




Acquiring the structure of a writing system is important in learning to read: a test of the character-word dual-focus approach in learning Chinese as a second language

Research Article



Lin Chen^{1,2,3} , Yi Xu⁴ and Charles Perfetti⁵

Cite this article: Chen, L., Xu, Y. and Perfetti, C. (2024). Acquiring the structure of a writing system is important in learning to read: a test of the character-word dual-focus approach in learning Chinese as a second language. *Bilingualism: Language and Cognition*, 1–12 <https://doi.org/10.1017/S1366728924000920>

Received: 06 March 2023
Revised: 16 September 2024
Accepted: 14 October 2024

Keywords:
word learning; word instruction; Chinese; character-word dual focus

Corresponding author:
Lin Chen;
Email: linchen8@illinois.edu

  This research article was awarded Open Data and Open Materials badges for transparent practices. See the Data Availability Statement for details.

¹Department of Educational Psychology, University of Illinois, Urbana-Champaign, Illinois, USA; ²Beckman Institute for Advanced Science and Technology, University of Illinois, Urbana-Champaign, Illinois, USA; ³Department of Linguistics, University of Illinois, Urbana-Champaign, Illinois, USA; ⁴Department of East Asian Languages and Literatures, University of Pittsburgh, Pennsylvania, USA and ⁵Learning Research & Development Center, University of Pittsburgh, Pennsylvania, USA

Abstract

An important question in literacy education is whether reading instruction should focus on whole words or subword constituents. We tested whether this question captures something general across writing systems by examining the functionalities of words and characters in learning Chinese. We introduce a character-word dual-focus instructional approach based on the Character-Word Dual Function model and test its predictions with American undergraduate students enrolled in a beginner-level Chinese course. One group learned new words through dual-focus instruction: characters for pronunciation and words for meaning. A second group followed typical word-focus instruction prevalent in classrooms, learning word-level pronunciation and meaning. Results indicated that while both approaches produced comparable levels of word pronunciation and meaning learning, the dual-focus instruction significantly enhanced character pronunciation and transfer to new word learning. The advantages of dual-focus instruction highlight the importance of learning the subword components through acquiring the systematic structure of the writing system in learning to read.

One of the enduring issues in learning to read is whether instructions for novices should focus attention on the whole-word level or on the parts of words that are structured by the writing system (Bowers, 2020; Castles et al., 2018; National Reading Panel, 2000; Rayner et al., 2001; Verhoeven & Perfetti, 2022). Although discussion of this issue has focused on alphabetic writing, where the component “parts” map onto phonemes, it is also relevant for writing systems that map orthography to higher units of language, such as syllables or morphemes. Such systems differ from alphabetic systems in how their words and word-constituent components relate, making it unclear whether constituent-focus versus word-focus captures something general across writing systems.

We address these questions from the perspective of learning Chinese as a second or foreign language (CFL) for two reasons. First, the Chinese morpho-syllabic writing system provides an interesting contrast with alphabetic languages. Unlike the grapheme-to-phoneme mappings in alphabetic languages, written Chinese maps to speak Chinese at the syllable morpheme level. A Chinese character, a writing unit, usually corresponds to a syllable that also functions as a meaningful morpheme, while a Chinese word generally consists of two or more characters. Chinese is typically considered to have a “deep” orthography (Katz & Frost, 1992; Tseng, 2002) that encodes meaning directly in its characters and, thus, in its words. Examining the roles of words and characters in learning Chinese can provide new evidence and a more universal perspective to the “whole versus parts” discussion that has been prominent in learning to read in alphabetic languages.

Second, current conclusions on learning to read words are mainly based on findings obtained from the study of native speakers. Their applications to second language word learning are unclear, given the different prior language exposure between native speakers and second-language learners. The task in learning to read for first-language children is to learn how their writing system encodes their language, whose basic structure they have largely learned prior to learning to read (Perfetti, 2003; Verhoeven & Perfetti, 2022). In contrast, adult learners of a second language in traditional classroom settings usually learn the written and spoken words nearly simultaneously. However, even with limited knowledge of the spoken language, second language learners also need to learn how the writing system works. We argue that instructional practices in second language reading vary in the extent they support learning the structure of the writing system. Just as an overemphasis on whole words can obscure the structure of mappings from written language (letter strings) to spoken language (phonemes) at the subword level

© The Author(s), 2024. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.



(Castles *et al.*, 2018; Rayner *et al.*, 2001), so too can an overemphasis on word instruction in Chinese.

In the following, we briefly review studies of word instruction in both alphabetic languages and Chinese to highlight the fact that, although the two writing systems are very different in how constituent parts map to whole words, they share the problem of whether instruction is constituent-focused or word-focused. Following this review, we discuss the Chinese writing system and explain how the Character-Word Dual Function (CWDF) model (L. Chen *et al.*, 2024) establishes the foundation for a dual-focus approach to learning to read Chinese. We then report the results of a study of classroom learners that tested this approach.

1. Word instructions in learning to read English

In alphabetic languages, especially in English, the arguments focus on whether instruction should explicitly focus on the sublexical components of words (i.e., systematic phonics instruction) or the whole word (whole-word instruction) (Bowers, 2020; Castles *et al.*, 2018; National Reading Panel, 2000; Rayner *et al.*, 2001). The basis for explicit phonics instruction is that reading an alphabetic language requires learning the systematic mappings of graphemes to phonemes, which facilitates reading new words. Proponents of whole-word instruction (and the whole-language approach) argue that instruction should focus on meaning. They also argue that the irregularities of letter-sound mappings make them unreliable cues to word identification, and the correspondences can be acquired implicitly as part of whole-word learning (Bowers, 2020; Goodman, 1967). Although a review of this issue is beyond the scope of this article, major reviews of the evidence over a period of years have concluded an early focus on grapheme-phoneme correspondences supports learning to read (Castles *et al.*, 2018; National Reading Panel, 2000; Rayner *et al.*, 2001).

This conclusion is supported by abundant empirical evidence as reviewed in the reports cited above. Establishing sublexical-level grapheme-phoneme decoding is foundational for the development of high-quality word representations that support reading fluency (Perfetti, 2007). Indeed, decoding is the major mechanism by which readers establish orthographic representations of words (Share, 1995, 1999) and become able to acquire new words (Byrne, 2005). More significantly, systematic phonics instruction continues to benefit children with less developed language skills even a decade after receiving the intervention (Blachman *et al.*, 2014).

Although two recent reviews have argued that the evidence supporting phonics instruction is not strong (Bowers, 2020; Wyse & Bradbury, 2022), their arguments face challenges, as their analyses are considered flawed (Brooks, 2023; Fletcher *et al.*, 2021). For example, Brooks (2023) argued that the meta-analyses conducted by Bowers (2020) lack a thorough assessment of the existing work on phonics instruction while disregarding relevant research that supports the effectiveness of systematic phonics instruction. Additionally, the quality of the studies selected for meta-analyses was not adequately considered; thus, some studies with problematic research designs were included, potentially leading to biased conclusions.

2. Chinese word instruction

In alphabetic writing systems, typical candidate units for instructional focus are words and subword letter-phoneme mappings. While the word-level focus remains a candidate for instruction in the Chinese writing system, its writing units are *characters* rather

than letters. A writing unit (character) usually corresponds to both a syllable and a morpheme. For example, the character “耳” is pronounced “ěr” and means “ear”.

Instruction in Chinese is based on either the character-focus approach, mirroring the structure of the writing system, or the word-focus approach. Interestingly, the typical choice differs for native Chinese speakers and nonnative Chinese learners. Character-focus instruction is widely used by children learning to read Chinese as their native language (Pine *et al.*, 2003), helping them establish mappings between the orthography of an individual character and its syllable phonology.

In contrast, word-focus instruction is prevalent in learning CFL (T. Li, 2005). Instruction draws the learner’s attention to the whole word and its pronunciation. The concept of instructing native speakers and CFL learners differently is magnificently reflected in the design of their textbooks. In textbooks for native Chinese readers, Pinyin (the Roman alphabet transcription of pronunciation) is usually put on the top of each word, and the correspondence between each constituent character and its pronunciation is made salient by adding spaces between them, such as, 耳机 (meaning: headphone). Pinyin is also added on top of each character (also with spaces) in the instructional texts to facilitate native beginners learning the character-level associations. In contrast, in textbooks for nonnative speakers, the pronunciation of a word is typically presented as a whole and the mappings between individual characters and their pronunciations are implicit, for example: 耳机 ěrjī (meaning: headphone).

The rationale for these contrasting approaches appears to be based on the assumption that CFL learners have lower exposure to spoken Chinese than native speakers do. Native speakers have developed an awareness of syllables and morphemes in their spoken language (McBride-Chang *et al.*, 2004; Shu *et al.*, 2006), which prepares them to learn the character-level associations between orthography and phonology. CFL learners, especially late CFL learners, can partially compensate for their lower knowledge of morphemes in spoken Chinese by establishing word-level correspondences between their first language and Chinese through translation equivalents, relying on the well-developed word representations in their native language. Thus, word-level correspondences are emphasized in CFL Chinese word learning, whereas character-level associations are not typically explicitly instructed (Li, 2005).

These word-emphasis instructional practices appear to favor the acquisition of whole-word knowledge over character knowledge for CFL learners (Bai *et al.*, 2008; Chen, 2015; L. Chen *et al.*, 2018; Shen *et al.*, 2012). For example, in a naming task, Chen (2015) found that CFL learners had difficulty naming individual characters, although they were able to name the words containing those characters. Other evidence comes from a study of the word superiority effect (Chen *et al.*, 2018): under conditions of very brief exposures (from 40 ms to 57 ms), the recognition of a character is facilitated when it appears as part of a real word, compared with when it appears in a nonword. In particular, Chen *et al.* (2018) found that CFL learners showed robust word superiority effects for high-frequency characters as well as low-frequency ones. This contrasts with skilled native readers, among whom the word superiority effects for high-frequency characters were smaller than for low-frequency characters. These findings suggest that native readers have acquired sufficient experience with high-frequency characters such that they can readily identify the characters without the context provided by the word. In contrast, CFL learners’ character recognition benefits from seeing the character in a word, even for characters that occur frequently. The dominance of word-level processing for CFL

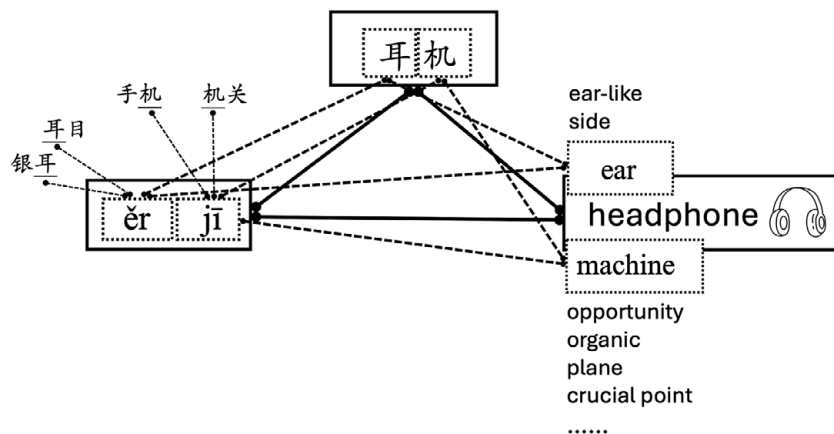


Figure 1 Illustration of the word (lexical) and character (sublexical) level mappings in a two-character Chinese word. (The solid black lines indicate lexical-level orthography, phonology and meaning mappings, while the dashed lines indicate sublexical-level mappings.)

learners has also been observed in text reading (Bai et al., 2013; Shen et al., 2012). In an eye-tracking study, Shen et al. (2012) inserted spaces between words to create word boundaries, despite Chinese texts being nonspaced, and found that total sentence reading times were significantly reduced for CFL learners from various first-language backgrounds (English, Korean, Japanese and Thai).

3. The functions of characters and words

Reconsidering the roles of words and characters can reveal general operating principles in learning to read Chinese for both native speakers and CFL learners. The dual function of characters – encoding both phonological and morphological information – is critical to our proposed instructional approach. In brief, Chinese orthography is consistent and direct in mapping to syllable-level phonology and indirect in coding morpheme-level meaning. We elaborate on this functional duality below.

The conventional view holds that the Chinese writing system has a “deep” orthography because the correspondence between orthography and language occurs at the syllable-size morpheme level rather than at the phoneme level (Katz & Frost, 1992), allowing meaning to be encoded directly in its characters. Although this description captures properties of Chinese that distinguish it from alphabetic writing, critical additions are needed to prevent unwarranted implications that written Chinese conveys meaning but not phonology.

The mappings from Chinese characters to syllables have a high degree of consistency in their pronunciations. Only 1,000 characters (out of the 13,000 characters in the Dictionary of Modern Chinese) are polyphonic. Among the 100 most commonly used polyphonic characters, the majority have a dominant pronunciation that applies to 95% of the words they appear (Zhang & Chu, 2009). Additionally, most Chinese words—70%—are two-character compounds (Lexicon of Common Words in Contemporary Chinese, 2009), where the individual character pronunciations typically combine to provide the pronunciation of the entire word. Thus, word-level phonology is highly predictable and transparent (Tones of a small number of words may change under certain circumstances.¹) For example, the word “耳机” (headphone) consists of two characters, “耳” pronounced /ěr/ and “机” pronounced /jī/ and is pronounced /ěr jī/

(See the left part in Figure 1. This section also presents additional words consisting of either “耳” or “机,” where all “耳” share the same pronunciation, as do all “机.”).

Characters also correspond to meaning-bearing morphemes. However, the majority of characters are bound morphemes that cannot stand alone as words (Yu et al., 1999; Yuan & Huang, 1998) and have less precise and distinctive meanings compared to free morphemes (Taft, 2003). Although most Chinese words are compounds of two or more characters, only the meanings of 29% of such words are completely transparent, that is, the meaning of the word is derived from the combination of its constituent characters’ meanings (e.g., 蓝莓 “blueberry”) (J. Li, 2011). More typical are words whose meanings are not directly inferred from the meanings of their constituent characters. For example, knowing the meanings of the individual character 耳 (“ear”) and 机 (“machine”) is not sufficient to infer the specific meaning of “headphone” for the Chinese word “耳机.” Further, one character often corresponds to multiple meanings (many of which are unrelated), and their interpretations tend to be highly word-dependent. For example, the character “耳” has three meanings “ear,” “ear-like,” and “side” and the character “机” has nine unrelated meanings, such as “machine,” “opportunity,” “organic,” “plane,” “crucial point,” etc (See the right part in Figure 1). Whether the meaning of “机” is interpreted as “machine” or “opportunity” depends on the compound word in which the character appears, e.g., “耳机” (headphone) or “时机” (timing). These features increase the difficulty of developing precise representations of character-morpheme meaning.

In summary, characters function as both orthographic and morphemic units, consistently encoding syllable-level phonology as primary orthographic units but less reliably representing meaning as morphemic units. These functional distinctions are the core assumptions of the CWDF model (L. Chen et al., 2024), as we explain below.

The character functions as the basic unit of orthography, providing orthographic information that supports the acquisition of high-quality lexical representations for Chinese reading. The development of orthographic awareness—including knowledge of radical forms and their positions within characters—strongly correlates with Chinese reading performance (H. Li et al., 2012). Establishing the mapping between a character and its pronunciation strengthens its orthographic representation, which in turn serves as an orthographic gateway to the written Chinese lexicon (Chen, Perfetti, & Leng, 2019; Chen, Perfetti, Fang, et al., 2019).

¹In words with two characters, both having the third tone (for example: 祖母 *mǔ*, grandma), one case of the tone sandhi rules involves changing the tone of the first character (祖母 *zǔ*) to the second tone (祖母 *zú*).

The most reliable conveyer of meaning is *word* meaning, which is both the result of word identification and the input for comprehension. Thus, word meaning plays a central role in linking these two subsystems of reading (Perfetti & Stafura, 2014). The reliable representation of word meaning provides a basis for rapid meaning retrieval and integration during reading comprehension (Perfetti & Helder, 2021). The importance of word meaning for integration (relative to morpheme meaning) in reading Chinese has been demonstrated in studies using both ERP (L. Chen *et al.*, 2017) and eye-tracking methods (Shen *et al.*, 2018).

In addition to specifying the primary functions of characters and words, the CWDF model further highlights the dependence of these functions on the quality of lexical representations developed through reading experience, thus linking the early stages of learning to read with more proficient reading. In learning to read Chinese, establishing character-level orthographic and word-level meaning representations is essential and central in the character-word dual-focus approach, as elaborated below. One important point to note is that the CWDF model does not dismiss the role of character meaning in reading Chinese. Rather, the model posits that the meaning function of characters is secondary in learning to read due to the previously mentioned features of characters regarding how they encode meaning. As reading experience increases, readers develop high-quality lexical representations, allowing characters and words to contribute efficiently to both orthographic and meaning-related processes (L. Chen *et al.*, 2024).

4. Character-word dual focus

The CWDF model, which distinguishes the functions of words and characters in reading Chinese, informs the development of a character-word dual-focus instructional approach for learning to read Chinese. This approach emphasizes a dual focus on both characters and words that respects their roles in the structure of written Chinese and the evidence supporting their differentiated functions in reading.

For orthographic learning, the instruction focuses on the character as a functional orthographic object to be mapped to a spoken syllable. Explicitly instructing the character-level mapping between orthography and phonology enables learners to develop precise orthographic representations at both the character and word levels. More importantly, by learning character-level orthography-phonology associations, learners can acquire the structure of the Chinese writing system, specifically its orthography-phonology mapping principles. This, in turn, facilitates new word learning and vocabulary development because learners can apply the learned character-level orthography-phonology mappings when encountering unfamiliar words.

Meaning instruction focuses on the word as a meaning unit, enabling learners to retrieve more precise meanings from its orthographic form. Reading instruction aims to provide learners with effective procedures for obtaining word meanings from the written form, for which the characters provide the orthographic input.

5. The present study

In the present study, we conducted a two-session classroom study with American college students enrolled in an introductory Chinese course to assess the effectiveness of the character-word dual-focus approach in Chinese word instruction. We compared the learning performance of students who received dual-focus instruction with

those who received conventional word-focus instruction. The dual-focus instruction emphasized two key elements. First, orthography-to-phonology mappings were taught at the character level to support the development of orthographic representations for both characters and words and the learning of new words. Second, orthography-to-meaning mappings were taught at the word level rather than the character level. Students who received word-focus instruction learned the same words by associating each two-character word with its pronunciation and meaning, rather than focusing on the individual characters. We predicted that the emphasis on establishing character-level orthographic-phonological representations in dual-focus instruction would lead learners to develop orthographic representations for both words and individual characters. These character-level representations will enable greater character-based generalizations, allowing learners to more effectively recognize these characters when they appear in unfamiliar words, compared to those receiving word-focus instruction. Additionally, the emphasis on word meaning in dual-focus instruction is expected to result in word-meaning learning comparable to that of word-focus instruction.

6. Method

6.1. Participants

44 CFL learners (average age 19.64, $SD = 1.46$, 22 females) from the introductory Chinese course at the University of Pittsburgh participated in the research after two full semesters in Chinese classes (7 hours of classroom instruction per week) that employed a word-emphasis approach. All participants reported English as their native language. Based on a background survey during the recruitment stage, none of the participants identified themselves as heritage language learners; nor did they have experience living in a Chinese-speaking country or region prior to the study. Chinese proficiency of the participants was measured based on the average scores of two Chinese tests (Test 1: $Mean = 84.62$, $SD = 11.25$ out of a total score of 100; Test 2: $Mean = 88.18$, $SD = 11.96$), covering reading, writing, speaking, and listening skills. The internal reliability of the two Chinese tests was Cronbach's $\alpha = 0.834$ (Taber, 2018).

Twenty-two learners were assigned to each of the two groups (dual-focus and word-focus condition). To ensure that participants' Chinese proficiency was well matched between the two groups, we used a "pseudo-randomized" design instead of a random design to prevent imbalanced covariant distribution in a relatively small sample size. We first ranked the scores of all the participants from high to low. Then, we assigned the participants in "ABBA" order (i.e., the first participant to the dual-focus condition, the second and third to the word-focus condition, the fourth and fifth to the dual-focus condition and so on). The language proficiency scores were not significantly different between the two groups (dual-focus: $mean = 87$, $SD = 10$; word-focus: $mean = 85.8$, $SD = 11.67$; $t_{(42)} = 0.36$, $p = .72$). Two participants were excluded from data analysis, one who completed only the first session of the experiment and one who had proficiency in Japanese, which shares a significant portion of characters with Chinese. Thus, 20 participants were left in the word-focus condition. Each participant signed the consent form before the experiment. All procedures were approved by the University of Pittsburgh Institutional Review Board.

6.1.1. Word and character knowledge test and results

We developed a test to assess the word and character knowledge of our participants, who received standard word-emphasis instruction in University classrooms prior to the experiment. We expect that

participants exposed to typical word-emphasis classroom instruction will demonstrate superior word knowledge compared to character knowledge. The discrepancy observed between word and character knowledge provides empirical support for this study's exploration of alternative instructional methods beyond word emphasis.

One week prior to the experiment, participating students completed a 10-minute test on 28 two-character words learned during previous Chinese classes (Table S1 in the Supplementary Material). Two versions of the test were used to prevent repetition of a word and its constituent characters. Each version contained 14 words to test word-level knowledge (word pronunciation and meaning) and 28 individual constituent characters from the other 14 words to test character-level knowledge (character pronunciation). The tested words and characters were listed in an inter-mixed format. Participants were asked to write down the word pronunciations and English equivalent meanings of the 14 words and the pronunciations of 28 individual characters. To avoid interference, two characters from the same word were not presented adjacent to one another (with an average of 15 words or other characters in between).

For scoring, we calculated the pronunciation accuracies for constituent characters and the whole word, as well as accuracies on word meaning for each participant. For each character, a pronunciation score of two was given when both syllable and tone were accurate; one was given when only the syllable was correct. The maximum score for a word pronunciation was thus four (two for each character), whether it was tested as a whole word or through its two constituent characters. T-test showed that participants performed better on retrieving the word pronunciation when a word was presented as a whole ($M = 2.40$, $SD = 0.54$) than when it was separated into constituent characters ($M = 1.78$, $SD = 0.60$): $t_{(27)} = 6.08$, $p < .001$. These results are consistent with a previous finding that CFL learners rely more on word knowledge than character knowledge in word identification (Chen, 2015), suggesting that a character learning disadvantage may be a general outcome of typical CFL instruction.

The meaning score was one if the meaning of the word was answered correctly. The average meaning score was 0.76 ($SD = 0.19$), significantly higher than zero, $t_{(27)} = 21.49$, $p < .001$, indicating that participants could remember the meanings of the majority of the words taught in their Chinese classes. Performance on word pronunciation, pronunciation of individual characters, and word meaning all showed high correlations with learners' Chinese proficiency. Learners with higher proficiency performed better on word pronunciation ($r = .51$, $p < .001$), character pronunciation ($r = .57$, $p < .001$) and word meaning ($r = .66$, $p < .001$).

Further, we analyzed the word and character knowledge of the two experimental groups to test whether they had similar levels of word and character knowledge before the experiment. We found that the two groups did not show differences in word pronunciation: $t_{(40)} = -0.22$, $p = .82$, character pronunciation: $t_{(40)} = 0.83$, $p = .41$ and word meaning performance: $t_{(40)} = 0.37$, $p = .71$. More importantly, both groups demonstrated better word knowledge compared to character knowledge: the dual-focus group: $t_{(21)} = 5.22$, $p < .001$; the word-focus group: $t_{(19)} = 8.06$, $p < .001$. Both groups had successfully learned word meanings as the word meaning performance was significantly higher than zero: for the dual-focus group ($t_{(21)} = 13.25$, $p < .001$) and the word-focus group ($t_{(19)} = 14.09$, $p < .001$).

6.2. Stimuli

The learning stimuli were 16 two-character words (see Table S2 in the Supplementary Material) chosen from the vocabulary syllabus

of the Chinese Proficiency Test (HSK). The HSK is a standardized test of Chinese language proficiency for nonnative speakers. To minimize the possibility that participants had learned the words and their constituent characters, we chose only words and characters that had not appeared in the Chinese textbooks the participants were using. The Chinese course instructor also reported that participants were unlikely to encounter the words and their characters in extracurricular reading based on their language proficiency. Further, the words were low in frequency (Range: 0.4–73/million, Mean = 16/million), and none were from the 1,000 commonly used words (Range: 90–50,155/million, Mean = 815/million) (Cai & Brysbaert, 2010). Additionally, according to a dataset of 19,716 Chinese words (Xu et al., 2021), the age of acquisition for those words² for native Chinese speakers ranges from 7.09 to 12.95 years old (Mean = 9.52, $SD = 1.95$), supporting that these words were unlikely encountered by our CFL learners. The number of strokes for each word ranges from 3 to 12 (Mean = 8.50, $SD = 2.44$). All words maintain the original tone of their constituent characters without requiring any tone alterations.

6.3. Procedure

Participants learned 16 two-character words through either word-focus or dual-focus instruction. Learning was distributed over 2 days to minimize experimental fatigue: 10 words were taught on the first day (day 1) and six words 3 days later (day 4). The words learned each day were consistent across all participants. Day 1 covered 10 words, while day 4 covered the remaining 6 words, with all participants learning the same words each day, presented randomly. Figure 2 outlines the two-day sessions, which included immediate tests after each day's learning phase, and day 4 began with a delayed test of day 1 learning. We conducted the delayed test 3 days later. This time period allows us to observe learners' memory consolidation, which typically occurs 1 day later after learning (Bakker et al., 2015). The two groups had the same learning times, with each daily session lasting 40 minutes and two practice trials provided to familiarize participants with the procedure.

Learning phase. Participants studied the words for 6 cycles in the learning phase, with every word appearing in each cycle.³ Each cycle contained an *exposure* component and a *retrieval* component (except the final cycle, which only contained the exposure component). We introduced the retrieval component to promote learning through testing. The exposure component showed comprehensive information (orthography, phonology and meaning) about each word to the participant on a computer screen, and the retrieval component asked participants to retrieve each word based on a varying prompt.

The arrangement of exposure and retrieval trials was varied across learning cycles to adapt to the expected progress made in learning. During early learning (first 3 cycles), participants were exposed to the studied words before they were asked to retrieve them. During the later stages (the fourth and fifth cycle), learners

²One word (弱项) in our stimuli was not included in the dataset.

³The design of six learning cycles was based on a pilot with two CFL learners from the experimental population. Both learners had learning accuracies above 0.8 in all learning cycles, indicating adequate learning. Additionally, both participants reported that learning 16 words in 80 minutes was similar to the effort in authentic classroom settings.

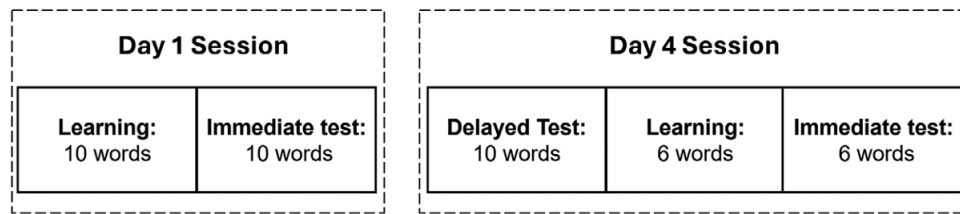


Figure 2 Outline of the two-day sessions.

Table 1. Six learning cycles in learning phase

Learning cycle	Learning procedure	Learning performance
1	Exposure+retrieval test	Retrieval accuracy
2	Exposure+retrieval test	Retrieval accuracy
3	Exposure+retrieval test	Retrieval accuracy
4	Retrieval test+exposure	Retrieval accuracy
5	Retrieval test+exposure	Retrieval accuracy
6	Exposure	N/A

were asked to retrieve the word first before they were exposed to it because they had studied each word multiple times. Retrieval accuracies of word pronunciation and word meaning were used to measure studying performance in the first five learning cycles but not the sixth cycle, in which participants were asked to only review all words once again without retrieval. **Table 1** shows the detailed learning procedures in each learning cycle and how learning performance was measured.

Across the six learning cycles, the word-focus group studied the word-level correspondence between the word (e.g., 朋友) and its pronunciation (e.g., péngyǒu) and its meaning (e.g., friend), following the typical word-focus instruction in the CFL classroom. The dual-focus group received explicit instruction on the mapping between each constituent character and its pronunciation while they learned the word meaning. The character-level correspondence was emphasized in both the exposure and retrieval trials for the dual-focus group. In the exposure trials, the correspondence between each character and its pronunciation was made visually salient by adding spaces between the characters (e.g., 朋 友). In the retrieval trials, participants were asked to produce the character-level pronunciation correspondence (e.g., “Type the Pinyin (pronunciation) for each character” instead of “Type the Pinyin for the word”). Complete details for each learning cycle are available in the [Supplementary Material](#).

Post-learning tests. The two groups took the same post-learning tests: an immediate test on both day 1 and day 4 and a delayed test on day 4 for day 1 learning. The delayed test was the same as the immediate test. Three subtests were designed to test learning from different aspects: (1) learning at the word level (word pronunciation and meaning), (2) learning at character level (character pronunciation), and (3) transfer to novel words (pronunciation). The word-level and character-level subtests employed both multiple-choice and recall tests. For the word-level subtest, the multiple-choice test required learners to choose the correct pronunciation and meaning from two word choices, both of which were from the learning phase. The recall test asked learners to type the pronunciation and meaning of each word. The character-level subtest followed the same

procedures, but character pronunciations were tested. In the character multiple-choice test, one choice was the pronunciation of the test character and the foil was the pronunciation of the other character from the same word. In the character recall test, learners were required to type the pronunciation for each character. In the transfer test, participants were asked to type the pronunciations of five new two-character words, each composed of characters from the 16 learned words (see [Table S3](#) in the Supplementary Material). For example, “虚” in “谦虚” (humble) and “弱” in “弱项” (disadvantage) composed a real unlearned word “虚弱” (weak).

7. Results

7.1. Results of learning phase

Word-learning performance on correctly retrieving the word pronunciation and word meaning was evaluated in the first five learning cycles. Because participants in the dual-focus group were asked to retrieve individual characters in a word separately, the pronunciation accuracy of a word was calculated based on the accuracy of its two constituent characters. Pronunciation accuracy of a whole word (two characters) received a score of one; word pronunciation accuracy was 0.5 when the pronunciation of one of the two characters was correct. The accuracy of word meaning was one if learners responded correctly, and zero otherwise. The average accuracy of word pronunciation and word meaning for all learning phases is listed in [Table 2](#).

We analyzed the data using mixed-effects modeling (Baayen *et al.*, 2008) with the accuracy of word pronunciation and word meaning as the respective dependent variables. The models were implemented in the lme4 packages in R (Bates *et al.*, 2015). Cumulative link mixed models were used to analyze the accuracy of word pronunciation because of the ordered distribution of its levels (0, 0.5, or 1). Logistic models were used to analyze the word meaning accuracy for its binominal distribution (0 or 1). Model comparisons determined whether the best-fit model included the random participant slope or the random item slope for the instruction group variable (word-focus versus dual-focus). The best-fit models for both pronunciation accuracy and meaning accuracy had the instruction group as a fixed effect and participant and item intercepts as random effects. (The data that support the findings of this study are available from: <https://osf.io/bta4d/>.)

The word-focus group showed higher retrieval accuracy of word pronunciation than the dual-focus group in learning cycles 2 (estimate = -1.36 , 95% CI: $[-2.27, -0.44]$, $SE = 0.47$, $z = -2.90$, $p < .01$), 3 (estimate = -1.29 , 95% CI: $[-2.54, -0.04]$, $SE = 0.64$, $z = -2.02$, $p = .04$), and 5 (estimate = -0.98 , 95% CI: $[-1.89, -0.07]$, $SE = 0.46$, $z = -2.11$, $p = .035$), respectively. In learning cycles 1 and 4, the groups did not show differences in retrieving word pronunciation, estimate = -0.09 , 95% CI: $[-0.77, 0.60]$, $SE = 0.35$,

Table 2. The retrieval accuracy (SD) of word pronunciation and meaning in five learning cycles

Group		Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5
Word pronunciation	Word-focus group	0.97 (0.15)	0.99 (0.10)	0.99 (0.06)	0.77 (0.40)	0.91 (0.27)
	Dual-focus	0.97 (0.15)	0.96 (0.15)	0.98 (0.10)	0.81 (0.36)	0.84 (0.28)
Word meaning	Word-focus group	0.92 (0.27)	0.95 (0.21)	0.97 (0.17)	0.81 (0.39)	0.93 (0.25)
	Dual-focus	0.90 (0.30)	0.85 (0.36)	0.87 (0.34)	0.75 (0.43)	N/A

Notes. To match the total learning times between the two groups, learners in the dual-focus group did not learn word meanings in learning cycle 5, while word-focus learners did. SDs reflected the distributions of raw data instead of the mean distribution of the participant.

$z = -0.25, p = .80$; estimate = 0.30, 95% CI: [-0.51, 1.11], $SE = 0.41, z = 0.74, p = .46$. (Note that the decline in retrieval performance in cycles 4 and 5 was expected, because participants were asked to retrieve a word before being exposed to it.)

The word-focus group showed higher accuracy in retrieving word meaning than the dual-focus group in learning cycles 2 (estimate = -1.34, 95% CI: [-2.34, -0.35], $SE = 0.51, z = -2.65, p < .01$) and 3 (estimate = -1.70, 95% CI: [-2.78, -0.62], $SE = 0.55, z = -3.08, p < .01$) but not in cycle 1 (estimate = -0.24, 95% CI: [-0.78, 0.30], $SE = 0.28, z = -0.87, p = .38$). This might be because the response to word meaning in the dual-focus group was placed between responses to the pronunciations of the two characters. The interleaving might interfere with word meaning retrieval from short-term memory. Nevertheless, reliable meaning learning differences between the two groups were not found in cycle 4: estimate = -0.42, 95% CI: [-1.01, 0.17], $SE = 0.30, z = -1.39, p = .16$.

Because the learning cycles were designed to encourage learners to learn words thoroughly (as they might in an authentic classroom setting), we observed high accuracy, which at first may seem to suggest a ceiling effect. However, the groups still showed statistically significant differences even at very high levels of accuracy in cycle 3 (0.99 vs. 0.98).

7.2. Results of post-learning tests

Results in the post-learning tests, including immediate tests and delayed tests, are the main research interests of the study. As described in the Method section, these tests assessed learners' retention of pronunciations and meanings of whole words, pronunciations of individual characters and pronunciation transfer to novel words. An accurate word pronunciation received a score of one; when the pronunciation of one of the two characters was correct, the score was 0.5. The accuracy of word meaning was one if learners responded correctly, and zero otherwise. So was the accuracy of character pronunciation. Cumulative link mixed models were used to analyze the accuracy of word pronunciation because of the ordered distribution of its levels (0, 0.5 or 1), and logistic models were used for the analysis of character pronunciation accuracy and word meaning accuracy because of their binomial distributions (0 or 1). The results of immediate tests and delayed tests are listed in Tables 3 and 4, respectively.

Table 3. Pronunciation and meaning retention accuracy (SD) at immediate tests

	Word pronunciation		Word meaning		Character pronunciation		
	Multiple-choice	Recall	Multiple-choice	Recall	Multiple-choice	Recall	Novel word pronunciation
Word-focus group	0.98(0.14)	0.91(0.28)	0.98(0.14)	0.97(0.18)	0.89(0.31)	0.73(0.45)	0.70(0.37)
Dual-focus group	0.99(0.11)	0.92(0.23)	0.96(0.20)	0.92(0.28)	0.97(0.17)	0.87(0.33)	0.89(0.23)

Notes. SDs reflected the distributions of raw data instead of the mean distribution of the participant.

7.2.1. Immediate tests

Test one: word pronunciation and word meaning. Immediately following learning, the two groups showed no differences in word pronunciation in either multiple-choice task (estimate = 0.51, 95% CI: [-0.90, 1.93], $SE = 0.72, z = 0.71, p = .48$), or recall task (estimate = 0.10, 95% CI: [-1.28, 1.47], $SE = 0.72, z = 0.14, p = .89$). Although the dual-focus group showed some degree of disadvantage during the word-level pronunciation assessments in the learning phase, these disadvantages disappeared in the immediate post-learning tests. The two groups did not differ significantly in word meaning accuracy in either the multiple-choice task, estimate = -0.87, 95% CI: [-1.90, 0.17], $SE = 0.53, z = -1.64, p = .1$ or in the recall task, estimate = -1.51, 95% CI: [-3.04, 0.02], $SE = 0.78, z = -1.93, p = .054$. To match the total learning times between the two groups, learners in the dual-focus group did not learn word meanings in learning cycle 5, while word-focus learners did. This may contribute to the tendency of higher meaning recall accuracy in the word-focus group.

Test two: character pronunciation. The dual-focus group showed higher accuracy of character pronunciation than the word-focus group in both the multiple-choice task (estimate = 1.48, 95% CI: [0.67, 2.30], $SE = 0.42, z = 3.56, p < .001$) and the recall task (estimate = 1.27, 95% CI: [0.50, 2.04], $SE = 0.39, z = 3.24, p < .01$).

Test three: pronunciation transfer to novel words. The dual-focus group outperformed the word-focus group in pronouncing novel words consisting of learned characters (estimate = 1.73, 95% CI: [0.75, 2.70], $SE = 0.50, z = 3.47, p < .001$).

The results of the immediate tests indicate that the dual-focus group performed as well as the word-focus group on word pronunciation and performed better on character pronunciation. Especially important is that the dual-focus group showed a stronger transfer of pronunciations to novel words. The word-focus group showed somewhat better immediate retention of word meanings, although this difference was not statistically reliable.

7.2.2. Delayed tests

Test one: word pronunciation and word meaning. The two groups did not show differences in the accuracy of word pronunciation in either the multiple-choice (estimate = -0.33, 95% CI: [-1.32, 0.67], $SE = 0.51, z = -0.64, p = .52$) or recall (estimate = 0.67, 95% CI: [-0.50, 1.84], $SE = 0.60, z = 1.13, p = .26$) tasks. They also did not

Table 4. Pronunciation and meaning retention accuracy (SD) at delayed tests

	Word pronunciation		Word meaning		Character pronunciation		Novel word pronunciation
	Multiple-choice	Recall	Multiple-choice	Recall	Multiple-choice	Recall	
Word-focus group	0.96(0.20)	0.79(0.38)	0.95(0.23)	0.87(0.34)	0.89(0.32)	0.66(0.47)	0.67(0.38)
Dual-focus group	0.95(0.23)	0.85(0.34)	0.89(0.32)	0.87(0.34)	0.91(0.29)	0.86(0.35)	0.88(0.25)

differ in meaning accuracy in the multiple-choice task (estimate = -0.89 , 95% CI: $[-1.98, 0.21]$, $SE = 0.56$, $z = -1.58$, $p = .11$) and the recall task (estimate = 0.19 , 95% CI: $[-0.90, 1.27]$, $SE = 0.55$, $z = 0.33$, $p = .74$). The results support that emphasizing learning constituent characters did not result in poorer learning of word-level pronunciation and meaning.

Test two: character pronunciation. The two groups did not differ in the accuracy of character pronunciation in the multiple-choice task, estimate = 0.22 , 95% CI: $[-0.34, 0.77]$, $SE = 0.28$, $z = 0.77$, $p = .44$. However, the dual-focus group showed better recall of character pronunciation, estimate = 1.61 , 95% CI: $[0.70, 2.51]$, $SE = 0.46$, $z = 3.48$, $p < .001$.

Test three: pronunciation transfer to novel words. The dual-focus group showed a better transfer of pronunciation to novel words than the word-focus group, estimate = 1.89 , 95% CI: $[0.77, 3.00]$, $SE = 0.57$, $z = 3.31$, $p < .001$.

The results of the delayed tests showed that the learning advantages of individual characters in the dual-focus group were retained. The dual-focus group was better at recalling the pronunciations of characters and performed better in transferring the pronunciations of learned words to novel words; their performance on learning word pronunciation and word meaning was not statistically different from that of the word-focus group.

The results for word learning, both word pronunciation and word meaning, were consistent with our expectation that dual-focus instruction would not disadvantage whole-word learning: The two groups did not differ significantly. Of course, a no-difference result resists straightforward interpretation. Ceiling effects or low power may make differences harder to detect. Ceiling

effects may well have been present in the easier multiple-choice tasks, where accuracy was high in both pronunciation and meaning, and the groups did not differ. However, in recall tasks that were challenging enough to produce low accuracy, the two groups again did not differ in either word meaning or pronunciation, first in immediate recall and then in delayed recall tasks. Thus, even in the absence of possible ceiling effects, there was no advantage in word learning for the word-focus group.

We addressed the second possibility, inadequate power, by calculating the effect size of each measure and conducting a power analysis to estimate the sample size needed to detect the effect of each measure. We used *t*-test results and the conventional benchmark for Cohen's *d* (Cohen, 1988; Sawilowsky, 2009) to estimate effect size because Generalized Linear Mixed-effects Models (GLMMs) present difficulties in defining the coefficient of determination due to their complexity and inherent heteroscedasticity (Nakagawa & Schielzeth, 2013; Schielzeth et al., 2017). Furthermore, the research literature on studies with our experimental design and population did not provide an appropriate effect size benchmark. We found the effect sizes of word pronunciation and meaning learning were small ($0.2 \leq$ Cohen's $d < 0.5$) across all tests, except for the medium effect in word meaning in the immediate tests (The effect size for each measure is in Table 5.) Further, we calculated the sample size required for each measure using G*Power (Erdfeider et al., 1996). To achieve a power of 0.80, a very large sample size was required to detect the effect with the observed effect size in every word-level test (see Table 5). The findings from effect size and power analyses suggest that the null results on word-level learning are more likely to be

Table 5. Results of *t*-test, effect size and desired sample sizes

			<i>t</i> value	<i>p</i>	Cohen's <i>d</i>	Desired sample size (power = 0.8)
Immediate tests	Word pronunciation	Multiple-choice task	0.79	.44	0.23	596
		Recall task	0.42	.68	0.13	1860
	Word meaning	Multiple-choice task	-1.71	.096	0.53	114
		Recall task	-1.64	.11	0.51	124
	Character pronunciation	Multiple-choice task	3.42	.001	1.06	30
		Recall task	3.15	.003	0.97	36
	Novel words	Recall task	3.56	<.001	1.10	30
Delayed tests	Word pronunciation	Multiple-choice task	-0.662	.51	0.24	548
		Recall task	0.898	.37	0.29	376
	Word meaning	Multiple-choice task	-1.58	.12	0.49	134
		Recall task	0.07	.95	0.02	78492
	Character pronunciation	Multiple-choice task	0.77	.45	0.24	548
		Recall task	3.51	.001	1.08	30
	Novel words	Recall task	3.47	.001	1.07	30

related to its relatively small effect size rather than lack of statistical power.

In sum, the lack of a difference in word learning should be interpreted as evidence for equal learning effects at the word level in the two groups. The advantages of dual-focus instruction were reflected in better pronunciation performance at the character level and in pronunciation transfer to novel words.

8. Discussion

Grounded in the CWDF model, we proposed a character-word dual-focus instructional approach for learning Chinese and tested its predictions in a study with students taking CFL. We predicted that the dual-focus instruction group would perform better in learning character-level orthographic-phonological correspondence and show better transfer of character pronunciation of novel words compared to the conventional word-focus instruction group. Furthermore, the dual-focus instruction with its focus on word meaning as well as character form should lead to word learning that is comparable to the word-focus instruction. As predicted, the dual-focus group generated better learning of character pronunciation and better transfer of character pronunciation to novel words and performed as well as the word-focus group in learning word pronunciation and word meaning. The advantages of dual-focus instruction are specifically on character learning, as expected, as evidenced especially through transfer to novel words. Importantly, there was no trade-off in learning at the word level. Thus, dual-focus instruction confers advantages in character learning while preserving the advantages of word learning.

8.1. Characters and words in learning to read Chinese

These results have direct implications for teaching and learning Chinese. In the CWDF model, words and characters are both functional units in reading, but with slightly different functions that can be leveraged in instruction. Learning character-level mappings of orthography and phonology supports learning the structure of the Chinese writing system and thus is important in the learning of new words. This learning process in Chinese is extensive if the 6,500 common characters are to be learned (State Language Commission of China, 2013). Although the character conveys meaning, the word that provides the more precise lexical meaning is central in the flow of reading processes from word identification to comprehension (Perfetti & Stafura, 2014). Thus, words, more than characters, provide the meaning units that readers continuously integrate into their mental model of the text (Yang et al., 2007). Thus, acquiring an orthographic lexicon that has characters as the basic orthographic units and words as the functional meaning units reflects properties of the Chinese writing system that are important in reading. The character-word dual-focus approach incorporates these properties into reading instruction.

This study demonstrates the application of the character-word dual-focus approach in one specific instructional approach in the CFL context. Because CFL learners can establish word-level correspondences between their first language (such as English) and Chinese through translation equivalents, they tend to rely more on word representations than character representations (Bai et al., 2008; Chen, 2015; Chen et al., 2018; Shen et al., 2012). Thus, sufficient attention, time, and practice to achieve character-level mappings is especially important for CFL learners. The study provides a successful example that emphasizes these character-level

mappings in learners' word-learning process. It should be noted that the progress of learning character-level mappings may vary among CFL learners with different first-language backgrounds. For example, for CFL learners whose first language is Japanese, prior knowledge of Kanji (Chinese characters in the Japanese language) may benefit their acquisition of character-level mappings.

The character-word dual-focus approach also holds implications for L1 Chinese reading instruction. In alignment with the dual-focus approach, L1 Chinese children receive explicit instruction on character-level orthography-phonology mapping, enabling them to develop orthographic representations of both characters and words and acquire the structure of the Chinese writing system for learning new words. In meaning instruction within L1 contexts, certain characters are taught alongside word meanings, particularly for words whose meanings have already been acquired in spoken language. This approach aligns with, rather than contrasts with, the dual-focus approach because meaning instruction for L1 children remains centered on words, with the character-meaning instruction being secondary.

The CWDF model, which forms the foundation of the dual-focus approach, acknowledges the meaning function of characters and their role in learning Chinese. According to the CWDF model, the function of character meanings in reading Chinese is contingent on the learner's reading experience (L. Chen et al., 2024), which directly impacts the effectiveness of character-meaning instruction. As a key indicator of this experience, vocabulary size plays a critical role because a larger vocabulary provides diverse word contexts for a character. The growth of vocabulary thus tunes the semantic representations of individual characters/morphemes, which can then support learning new words and enriching the meanings of known words (T. Chen, 2018; McBride-Chang et al., 2008). While learners use character *form* to develop high-quality orthographic forms to serve reading through word identification, vocabulary is important for acquiring the functional (in-context) value of character *meaning*. Thus, incorporating certain characters into meaning instruction represents a dual-focus approach for more proficient learners, applicable to both native Chinese children and CFL learners. For L1 children, character-meaning instruction is involved because their already acquired spoken vocabulary enables them to use rich word contexts to infer and acquire the meanings of certain characters (McBride-Chang et al., 2004; Shu et al., 2006). Similarly, as CFL learners' vocabulary expands, character-meaning instruction is expected to become increasingly relevant.

While detailed instructional procedures for individual character meanings are beyond the scope of this study, we note their dependence on various factors. These factors include, among others, the semantic relationship between the character and the word (Mok, 2009), the syntactic structure of the word (Tang & Liang, 2020), and the family size of the character (Liu et al., 2017). In general, three fundamental principles can inform character-meaning instruction: First, characters often have multiple meanings, and their interpretations heavily rely on the word in which they appear. Therefore, instruction on character meaning should be grounded in word context. Second, most characters are bound morphemes, whose meanings are typically less precise. Providing sufficient word examples containing the same character with consistent meanings is critical for learners to obtain the character's meaning. Third, many Chinese words lack direct mappings between the meanings of a word and its constituent characters. Therefore, character-meaning instruction should aim to bridge these discrepancies.

It is worth recalling that our participants were students in a whole-word instruction classroom; thus, the word-focus

participants were using a method more similar to their experience than the dual-focus participants. This point seems relevant to two interesting results of the study: First, during the learning phase of the study, the word-focus group showed higher accuracy in learning word pronunciation than the dual-focus group. This suggests that learning a word pronunciation through its constituent characters, as the dual-focus group did, might be more effortful compared to learning a word as a whole. Second, however, during the post-learning tests, the dual-focus group performed as well as the word-focus group on word pronunciation, while outperforming the word-focus group in character pronunciations and transfer to pronunciations of new words. It supports the idea that learning difficulty did not translate to poor final learning (Bjork, 1994; Schmidt & Bjork, 1992). On the contrary, the initial difficulty of retrieving character knowledge to assemble into word knowledge might have engaged more substantial retrieval operations that can support learning.

While word-focus instruction is prevalent in CFL learning, learners may not have the knowledge about the effective dual-focus learning, which requires explicit instructions as in our experiment. It is also possible that CFL learners are not comfortable with the dual-focus instruction, which they perceive as more exerting during learning the character-level mappings. In fact, learners can misinterpret the difficulty or mental effort they experience in learning, associating greater effort with poor learning performance and avoiding the strategies that need more effort (Kirk-Johnson *et al.*, 2019). As shown in this study, the lower accuracy of the dual-focus instruction in the learning phase may discourage learners from adopting this learning strategy. However, our results indicate that explicit instruction on learning orthographic-phonological representations of characters is essential to implementing word instruction for beginning learners. Thus, while the dual-focus instruction shows benefits over conventional word-focus instruction, we suggest that additional investigation is needed in order to identify and mediate the impact of these affective factors on learning outcomes.

8.2. Word learning across writing systems

At a more general level, the findings suggest the importance of learning the systematic structure of written Chinese, its words and word components, as in other writing systems. In Chinese, the value of learning the character-level orthographic-phonological mappings echoes that of learning the systematic phoneme-grapheme mappings of alphabetic writing (Caravolas *et al.*, 2012). Although writing systems vary in how they map orthography to phonology, learning to read in any system is supported by learning the component structures (Verhoeven & Perfetti, 2017). Learning the orthography-phonology mappings of a writing system facilitates the acquisition of subword knowledge, which is crucial for building high-quality word representations. More importantly, acquiring the structure of a writing system supports new word learning by enabling the transfer of shared components from known words, thus expanding vocabulary.

In alphabetic languages, even in English, where grapheme-phoneme mappings are less consistent, learning these grapheme-phoneme mappings has been found to be more productive than learning a word as a whole (Byrne *et al.*, 1992; Freebody & Byrne, 1988). Freebody and Byrne (1988) found that among students with below-average word reading performance, students who were instructed using word-specific association strategies showed improvement in reading exceptional words. In contrast, students

who were instead instructed using spelling-sound rules showed improvement in both regular and exceptional word reading. The advantages of learning word components are also found in adults learning an artificial alphabetic language. Instruction in the correspondence between component symbols and their sounds generates more transfer to new items than does whole-word instruction (Bitan & Booth, 2012; L. Brooks, 1977; Spring, 1978). Consistent with this advantage in English, our results also show that learning constituent characters facilitates learning novel Chinese words.

Beyond the generalizations across writing systems are some writing system-specific features to consider. Children from alphabetic languages and adult learners learning an artificial alphabetic language commonly have difficulty in implicitly learning the systematic orthography-phonology mapping with the whole-word method (Byrne, 1984; Byrne & Carroll, 1989; Seymour & Elder, 1986). By contrast, in the present study, learners of Chinese in the word-focus group implicitly learned the pronunciations of some characters, although their performance was not as good as the dual-focus group who explicitly learned the characters. In the post-learning tests, the recall accuracy of the character pronunciation of the word-focus group (0.73 in the immediate tests and 0.66 in the delayed tests) was significantly higher than the chance but lower than the dual-focus group. Further, implicit learning appears to have occurred during students' word-focus in-class instruction: On the test of word and character knowledge, students' character pronunciation accuracy showed a significant correlation with learners' Chinese language proficiency, even though the characters were not taught explicitly.

The differences between learning to read Chinese and alphabetic languages are arguably related to the spoken units required by the orthography-phonology mapping. Learners of alphabetic languages must map graphs to lower-level phoneme units, whereas Chinese learners map higher-level syllable units, which are more accessible than phonemes. In the specific case of Chinese, access to syllable units may be strengthened by the addition of their meaning function. Studies on the development of phonological knowledge suggest that compared to other levels of phonological knowledge, the knowledge of syllables is easy to obtain because syllables have a larger acoustic salience (Treiman & Zukowski, 1996). Most children across cultures develop syllable awareness naturally by 5 or 6 years old (Treiman & Zukowski, 1991). Thus, implicitly learning the phonological structure with syllables can be achievable even without explicit instruction. For our study specifically, participants had studied Chinese for a year. It is very likely that they have implicitly learned the Chinese phonological structure with syllables. In contrast, for children who natively speak an alphabetic language, who have not mastered the phonological structure and for adult learners learning an artificial alphabetic language just for a short period, explicit instruction might be critical for learners to gain the finer-grained representations of the structure (Byrne, 1992). Additionally, the salient orthographic forms of characters in Chinese provide visual cues to facilitate CFL learners establishing the ortho-phonological mapping between individual characters and their corresponding pronunciations.

9. Conclusion

We tested the character-word dual-focus instructional approach by comparing the learning outcomes of dual-focus instruction with those of conventional word-focus instruction. The dual-focus instruction produced performance on learning word pronunciation and word

meaning that was comparable to the word-focus instruction, while demonstrating advantages in learning individual characters pronunciations and transferring this knowledge to new word learning. The results reflect the value of learning the functional properties of a writing system's structure in both nonalphabetic and alphabetic writing.

Supplementary material. To view supplementary material for this article, please visit <http://doi.org/10.1017/S1366728924000920>.

Data availability statement. The data that support the findings of this study are openly in OSF and are available at <https://osf.io/bta4d/>.

Competing interest. No potential conflict of interest was reported by the author(s).

References

- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. <https://doi.org/10.1016/j.jml.2007.12.005>
- Bai, X., Liang, F., Blythe, H. I., Zang, C., Yan, G., & Liversedge, S. P. (2013). Interword spacing effects on the acquisition of new vocabulary for readers of Chinese as a second language. *Journal of Research in Reading*, *36*(Suppl.1), S4–S17. <https://doi.org/10.1111/j.1467-9817.2013.01554.X>
- Bai, X., Yan, G., Liversedge, S. P., Zang, C., & Rayner, K. (2008). Reading spaced and unspaced Chinese text: evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1277–1287. <https://doi.org/10.1037/0096-1523.34.5.1277>
- Bakker, I., Takashima, A., Hell, J. Van, Janzen, J., & McQueen, J. M. (2015). Tracking lexical consolidation with ERPs: Lexical and semantic-priming effects on N400 and LPC responses to newly-learned words. *Neuropsychologia*, *79*, 33–41. https://www.sciencedirect.com/science/article/pii/S0028393215301950?casa_token=CgE946N0mhUAAAAA:Q-piLsq1VhaYa-m7RzIqm7J4m3li b8OEc3wh7p_4odCFBzbq7NVdMIRPRJdspKzccqWYUFX1vw
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. <https://arxiv.org/abs/1406.5823v1>
- Bitan, T., & Booth, J. R. (2012). Offline improvement in learning to read a novel orthography depends on direct letter instruction. *Cognitive Science*, *36*(5), 896–918. <https://doi.org/10.1111/j.1551-6709.2012.01234.x>
- Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–206). The MIT Press.
- Blachman, B. A., Schatsneider, C., Fletcher, J. M., Murray, M. S., Munger, K. A., & Vaughn, M. G. (2014). Intensive reading remediation in grade 2 or 3: Are there effects a decade later? *Journal of Educational Psychology*, *106*(1), 46–57. <https://doi.org/10.1037/A0033663>
- Bowers, J. S. (2020). Reconsidering the evidence that systematic phonics is more effective than alternative methods of reading instruction. *Educational Psychology Review*, *32*(3), 681–705. <https://doi.org/10.1007/S10648-019-09515-Y>
- Brooks, G. (2023). Disputing recent attempts to reject the evidence in favour of systematic phonics instruction. *Review of Education*, *11*(2). <https://doi.org/10.1002/rev3.3408>
- Brooks, L. (1977). Visual pattern in fluent word identification. In A. S. Reber, & D. L. Scarborough (Eds.), *Toward a psychology of reading* (pp. 143–181). Hillsdale, NJ: Erlbaum.
- Byrne, B. (1984). On teaching articulatory phonetics via an orthography. *Memory & Cognition*, *12*(2), 181–189. <https://doi.org/10.3758/BF03198432>
- Byrne, B. (1992). Studies in the acquisition procedure for reading: Rationale, hypotheses, and data. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 1–34). Routledge. <https://doi.org/10.4324/9781351236904-1>
- Byrne, B. (2005). Theories of learning to read. In M. J. Snowling, & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 104–119). Malden, MA: Blackwell.
- Byrne, B., & Carroll, M. (1989). Learning artificial orthographies: Further evidence of a nonanalytic acquisition procedure. *Memory & Cognition*, *17*(3), 311–317. <https://doi.org/10.3758/BF03198469>
- Byrne, B., Freebody, P., & Gates, A. (1992). Longitudinal data on the relations of word-reading strategies to comprehension, reading time, and phonemic awareness. *Reading Research Quarterly*, *27*(2), 140–151. <https://doi.org/10.2307/747683>
- Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS ONE*, *5*(6). <https://doi.org/10.1371/JOURNAL.PONE.0010729>
- Caravolas, M., Lervåg, A., Mousikou, P., Efrim, C., Litavský, M., Onochie-Quintanilla, E., Salas, N., Schöffelová, M., Defior, S., Mikulajová, M., Seidlová-Málková, G., & Hulme, C. (2012). Common patterns of prediction of literacy development in different alphabetic orthographies. *Psychological Science*, *23*(6), 678–686. <https://doi.org/10.1177/0956797611434536>
- Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, *19*(1), 5–51. <https://doi.org/10.1177/1529100618772271>
- Chen, L. (2015). The role of morpheme in recognizing Chinese compound words for beginning Chinese learners (In Chinese 语素和整词在初级汉语二语者合成词语音识别中的作用研究). *TCSOL Studies (华文教学与研究)*, *3*, 1–5. <https://zhuanfou.com/thesis/00da06beac73>
- Chen, L., Fang, X., & Perfetti, C. A. (2017). Word-to-text integration: ERP evidence for semantic and orthographic effects in Chinese. *Journal of Neurolinguistics*, *42*, 83–92. <https://doi.org/10.1016/J.JNEUROLING.2016.11.010>
- Chen, L., Perfetti, C. A., Fang, X., Chang, L.-Y., & Fraundorf, S. (2019). Reading Pinyin activates sublexical character orthography for skilled Chinese readers. *Language, Cognition and Neuroscience*, *34*(6), 736–746. <https://doi.org/10.1080/23273798.2019.1578891>
- Chen, L., Perfetti, C. A., & Leng, Y. (2019). Reading Pinyin activates character orthography for highly experienced learners of Chinese. *Bilingualism: Language and Cognition*, *22*(1), 103–111. [10.1017/S136672891700058X](https://doi.org/10.1017/S136672891700058X)
- Chen, L., Perfetti, C. A., Leng, Y., & Li, Y. (2018). Word superiority effect for native Chinese readers and low-proficiency Chinese learners. *Applied Psycholinguistics*, *39*(6), 1097–1115. <https://doi.org/10.1017/S0142716418000255>
- Chen, L., Xu, Y., & Perfetti, C. (2024). A character-word dual function model of reading Chinese: evidence from reading Chinese compounds. *Reading and Writing*, *37*, 2429–2455. <https://doi.org/10.1007/S11145-023-10478-4/TABLES/8>
- Chen, T. (2018). The contribution of morphological awareness to lexical inferencing in L2 Chinese: Comparing more-skilled and less-skilled learners. *Foreign Language Annals*, *51*(4), 816–830. <https://doi.org/10.1111/FLAN.12365>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (J. Cohen, Ed., 2nd ed.). Hillsdale.
- Erdfelder, E., Faul, F., & Buchner, A. (1996). Gpower: A general power analysis program. *Behavior Research Methods, Instruments, & Computers*, *28*(1), 1–11. <https://psycnet.apa.org/record/1996-04345-001>
- Fletcher, J. M., Savage, R., & Vaughn, S. (2021). A commentary on bowers (2020) and the role of phonics instruction in reading. *Educational Psychology Review*, *33*(3), 1249–1274. Springer. <https://doi.org/10.1007/s10648-020-09580-8>
- Freebody, P., & Byrne, B. (1988). Word-reading strategies in elementary school children: relations to comprehension, reading time, and phonemic awareness. *Reading Research Quarterly*, *23*(4), 441–453. <https://doi.org/10.2307/747642>
- Goodman, K. S. (1967). Reading: A psycholinguistic guessing game. *Literacy Research and Instruction*, *6*(4), 126–135. <https://doi.org/10.1080/19388076709556976>
- Katz, L., & Frost, R. (1992). The Reading Process is different for different orthographies: The orthographic depth hypothesis. *Advances in Psychology*, *94*, 67–84. <https://files.eric.ed.gov/fulltext/ED359575.pdf#page=157>
- Kirk-Johnson, A., Galla, B. M., & Fraundorf, S. H. (2019). Perceiving effort as poor learning: The misinterpreted-effort hypothesis of how experienced effort and perceived learning relate to study strategy choice. *Cognitive Psychology*, *115*, 101237. <https://doi.org/10.1016/J.COGLING.2019.101237>
- Li, H., Shu, H., McBride-Chang, C., Liu, H., & Peng, H. (2012). Chinese children's character recognition: Visuo-orthographic, phonological processing and morphological skills. *Journal of Research in Reading*, *35*(3), 287–307. <https://doi.org/10.1111/j.1467-9817.2010.01460.X>

- Li, J. (2011). Investigation on the transparency of words in the Modern Chinese Dictionary (In Chinese 《现代汉语词典》的词义透明度考察). *Chinese Linguistics* (汉语学报), 3, 54–62.
- Li, T. (2005). The three major arguments in teaching Chinese as a foreign language in the past ten years (In Chinese 近十年对外汉语词汇教学研究中的三大流派). *Applied Linguistics* (语言文字应用), 51, 9–11.
- Liu, D., Li, H., & Wong, K. S. R. (2017). The anatomy of the role of morphological awareness in Chinese character learning: The mediation of vocabulary and semantic radical knowledge and the moderation of morpheme family size. *Scientific Studies of Reading*, 21(3), 210–224. <https://doi.org/10.1080/10888438.2017.1278764>
- McBride-Chang, C., Bialystok, E., Chong, K. K. Y., & Li, Y. (2004). Levels of phonological awareness in three cultures. *Journal of Experimental Child Psychology*, 89(2), 93–111. <https://doi.org/10.1016/j.jecp.2004.05.001>
- McBride-Chang, C., Tardif, T., Cho, J. R., Shu, H., Fletcher, P., Stokes, S. F., Wong, A., & Leung, K. (2008). What's in a word? Morphological awareness and vocabulary knowledge in three languages. *Applied Psycholinguistics*, 29(3), 437–462. <https://doi.org/10.1017/S014271640808020X>
- Mok, L. W. (2009). Word-superiority effect as a function of semantic transparency of Chinese bimorphemic compound words. *Language and Cognitive Processes*, 24(7–8), 1039–1081. <https://doi.org/10.1080/01690960902831195>
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4(2), 133–142. <https://doi.org/10.1111/j.2041-210X.2012.00261.X>
- National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development.
- Perfetti, C. A. (2003). The universal grammar of reading. *Scientific Studies of Reading*, 7(1), 3–24. https://doi.org/10.1207/S1532799XSSR0701_02
- Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11(4), 357–383. <https://doi.org/10.1080/10888430701530730>
- Perfetti, C. A., & Helder, A. (2021). Discourse processes incremental comprehension examined in ERPs: Word-to-text integration and structure building. *Discourse Processes*, 58(1), 2–21. <https://doi.org/10.1080/0163853X.2020.1743806>
- Perfetti, C. A., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, 18(1), 22–37. <https://doi.org/10.1080/10888438.2013.827687>
- Pine, N., Ping'an, H., & Ren Song, H. (2003). Decoding strategies used by Chinese primary school children. *Journal of Literacy Research*, 35(2), 777–812. https://doi.org/10.1207/S15548430JLR3502_5
- Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest*, 2(2), 31–74. <https://doi.org/10.1111/1529-1006.00004>
- Sawilowsky, S. S. (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods*, 8(2), 597–599. <https://core.ac.uk/download/pdf/56683105.pdf>
- Schielzeth, H., Nakagawa, S., & Johnson, P. C. D. (2017). The coefficient of determination R² and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface*, 14(134), 1–11. <https://doi.org/10.1098/rsif.2017.0213>
- Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, 3(4), 207–217. <https://doi.org/10.1111/j.1467-9280.1992.tb00029.x>
- Seymour, P. H. K., & Elder, L. (1986). Beginning reading without phonology. *Cognitive Neuropsychology*, 3(1), 1–36. <https://doi.org/10.1080/02643298608252668>
- Share, D. L. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition*, 55(2), 151–218. [https://doi.org/10.1016/0010-0277\(94\)00645-2](https://doi.org/10.1016/0010-0277(94)00645-2)
- Share, D. L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, 72(2), 95–129. <https://doi.org/10.1006/JECP.1998.2481>
- Shen, D., Liversedge, S. P., Tian, J., Zang, C., Cui, L., Bai, X., Yan, G., & Rayner, K. (2012). Eye movements of second language learners when reading spaced and unspaced Chinese text. *Journal of Experimental Psychology: Applied*, 18(2), 192–202. <https://doi.org/10.1037/A0027485>
- Shen, W., Li, X., & Pollatsek, A. (2018). The processing of Chinese compound words with ambiguous morphemes in sentence context. *Quarterly Journal of Experimental Psychology*, 71(1 Special Issue), 131–139. <https://doi.org/10.1080/17470218.2016.1270975>
- Shu, H., McBride-Chang, C., Wu, S., & Liu, H. (2006). Understanding Chinese developmental dyslexia: Morphological awareness as a core cognitive construct. *Journal of Educational Psychology*, 98(1), 122–133. <https://doi.org/10.1037/0022-0663.98.1.122>
- Spring, C. (1978). Automaticity of word recognition under phonics and whole-word instruction. *Journal of Educational Psychology*, 70(4), 445–450. <https://doi.org/10.1037/0022-0663.70.4.445>
- Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48(6), 1273–1296. <https://doi.org/10.1007/S11165-016-9602-2>
- Taft, M. (2003). Morphological representation as a correlation between form and meaning. In E. M. H. Assink & D. Sandra (Eds.), *Reading complex words* (pp. 113–137). Springer, Boston, MA. https://doi.org/10.1007/978-1-4757-3720-2_6
- Tang, X., & Liang, S. (2020). Study on semantic transparency of Chinese compounds based on word embedding. *International Conference on Asian Language Processing*, 130–134. IEEE, Kuala Lumpur, Malaysia. <https://doi.org/10.1109/IALP51396.2020.9310483>
- Treiman, R., & Zukowski, A. (1991). Levels of phonological awareness. In S. A. Brady, I. Y. Liberman, & D. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp. 67–83). Hillsdale, NJ: Lawrence Erlbaum.
- Treiman, R., & Zukowski, A. (1996). Children's sensitivity to syllables, onsets, rimes, and phonemes. *Journal of Experimental Child Psychology*, 61(3), 193–215. <https://doi.org/10.1006/jecp.1996.0014>
- Tseng, O. (2002). Current issues in learning to read in Chinese. In W. Li, J. Gaffney, & J. Packard (Eds.), *Chinese children's reading acquisition: Theoretical and pedagogical issues* (pp. 3–16). Dordrecht: Kluwer.
- Verhoeven, L., & Perfetti, C. A. (2017). *Learning to read across languages and writing systems*. Cambridge University Press.
- Verhoeven, L., & Perfetti, C. A. (2022). Universals in learning to read across languages and writing systems. *Scientific Studies of Reading*, 26(2), 150–164. <https://doi.org/10.1080/10888438.2021.1938575>
- Wyse, D., & Bradbury, A. (2022). Reading wars or reading reconciliation? A critical examination of robust research evidence, curriculum policy and teachers' practices for teaching phonics and reading. *Review of Education*, 10(1). <https://doi.org/10.1002/rev3.3314>
- Xu, X., Li, J., & Guo, S. (2021). Age of acquisition ratings for 19,716 simplified Chinese words. *Behavior Research Methods*, 53(2), 558–573. <https://doi.org/10.3758/S13428-020-01455-8/TABLES/7>
- Yang, C. L., Perfetti, C. A., & Schmalhofer, F. (2007). Event-related potential indicators of text integration across sentence boundaries. *Journal of Experimental Psychology: Learning, memory, and cognition*, 33(1), 55–89. <https://doi.org/10.1037/0278-7393.33.1.55>
- Yu, S., Zhu, X., & Li, F. (1999). Development and application of a dataset of modern Chinese morphemes (In Chinese 现代汉语语素库的开发及应用). *Chinese Teaching in the World* (世界汉语教学), 2, 39–46.
- Yuan, C., & Huang, C. (1998). Chinese morpheme and word formation based on a Chinese morpheme corpus (In Chinese 基于语素数据库的汉语语素及构词研究). *Chinese Teaching in the World* (世界汉语教学), 2, 7–12.
- Zhang, Z., & Chu, M. (2009). A statistical approach for grapheme to phoneme conversion in Chinese (In Chinese 解决多音字字音转换的一种统计学习方法). *Journal of Chinese Information Processing* (中文信息学报), 16(3), 39–45.