming procedure as shown in the immediate test, we removed 4.61% outliers and included 4,736 trials for final analyses. An additional 2.47% outliers were excluded to remove datapoints as standardized residuals beyond ± 2.5 standard deviations. The mean RTs of the four conditions were presented in **Table 4**.

TABLE 4. Mean RT in ms (SD) of the Four Training Conditions in Delayed Tests Analysis.

	Identical Stroke	Different Stroke
One Character, Color-Coded	2981 (1946)	2706 (1743)
One Character, No Color	2300 (1456)	2087 (1337)
Two Characters, Color-Coded	2771 (1638)	2338 (1307)
Two Characters, No Color	2533 (1597)	2347 (1476)

A linear mixed-effects model was fitted, including a random intercept and slope for stroke number by subject and a random intercept by item in (4).

$$\begin{aligned} \text{mixed}(\text{logrt} \sim & \text{color} * \text{presentation} * \text{stroke_num} + (\text{stroke_num} | \text{subj}) + \\ & (\text{presentation} | \text{itemN}) \end{aligned} \tag{4}$$

Similar to the Immediate test analysis, the model revealed a main effect of colour coding ($\beta = 0.11$, $SE = 0.04$, $t = 2.78$, $p = .006$), showing that color coding led to slower reaction times. Additionally, there was a main effect of stroke number ($\beta = -0.04$, $SE = 0.01$, $t = -3.39$, $p = .001$), showing that different-stroke pairs resulted in faster reaction times. Presentation style did not show a main effect ($\beta = 0.01$, $SE = 0.04$, $t = 0.28$, $p = .78$). No interaction was observed between any two variables (all p 's > 0.14).

Further pairwise analyses revealed that characters with different strokes were significantly quicker to recognize than those with identical strokes in the following conditions: One Character, Color-Coded ($\beta = -0.07$, $SE = 0.03$, $z = -2.04$, $p = .04$); One Character, No Color ($\beta = -0.08$, $SE = 0.03$, $z = -2.29$, $p = .02$); and Two Characters, Color-Coded ($\beta = -0.13$, $SE = 0.04$, $z = -3.46$, $p < .01$). However, this difference was not significant in the Two Characters, No Color condition ($\beta = -0.05$, $SE = 0.04$, $z = -1.35$, $p = .18$).

In the massed presentation condition, color coding significantly increased RTs compared to no color coding ($\beta = -0.35$, $SE = 0.11$, $z = -3.03$, $p < .01$). However, in the simultaneous presentation condition, no significant difference in RTs was observed between color coding and no color coding ($\beta = -0.11$, $SE = 0.12$, $z = -0.92$, $p = .36$).

4. Discussion

This study aims to determine whether visual complexity of characters with identical stroke number or different stroke number poses the same level of challenge to L2 Chinese learners and understand an optimal way to present input (i.e., Chinese characters) by empirically testing two strategies: color coding and presentation style. Specifically, we aim to understand whether and how these strategies can be applied to visually similar Chinese characters during the early stage of character learning. We employed an online training and testing approach to examine how L2 Chinese beginning learners retain visually similar words under different training conditions: with or without color coding and presented either simultaneously or in a massed format. The current study also compared the retention of visually similar word pairs of identical strokes to those of different strokes. Results from both Immediate tests and Delayed tests revealed a consistent pattern: word pairs of different strokes were better retained than word pairs of identical strokes, as evidenced by both accuracy and RT analyses. The presentation style showed no main effect either in Immediate tests or Delayed tests analysis. However, in the accuracy analysis, the interaction effects showed that words with word pairs of a different-stroke number were recognized more accurately than identical-stroke words in the massed presentation condition, but not in the simultaneous condition, which suggested that simultaneous presentation enhanced the learning of identical-stroke words more than different-stroke words and leveled the difference between identical-stroke characters and different-stroke characters. This is because simultaneous presentation encourages learners to compare the stroke structures of identical-stroke words, which is consistent with the structure mapping theory (Gentner, 1983; Namy & Gentner, 2002; Thompson & Opfer, 2010). Color coding showed no main effects in accuracy analysis in either immediate tests or delayed tests. However, the RT analysis showed that the use of color cues slowed down word recognition in both immediate tests and delayed tests.

To address the first research question, we examined which type of visually similar word pairs—identical-stroke or different-stroke—is easier to learn. By logic, the addition of a stroke increases the overall stroke count, which, according to the conventional measure of visual complexity, should make these words more difficult to learn and retain. However, the results contradicted this expectation, revealing that words with identical stroke numbers were harder to remember. Shen (2013) suggests that beginning L2 learners often struggle to recognize the internal structures of characters and require sufficient time and practice to segment compound characters into their constituent parts effectively. The current study further explores how well beginners can identify the stroke counts and internal structures of characters, through exposure and practice such that they could visualize and familiarize themselves with basic character components. As Reder et al. (2016) have noted, novice learners can accurately decode characters with minimal training, especially in controlled laboratory settings. This ability may enhance their capacity to differentiate characters with different stroke counts, as these characters share most of their stroke structure except for one stroke. In contrast, for words with identical strokes, the stroke structures are different even though they share the same number of strokes. Our results indicated that word pairs with identical strokes, but different stroke orientations were harder to learn than those with additional-stroke differences. This difficulty might stem from the more subtle differences in identical-stroke words, which require

higher visual processing skills. As Wang, Koda, and Perfetti (2003) have highlighted, the complex, square-shaped structure and high spatial frequency of Chinese characters place significant demands on spatial memory. This may impose an additional processing burden when learning and retaining identical-stroke words compared to different-stroke words. McBride-Chang et al (2005) found a positive correlation between visual-spatial abilities and the acquisition of Chinese characters. Thus, visual working memory is crucial for learning Chinese characters, as it allows for the temporary storage of approximately three to four visual objects (e.g., Luck, 2008; Luck & Vogel, 2013). This implies that the average capacity for remembering items or features is quite limited. In our study, the characters selected averaged 3-4 strokes, which falls within the capacity of visual working memory. According to Zimmer & Fisher (2020), even native Chinese speakers rely on visual working memory rather than semantic information when memorizing trained characters and pseudo-characters whose one radical were altered from real characters. Zimmer and Fisher (2020) found in Experiment 4 that highly conceptually similar distractors did not increase memory errors, while visually similar distractors weakened memory performance. It can be inferred that beginners would primarily use visual-spatial memory to memorize Chinese characters. The current study is the first to show that characters of the same number of strokes, but different spatial orientations are more challenging to learn compared to visually similar characters of different stroke numbers.

Second, presentation style showed no main effects. However, the interaction effect indicated that simultaneous presentation facilitated learning identical stroke characters more than the different stroke characters. That the highest accuracy rate was observed in the massed presentation of characters with different strokes, whereas the lowest accuracy rate was in the simultaneous presentation with identical strokes. To elaborate, when characters were presented one by one, learning different-stroke pairs resulted in higher accuracy compared to identical-stroke pairs. Conversely, in the simultaneous presentation condition, no significant difference was found between identical-stroke and different-stroke pairs. This suggests that the simultaneous presentation format helped reduce the difficulty associated with learning characters of identical strokes and thus reduced the difference between identical-stroke characters and different-stroke characters. As previously noted, simultaneous presentation allows learners to directly compare the stroke patterns of two characters, enhancing learning efficiency. In contrast, massed presentation may increase cognitive load by requiring learners to process two elements presented in random order and integrate them later, which can strain working memory (Chandler & Sweller, 1992, 1996; Kalyuga et al., 1999, 2000). Shen (2005, p. 56) identified learners' common strategies such as "paying attention to graphic structures" and "visualizing the graphic structure of the character". This study provides strong evidence that simultaneous presentation significantly improves beginning learners' ability to learn the identical-stroke words via comparing the graphic structures. Next, in the massed presentation condition, the RT analysis revealed that colored characters led to slower responses compared to the no-color condition. In contrast, during simultaneous presentation, there were no significant differences in RT between the colored and no-color conditions, irrespective of whether the tests were immediate or delayed. The analysis also suggested that while color exerted a slight influence on response times, simultaneous presentation seemed to mitigate the effects of color on response times. Additionally, the results of our study were consistent with a large number of studies that have shown that viewing multiple instances simultaneously enhances generalization (e.g., Gentner, et al., 2009; Oakes & Ribar, 2005; Vlach et al., 2012). For example, Vlach et al. (2012) found that 2-year-old children who were exposed to stimuli simultaneously outperformed their peers who were in spaced or massed conditions on a novel noun generalization task when tested immediately afterward. The researchers suggested that the simultaneous presentation condition likely reduced forgetting and memory load, allowing learners to engage more readily in comparative mental processes during learning. Since all instances were visible throughout the learning phase, children in the simultaneous condition didn't need to mentally retrieve previous instances they had observed. In summary, simultaneous presentation might help beginners retain the memory of visually similar characters, but it is inconclusive to draw this conclusion because the main effect was not significant.

Finally, we applied a color cue to highlight the stroke distinguishing the paired-characters and found that the color cue increased recognition reaction times in both the immediate tests and delayed tests. The design of the color cues aimed to guide learners’ visual focus towards the differential stroke structure within characters. Nonetheless, the use of color might inadvertently introduce a negative effect by acting as an additional distraction. The intricate visual information could potentially compel participants to engage in an extra process of matching radicals with colors, resulting in a divided attention effect. As a result, this could adversely affect learners' character acquisition process. These findings align with Hou and Jiang (2022), who reported that radical marking using color or animation increased RT and reduced recognition accuracy in character recognition tasks. They suggested that providing radical and stroke information might interfere, rather than facilitate character learning. Our study supports this view, indicating that excessive visual information introduced during the learning process may increase cognitive load for L2 learners, thereby reducing learning effectiveness (Baddeley, 1992). When learners process multiple elements of information simultaneously, they must divide their attention across all elements. This simultaneous presentation of information can result in the split attention effect (Chandler & Sweller, 1991, 1992; Owens & Sweller, 2008), further straining working memory capacity. Thus, the color cue did not facilitate learning characters in our study, which may have important pedagogical implications for L2 instruction.

5. Conclusion

This study uncovered an intriguing finding: characters with identical strokes were more challenging to learn than those with different strokes, despite the difference being just one stroke. When learning such word pairs, simultaneous presentation could assist L2 beginners in comparing the stroke structures, as it reduced the differences of RT and accuracy rate between identical-stroke characters and different-stroke characters. However, color coding may interfere with the learning process. When guiding learners to analyze the internal components of Chinese characters, such as strokes and radicals, it is crucial to minimize the cognitive load imposed by instructional materials and educational software.

Author Contributions: removed for peer review.
Funding: removed for peer review
Institutional Review Board Statement: removed for peer review
Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.
Data availability statements: Research data and the detailed analysis scripts for this article were available at https://osf.io/jt6f4/?view_only=7dbe6820f2ae4fecbbac5782723ec5ee
Conflicts of Interest: The authors declare no conflict of interest.

Appendix

identical-stroke characters		different-stroke characters	
未	末	目	自
犬	太	木	禾
天	夫	日	旦
手	毛	今	令
刀	力	火	灭
土	士	月	用
干	千	白	百
开	井	中	申
午	牛	尸	户

元	无	从	丛
己	已	王	主
人	入	了	子

References

Anderson, R. C., Ku, Y.-M., Li, W., Chen, X., Wu, X., & Shu, H. (2013). Learning to See the Patterns in Chinese Characters. *Scientific Studies of Reading*, 17(1), 41–56. <https://doi.org/10.1080/10888438.2012.689789>

Baddeley, A. (1992). Working memory. *Science*, 255, 556-559. <https://doi.org/10.1126/science.1736359>

Chandler, P., and Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition & Instruction*, 8, 293–332. https://doi.org/10.1207/s1532690xci0804_2

Chandler, P., and Sweller, J. (1992). The split attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, 62, 233–246. <https://doi.org/10.1111/j.2044-8279.1992.tb01017.x>

Chen, Y. P., Allport, D.A. & Marshal, J. C. (1996). What Are the Functional Orthographic Units in Chinese Word Recognition: The Stroke or the Stroke Pattern? *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, 49(4), 1024–1043. <https://doi.org/10.1080/713755668>

Chung, K. K. H. (2002). Effective use of hanyu pinyin and English translation as extra stimulus prompts on learning of Chinese characters. *Educational Psychology*, 22, 149–64.

Chung, K. K. H. (2007). Presentation Factors in the Learning of Chinese Characters: The order and position of Hanyu pinyin and English translations. *Educational Psychology (Dorchester-on-Thames)*, 27(1), 1–20. <https://doi.org/10.1080/01443410601061306>

Davis, M. H., Di Betta, A. M., Macdonald, M.J. E., & Gaskell, M. G. (2009). Learning and consolidation of novel spoken words. *Journal of Cognitive Neuroscience*, 21(4), 803–820. <https://doi.org/10.1162/jocn.2009.21059.Learning>

Davis, M. H.,&Gaskell, M. G. (2009). A complementary systems account of word learning: Neural and behavioural evidence. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 364(1536), 3773–3800. <https://doi.org/10.1098/rstb.2009.0111>

Donderi, D. C. (2006). Visual Complexity. *Psychological Bulletin*, 132 (1), 73-97.

Gao, L, & Ling, M. (2000). The role of the phonetic code and orthographic code of Chinese character recognition by foreign learners 外国留学生汉语阅读中音，形信息对汉字辨认的影响. *Chinese Teaching in the World*, 4, 67–76.

Gentner D, Loewenstein J, Thompson L, & Forbus KD. (2009). Reviving inert knowledge: Analogical abstraction supports relational retrieval of past events. *Cognitive Science: A Multidisciplinary Journal*. 33:1343–1382. <https://doi.org/10.1111/j.1551-6709.2009.01070.x>

Ho, C. S.-H., & Bryant, P. (1999). Different visual skills are important in learning to read English and Chinese. *Educational and Child Psychology*, 16, 4–14. <https://doi.org/10.53841/bpsecp.1999.16.4.4>

Hou, F., & Jiang, X. (2022). Interference effects of radical markings and stroke order animations on Chinese character learning among L2 learners. *Frontiers in Psychology*, 13, 783613–783613. <https://doi.org/10.3389/fpsyg.2022.783613>

Huang, Z., & Chen, Y. (1988). *Introductory Chinese: Reading comprehension*. Beijing: Sinolingua.

Jackson, N., Everson, M., & Ke, C. (2003). Beginning readers’ awareness of the orthographic structure of semantic-phonetic compounds: lessons from a study of learners of Chinese as a foreign language. *Reading Development in Chinese Children*, ed. by Catherine McBride-Chang and Hsuan-chih Chen, 141–53. London: Praeger Publishers.

Lee, C. H., & Kalyuga, S. (2011a). Effectiveness of on-screen pinyin in learning Chinese: an expertise reversal for multimedia redundancy effect. *Computer in Human Behavior*, 27, 11–5.

Lee, C. H., & Kalyuga, S. (2011b). Effectiveness of Different Pinyin Presentation Formats in Learning Chinese Characters: A Cognitive Load Perspective. *Language Learning*, 61(4), 1099–1118. <https://doi.org/10.1111/j.1467-9922.2011.00666.x>

Li, L., Liu, H., and Liu, X. (2005). Effects of characters construction on basic processing unit of Chinese character recognition(汉字结构对汉字识别加工的影响). *Psycholog. Exploration(心理学探新)* 25, 23–27.

- Liversedge, S. P., Zang, C., Zhang, M., Bai, X., Yan, G., & Drieghe, D. (2014). The effect of visual complexity and word frequency on eye movements during Chinese reading. *Visual Cognition*, 22, 441–457. doi:10.1080/13506285.2014.889260
- Luck, S. J. (2008). "Visual short-term memory," in *Visual Memory*, eds S. J. Luck, and A. Hollingworth, (Oxford: Oxford University Press), 43–85.
- Luck, S. J., and Vogel, E. K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends Cogn. Sci.* 17, 391–400. <https://doi.org/10.1016/j.tics.2013.06.006>
- McBride-Chang, C., Chow, B. W. Y., Zhong, Y., Burgess, S., and Hayward, W. G. (2005). Chinese character acquisition and visual skills in two Chinese scripts. *Reading and Writing*, 18, 99–128. <https://doi.org/10.1007/s11145-004-7343-5>
- Majaj N. J., Pelli D. G., Kurshan P., & Palomares M. (2002). The role of spatial frequency channels in letter identification. *Vision Research*, 42, 1165–1184. [https://doi.org/10.1016/S0042-6989\(02\)00045-7](https://doi.org/10.1016/S0042-6989(02)00045-7)
- Oakes, L.M., & Ribar, R. J. (2005). A comparison of infant's categorization in paired and successive presentation familiarization tasks. *Infancy*. 7, 85–98. https://doi.org/10.1207/s15327078in0701_7.
- Owens, P., and Sweller, J. (2008). Cognitive load theory and music instruction. *Educ. Psychol.* 28, 29–45. <https://doi.org/10.1080/01443410701369146>
- Pelli D. G. Burns C. W. Farell B. Moore-Page D. C. (2006). Feature detection and letter identification. *Vision Research*, 46, 4646–4674. <https://doi.org/10.1016/j.visres.2006.04.023>
- Peng, D. L., and Wang, C. M. (1997). Basic processing unit of Chinese character recognition: evidence from stroke number effect and radical number effect(汉字加工的基本单元:来自笔画数效应和部件数效应). *Acta Psychologica Sinica*(心理学报) 29, 8–16.
- Shen, H. H. (2004). Level of Cognitive Processing: Effects on Character Learning Among Non-Native Learners of Chinese as a Foreign Language. *Language and Education*, 18(2), 167–182. <https://doi.org/10.1080/09500780408666873>
- Shen, H. H. (2005). An investigation of Chinese-character learning strategies among non-native speakers of chinese. *System*, 33(1), 49–68. <https://doi.org/10.1016/j.system.2004.11.001>
- Shen, H. H. (2013). Chinese L2 Literacy Development: Cognitive Characteristics, Learning Strategies, and Pedagogical Interventions. *Language and Linguistics Compass*, 7(7), 371–387. <https://doi.org/10.1111/lnc3.12034>
- Shu, H., Chen, X., Anderson, R. C., Wu, N., & Xuan, Y. (2003). Properties of School Chinese: Implications for Learning to Read. *Child Development*, 74(1), 27–47. <https://doi.org/10.1111/1467-8624.00519>
- Songsangkaew, P., Yodchim, S., Sukamolson, S., & Person, K. R. (2023). English Morphological Awareness in Complex Words through Color Coding for Reading Comprehension. *Journal of Namibian Studies*, 34. <https://doi.org/10.59670/jns.v34i.1409>
- Taylor, I., & Taylor, M. M. (1995). *Writing and literacy in Chinese, Korean, and Japanese*. Philadelphia: John Benjamins Publishing.
- Tong, X., & McBride, C. (2014). Chinese Children's Statistical Learning of Orthographic Regularities: Positional Constraints and Character Structure. *Scientific Studies of Reading*, 18(4), 291–308. <https://doi.org/10.1080/10888438.2014.884098>
- Tortorelli, L. S., Lupo, S. M., & Wheatley, B. C. (2021). Examining Teacher Preparation for Code-Related Reading Instruction: An Integrated Literature Review. *Reading Research Quarterly*, 56(S1), S317–S337. <https://doi.org/10.1002/rrq.396>
- Vlach, H. A. , Ankowski, A. A. & Sandhofer, C. M. (2012). At the Same Time or Apart in Time? The Role of Presentation Timing and Retrieval Dynamics in Generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38 (1), 246-254. <https://doi.org/10.1037/a0025260>.
- Vlach, H. A., Kaul, M., Hosch, A., Lazaroff, E., & Wang, Q. (2022). Attending Less and Forgetting More: Dynamics of Simultaneous, Massed, and Spaced Presentations in Science Concept Learning. *Journal of Applied Research in Memory and Cognition*, 11(3), 361–373. <https://doi.org/10.1016/j.jarmac.2021.10.007>
- Wang, M., Koda, K., & Perfetti, C. A. (2003). Alphabetic and nonalphabetic L1 effects in English word identification: A comparison of Korean and Chinese English L2 learners. *Cognition*, 87, 129–149.
- Wang, H., He, X., & Legge, G. E. (2014). Effect of pattern complexity on the visual span for Chinese

- and alphabet characters. *Journal of Vision*, 14(8):6, 1–17. <https://doi.org/10.1167/14.8.6>
- Wong, Y. K. (2017). The role of radical awareness in Chinese-as-a-second-language learners' Chinese character reading development. *Language Awareness*, 26(3), 211–225. <https://doi.org/10.1080/09658416.2017.1400039>
- Wright, C., & Wang, J. (2023). The role of visual processing in learning Mandarin characters. *Journal of the European Second Language Association*, 7(1), 31–45. <https://doi.org/10.22599/jesla.92>
- Yeh, S.-L., Li, J.-L., Takeuchi, T., Sun, V., & Liu, W.-R. (2003). The role of learning experience on the perceptual organization of Chinese characters. *Visual Cognition*, 10(6), 729–764. <https://doi.org/10.1080/13506280344000077>
- Yin, L., & McBride, C. (2015). Chinese Kindergartners Learn to Read Characters Analytically. *Psychological Science*, 26(4), 424–432. <https://doi.org/10.1177/0956797614567203>
- Yu, L., Zhang, Q., Priest, C., Reichle, E. D., & Sheridan, H. (2018). Character-complexity effects in Chinese reading and visual search: A comparison and theoretical implications. *Quarterly Journal of Experimental Psychology* (2006), 71(1), 140–151. <https://doi.org/10.1080/17470218.2016.1272616>
- Zhang J. Y. Zhang T. Xue F. Liu L. Yu C. (2007). Legibility of Chinese characters and its implications for visual acuity measurement in Chinese reading population. *Investigative Ophthalmology & Visual Science*, 48 (5), 2383–2390, <http://www.iovs.org/content/48/5/2383>.
- Zhang, Jingxian. 1992. *Modern Chinese Character Course* (现代汉字教程) . Beijing: Modern Publishing House.
- Zimmer, H. D. & Fischer, B. (2020). Visual working memory of chinese characters and expertise: the expert's memory advantage is based on long-term knowledge of visual word forms. *Frontiers in Psychology*, 11, 516.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.