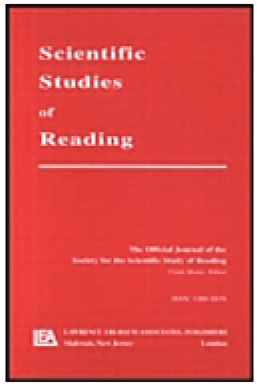


Is Character Position Encoding in Chinese Modulated by Morpheme-Level Semantics? Evidence From Lexical Decision and Sentence Reading

Journal:	*Scientific Studies of Reading*
Manuscript ID	HSSR-2025-0140.R2
Manuscript Type:	Research Article
Keywords:	eye tracking, lexical processing, semantics, sentence processing

Abstract

Purpose

As a logographic writing system, Chinese characters are not only orthographic symbols, but also function as morphemes that carry independent semantic content. The present study investigated whether character position encoding in Chinese is modulated by morpheme-level semantic information.

Method

Two experiments were carried out using complementary paradigms: an unprimed lexical decision task (Experiment 1, N = 70; mean age 21.8; 58.6% female; all native speakers of Chinese) and sentence reading with eye-tracking (Experiment 2, N = 70, mean age 22.7; 64.3% female, all native speakers of Chinese) to investigate this issue. In both experiments, we adopted the classic paradigm for studying position encoding—the transposed-character effect (TC effect)—where transposed-character (TC) nonwords (e.g., "震地") are more similar to base words (e.g., "地震" [earthquake]) than substituted-character (SC) nonwords (e.g., "牲缉"). Critically, we manipulated the semantic relatedness between the two constituent characters of disyllabic words and tested whether it modulates the TC effect.

Results

Results showed a consistent TC effect. In lexical decision, TC nonwords were harder to reject than SC nonwords; in sentence reading, TC nonwords elicited shorter

fixation times than SC nonwords. Critically, neither experiment showed any differences in the TC effect between semantically related and unrelated morpheme pairs. This null effect was supported by Bayes factor analyses, which provided quantitative evidence that morpheme-level semantic relatedness does not modulate transposition effects.

Conclusion

The present results provide no evidence that morpheme-level semantic relatedness modulates the transposed-character effect in Chinese.

Is Character Position Encoding in Chinese Modulated by Morpheme-Level Semantics? Evidence From Lexical Decision and Sentence Reading

In alphabetic languages, readers must encode the positions of letters within a word while reading; otherwise, they will struggle to distinguish between words composed of the same letters (e.g., *stop*, *spot*, *pots*, *post*, *tops*, *opts*; see Davis, 2010; Gómez et al., 2008; Grainger & Van Heuven, 2004; Logan, 2021; Snell, 2025; Whitney, 2001). However, letter position encoding is not strictly rigid but allows for a certain degree of flexibility. For instance, readers can fluently read the sentence "The adroable dog ran towrads Mike", in which some letters are in the wrong position (Rayner et al., 2006). This phenomenon is further exemplified by the transposed-letter effect (TL effect), where transposed-letter (TL) nonwords (e.g., "jugde", where the letters "d" and "g" are transposed) are perceived as more similar to their base words (e.g., "JUDGE") than substituted-letter (SL) nonwords (e.g., "jupte", where the letters "d" and "g" are replaced with "p" and "t", respectively), indicating that letter position encoding allows for a degree of flexibility (Johnson et al., 2007; Kinoshita & Norris, 2009; Lupker et al., 2008; Perea & Carreiras, 2006; Perea & Lupker, 2004; Rayner et al., 2006). Importantly, this phenomenon is robust and has been observed across multiple languages, including Chinese (e.g., English, Perea & Lupker, 2003; Spanish, Perea & Lupker, 2004; Thai, Perea et al., 2012; Arabic, Boudelaa et al., 2019; Hebrew, Kinoshita et al., 2012; Korean, Lee et al., 2021; Japanese kana, Perea et al., 2011; Chinese, Gu et al., 2015). In Chinese, where words are composed of characters rather than letters, the analogous phenomenon is typically referred to as the transposed-character effect (TC effect; see Gu et al., 2015).

Although transposition effects have been extensively studied in alphabetic scripts, their implications in non-alphabetic writing systems (e.g., Chinese) remain less well understood. Word processing is usually assumed to involve orthographic, phonological, and semantic processing (Harm & Seidenberg, 2004; Seidenberg & McClelland, 1989), and most models of letter position encoding in alphabetic languages assume that this process occurs at the orthographic level (Davis, 2010; Gómez et al., 2008; Grainger & Van Heuven, 2004; Norris & Kinoshita, 2012; Snell, 2025; Whitney, 2001).

However, Chinese, a logographic writing system, differs from alphabetic languages in that the characters are not only orthographic symbols that encode phonological information but also, in most cases, function as morphemes that carry independent semantic content (Pan et al., 2023; Perfetti & Liu, 2006). For instance, the two-character word "地震" (meaning "earthquake") is composed of "地" ("ground") and "震" ("quake"), both of which carry independent meanings. Therefore, transposing two characters potentially changes both character order (i.e., an orthographic relation) and the reordering of two meaning-bearing morphemes. This raises a possibility that the position encoding in Chinese may occur at the semantic level.

In light of these considerations, an open question remains: Is character position encoding in Chinese shaped by the morpheme-level semantic relationships carried by characters? To address this question, the present study examined whether the semantic relatedness between constituent characters modulates the TC effect during reading. We conducted two experiments, one with isolated word presentation and another with natural sentence contexts, in which we systematically varied the semantic relationship between the morphemes (i.e., characters) in two-character compound words. In what follows, we first review relevant evidence on position encoding in alphabetic scripts, then consider the specific properties of Chinese orthography and morphology. Subsequently, we review morpheme position encoding in alphabetic scripts and several relevant studies on this issue in Chinese. We then present the rationale for our two experiments, each designed to test whether morpheme-level semantic relationships between characters modulate the TC effect.

Letter Position Encoding in Alphabetic Languages

Research on position encoding in alphabetic languages has repeatedly reported that letter position coding is flexible rather than strictly slot-based. Early models implementing rigid position slots (e.g., McClelland & Rumelhart, 1981; see also Coltheart et al., 2001; Grainger & Jacobs, 1996) predict a comparable reading cost for transposed-letter (TL) nonwords (e.g., "*jugde"*) and substituted-letter (SL) nonwords (e.g., "*jupte"*). Empirically, however, TL pseudowords are consistently processed as more similar to their base words than SL controls, both in lexical decision and in sentence reading (e.g., Perea & Carreiras, 2006; Perea & Lupker, 2004; Rayner et al., 2006). This pattern motivates the idea that positional information is encoded with graded uncertainty, such that pseudowords created from letter transpositions preserve more "wordlikeness" from the base word than those created from letter substitutions.

To account for this effect, a range of models have been proposed (e.g., Davis, 2010; Gómez et al., 2008; Grainger & Van Heuven, 2004; Norris & Kinoshita, 2012; Snell, 2025). Critically, despite their differences, these accounts converge that the locus of letter transposition effects is primarily orthographic. The present set of experiments takes this orthographic consensus as a starting point and examined whether it generalizes to Chinese, where characters are meaning-bearing morphemes. In particular, character transpositions swap morpheme order, providing a test of whether the TC effect in Chinese is modulated by morpheme-level semantic relationships between constituent characters.

Character Position Encoding in Chinese

Chinese, as a logographic writing system, differs significantly from alphabetic scripts, potentially resulting in unique position encoding mechanisms. First, although both Chinese characters and alphabetic letters serve as orthographic symbols carrying phonological information, most Chinese characters also function as morphemes with independent semantic content—unlike alphabetic letters, which are generally meaningless in isolation (Mattingly, 1992; Perfetti & Liu, 2006; Zhang et al., 2024). Second, with over 5,000 commonly used Chinese characters compared to only 26 letters in the English alphabet, Chinese has significantly fewer anagrams. Third, compared to alphabetic languages, Chinese words are on average shorter and have less variability in length, with 72% of them consisting of only two characters (Lexicon of Common Words in Contemporary Chinese Research Team, 2008). For example, the two-character word "地震" ("earthquake") does not form a new word when the two characters are transposed ("震地"). Consequently, transposing two characters in a two-character word may be more salient than transposing two letters within longer alphabetic words. These linguistic differences raise the possibility that models of letter position encoding developed for alphabetic scripts may not fully generalize to Chinese.

Evidence concerning character position encoding in Chinese comes from the transposed-character effect (TC effect), first demonstrated by Gu et al. (2015). Using a masked priming lexical decision task and the boundary paradigm during sentence reading, Gu et al. compared three prime/preview conditions: identity (e.g., "地震" – "地震" [earthquake]), TC (e.g., "震地" – "地震"), and substituted-character (SC; e.g., "牲缉" – "地震"). Results showed that TC primes/previews led to shorter response latencies or shorter reading times than SC ones, while the identity condition yielded the greatest facilitation overall. In the present paper, we use the "TC effect" to refer specifically to the advantage for the TC condition relative to the SC condition, and "transposition cost" to refer to the cost of a transposition relative to the intact identity condition. The difference between the identity and TC conditions indicates that transposing the two characters disrupts recognition relative to the intact word. The difference between the TC and SC conditions indicates that transposed-character strings preserve greater similarity to the base word than substituted-character controls. These findings suggest that Chinese readers tolerate character transpositions to some extent. Later studies have consistently replicated the TC effect in both isolated word recognition and natural sentence reading (Gu et al., 2023; Gu & Li, 2015; Liu et al., 2025; Su et al., 2024; Yang et al., 2022).

The encoding of character position in Chinese involves unique properties, differing from letter order encoding in alphabetic systems. First, the basic orthographic units are not equivalent: letters are minimal, meaningless units, while Chinese characters are complex units representing a syllable, a morpheme, and are themselves composed of sub-lexical radicals (Li, Marinus, et al., 2021; Li et al., 2022; Tong et al., 2020; Yao et al., 2022). Second, alphabetic writing like English is linear, relying heavily on sequence memory, whereas Chinese character recognition additionally requires encoding internal spatial structure (i.e., the arrangement of strokes and radical), placing different cognitive demands on positional encoding. Consequently, character order encoding cannot be directly equated with letter order encoding in alphabetic systems. Moreover, empirical findings challenge the analogy between Chinese radical position encoding and letter position encoding in English as well. Taft et al. (1999) found that transposing radicals within a character (e.g., "杏" vs. "呆") causes minimal interference, indicating position-specific radical representations that are less flexible than letter position encoding in English. Furthermore, Tong et al. (2020) demonstrated that radical processing is semantically mediated. They found that Chinese readers were more likely to omit the radical "口" when it appeared in top/inside positions with low semantic relatedness, but also when it was on the right with high semantic relatedness. This influence of meaning at the radical level has no parallel in English, where letters do not carry semantic information themselves. Therefore, the cognitive mechanisms for encoding character position in Chinese are likely unique, necessitating language-specific models of word recognition and reading. In sum, Chinese shows robust transposition effects akin to those in alphabetic languages. However, given the structural differences in the writing system—and the fact that character transpositions also entail morphemic shifts—position encoding in Chinese may be influenced by morpheme-level semantic relationships.

Morpheme Transposition in Alphabetic Writing Systems

A number of studies have investigated morpheme encoding in alphabetic languages by manipulating morpheme positions, suggesting that position encoding may vary depending on the type of morphological structure. Morphologically complex words can be broadly categorized into three types: inflectional, derivational, and compound words (Booij, 2012). Derivational and inflectional words modify the grammatical or meaning function of stems through affixes (e.g., "kindness"), while compound words are created by combining two free stem morphemes (e.g., "bookshop").

Previous studies suggest that morpheme position encoding is stricter in affixed words than in compound words. Using an unprimed lexical decision task, Crepaldi et al. (2010, Experiment 3) found no significant difference in rejection times between transposed derivational suffixed nonwords (e.g., “nesskind”) and their control nonwords (e.g., “nusskind”), suggesting a strict position encoding. In contrast, morpheme position encoding in compound words is more flexible. Crepaldi et al. (2013, Experiment 1) found that transposed compound nonwords (e.g., “moonhoney”) took longer to reject than their orthographic controls (e.g., “moonbasin”). Further, using masked priming lexical decision task (Experiments 2 and 3), they found transposed compound nonword primes (e.g., “firecamp” – “CAMPFIRE”) significantly reduced RTs for their base words compared to control primes (e.g., “jlvkxbwu” – “CAMPFIRE”). This indicated morphemes in compound words can be activated regardless of their position. Spencer et al. (2024) directly compared derivational and compound structures using unprimed lexical decision. They replicated the finding that transposed compound nonwords (e.g., “portair”) were harder to reject than controls (e.g., “handair”), whereas transpositions in derivational words (e.g., “fulhelp” vs. “tichelp” or “agreedis” vs. “agreenon”) yielded no effect, again supporting the idea that morpheme order is more rigid in affixed forms.

Why is morpheme encoding more constrained in derivational words? Spencer et al. argued that derivational affixes typically serve clear grammatical or semantic functions (e.g., -*er* in *teacher* indicates agent; *dis*- in *disagree* signals negation), which depend on their canonical positions. Only when affixes appear in their expected locations can their roles be reliably parsed. In contrast, the free morphemes that form compounds can stand alone and carry independent meanings, allowing more positional flexibility.

In sum, the flexibility of morpheme position encoding appears to depend on the semantic autonomy of the morphemes involved. This suggests that position encoding can be shaped by semantic structure. Given the morphemic nature of Chinese characters, it is plausible that character position encoding in Chinese engages morpheme-level semantic representations.

Models of Compound Word Processing in Chinese Reading

Different models of compound word processing made different assumptions regarding how semantic information influence character position encoding. Taft's model of compound word processing provides a strong rationale for our approach (Taft & Forster, 1975; Taft, 2023). Applied to Chinese, the model holds that character morphemes are activated at the lemma level and combined to access the word's meaning for both transparent and opaque compound words, although the character morphemes are not activated for monomorphemic words (Li & Taft, 2020; Taft, Zhu & Peng, 1999; Wei et al., 2023; See Figure 1). A key implication is that any semantic activation carried by the constituent morphemes can, in principle, enter into the combinatorial stage at which constituent activation is mapped onto a coherent compound representation. Thus, on this account, semantic relatedness between constituents is expected to modulate the similarity between a transposed-character string and its base words (i.e., a transposition would preserve "which morphemes are present", while altering their order), and the semantic consequences of this transposition would depend on how closely related the two morphemes are. In contrast, other models posit that the meaning of a compound word is not necessarily composed from its constituent characters (Zhou & Marslen-Wilson, 1995, 2000, 2009). In this framework, compound words and their constituents are represented at orthographic (O), phonological (P), and semantic (S) levels, with direct connections between these levels. Critically, the representation of the meaning of compound word is not mandatorily derived via composition of its constituent morphemes; instead, the semantic representations of both the whole word and its morphemes are activated in parallel from their form representations. Consequently, Zhou and Marslen-Wilson's model predicts that transposing two characters will cause a similar level of disruption, regardless of the semantic relatedness between them.

---Insert Figure 1 about here---

The Chinese Reading Model (CRM) proposed by Li and Pollatsek (2020) provides another perspective on the mechanism of Chinese compound word processing. In the word recognition module of CRM, characters within the perceptual span are activated in parallel at the character level, and then they activate possible words containing these characters. All activated words compete for recognition, and the word with the highest activation wins. When CRM processes compound words, both single and multi-character words within the perceptual span are activated, and they compete for a winner. Compound words win most of the time because they receive activation from all constituent characters, and their activation value increases faster than the embedded single-character words. Therefore, CRM predicts that compound words are ultimately identified as a whole. CRM inherits the Interactive Activation Model (IAM; McClelland & Rumelhart, 1981) orthographic front-end and assumes rigid character position encoding. Therefore, it lacks a mechanism to simulate flexible character order encoding in its present implementation.

Previous Studies on Semantic Influences in Character Position Encoding

Previous research has preliminarily explored whether character position encoding in Chinese engages morpheme-level semantic processing. Two main approaches have been used.

First, several studies compared the TC effect between single-morpheme words and multiple-morpheme words (Gu et al., 2015; Yang et al., 2022). In single-morpheme words, the constituent characters either lack independent meaning (e.g., "哆嗦" [tremble], "哆" and "嗦" are meaningless in isolation) or have meanings unrelated to the whole word (e.g., "烂漫" [brilliant], with "烂" [bad] and "漫" [overflow]). Yang et al. (2022) further classified these as simple (no meaningful constituents) and complex (semantically divergent constituents). In contrast, multiple-morpheme words consist of characters that each carry independent, compositional meaning (e.g., "地震" [earthquake], with "地" [ground] and "震" [quake]). If semantic information contributes to position encoding, the TC effect should be larger for multiple-morpheme words due to contributions from both orthographic and semantic factors.

Gu et al. (2015) tested this hypothesis using masked priming lexical decision and the boundary paradigm, but found no significant difference in the TC effect between word types. Yang et al. (2022, Experiment 1) replicated these findings using the same materials, procedure and primed lexical decision task under more standardized parameters (with a monitor of 60 Hz refresh rate and 50 ms prime duration instead of 150 Hz and 60 ms). Notably, 88% of their single-morpheme words were of the simple type. In a later experiment, Yang et al. directly compared the TC effect across single-morpheme simple words (e.g., "萧瑟", meaning *rustle in the air*; an English parallel word would be "practice"), single-morpheme complex words (e.g., "烂漫", meaning *blossoming*; an English parallel word would be "carpet" [car/pet]), and multiple-morpheme words (e.g., 谦逊, meaning *modest*; an English parallel word would be "earthquake" [earth/quake]). Again, TC effect was comparable for simple and multiple morpheme words. Even when Yang et al. directly compared simple, complex, and multiple-morpheme words, the magnitude of the TC effect remained equivalent. These converging null results suggest that the mere presence of independent character meaning does not modulate transposition costs.

Second, Yang et al. (2020) examined position encoding mechanisms using backward priming—a design in which reversed four-character primes (e.g., "同不所有") facilitated recognition of canonical target words such as "有所不同" (pronouncing /yǒu suǒ bù tóng/). In Experiments 1-2, they tested phonological contributions using both backward and forward phonological primes (e.g., "佟步锁友", pronounced as /tóng bù suǒ yǒu/ or "友锁步佟", pronounced as /yǒu suǒ bù tóng/) in a masked priming lexical decision task. However, they found no priming effects compared to phonologically unrelated controls (e.g., "探话养咿", pronounced as /tàn huà yǎng yī/), thereby ruling out phonological overlap as the source of facilitation. To test the role of morpheme-level semantic effects, Experiments 3-4 employed a masked priming same-different task, which minimizes semantic processing by requiring participants to decide whether a briefly primed target is identical to a previously shown reference stimulus. In Experiment 3, using materials from Gu et al. (2015), the primes were either transposed-character nonwords (e.g., "震地" for "地震") or substituted-character nonwords (e.g., "肴肿"). The results showed no modulation of the TC effect by morpheme type, suggesting that morphemic information did not modulate character position encoding. In Experiment 4, they applied the same paradigm to four-character compounds and again observed backward priming effects equivalent in magnitude to those found in lexical decision. Given that the same-different task is designed to reduce semantic processing, the convergence of results across tasks indicates that the TC effect in Chinese primarily reflects orthographic processing, with no reliable evidence for morpheme level semantic involvement.

In sum, current evidence provides no compelling support for morpheme-level semantic processing in Chinese character position encoding. Although some studies have manipulated morphemic structure and conducted cross-paradigm comparisons, results have been consistently null. At the same time, these conclusions rest on indirect comparisons and null interactions, and therefore benefit from further confirmatory tests that isolate morpheme-level semantic effects while holding orthographic overlap constant. This idea was behind the logic of the manipulation of the present experiments.

The Present Study

In the present study, we developed a novel manipulation to test whether character position encoding in Chinese is modulated by morpheme-level semantic information. Participants read two-character Chinese words either in isolation (Experiment 1) or embedded in sentences (Experiment 2), and we examined how character transpositions affected reading performance.

Crucially, unlike previous studies, we manipulated the semantic relatedness between the two constituent characters of the target words. In the semantic-related condition, the two characters were semantically related (e.g., "磨蹭" [dawdle], where "磨" [rub] and "蹭" [rub] share a related meaning); in the semantic-unrelated condition, the characters were not meaningfully related (e.g., "浮夸" [exaggeration], where "浮" [float] and "夸" [praise] have unrelated meanings).

To avoid confounds with semantic transparency, we only selected words with low transparency, where the meaning of the whole word could not be easily inferred from its parts. This control was essential, as semantic relatedness between characters is typically correlated with the transparency of the full word: semantically related characters often form transparent words, whereas semantically unrelated characters tend to form opaque ones. Since semantically unrelated characters cannot form semantically transparent words, restricting the stimuli to low-transparency items ensured a fair comparison between conditions.

Thus, these two experiments test whether character position encoding in Chinese is modulated by morpheme-level semantic relationships. If morpheme semantics contributes to the computation of character position encoding (i.e., beyond orthographic overlap), as proposed by Taft (2023), transpositions of two characters should be more word like when the constituent morphemes are semantically close, yielding a larger TC effect in the semantically related condition than in the semantically unrelated condition. In contrast, if character position encoding is effectively insulated from morpheme semantics, as predicted by Zhou and Marslen-Wilson's model (2009), the size of the TC effect should be comparable across semantic conditions, regardless of whether the characters are semantically related.

Experiment 1

Experiment 1 employed an unprimed lexical decision task to investigate whether character position encoding in Chinese is modulated by morpheme-level semantic relationships. The lexical decision task is widely used in visual word recognition research and is known to be sensitive to a range of linguistic factors, including character position. Using this paradigm, many studies have provided valuable evidence on how letter order is processed in alphabetic writing systems (Lupker et al., 2008; Perea & Lupker, 2004; Romero-Ortells et al., 2024). Thus, the present experiment offers a controlled test of whether semantic similarity between constituent morphemes shapes the tolerance to character transpositions.

Method

Participants

Seventy Chinese native speakers (41 females; aged 18-30 years, M = 21.8, SD = 2.97, all native speakers of Chinese) with normal or corrected-to-normal vision were recruited from [name deleted to maintain the integrity of the review process]. They were paid 10 Yuan (approximately USD $1.6) to compensate for the time of participating in the experiment. With 23 trials per condition, the sample size provided 1,610 observations per condition, satisfying Brysbaert and Stevens' (2018) recommendation of 1,600 observations for a well-powered experiment.

Materials and Design

The experiment employed a 2 (semantic relatedness: related or unrelated) × 3 (presentation type: Intact, TC or SC) within-participant design (See Table 1 for a description of the stimuli properties). We selected 138 two-character Chinese words from the SUBTLEX-CH corpus (Cai & Brysbaert, 2010). Half of the words (n = 69) consisted of semantically related character pairs (e.g., "磨蹭", [meaning dawdle], "磨", [meaning rub], "蹭", [meaning rub]), while the other half (n = 69) consisted of semantically unrelated character pairs (e.g., "浮夸", [meaning exaggeration], "浮", [meaning float], "夸" [meaning praise]). Forty participants who did not participate in the main experiment rated the semantic relatedness on a 7-point Likert scale (1 = *completely unrelated* to 7 = *highly related*) between the two constituent characters. The semantic relatedness between the two characters was higher for semantic-related words (M = 5.03, SD = 0.51, range: 4.6-6.5) than semantic-unrelated words (M = 1.78, SD = 0.43, range: 1.1-2.9), t (68) = 40.37, p < .001. To make sure that the selected words are opaque words, the same group of participants also rated the semantic relatedness between each character and the whole word. The semantic transparency of each word was calculated as the average of the two character-word relatedness ratings. The results showed that semantic transparency was not significantly different between semantic-related words (M = 2.72, SD = 0.52, range: 1.6-3.85) and semantic-unrelated words (M = 2.62, SD = 0.62, range: 1.3-3.85), t (68) = 1.23, p = .23.

---Insert Table 1 about here---

Each target word was presented in three conditions: Intact condition, TC condition, and SC condition. In the intact condition, the stimuli were presented as the word itself (e.g., "马虎" for semantic-related words; "浮夸" for semantic-unrelated words). In the TC condition, the stimuli were formed by transposing two characters (e.g., "虎马" for semantic-related words; "夸浮" for semantic-unrelated words). In the SC condition, the stimuli were formed by substituting two characters (e.g., "身衣" for semantic-related words; "州克" for semantic-unrelated words). In the SC condition, two semantically related characters (e.g., "身衣" [a nonword], "身" [meaning body], "衣" [meaning clothes]) were chosen in the semantic-related condition, and two semantically unrelated characters (e.g., "州克" [a nonword], "州" [meaning state], "克" [meaning gram]) were chosen in the semantic-unrelated condition. This design ensures that the SC condition is matched to the TC condition for constituent-level semantic relatedness within each semantic class, so that the TC–SC contrast provides an index of the standard TC effect while minimizing differences in constituent-level semantic similarity. The semantic relatedness between the two characters, as measured by the ratings of forty participants, indicated no difference between TC and SC nonwords in the semantic-related condition (TC: *M* = 5.03, *SD* = 0.51; SC: *M* = 5.10, *SD* = 0.72; *t* [68] = -0.63, *p* = .53) and semantic-unrelated condition (TC: *M* = 1.78, *SD* = 0.43; SC: *M* = 1.66, *SD* = 0.53; *t* [68] = 1.44, *p* = .16).

To reduce the influence of other factors, we matched word frequency, total stroke count, first/second character frequency, first/second character stroke count between semantic-related and semantic-unrelated words (*ps* > .05). In addition, we ensured that the TC and SC conditions were matched for both word types in terms of the first/second character frequency, first/second character stroke (*p*s > .05) (all metrics derived from Cai & Brysbaert, 2010; see Table 1 for descriptive statistics). To control for potential confounding effects of shared radicals (e.g., "氵" in "流浪"), neither character in word/nonword stimuli shared a radical component. Additionally, 46 words were included as fillers to balance the number of words and nonwords in the unprimed lexical decision task.

There were three counterbalanced lists, ensuring that each item was presented under only one condition (Intact, TC, or SC) across the experiment. Each list contained an equal proportion of word and nonword items, and each participant completed only one list.

Apparatus

All stimuli were displayed in black 60-pt Song font on a white background at the center of a 14-inch LCD monitor with 2560×1600-pixel resolution and a refresh rate of 60Hz. The experiment was controlled using the E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA; see Schneider et al., 2002).

Procedure

Participants performed an unprimed lexical decision task in this experiment. Each trial began with a central fixation cross (500 ms), followed by a two-character stimulus presented until response. If the participants did not respond within 3,000 ms, the stimulus disappeared and the next trial started. Participants were asked to press "F" or "J" to judge whether the stimulus was a word or nonword as quickly and accurately as possible, with key assignments counterbalanced across participants. The stimulus immediately disappeared upon response. After that, a blank screen was displayed for 1,000 ms. Participants first completed 12 practice trials to familiarize themselves with the task. Afterward, they proceeded to the main experiment, which was divided into two blocks with randomized trial orders within blocks. Participants could have brief rests between blocks. RTs were recorded from stimulus onset until participants made a response. The entire session lasted approximately 10 minutes.

Data Analysis

Error rates (ERs) were analyzed using generalized linear mixed-effects models (GLMMs), and RTs were analyzed using linear mixed-effects models (LMMs), implemented in R (v4.3.1; R Core Team, 2023) with the *lmerTest* package (v3.1.3; Kuznetsova et al., 2020), both optimized using the *bobyqa* algorithm. Word fillers were excluded from the analysis. The mean accuracy of the lexical decisions was 93.0%, suggesting participants performed the task attentively. Trials with RTs shorter than 200 ms, longer than 3,000 ms (0.01%), or exceeding 3 standard deviations from the mean for each condition and participant (1.58%) were excluded, along with trials with incorrect responses (8.18%). RTs were log-transformed to meet the normal distribution assumptions of LMMs. The descriptive statistics are presented in Table 2.

Given that participants were required to make a “yes” response to intact words, and a “no” response to both TC and SC nonwords, it is not feasible to conduct an interaction analysis involving response type. Therefore, we directly used the following five customized contrasts to test the theoretical hypotheses as suggested by Schad et al. (2020): (1) the difference between the TC and intact conditions, which aimed to examine the cost of character transposition relative to the intact word; (2) the difference between the TC and SC conditions, which indexed the standard TC effect; (3) the difference in the TC effect between the semantic-related and semantic-unrelated conditions, which constituted the critical test of whether morpheme-level semantic relatedness modulates the TC effect; (4) the difference between semantic-related and semantic-unrelated words in the intact condition, which aimed to examine the impact of semantic relatedness of characters on lexical decision; (5) the main effect of semantic relatedness in the TC and SC conditions, which aimed to examine the impact of semantic relatedness of two characters on nonwords lexical decision. For clarity, the critical test of semantic modulation is Contrast (3) (i.e., whether the TC-SC difference is modulated by semantic relatedness). Data from both experiments satisfied the LMM and GLMM assumptions of linearity, homoscedasticity, and normally distributed residuals (Meteyard & Davies, 2020).

We specified the participants and items as crossed random effects, including intercepts and slopes. Following Barr et al. (2013), we used the maximal model that could converge. We first constructed a model with a maximal random factor structure. When the maximal model failed to converge, we used a zero-correlation parameter model and removed the random components that generated the smallest variances. We report the regression coefficients (*b*), standard errors (*SE*), *t*/*z*-values, and *p* values of the optimal model (see Supplement 1 for the final models and model outputs).

Results

Error Rate

TC nonwords resulted in more errors than the intact words, $b = 1.31$, $SE = 0.23$, $z = 5.72$, $p < .001$, indicating a transposition cost (see Table 2 for detailed results). Furthermore, TC nonwords resulted in more errors than SC nonwords, $b = 1.00$, $SE = 0.20$, $z = 5.13$, $p < .001$, indicating the typical TC effect. Importantly, there was no difference in the TC effect between the semantic-related and semantic-unrelated conditions, $b = 0.05$, $SE = 0.32$, $z = 0.15$, $p = .878$, providing no evidence that morpheme-level semantic relatedness modulates the TC effect.

Additionally, we found an effect of semantic relatedness on lexical decision. Specifically, in the TC and SC conditions, participants made more errors for semantic-related nonwords than semantic-unrelated nonwords, $b = 0.74$, $SE = 0.16$, $z = 4.75$, $p < .001$, indicating semantic-related nonwords were harder to reject. In the intact condition, there was a trend that participants made more errors for semantic-unrelated words than semantic-related words, $b = -0.54$, $SE = 0.29$, $z = -1.84$, $p = .066$.

---Insert Table 2 about here---

Reaction Time

TC nonwords elicited longer RTs than intact words, $b = 0.23$, SE = 0.02, $t = 14.29$, p < .001, indicating a transposition cost. TC nonwords elicited longer RTs than SC nonwords, $b = 0.07$, $SE = 0.01$, $t = 5.02$, $p < .001$, indicating a typical TC effect. Importantly, there was no difference in the TC effect between the semantic-related and semantic-unrelated conditions, $b = -0.02$, $SE = 0.02$, $t = -0.97$, $p = .335$, providing no evidence that morpheme-level semantic relatedness modulates the TC effect.

Similar to the results of error rates, we also found an effect of semantic relatedness on lexical decision in the RT analysis. Specifically, in the TC and SC conditions, participants showed longer RTs for semantic-related nonwords than semantic-unrelated nonwords, $b = 0.04$, $SE = 0.01$, $t = 4.37$, $p < .001$. In the intact condition, RTs showed no significant difference between the semantic-related and semantic-unrelated conditions, $b = -0.01$, $SE = 0.02$, $t = -0.88$, $p = .384$.

Bayes Factor Analysis

Error rate and RT analyses did not reveal a difference in the TC effect between semantic-related and semantic-unrelated conditions. To quantify evidence for the absence of this semantic modulation, we conducted Bayes factors using the *BayesFactor* package in R (version 0.9.12-4.7; Morey & Rouder, 2024). Specifically, we compared a model that excluded the critical semantic-relatedness × (TC–SC) term with an otherwise identical model that included this term. Bayes factors are reported as BF_{01} (i.e., evidence for the null hypothesis relative to the alternative hypothesis), where values greater than 1 favor the null hypothesis. Values exceeding 3, 10, and 30 respectively provide “moderate”, “strong”, and “very strong” evidence for the null hypothesis (Lee & Wagenmakers, 2013; see also Jeffreys, 1961). For error rates, we found BF_{01} of 2.88, indicating anecdotal-to-moderate evidence for the null hypothesis; for the RT analyses, $BF_{01} = 12.11$, indicating strong evidence for the null hypothesis[1]. Thus, in both cases, the Bayes factors favors a model that excludes a semantic modulation of the TC effect versus an otherwise identical model that includes this

[1] For concision, we only report Bayes factors related to semantic modulation of the TC effect here. See Table S1 for Bayes factors for all of the other contrasts in the model.

modulation term.

Discussion

The present experiment tested whether character position encoding in Chinese is modulated by semantic information. Results showed the typical TC effect: TC nonwords were harder to reject than SC nonwords. However, there was no significant difference in the TC effect between semantic-related and semantic-unrelated conditions. Thus, under the present conditions, morpheme-level semantic relatedness did not modulate the TC effect. These findings are consistent with prior work (Gu et al., 2015; Yang et al., 2020, 2022), suggesting that character order encoding in Chinese does not reflect the modulation of morpheme-level semantics.

At the same time, participants were slower and less accurate in rejecting nonwords formed by semantically related character pairs, indicating that semantic relatedness can interfere with lexical decisions independently of position encoding. This effect may reflect the increased difficulty of rejecting nonwords that are semantically related at the morpheme level. Importantly, because semantic relatedness affected the overall nonword rejection rate without altering the TC effect, the present pattern is consistent with a semantic contribution to the overall wordlikeness in lexical decision, and morphemic semantic factors are unrelated to the positional uncertainty (i.e., the TC effect).

Experiment 2

Experiment 1 found no evidence that semantic relatedness at the morpheme level modulates character position encoding in an isolated word recognition task. To extend this investigation to a more naturalistic context, Experiment 2 used eye-tracking to assess how character transpositions and semantic relatedness interact in a more natural sentence reading task.

Method

Participants

Seventy Chinese native speakers (45 females; aged 18-29 years, *M* = 22.7, *SD* = 2.56, all native speakers of Chinese) with normal or corrected-to-normal vision were recruited from [name deleted to maintain the integrity of the review process]. They were paid 30 Yuan (approximately USD $5) to participate in the experiment. None of them participated in Experiment 1. With 23 trials per condition, the sample size provided 1,610 observations per condition, satisfying Brysbaert and Stevens' (2018) recommendation of 1,600 observations per condition.

Apparatus

Eye-movement data were recorded using an SR Research Eyelink 2000 eye-tracking system with a sampling rate of 1,000 Hz. Stimuli were presented on a 21-inch CRT monitor (resolution: 1,024 × 768 pixels; refresh rate: 150 Hz). The viewing distance was maintained at 58 cm between the monitor and participants' eyes, with head movements minimized using chin and forehead rests. For each participant, the viewing was binocular, but only the right eye was monitored. All experimental materials were displayed in a single line using 28-pt Song font with white (RGB: 255, 255, 255) characters on a grey (RGB: 128, 128, 128) background, subtending approximately 1° of visual angle per character. The movements of the participant's right eye were recorded throughout the experiment. The experimental procedure was programmed using the EyeTrack software developed by the UMASS Eye Tracking Lab, and data extraction was directly from the experimental software.

Materials and Design

The stimuli from Experiment 1 were embedded into sentences for Experiment 2. The target words and design were identical to those in Experiment 1. Examples of the experimental materials are shown in Table 3. Twenty-one participants rated the plausibility of the sentences on a 7-point scale (1 = *highly implausible*, 7 = *highly plausible*) to ensure they were plausible and that there was no difference in semantic-related ($M = 6.12$, $SD = 0.43$) and semantic-unrelated conditions ($M = 6.05$, $SD = 0.46$), $t(68) = 0.87$, $p = .387$. Another 20 participants completed a predictability test in which they were presented with text before the target words and were asked to provide the most likely subsequent word. Predictability was close to 0 in both conditions, confirming that the target words were not predictable. The target words were in the middle of the sentence so that they were at least six characters away from the beginning and the end of the sentence.

---Insert Table 3 about here---

Procedure

Upon arrival at the laboratory, participants first read a brief introduction of the experimental procedures and equipment. They were then verbally instructed on the tasks involved. The eye tracker was calibrated using a three-point calibration procedure at the beginning of the experiment, with validation errors maintained below 0.5°. Recalibration was performed as needed during the experiment. Each trial began with a drift check procedure, during which participants were required to fixate on a central circle displayed on the screen. Following successful drift check, a white fixation box appeared at the position corresponding to the first character of the sentence. A sentence was displayed immediately upon detection of the participant's fixation on the box. Participants first read eight practice sentences to familiarize themselves with the procedure, and then read 138 experimental sentences and 46 filler sentences in a random order. Comprehension questions related to the sentences were presented after approximately 30% of the sentences to ensure participants read the sentences carefully. The entire experiment lasted approximately 30 minutes.

Data Analysis

Mean accuracy of the comprehension questions was 97%, indicating that participants understood the sentences well. Trials were excluded if participants blinked more than three times during sentence reading or blinked on the target region, resulting in the exclusion of 5.32% of the trials. Three participants were excluded because more than one third of their trials were removed. Extremely short (< 80 ms) fixations and extremely long (> 1,000 ms) fixations (approximately 2.35%) were excluded from the data set prior to analyses.

We analyzed the following eye-movement measures for the target word region: (1) First fixation duration (the duration of the first fixation on the target region during the first-pass reading); (2) Gaze duration (the summed duration of all first-pass fixations on the target region before moving on to other words); (3) Go-past time (the summed duration starting when entering the target region until this region's right boundary is crossed); (4) Total reading time (the sum of all fixations on the target word, including regressions); (5) saccade landing position (landing position of the first forward saccades onto the target word, coded as 0 for the first character and 1 for the second character). The descriptive statistics are presented in Table 4.

All measures were analyzed using linear mixed-effects models (LMMs) in *R* (v4.3.1; R Core Team, 2023) with the *lmerTest* package (v3.1.3; Kuznetsova et al., 2020), optimized using the *bobyqa* algorithm. All measures were log-transformed to meet LMMs' normality assumptions. As in Experiment 1, we directly used the five customized contrasts to test the theoretical hypotheses, with the critical constraint being whether the TC effect varies as a function of semantic relatedness. In addition, because the present design forms a full 3 x 2 factorial design, in Appendix 1 we also report the corresponding fixed effects for semantic relatedness, presentation type, and their interaction within a single model. Following Barr et al. (2013), we used the maximal model that could converge. We first constructed a model with a maximal random factor structure. When the maximal model failed to converge, we used a zero-correlation parameter model and removed the random components that generated the smallest variances. We report regression coefficients (b), standard errors (SE), t/z-values, and p values of the optimal model (see Supplement 1 for the final models and outputs).

Results

The main effect of presentation type was significant in all measures (See Table 4 for detailed results). Reading times were longer in the TC condition than in the intact condition, as shown in first fixation duration (b = -0.07, SE = 0.01, t = -6.89, p < .001), gaze duration (b = -0.11, SE = 0.02, t = -7.37, p < .001), go-past time (b = 0.12, SE = 0.02, t = 7.09, p < .001) and total reading time (b = -0.23, SE = 0.02, t = -10.38, p < .001), indicating a transposition cost. Reading times were shorter in the TC condition than in the SC condition, as shown in first fixation duration (b = 0.10, SE = 0.01, t = 7.98, p < .001), gaze duration (b = -0.17, SE = 0.02, t = -10.15, p < .001), go-past time (b = -0.25, SE = 0.02, t = -11.14, p < .001) and total reading time (b = -

0.39, *SE* = 0.03, *t* = -14.07, *p* < .001), indicating the typical TC effect.

---Insert Table 4 about here---

Importantly, there was no difference in the TC effect between the semantic-related and semantic-unrelated conditions, as shown in first fixation duration (*b* = -0.03, *SE* = 0.02, *t* = -1.22, *p* = .227), gaze duration (*b* = -0.04, *SE* = 0.03, *t* = -1.49, *p* = .142), go-past time (*b* = -0.06, *SE* = 0.03, *t* = -1.52, *p* = .132) and total reading time (*b* = -0.04, *SE* = 0.03, *t* = -1.35, *p* = .183) providing no evidence that morpheme-level semantic relatedness modulates the TC effect.

Semantic relatedness nevertheless had an effect in some eye movement measures. Specifically, in the intact condition, fixation durations were significantly shorter in the semantic-related condition than the semantic-unrelated condition for go-past time (*b* = -0.06, *SE* = 0.03, *t* = -2.26, *p* = .024) and total reading time (*b* = -0.08, *SE* = 0.03, *t* = -3.23, *p* = .001), but was not significant for first fixation duration (*b* = -0.02, *SE* = 0.01, *t* = -1.15, *p* = .251) and gaze duration (*b* = -0.02, *SE* = 0.02, *t* = -1.25, *p* = .221). In the TC and SC conditions, participants showed longer fixation durations for semantic-related nonwords than semantic-unrelated nonwords, an effect that reached significance only in first fixation duration (*b* = -.03, *SE* = 0.01, *t* = -2.57, *p* =.012), but the effect was not significant in gaze duration (*b* = -0.02, *SE* = 0.02, *t* = -1.55, *p* = .123) , go-past time (*b* = -0.04, *SE* = 0.02, *t* = -1.79, *p* = .077) or total reading time (*b* = -0.04, *SE* = 0.02, *t* = -1.80, *p* = .074). The analyses of saccade landing position did not show any significant effects. Critically, as shown in Appendix 1, the semantic-relatedness × presentation-type interaction was not significant in any eye-movement measure.

Bayes Factor Analysis

As in Experiment 1, we used Bayes factors to quantify the statistical evidence for the absence of a semantic modulation of the TC effect, this time in sentence reading. For each dependent variable, we compared a model without the TC–SC × semantic-relatedness term to an otherwise identical model with this term (Bayes factor package; the default prior) and report BF_{01}[2]. The resulting value of BF_{01} was 6.34 for first fixation duration, 4.83 for gaze duration, 3.19 for go-past time, 7.82 for total reading time, 9.13 for landing position; thus, all indexes provide consistent evidence in favor of a null interaction between semantic relatedness and presentation type.

Discussion

This eye movement experiment examined whether semantic relatedness between characters modulates the TC effect during natural sentence reading. As in Experiment 1, we found the classic TC effect in all measures: TC nonwords were read more easily than SC nonwords. However, results revealed no TC–SC × semantic-relatedness interaction across any eye-movement measure. This absence of interaction was supported by Bayesian analyses. Overall, these findings extend the conclusion that morpheme-level semantic relatedness did not modulate the TC effect in sentence reading.

Notably, consistent with Experiment 1, we observed main effects of morpheme-level semantic relatedness: words composed of semantically related characters were read more quickly than those with unrelated characters. This pattern suggests that semantic relatedness between morphemes can facilitate lexical processing, though it

[2] For concision, we only report Bayes factors related to semantic modulation of the TC effect here. See Table S2 for Bayes factors for all of the other contrasts in the model.

did not modulate the TC effect in the present experiments. We return to the implications of this dissociation in the General Discussion.

General Discussion

The present study investigated whether position encoding in Chinese is modulated by morpheme-level semantic relationships. To this end, we carried out two experiments using complementary paradigms: an unprimed lexical decision task (Experiment 1) and sentence reading with eye-tracking (Experiment 2). In both experiments, we manipulated the semantic relatedness between the two constituent characters of disyllabic words, while controlling for semantic transparency. Results showed sizable transposition effects: in lexical decision, TC nonwords produced more errors and slower responses than intact words, and were harder to reject than SC nonwords; in sentence reading, TC nonwords produced longer fixation times than intact words, and shorter fixation times than SC nonwords. In other words, using two experimental tasks that differ markedly in decision demands and time-course sensitivity, transposed-character stimuli consistently behaved as more “wordlike” than substituted-character controls.

Crucially, however, neither experiment showed any differences in the TC effect between morpheme-level semantic-related versus -unrelated words, providing no evidence that morpheme-level semantic relatedness modulates the TC effect in Chinese. These findings are consistent with the predictions of Zhou and Marslen-Wilson (2009), but are not consistent with the predictions of Taft and colleagues (Taft, 2023; Wei et al., 2023). Zhou and Marslen-Wilson proposed that the semantic representations of both the whole word and its morphemes are activated in parallel from their form representations, and the representation of the meaning of compound word is not necessarily derived via the composition of its constituent morphemes. Therefore, semantic relatedness between constituent characters is not expected to alter positional coding. In contrast, Taft's model holds that character morphemes are activated at the lemma level and combined to access the word's meaning for both transparent and opaque compound words. A key prediction (under a lemma-composition implementation in which constituent semantics contributes to combinatorial coherence) is that a character transposition should be closer to its base word when the constituent morphemes are semantically related than when they are not, yielding a larger TC advantage (TC–SC) for semantically related pairs. The null effect of semantic relatedness observed in the present study is therefore more naturally accommodated by the model proposed by Zhou and Marslen-Wilson (2009). At the same time, we note that Taft's model makes multiple claims about the locus and timing of lemma-level composition, and the present results constrain—rather than exhaust—those claims. Under the present conditions, we found no evidence that constituent semantic relatedness involves the computation that yields the TC effect.

These findings replicate and extend prior findings on character position flexibility in Chinese (Gu et al., 2015, 2023; Liu et al., 2025; Su et al., 2024; Yang et al., 2022). Importantly, while the magnitude of the TC effect was unaffected by constituent semantic relatedness, words composed of semantically related characters produced shorter reading times (Experiment 2), and their transposed forms were more difficult to reject (Experiment 1), indicating that morpheme-level semantic relatedness can influence lexical processing independent of character position coding. The absence of a difference in the TC effect between words composed of morpheme-level semantically related and unrelated characters provides no evidence that morpheme-level semantic relatedness modulates the processes indexed by the TC effect in Chinese, converging with earlier findings (Gu et al., 2015; Yang et al., 2020, 2022). Previous experiments approached this question using cross-paradigm comparisons (Yang et al., 2020) or contrasts between single- and multiple-morpheme words (Gu et al., 2015; Yang et al., 2022), but these designs were limited in number and often involved potential confounds. The present experiments provide a more direct test constituent-to-constituent semantic relatedness while keeping semantic transparency low and orthographic overlap constant. In Taft et al. (2023) model, the constituent morphemes would be activated and combined at a combinatorial stage, and semantic similarity between the morphemes should increase the coherence of a transposed-character string (i.e., preserve which morphemes are present while altering their order), predicting a larger TC effect for semantically related pairs. In contrast, for parallel-activation model (e.g., Zhou & Marslen-Wilson, 2009), whole-word and constituent meanings would be activated from form in parallel, and the model predicts little or no modulation of the TC effect by constituent relatedness. Consistent with the latter account, semantic relatedness did not modulate the TC effect in either task (lexical decision, sentence reading), even though it influenced overall lexical processing (see below).

While the absence of the difference in the TC effect could, in a frequentist framework, be seen as inconclusive, our use of Bayes Factors provides a stronger basis for inference. In both experiments, Bayes Factors consistently supported the null hypothesis, offering quantitative evidence that semantic relatedness does not modulate transposition effects. Taken together, the Bayes factors strengthen the conclusion that, under the present scenario, morpheme-level semantic relatedness did not modulate the TC effect. We acknowledge that the results of Bayes Factors must be interpreted with caution due to the phenomenon of Bayesian Occam's Razor (Jeffreys & Berger, 1992). This mechanism inherently penalizes more complex models that possess more free parameters or parameters with wider plausible ranges. Consequently, Bayes Factors should be interpreted jointly with evidence derived from the frequentist framework. Accordingly, we treat the Bayes factors as quantifying relative evidence for models with versus without the semantic-modulation term, rather than as a categorical "proof" of no semantic involvement in the TC effect.

In the experimental design of both experiments, we deliberately employed semantically opaque words, thereby minimizing the influence of whole-word semantic feedback on the TC effect. Previous research has suggested that the processing difficulty associated with transposed morpheme nonwords—compared to substituted morpheme nonwords—may arise from two distinct mechanisms: flexible position encoding and activation of whole-word semantics. For instance, Spencer et al. (Experiment 3) used an unprimed lexical decision task and found that TC nonwords were harder to reject than their controls, especially for semantically transparent compounds (e.g., "teacup") compared to opaque ones (e.g., "honeymoon"). They argued that orthographic activation of morphemes can trigger semantic activation, which in turn can partially activate the base word's meaning, making transposed items seem more plausible. Most prior studies of morpheme transposition in both Chinese and alphabetic scripts did not control for this factor. Our use of semantically opaque compounds thus helped minimize this confound. The consistent TC effect found in both experiments suggests that the effect may primarily reflect orthographic position flexibility rather than morpheme-level semantic activation. In this respect, the present findings provide support for models that assume character position coding is primarily orthographic in Chinese visual word recognition.

The findings of the present study have important implications for future implementations of computational models of Chinese word recognition. The most recent model—the Chinese Reading Model (CRM; Li & Pollatsek, 2020)—inherits the Interactive Activation Model (IAM) architecture and assumes rigid character position encoding. In light of our results, this assumption warrants revision. We suggest three key directions for refining the CRM. First, the model should specify more clearly how orthographic, phonological, and semantic information interact, drawing inspiration from frameworks such as the Triangle Model (Plaut et al., 2020), to better reflect the multimodal nature of Chinese reading. This clarification is a prerequisite for incorporating flexible position coding mechanisms. Second, CRM should integrate flexible position encoding, potentially adapting principles from models developed for alphabetic scripts, such as the Overlap Model (Gómez et al., 2008) and the Spatial Coding Model (Davis, 2010), which assign graded positional activations and could be extended to account for the flexibility observed in Chinese character position encoding. Third, parameter tuning and empirical validation are needed to quantify the degree of positional flexibility—e.g., through measures of activation dispersion—using behavioral and eye-tracking data. These modifications would improve the model's ability to account for the unique properties of Chinese visual word recognition and provide a more realistic account of the cognitive mechanisms underlying reading in a non-alphabetic script. A further desideratum is that future implementations of the CRM explicitly separate positional coding at the between-character level (the focus of the present study) from positional coding within characters (e.g., radical structure), given accumulating evidence that these two levels can behave differently.

Although morpheme-level semantic relatedness between component characters did not modulate the TC effect, it did affect the lexical processing of semantically opaque compound words. Specifically, compounds composed of morpheme-level semantically related characters produced shorter first fixation durations, go-past times, and total reading times, and their transposed nonword counterparts were more difficult to reject. Importantly, the exclusion of items that shared radicals (e.g., “流浪”) allowed us to control for potential confounds arising from orthographic similarity. Overall, our results suggest that morpheme-level semantic relatedness between constituent characters facilitates word recognition, shedding light on the processing mechanisms of compound words. This issue has long been debated. According to the holistic view (Giraudo & Grainger, 2000; Hyönä & Olson, 1995; Kuperman et al., 2008), compound words are stored and retrieved as whole lexical units. In contrast, decomposition accounts (Taft & Forster, 1975, 1976; Zhang & Peng, 1992) propose that compound processing involves morphemic segmentation followed by semantic integration. This debate extends to Chinese. Our findings (i.e., that morpheme-level semantic relatedness between characters facilitates recognition even in opaque compounds) support the decomposition view, as whole-word recognition appears to be modulated by the semantics of individual morphemes. Taken together, this evidence favors a two-stage process: initial decomposition of morphemes followed by integration into a coherent whole-word meaning (Fiorentino et al., 2014; Taft, 2004).

Why, then, does morpheme-level semantic relatedness between two characters facilitate lexical recognition? We propose two possible explanations. First, during the morpheme decomposition stage, semantically related constituent characters may accelerate the extraction of morpheme-level meaning due to their closer semantic distance (Kumar, 2021; Trott & Bergen, 2023). In contrast, semantically unrelated characters—being farther apart in meaning—may require more processing effort, thereby slowing overall word recognition. Second, morpheme-level semantic relatedness may also influence the subsequent stage of semantic integration. Prior research suggests that in order to correctly recognize semantically opaque words, whole-word semantics must inhibit the activation of morpheme-level semantics (Zhou & Marslen-Wilson, 2000). In this context, two semantically related characters are more likely to combine into a coherent semantic unit, potentially making it easier to suppress individual morpheme meanings and facilitating whole-word recognition. However, it is also possible that such combinations might generate increased competition with the whole-word meaning, making inhibition more difficult. If the benefit from faster morpheme extraction outweighs any cost from semantic competition, the facilitation pattern we observed could still emerge. While the current experiments cannot distinguish between these alternatives, further investigation into the interaction between morpheme-level and whole-word semantics is warranted.

While the results of Experiments 1 and 2 may appear discrepant—with TC nonwords producing longer RTs in a lexical decision task but shorter reading times in a natural reading task——this pattern is readily explained by the different cognitive processes each task engages, and it closely parallels what has been reported for transposed-letter stimuli in alphabetic languages (e.g., see Perea & Lupker, 2004, Rayner et al., 2006). The lexical decision task required participants to reject nonwords, a process hindered by the high word-likeness of TC pseudowords. In contrast, the natural reading task capitalized on this same word-likeness to facilitate efficient word identification. Consequently, both findings converge on the same conclusion: TC nonwords possess a higher degree of perceptual and lexical similarity to real words than do SC nonwords. Therefore, these findings are not in conflict but rather reveal the same underlying principle across different cognitive tasks. This task-dependent reversal is precisely what one would expect if TC stimuli activate more strongly lexical candidates than SC stimuli.

We acknowledge several limitations in the present experiments. First, we exclusively selected words with low semantic transparency to minimize whole-word semantic activation and isolate effects at the morpheme level. While this choice increases internal validity, it also means that the role of whole-word meaning in character position encoding remains unexplored. Future research could address this by systematically manipulating whole-word semantic transparency and controlling for morpheme-level semantics of the constituent characters, thereby clarifying how different levels of meaning interact during word recognition. Second, our use of an unprimed lexical decision task and a sentence reading paradigm was designed to capture both isolated and naturalistic reading processes. However, these paradigms may have made the experimental manipulations more detectable to participants, potentially altering their reading strategies. To mitigate this, future studies could incorporate primed lexical decision tasks, which allow finer control over pre-activation effects, or boundary paradigms, which better preserve natural reading by manipulating parafoveal input. Although semantic effects are typically weak in these paradigms, using converging evidence from multiple paradigms would increase the robustness and generalizability of the findings. A further limitation is that we operationalized "morpheme-level semantics" via constituent semantic relatedness ratings; future studies could triangulate this construct using distributional-semantic measures or corpus-based estimates of constituent co-occurrence.

Conclusion

The present experiments systematically investigated whether position encoding in Chinese is modulated by semantic information. Across two complementary experiments—unprimed lexical decision and sentence reading—we consistently found TC effects that were not modulated by morpheme-level semantic relatedness. Thus, our findings do not provide evidence that morpheme-level semantic relatedness modulates the transposed-character effect in Chinese. At the same time, semantic relatedness between constituent characters facilitated lexical processing, even for semantically opaque compounds. Taken together, these results indicate that morpheme-level semantic relatedness affects lexical processing, but not the magnitude of the TC effect, and they provide constraints for future models of Chinese visual word recognition. More broadly, our findings contribute to our understanding of position encoding in logographic scripts and may help refine cross-linguistic models of visual word recognition.

Funding

This research was supported by a grant from the National Natural Science Foundation of China (32371156).

Disclosure statement.

The authors report there are no competing interests to declare.

Data availability statement.

The materials, data and code that support the findings of this study are openly available in OSF at https://osf.io/gb6aw/overview?view_only=22d40e343ae349f2a9d0955e4cd64c9f.

References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Booij, G. (2012). *The grammar of words: An introduction to linguistic morphology*. Oxford

Boudelaa, S., Norris, D., Mahfoudhi, A., & Kinoshita, S. (2019). Transposed letter priming effects and allographic variation in Arabic: insights from lexical decision and the same-different task. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(6), 729–757. https://doi.org/10.1037/xhp0000621

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, *1*(1), 9. https://doi.org/10.5334/joc.10

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PloS One*, *5*(6), e10729. https://doi.org/10.1371/journal.pone.0010729

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–256. http://doi.org/10.1037//0033-295x.108.1.204

Crepaldi, D., Rastle, K., & Davis, C. J. (2010). Morphemes in their place: Evidence for position-specific identification of suffixes. *Memory & Cognition*, *38*(3), 312–321. https://doi.org/10.3758/MC.38.3.312

Crepaldi, D., Rastle, K., Davis, C. J., & Lupker, S. J. (2013). Seeing stems everywhere: Position-independent identification of stem morphemes. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(2), 510–525. https://doi.org/10.1037/a0029713

Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review*, *117*(3), 713–758. https://doi.org/10.1037/a0019738

Fiorentino, R., Naito-Billen, Y., Bost, J., & Fund-Reznicek, E. (2014). Electrophysiological evidence for the morpheme-based combinatoric processing of English compounds. *Cognitive Neuropsychology*, *31*(1−2), 123−146. https://doi.org/10.1080/02643294.2013.855633

Giraudo, H., & Grainger, J. (2000). Effects of prime word frequency and cumulative root frequency in masked morphological priming. *Language and Cognitive Processes*, *15*(4–5), 421–444. https://doi.org/10.1080/01690960050119652

Gómez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review*, *115*(3), 577–600. https://doi.org/10.1037/a0012667

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review,103*(3), 518–565. http://doi.org/10.1037//0033-295X.103.3.518

Grainger, J., & Van Heuven, W. J. B. (2004). Modeling letter position coding in printed word perception. In P. Bonin (Ed.), *The mental lexicon: "Some words to talk about words"* (pp. 1–23). Nova Science.

Gu, J., & Li, X. (2015). The effects of character transposition within and across words in Chinese reading. *Attention, Perception, & Psychophysics*, *77*(1), 272-281. https://doi.org/10.3758/s13414-014-0749-5

Gu, J., Li, X., & Liversedge, S. P. (2015). Character order processing in Chinese reading. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(1), 127-137. https://doi.org/10.1037/a0038639

Gu, J., Zhou, J., Bao, Y., Liu, J., Perea, M., & Li, X. (2023). The effect of transposed-character distance in Chinese reading. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *49*(3), 464-476. https://doi.org/10.1037/xlm0001180

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, *111*(3), 662-720. https://doi.org/10.1037/0033-295X.111.3.662

Hyönä, J., & Olson, R. (1995). Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(6), 1430–1440. https://doi.org/10.1037/0278-7393.21.6.1430

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). The Clarendon Press, Oxford.

Jeffreys, W. H., & Berger, J. O. (1992). Ockham's razor and Bayesian analysis. *American Scientist*, *80*(1), 64-72.

Johnson, R. L., Perea, M., & Rayner, K. (2007). Transposed-letter effects in reading: Evidence from eye movements and parafoveal preview. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(1), 209–229.

https://doi.org/10.1037/0096-1523.33.1.209

Kinoshita, S., & Norris, D. (2009). Transposed-letter priming of prelexical orthographic representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(1), 1–18. https://doi.org/10.1037/a0014277

Kinoshita, S., Norris, D., & Siegelman, N. (2012). Transposed-letter priming in Hebrew in the same-different task. *Quarterly Journal of Experimental Psychology*, *65*, 1296-1305. https://doi.org/10.1080/17470218.2012.655749

Kumar, A. A. (2021). Semantic memory: A review of methods, models, and current challenges. *Psychonomic Bulletin & Review*, *28*(1), 40-80. https://doi.org/10.3758/s13423-020-01792-x

Kuperman, V., Bertram, R., & Baayen, R. H. (2008). Morphological dynamics in compound processing. *Language and Cognitive Processes*, *23*(7–8), 1089–1132. https://doi.org/10.1080/01690960802193688

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2020). *lmerTest: Tests in linear mixed effects models*. R Package Version 3.1.3. http://CRAN.Rproject.org/package=lmerTest

Lee, C. H., Lally, C., & Rastle, K. (2021). Masked transposition priming effects are observed in Korean in the same–different task. *Quarterly Journal of Experimental Psychology*, *74*(8), 1439–1450. https://doi.org/10.1177/1747021821997336

Lee, M. D., & Wagenmakers, E. J. (2013). *Bayesian data analysis for cognitive science: A practical course*. Cambridge University Press.

Lexicon of Common Words in Contemporary Chinese Research Team. (2008).

Lexicon of common words in contemporary Chinese. Commercial Press.

Li, L., Marinus, E., Castles, A., Hsieh, M. L., & Wang, H. C. (2021). Semantic and phonological decoding in children's orthographic learning in Chinese. *Scientific Studies of Reading, 25*(4), 319–334. https://doi.org/10.1080/10888438.2020.1781863

Li, X., Huang, L., Yao, P., & Hyönä, J. (2022). Universal and specific reading mechanisms across different writing systems. *Nature Reviews Psychology*, *1*(3), 133-144. https://doi.org/10.1038/s44159-022-00022-6

Li, X., & Pollatsek, A. (2020). An integrated model of word processing and eye-movement control during Chinese reading. *Psychological Review*, *127*(6), 1139–1162. https://doi.org/10.1037/rev0000248

Li, J., & Taft, M. (2020). The processing of English prefixed words by Chinese-English bilinguals. *Studies in Second Language Acquisition, 42(*1), 239-249. https://doi.org/10.1017/S0272263119000172

Liu, Z., Li, Y., & Wang, J. (2025). Exploring the flexibility of word position encoding in Chinese reading: the role of transposition effects. *Language, Cognition and Neuroscience*, *40*(2), 263-269. https://doi.org/10.1080/23273798.2024.2417430

Logan, G. D. (2021). Serial order in perception, memory, and action. *Psychological Review*, *128*(1), 1–44. https://doi.org/10.1037/rev0000253

Lupker, S. J., Perea, M., & Davis, C. J. (2008). Transposed-letter priming effects: Consonants, vowels and letter frequency. *Language and Cognitive Processes*, *23*(1), 93–116. https://doi.org/10.1080/01690960701579714

Mattingly, I. G. (1992). Linguistic awareness and orthographic form. In R. Frost, & L.

Katz (Eds.). *Orthography, phonology, morphology, and meaning* (pp. 11–26). Elsevier.

McClelland, J., & Rumelhart, D. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, *88*(5), 375–407. https://doi.org/10.1037/0033-295X.88.5.375

Meteyard, L., & Davies, R. A. I. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language, 112*, 104092. https://doi.org/10.1016/j.jml.2020.104092

Morey, R. D., & Rouder, J. N. (2024). *BayesFactor: Computation of Bayes factors for common designs*. R Package Version 0.9.12-4.7. https://CRAN.Rproject.org/package=BayesFactor

Norris, D., & Kinoshita, S. (2012). Reading through a noisy channel: Why there's nothing special about the perception of orthography. *Psychological Review*, *119*(3), 517–545. https://doi.org/10.1037/a0028450

Pan, J., Wang, A., McBride, C., Cho, J.-R., & Yan, M. (2023). Online assessment of parafoveal morphological processing/awareness during reading among Chinese and Korean adults. *Scientific Studies of Reading, 27*(3), 232–252. https://doi.org/10.1080/10888438.2022.2149335

Perea, M., & Carreiras, M. (2006). Do transposed-letter effects occur across lexeme boundaries? *Psychonomic Bulletin & Review*, *13*(3), 418–422. https://doi.org/10.3758/BF03193863

Perea, M., & Lupker, S. J. (2003). Transposed-letter confusability effects in masked form priming. In S. Kinoshita and S. J. Lupker (Eds.), *Masked priming: State of* *the art* (pp.97-120). Psychology Press

Perea, M., & Lupker, S. J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter transpositions. *Journal of Memory and Language*, *51*(2), 231–246. https://doi.org/10 .1016/j.jml.2004.05.005

Perea, M., Nakatani, C., & van Leeuwen, C. (2011). Transposition effects in reading Japanese Kana: Are they orthographic in nature? *Memory & Cognition*, *39*(4), 700–707. https://doi.org/10.3758/s13421-010-0052-1

Perea, M., Winskel, H., & Ratitamkul, T. (2012). On the flexibility of letter position coding during lexical processing: The case of Thai. *Experimental Psychology*, *59*(2), 68–73. https://doi.org/10.1027/1618-3169/a000127

Perfetti, C. A., & Liu, Y. (2006). Reading Chinese characters: Orthography, phonology, meaning, and the lexical constituency model. *Handbook of East Asian psycholinguistics* (pp. 225–236). Cambridge University Press.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (2020). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. In R. Ellis & G. W. Humphreys (Eds.), *Connectionist Psychology* (pp. 367-454). Psychology Press. https://doi.org/10.4324/9781315784779-13

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rayner, K., White, S. J., Johnson, R. L., & Liversedge, S. P. (2006). Raeding wrods with jubmled lettres: There is a cost. *Psychological Science*, *17*(3), 192–193. https://doi.org/10.1111/j.1467-9280.2006.01684.x

Romero-Ortells, I., Baciero, A., Marcet, A., Perea, M., & Gómez, P. (2024). A stringent test of visuospatial position uncertainty accounts of letter position coding. *Language, Cognition and Neuroscience*, *39*(10), 1278-1290. https://doi.org/10.1080/23273798.2024.2384045

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, *110*, 104038. https://doi.org/10. 1016/j.jml.2019.104038

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-prime (version 2.0). *Computer software and manual]. Psychology Software Tools Inc*.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*(4), 523–568. https://doi.org/10.1037//0033-295X.96.4.523

Snell, J. (2025). PONG: A computational model of visual word recognition through bihemispheric activation. *Psychological Review*, *132*(3), 505-527. https://doi.org/10.1037/rev0000461

Spencer, J., Kahraman, H., & Beyersmann, E. (2024). Positional encoding of morphemes in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *50*(5), 712-739. https://doi.org/10.1037/xlm0001296

Su, X., Li, X., Li, R., Zhao, C., Cui, L. (2024). Morphological structures of two-character words influence character position encoding. *Acta Psychologica Sinica*, *56*(4), 383-393. https://doi.org/10.3724/SP.J.1041.2024.00383

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *Quarterly Journal of Experimental Psychology Section A*, *57*(4), 745–765.

https://doi.org/10.1080/02724980343000477

Taft, M. (2023). Localist lexical representation of polymorphemic words: The AUSTRAL model. In D. Crepaldi (Ed.), *Linguistic morphology in the mind and brain* (pp. 152-166). Routledge.

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 638-647. https://doi.org/10.1016/S0022-5371(75)80051-X

Taft, M., & Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior*, *15*(6), 607–620. https://doi.org/10.1016/0022-5371(76)90054-2

Taft, M., Zhu, X., & Peng, D. (1999). Positional specificity of radicals in Chinese character recognition. *Journal of Memory and Language*, *40*(4), 498-519. https://doi.org/10.1006/jmla.1998.2625

Tong, X., Deng, Q., Deacon, H., Saint-Aubin, J., & Wang, S. (2020). To see or not to see: The roles of item properties and language knowledge in Chinese missing logographeme effect. *Applied Psycholinguistics*, *41*(5), 1113-1139. https://doi.org/10.1017/S0142716420000466

Trott, S., & Bergen, B. (2023). Word meaning is both categorical and continuous. *Psychological Review*, *130*, 1239–1261. https://doi.org/10.1037/rev0000420

Wei, Y., Niu, Y., Taft, M., & Carreiras, M. (2023). Morphological decomposition in Chinese compound word recognition: Electrophysiological evidence. *Brain and Language*, *241*, 105267. https://doi.org/10.1016/j.bandl.2023.105267

Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin & Review*, *8*(2), 221–243. https://doi.org/10.3758/BF03196158

Yang, H., Hino, Y., Chen, J., Yoshihara, M., Nakayama, M., Xue, J., & Lupker, S. J. (2020). The origins of backward priming effects in logographic scripts for four-character words. *Journal of Memory and Language*, *113*, 104107. https://doi.org/10.1016/j.jml.2020.104107

Yang, H., Taikh, A., & Lupker, S. J. (2022). A reexamination of the impact of morphology on transposed character priming effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48*(6), 785–797. https://doi.org/10.1037/xlm0001119

Yao, P., Alkhammash, R., & Li, X. (2022). Plausibility and syntactic reanalysis in processing novel noun-noun combinations during Chinese reading: Evidence from native and non-native speakers. *Scientific Studies of Reading, 26*(5), 390–408. https://doi.org/10.1080/10888438.2021.2020796

Zhang, B., & Peng, D. L. (1992). Decomposed storage in the Chinese lexicon. In H.-C. Chen & O. Tzeng (Eds.), *Language processing in Chinese* (pp. 131–149). North-Holland. https://doi.org/10.1016/S0166-4115(08)61890-7

Zhang, Q., Huang, K. J., & Li, X. (2024). Competition between parts and whole: A new approach to Chinese compound word processing. *Journal of Experimental Psychology: Human Perception and Performance*, *50*(5), 479-497. https://doi.org/10.1037/xhp0001198

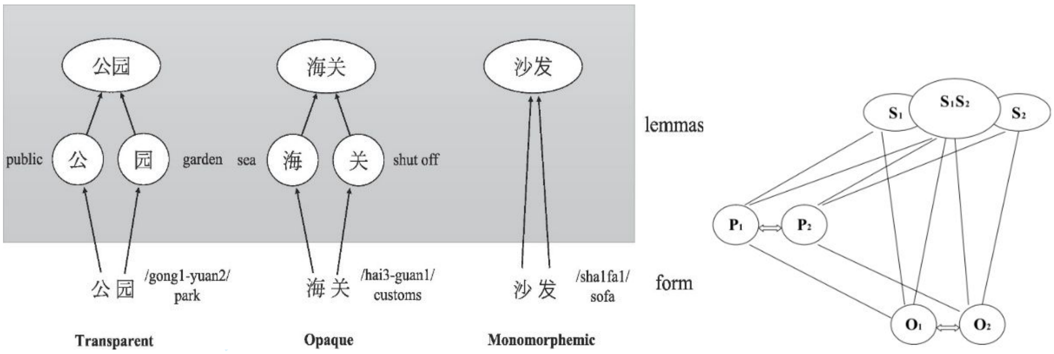
Zhou, X., & Marslen-Wilson, W. (1995). Morphological structure in the Chinese mental lexicon. *Language and Cognitive Processes*, *10*(6), 545-600. https://doi.org/10.1080/01690969508407114

Zhou, X. & Marslen- Wilson, W (2000). The relative time course of semantic and phonological activation in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. *26*(5), 1245–1265. https://doi.org/10.1037/0278-7393.26.5.1245

Zhou, X., & Marslen-Wilson, W. D. (2009). Pseudohomophone effects in processing Chinese compound words. *Language and Cognitive Processes, 24*(7-8), 1009-1038. https://doi.org/10.1080/01690960802174514

Figure 1

Models of Chinese Compound Word Processing



Note: A. The Lemma model of multiple-character word processing as shown in Wei et al. (2023); B. The framework proposed by Zhou & Marslen-Wilson (2009). P_s, O_s, and S_s are phonological, orthographic and semantic representations of constituent characters respectively. S1S2 is the semantic representation of the compound word.

Table 1

Word and Nonword Properties in Experiment 1

Properties	Semantic-related words	Semantic-unrelated words	Semantic-related SC nonwords	Semantic-unrelated SC nonwords
Example	磨蹭 [dawdle] 磨 [rub], 蹭 [rub]	浮夸 [exaggeration] 浮 [float], 夸[praise]	身衣 身 [body], 衣 [clothing]	州克 州 [state],克 [gram]
Semantic relatedness	5.03 (0.51)	1.78 (0.43)	5.01 (0.72)	1.66 (0.53)
Semantic transparency	2.72 (0.52)	2.62 (0.62)	N/A	N/A
Word Frequency	8.24 (12.63)	8.34 (15.25)	N/A	N/A
C1 Character Frequency	495.90 (595.97)	493.21 (501.01)	303.11 (408.74)	314.37 (368.30)
C2 Character Frequency	316.93 (386.71)	339.30 (327.93)	497.72 (659.17)	487.85 (450.54)
Word Stroke Number	16.58 (5.55)	16.39 (4.54)	17.19 (4.28)	16.55 (3.44)
C1 Stroke Number	7.38 (3.15)	7.32 (2.61)	9.58 (2.68)	9.01 (2.42)
C2 Stroke Number	9.20 (3.49)	9.07 (3.28)	7.61 (3.17)	7.54 (2.77)

Note. Frequencies (per million counts) were based on SUBTLEX-CH (Cai & Brysbaert, 2010). Standard Deviations were presented in parentheses. SC stands for substituted-character. For words, C1 is the first character and C2 is the second character. For the transposed nonwords, the characters are transposed from real words so that C2 was presented on the left of C1.

Table 2

Mean Correct RTs and Error Rates in Experiment 1

	Semantic-related			Semantic-unrelated		
	Intact	TC	SC	Intact	TC	SC
RT (ms)	708 (223)	927 (305)	888 (332)	709 (215)	897 (291)	840 (311)
Error %	3.2 (4.2)	14.0 (12.4)	6.8 (8.1)	5.8 (6.1)	8.0 (8.6)	4.0 (7.2)

Note: Standard Deviations were presented in parentheses. TC stands for transposed-character condition, SC stands for substituted-character condition.

Table 3

Sample Experimental Sentence for Experiment 2

Semantic relatedness	Presentation type	Example
Semantically related	Intact	王雨在学校大门口**磨蹭**着死活都不进去 Wang Yu *dawdled* at the school gate, refusing to go in
	TC	王雨在学校大门口**蹭磨**着死活都不进去
	SC	王雨在学校大门口**驾载**着死活都不进去
Semantically unrelated	Intact	邻居对他说不用那么**见外**后他便放松了 After the neighbor told him not to be so *formal*, he felt relaxed
	TC	邻居对他说不用那么**外见**后他便放松了
	SC	邻居对他说不用那么**松未**后他便放松了

Note: TC stands for transposed-character condition, SC stands for substituted-character condition.

Table 4

Eye-Movement Measures on the Target Region (duration measures, in ms)

	Semantically-related			Semantically-unrelated		
	Intact	TC	SC	Intact	TC	SC
FFD	259 (36)	275 (40)	311 (60)	264 (41)	287 (43)	315 (54)
GD	276 (46)	318 (74)	391 (110)	284 (51)	329 (82)	389 (103)
GP	305 (70)	360 (105)	492 (183)	333 (91)	381 (118)	493 (177)
TT	337 (99)	453 (194)	689 (309)	366 (113)	481 (186)	703 (319)
LP	.52 (.13)	.50 (.14)	.46 (.14)	.49 (.16)	.48 (.12)	.46 (.13)

Note: Standard Deviations are presented in parentheses. FFD = First fixation duration; GD = Gaze duration; GP= Go-past time; TT = Total reading time; LP = Saccade landing position. TC stands for transposed-character condition, SC stands for substituted-character condition

Supplement 1

Final models and outputs

Experiment 1

Planed Contrasts

H1:TC vs identity

H2:TC vs SC

H3:interaction

4:identity: related vs unrelated

H5:TC+SC:related vs unrelated

Error Rate

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']

Family: binomial (logit)

Formula: error ~ H1 + H2 + H3 + H4 + H5 + (1 + H1 + H2 + H3 + H4 || id) + (1 + H1 + H3 + H4 + H5 || item)

Data: error

Control: glmerControl(optimizer = "bobyqa")

AIC	BIC	logLik	-2*log(L)	df.resid
4363.6	4478.4	-2165.8	4331.6	9643

Scaled residuals:

Min	1Q	Median	3Q	Max
-1.4855	-0.2662	-0.1772	-0.1201	8.3460

Random effects:

Groups	Name	Variance	Std.Dev.
item	H5	0.1004	0.3169
item.1	H4	1.4738	1.2140
item.2	H3	0.4691	0.6849
item.3	H1	0.8404	0.9168
item.4	(Intercept)	0.5833	0.7637
id	H4	0.1059	0.3255
id.1	H3	0.3930	0.6269
id.2	H2	0.7580	0.8706
id.3	H1	0.3369	0.5804
id.4	(Intercept)	0.4913	0.7009

Number of obs: 9659, groups: item, 414; id, 70

Fixed effects:

```
                Estimate Std. Error z value Pr(>|z|)
(Intercept) -3.40240    0.12583 -27.039  < 2e-16 ***
H1           1.30758    0.22869   5.718 1.08e-08 ***
H2           1.00169    0.19531   5.129 2.92e-07 ***
H3           0.04918    0.31961   0.154   0.8777
H4          -0.54219    0.29430  -1.842   0.0654 .
H5           0.73988    0.15564   4.754 2.00e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
   (Intr) H1     H2     H3     H4
H1 -0.264
H2 -0.109  0.213
H3  0.022 -0.029 -0.114
H4  0.042 -0.066  0.000  0.000
H5 -0.099 -0.053  0.047 -0.162  0.000
```

Reaction Time

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: logRT ~ H1 + H2 + H3 + H4 + H5 + (1 + H1 + H2 + H4 || id) + (-1 +          H
3 + H4 || item)
   Data: rt
Control: lmerControl(optimizer = "bobyqa")

REML criterion at convergence: -1066.7

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.7181 -0.6600 -0.1082  0.5489  4.6350

Random effects:
 Groups   Name        Variance Std.Dev.
 item     H4          0.026537 0.16290
 item.1   H3          0.052819 0.22982
 id       H4          0.001237 0.03517
 id.1     H2          0.007503 0.08662
 id.2     H1          0.011112 0.10541
 id.3     (Intercept) 0.035401 0.18815
 Residual             0.046014 0.21451
Number of obs: 8839, groups:  item, 414; id, 70

Fixed effects:
```

```
              Estimate  Std. Error         df  t value  Pr(>|t|)
(Intercept)   6.668109    0.022840  71.824857  291.943   < 2e-16 ***
H1            0.231648    0.016213 120.936391   14.288   < 2e-16 ***
H2            0.068783    0.013693 112.117862    5.023  1.93e-06 ***
H3           -0.017299    0.017896 260.027169   -0.967     0.335
H4           -0.014389    0.016453 123.156353   -0.875     0.384
H5            0.039061    0.008948 260.030084    4.365  1.83e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
   (Intr) H1     H2     H3     H4
H1 -0.020
H2  0.001  0.184
H3  0.001  0.006  0.010
H4  0.000  0.002  0.000  0.000
H5  0.002  0.006  0.004  0.017  0.000
```

Experiment 2

Planned contrasts

H1:TC vs identity

H2:TC vs SC

H3:interaction

H4:identity: related vs unrelated

H5:TC+SC:related vs unrelated

First Fixation Duration

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: log(DV) ~ H1 + H2 + H3 + H4 + H5 + (1 + H1 + H2 + H3 + H5 ||      id) +
  (1 | item)
   Data: ffd
Control: lmerControl(optimizer = "bobyqa")

REML criterion at convergence: 3215.4

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.4041 -0.6426 -0.0669  0.5830  3.4385

Random effects:
 Groups   Name        Variance  Std.Dev.
```

```
 item      (Intercept) 0.0010405 0.03226
 id        H5          0.0006599 0.02569
 id.1      H3          0.0062186 0.07886
 id.2      H2          0.0045700 0.06760
 id.3      H1          0.0012657 0.03558
 id.4      (Intercept) 0.0141314 0.11888
 Residual              0.0898012 0.29967
Number of obs: 6647, groups:  item, 138; id, 67

Fixed effects:
             Estimate Std. Error        df t value Pr(>|t|)
(Intercept)   5.59630    0.01526  69.98652 366.825  < 2e-16 ***
H1            0.06973    0.01013  74.51435   6.885 1.56e-09 ***
H2           -0.09791    0.01228  80.58927  -7.976 8.62e-12 ***
H3           -0.02502    0.02049  61.57847  -1.221   0.2268
H4           -0.01609    0.01401 693.41741  -1.148   0.2513
H5           -0.02836    0.01102  85.31064  -2.573   0.0118 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
   (Intr) H1     H2     H3     H4
H1  0.000
H2  0.004  0.334
H3  0.000 -0.002  0.000
H4  0.001 -0.004  0.000  0.001
H5  0.000 -0.002 -0.002  0.016  0.196
```

Gaze Duration

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: log(DV) ~ H1 + H2 + H3 + H4 + H5 + (1 + H1 + H2 + H3 || id) +     (1 | it
em)
   Data: gd
Control: lmerControl(optimizer = "bobyqa")

REML criterion at convergence: 6753.7

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.9777 -0.6427 -0.1181  0.5724  3.5680

Random effects:
 Groups   Name        Variance Std.Dev.
```

```
 item      (Intercept)  0.003041  0.05514
 id         H3           0.009544  0.09769
 id.1      H2           0.009041  0.09508
 id.2      H1           0.006674  0.08169
 id.3      (Intercept)  0.027303  0.16524
 Residual               0.152133  0.39004
Number of obs: 6625, groups:  item, 138; id, 67

Fixed effects:
              Estimate  Std. Error       df  t value  Pr(>|t|)
(Intercept)   5.69791    0.02130   72.35576  267.450   < 2e-16 ***
H1             0.11479    0.01559   72.82872    7.365  2.21e-10 ***
H2            -0.16850    0.01660   71.79378  -10.148  1.63e-15 ***
H3            -0.03928    0.02643   68.49125   -1.486     0.142
H4            -0.02405    0.01925  553.66154   -1.249     0.212
H5            -0.02327    0.01504  212.43515   -1.547     0.123
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
   (Intr)  H1      H2      H3      H4
H1  -0.001
H2   0.003   0.273
H3   0.000  -0.001  -0.002
H4   0.001  -0.003   0.000   0.001
H5   0.000  -0.002  -0.002   0.009   0.305
```

Go-Past Time

```
> summary(gp2)
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: log(DV) ~ H1 + H2 + H3 + H4 + H5 + (1 + H1 + H2 + H3 + H5 ||      id) +
 (1 | item)
   Data: gp
Control: lmerControl(optimizer = "bobyqa")

REML criterion at convergence: 9037.4

Scaled residuals:
    Min       1Q   Median       3Q      Max
-3.0040  -0.6549  -0.1491   0.5700   3.9551

Random effects:
```

```
 Groups    Name        Variance  Std.Dev.
 item      (Intercept) 0.008105  0.09003
 id        H5          0.001530  0.03912
 id.1      H3          0.019155  0.13840
 id.2      H2          0.019743  0.14051
 id.3      H1          0.006798  0.08245
 id.4      (Intercept) 0.044620  0.21123
 Residual              0.212447  0.46092
Number of obs: 6619, groups:  item, 138; id, 67

Fixed effects:
             Estimate  Std. Error        df  t value  Pr(>|t|)
(Intercept)   5.80928    0.02755  76.90298  210.882   < 2e-16 ***
H1            0.12314    0.01738  67.88015    7.087  9.98e-10 ***
H2           -0.24717    0.02219  69.66324  -11.140   < 2e-16 ***
H3           -0.04970    0.03265  69.61590   -1.522    0.1325
H4           -0.05691    0.02513 401.35020   -2.264    0.0241 *
H5           -0.03797    0.02126 122.99195   -1.786    0.0765 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
   (Intr) H1     H2     H3     H4
H1 -0.001
H2  0.003  0.257
H3 -0.001 -0.003  0.000
H4  0.001 -0.003  0.000  0.001
H5  0.000 -0.003 -0.003  0.009  0.440
```

Total Time

```
summary(tt2)
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: log(DV) ~ H1 + H2 + H3 + H4 + H5 + (1 + H1 + H2 + H3 || id) +     (1 | it
em)
   Data: tt
Control: lmerControl(optimizer = "bobyqa")

REML criterion at convergence: 10898.9

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.8021 -0.6719 -0.0372  0.6764  3.2229
```

```
Random effects:
 Groups   Name        Variance  Std.Dev.
 item     (Intercept) 0.010050  0.10025
 id       H3          0.005361  0.07322
 id.1     H2          0.039523  0.19880
 id.2     H1          0.019463  0.13951
 id.3     (Intercept) 0.088189  0.29697
 Residual             0.231927  0.48159
Number of obs: 7464, groups:  item, 138; id, 67

Fixed effects:
             Estimate Std. Error        df t value Pr(>|t|)
(Intercept)   6.00638    0.03770  73.13018 159.330  < 2e-16 ***
H1            0.22881    0.02204  68.61322  10.381 1.04e-15 ***
H2           -0.39230    0.02787  68.73674 -14.074  < 2e-16 ***
H3           -0.03854    0.02862  64.51985  -1.347  0.18273
H4           -0.08429    0.02610 360.70935  -3.230  0.00135 **
H5           -0.03919    0.02182 177.76347  -1.796  0.07412 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
   (Intr) H1     H2     H3     H4
H1 -0.001
H2  0.002  0.157
H3  0.000 -0.001  0.000
H4  0.001 -0.005  0.000  0.000
H5  0.000 -0.001 -0.001  0.016  0.512
```


Landing position

```
> summary(lp2)
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
 Family: binomial  ( logit )
Formula: DV ~ H1 + H2 + H3 + H4 + H5 + (1 + H3 + H4 + H5 || id) + (1 |      item)
   Data: lp
Control: glmerControl(optimizer = "bobyqa")

     AIC      BIC   logLik -2*log(L)  df.resid
  9324.9   9399.9  -4651.4    9302.9      6756
```

```
Scaled residuals:
    Min      1Q  Median      3Q     Max
-1.3560 -0.9390 -0.7395  0.9998  1.6315

Random effects:
 Groups Name        Variance Std.Dev.
 item   (Intercept) 0.055764 0.23614
 id     H5          0.002078 0.04558
 id.1   H4          0.056137 0.23693
 id.2   H3          0.084971 0.29150
 id.3   (Intercept) 0.054184 0.23277
Number of obs: 6767, groups:  item, 138; id, 67

Fixed effects:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.06879    0.04297  -1.601   0.1094
H1          -0.06457    0.06136  -1.052   0.2927
H2           0.11446    0.06019   1.902   0.0572 .
H3           0.10992    0.12566   0.875   0.3817
H4           0.14285    0.10053   1.421   0.1553
H5           0.05021    0.07263   0.691   0.4894
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
   (Intr) H1     H2     H3     H4
H1 -0.006
H2  0.015  0.503
H3 -0.003 -0.005 -0.004
H4  0.003 -0.004  0.000  0.000
H5 -0.001 -0.005 -0.006  0.022  0.222
```

Supplement 2

Table S1

Bayes Factor Between Models Without and With an Effect in Experiment 1

	Error rate	Reaction Times
TC vs Identity	$4.50\text{x}10^{-9}$	$7.54\ \text{x}10^{-65}$
TC vs SC	$6.86\ \text{x}10^{-7}$	$1.37\ \text{x}10^{-7}$
Interaction	2.88	12.11
Identity: related vs unrelated	1.90	9.633
TC+SC: Related vs unrelated	$4.26\ \text{x}10^{-4}$	$6.53\ \text{x}10^{-3}$

Supplement 3

Additional Analyses of Experiment 2

Because the design of Experiment 2 forms a full 3 x 2 factorial design, we also report the corresponding fixed effects for semantic relatedness, presentation type, and their interaction within a single model. This analysis enabled the model to fully incorporate the interrelationships between all factors. All measures were analyzed using linear mixed-effects models (LMMs) in R (v4.3.1; R Core Team, 2023) with the lmerTest package (v3.1.3; Kuznetsova et al., 2017), optimized using the bobyqa algorithm. All measures were log-transformed to meet LMMs' normality assumptions. In the model, fixed effects included semantic relatedness (semantic-related, semantic-unrelated) and presentation type (Intact, TC, SC), both simple-contrast coded with semantic-related and TC condition as reference levels, respectively. We specified the participants and items as crossed random effects, including intercepts and slopes. Following Barr et al. (2013), we used the maximal model that could converge. We first constructed a model with a maximal random factor structure. When the maximal model failed to converge, we used a zero-correlation parameter model and removed the random components that generated the smallest variances. We report regression coefficients (*b*), standard errors (*SE*), *t*/*z*-values, and *p* values of the optimal model.

The main effect of presentation type was significant in all measures. Reading times were longer in the TC condition than in the intact condition, as shown in first fixation duration (b = -0.07, SE = 0.01, t = -7.36, p < .001), gaze duration (b = -0.11, SE = 0.02, t = -7.41, p < .001), go-past time (b = -0.12, SE = 0.02, t = -7.14, p < .001) and total reading time (b = -0.23, SE = 0.02, t = -10.27, p < .001), indicating a transposition cost. Reading times were shorter in the TC condition than in the SC condition, as shown in first fixation duration (b = 0.10, SE = 0.01, t = 7.00, p < .001), gaze duration (*b* = 0.17, *SE* = 0.02, *t* = 8.71, *p* < .001), go-past time (*b* = 0.25, *SE* = 0.02, *t* = 10.17, *p* < .001) and total reading time (*b* = 0.39, *SE* = 0.02, *t* = 13.34, *p* < .001), indicating the typical TC effect.

Eye-movement measures were shorter for the semantic-related condition than the semantic-unrelated condition, as indicated by significantly shorter first fixation duration (*b* = 0.02, *SE* = 0.01, *t* = 2.59, *p* = .011), go-past time (*b* = 0.04, *SE* = 0.02, *t* = 2.31, *p* = .023) and total reading time (*b* = 0.05, *SE* = 0.02, *t* = 2.65, *p* = .009). These results are consistent with the results in Experiment 1, suggesting that the semantic relatedness between two constitute characters has an influence on whole word recognition.

However, there was no interaction between semantic relatedness and presentation type in any eye-movement measures. The comparison between the intact and TC condition, first fixation duration (*b* = -0.02, *SE* = 0.02, *t* = -1.23, *p* = .221), gaze duration (*b* = -0.02, *SE* = 0.02, *t* = -0.77, *p* = .439), go-past time (*b* = -0.01, *SE* = 0.03, *t* = -0.25, *p* = .804) and total reading time (*b* = 0.03, *SE* = 0.03, *t* = 0.91, *p* = .366) did not showed any interaction. The comparison between the SC and TC conditions, first fixation duration (*b* = -0.02, *SE* = 0.02, *t* = -1.07, *p* = .289), gaze duration (*b* = -0.04, *SE* = 0.03, *t* = -1.20, *p* = .234), go-past time (*b* = -0.05, *SE* = 0.04, *t* = -1.38, *p* = .171) and total reading time (*b* = -0.04, *SE* = 0.03, *t* = -1.14, *p* = .256) also did not show any clear signs of an interaction, providing no evidence that semantic relatedness modulates character position encoding.

Bayes Factor Analysis

We also conducted a Bayes Factor analysis on all eye-movement measures to quantify the statistical evidence against an interaction between semantic relatedness and presentation type. We used default priors to ensure a neutral starting point for the analysis. Bayes Factor analysis compares the relative likelihood of a model without the interaction (null hypothesis H0) to a model with the interaction (alternative hypothesis H1). The value of BF01 was 80.37 ±2.97% for first fixation duration, 81.36 ±3.2% for gaze duration, 34.92 ±2.82% for go-past time, and 18.24 ±21.95% for total reading time; thus, all indexes provide strong or very strong evidence in favor of a null interaction between *se*mantic relatedness and pre*se*ntation type.

Supplement 4

Table S2

Bayes Factor Between the Model Without vs with an Effect in Experiment 2

	FFD	GD	GP	TT	LP
TC vs Identity	2.62×10^{-12}	4.39×10^{-21}	8.07×10^{-17}	3.326×10^{-58}	2.19
TC vs SC	1.99×10^{-26}	2.06×10^{-46}	9.26×10^{-71}	1.681×10^{-168}	2.19
Interaction	6.34	4.83	3.19	7.82	9.13
Identity: related vs unrelated	7.99	6.90	1.25	0.155	4.57
TC+SC: Related vs unrelated	1.00	4.57	2.52	3.93	7.95