



Meanings within meanings: skilled readers activate irrelevant meanings of radicals in Chinese

Tianlin Wang¹ · Matt Cooper Borkenhagen² · Madison Barker³ · Mark S. Seidenberg²

Accepted: 19 January 2022

© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

Many characters in written Chinese incorporate components (radicals) that provide cues to meaning. These cues are often partial, and some are misleading because they are unrelated to the character's meaning. Previous studies have shown that radicals influence the reader's processing of the characters in which they occur (e.g., Feldman and Siok in *J Memory Language* 40(4):559–576, 1999). We investigated whether readers automatically activate the semantic information associated with a radical even when it is irrelevant to the character's meaning, using a modified version of the Van Orden (*Memory Cogn* 15(3):181–198, 1987) task with auditory semantic relatedness ratings on test items. Fifty-one Mandarin speakers participated in the study. On each trial they saw a reference category such as “animal” prior to seeing a character then indicated whether the target character was a member of that category. Decisions were slower and less accurate when a target that is not a member of the target category contained a radical that is. For example, if the category is “found in the kitchen,” the answer for the target 券 “ticket” is no; however the character contains the misleading radical 刀 “knife”. These patterns suggest that readers process the semantics of the radical even when it is not relevant to the meaning of the character. The results further verify the role of radical semantics in character processing and raise questions as to whether repetitions of experience with the idiosyncrasies of the script may result in some of the irrelevant semantics influencing the meaning of the character.

Keywords Chinese writing system · Chinese characters · Reading · Script relativism

✉ Tianlin Wang
twang23@albany.edu

¹ Department of Educational and Counseling Psychology, University at Albany, SUNY, Albany, NY 12222, USA

² University of Wisconsin-Madison, Madison, WI, USA

³ University of California, Davis, Davis, CA, USA

Introduction

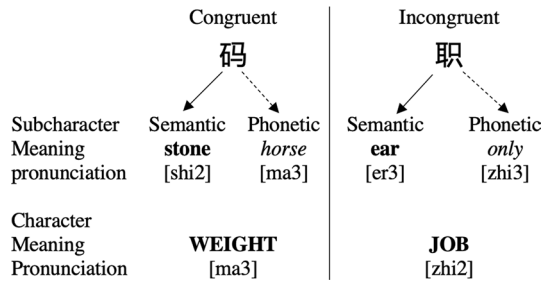
Psycholinguistic theories of reading attempt to specify the types of knowledge and processing mechanisms used in comprehending text. Although most reading research has been conducted on alphabetic systems, there is a growing body of research on nonalphabetic systems as in Chinese and Japanese (Pae, 2020). Studying different types of writing systems allows the identification of commonalities in how reading occurs and also differences that are conditioned by properties of a given script (Frost, 2012; Pae, 2020; Seidenberg, 2011).

As the Chinese writing system uses characters rather than the letters found in alphabetic systems, many researchers (e.g., DeFrancis, 1984; McBride-Chang et al., 2004) have described it as a morphosyllabic system, where one character represents one morpheme, which then maps onto one syllable phonologically. A dominant assumption has been that characters are the main units of processing and the level at which semantics are computed (i.e., mono-morphemic words; e.g., Cheng, 1981; Hoosain, 1991; Perfetti & Zhang, 1991; Tan et al., 2001). According to this view, sub-character structures such as radicals are not relevant because characters are the units of recognition and bearers of meaning. Consistent with this view, Perfetti (2020) described the Chinese writing system as “a character system that places spaces between meaning-bearing syllables [i.e., morphemes].”

Previous studies, however, have also attempted to examine whether the components of characters are processed in recognition. It has been shown that readers are sensitive to a wide range of characteristics associated with sub-character components, such as the positional specificity of a radical (e.g., how frequently a radical appears on the left side of a character; Taft et al., 1999) and the frequency of a semantic radical in relation to its positional consistency (e.g., 犭 and 牛 are radicals that share the same semantic meaning, but the former only occurs on the left side of a character while the latter only on the right side; Taft & Zhu, 1997). Though findings suggest that radicals may be a psychological entity in the processing of Chinese words (see Taft, 2006 for a review), Yu and Reichle (2017) points out that there is a lack of consensus regarding the level at which meaning is represented in Chinese; moreover, concepts such as *word*, *radical*, and *character* used in describing Chinese do not have exact equivalence to the usage of terms such as *words* and *morphemes* when describing English. It has been estimated that 60–90% of all characters consist of at least two components (Norman, 1988; Shu, 2003): a phonetic component, which provides a cue to pronunciation, and a semantic component (radical), which provides a cue to meaning. Since most of the phonetic components and some of the radicals are also themselves standalone characters with their own meanings, a fundamental question is whether the meanings of these components when assembled together to produce combinatorial characters affect the recognition of characters in which they occur. This possibility has been for the most part rejected by the general consensus that characters are recognized as holistic units (Wang, 2018).

Several previous studies have indeed shown that the phonetic components do influence the processing of the characters in which they occur (e.g., Zhou &

Fig. 1 Examples of radicals with meanings that are congruent or incongruent with the meaning of the combinatorial character in which they occur



Marslen-Wilson, 1999); however, the possibility that radicals may also play a significant role in character recognition has received far less support (Yu & Reichle, 2017). It is surprising that this is the case, given the fact that the role of radical is highlighted during learning and instruction and that beginning readers of Chinese are often explicitly taught that radicals suggest the meaning of the combinatorial character in which they occur and are instructed to rely on them in reading¹ (Wu et al., 1999), and studies show that relying on semantic radicals is related to better reading comprehension (e.g., Ho et al., 2003; Packard et al., 2006). Yet it remains the case that a radical is sometimes not related to the meaning of the character in which it occurs (e.g., Shu & Anderson, 1997), at least not in any obvious or intuitive way. As illustrated in Fig. 1, whereas the radical 石 “stone” provides a reliable semantic cue to the overall meaning of the combinatorial character 码 “weight”, the radical 耳 “ear” in 职 “job” does not.

Among the few studies that have investigated the role of radicals, Ju and Jackson (1995) used a backward-masking paradigm and found that target character recognition was facilitated when the target and the mask shared the same radical (e.g., it was easier to recognize the target 河 “river” when the mask 法 “law” shared the same radical 氵 “water”). Feldman and Siok (1999) further demonstrated that sub-lexical processing (i.e., sub-character, at the level of radicals and phonetics) occurs in reading Chinese combinatorial characters by systematically manipulating the relationship between the semantics of the radical and the character. Using a primed lexical decision task, they found that when the prime 评 “comment” and the target 论 “review” shared the same radical (讠 “say”), target identification was facilitated compared to when the prime 竿 “stick” (with radical “bamboo”) and the target 论 “review” did not share a radical. Critically, recognition of the target character 论 “review” was inhibited when the prime 诸 “some” shared a radical with the target but was not semantically related to the target. Neuroimaging studies have also shown that radicals play an important role in early stages of visual word recognition (e.g., Hsu et al., 2009). In addition, distributional information of radicals, including both radical and character frequency as well as radical positional specificity, also modulate the activation of character-level semantics. For example, Taft and

¹ This practice dates back to the Eastern Han Dynasty (25–189). In an ancient Chinese dictionary titled *Shuowen Jiezi*, Xu Shen (c. 58-c. 148 CE) categorized Chinese characters and discussed the function of radicals.

Zhu (1997) demonstrated that though radical frequency impacts character decision responses, this effect is mainly constrained to radicals that appear on the right side of a character. Ding et al. (2004) showed that while a facilitation effect occurs during a priming task when the target contains the prime (as a radical) for low frequency targets, an inhibitory effect is observed for targets that contain the prime but in different positions. Thus, several types of evidence suggest the following conclusion: the semantic information associated with radicals is in fact accessed early in character processing; radicals whose meanings are congruent with that of the combinatorial characters and are positionally consistent across the characters in which they occur can facilitate the activation of character-level semantics (see Yu & Reichle, 2017, for a review).

The above-discussed studies mainly focused on cases where the meaning of a given radical is consistent with the meaning of the character, and thus provided a valid, if only partial, cue. Yet it remains the case that there are frequent examples where the radical's meaning is not a valid cue to character meaning, as in the 耳 “ear” (radical) 职 “job” (character) example discussed above. In such instances, there is no obvious connection between the semantic information within the radical, itself functioning as a word in its own right when written as a standalone character, and the combinatorial character of which it is a seemingly random component. If we accept that semantically congruent radicals do in fact facilitate character processing, the question, then, is: do radicals that are deemed misleading or irrelevant interfere with, or otherwise condition, character processing?

The basic written composition of many, if not most, Chinese characters has changed little over the course of the last 2000 years (Boltz, 1999). Nonetheless, lexical semantics is subject to diachronic change. This leads to characters possessing artifacts of older meanings, which has an important sociological implication: cultural knowledge can be passed on through properties of the writing system itself thus influencing cognition in a way that is at least partially independent of oral communication among speakers of the language. Examples include words that can be quite important in terms of the cultural information they possess. Take, for example, 妒 “jealousy” and 奸 “evil, wicked”. Both characters have the 女 “female” radical and are quite negative for their meaning. Additionally, consider the character 狄 (ethnonym for a minority group) and 蛮 “savage”. Both employ radicals that denote nonhuman animals. These characters and their associated semantics when considering their orthographic composition may encode what would today be considered misogynist or ethnocentric attitudes. These are just a few examples among many.

Investigating the role of radicals that offer incongruent semantic meanings to the overall characters, Leck et al. (1995) used a semantic categorization task in English and tested fourteen native Mandarin speakers and showed that participants produced a higher false positive response rate when the radical incongruent character and the target shared the same radical (e.g., it was harder to decide that 狐 “fox” and 猜 “guess” are not semantically related than 狐 and 弧 “curve” because the first two share the radical 犭 “dog”). While Leck et al. (1995) provided preliminary evidence that incongruous radicals can impact the processing of character-level semantics, their design did not control for the possibility that the radical-incongruous characters may differentially relate to the semantics of the target exemplar than the other

characters used in the experiment—that the results capture differences of the character-level semantics rather than effects induced by orthography. For instance, in English, it might be more difficult to decide that FUR is not an animal than to decide the same of CLOUD due to the fact that many animals have fur. Since the two concepts ANIMAL and FUR are intrinsically more related, it would not be surprising to observe differences related to decision difficulty when comparing FUR-ANIMAL versus CLOUD-ANIMAL. This difference, though unrelated to orthography, can be reflected by a semantic category judgment task such as the one used by Leck et al. In this case, it would be misleading to conclude that the observed differences are due to orthographical idiosyncrasies. Therefore, in order to verify that the effects captured by Leck et al. (1995) were indeed an artifact of Chinese orthography, it would be essential to first establish that concepts of FUR and CLOUD are equally distant from that of ANIMAL in a non-visual domain (in their example, the amount of relatedness of 狐-猜 FOX-GUESS should be comparable to that of 狐-弧 FOX-CURVE). One way to address this concern is to examine how listeners might rate their relatedness in an auditory semantic relatedness task. Without such information, it remains difficult to draw conclusive interpretations from Leck et al. without independent auditory semantic relatedness ratings on the stimuli.

Taken together, these studies suggest that radical semantics do impact character level processing. The results from Leck et al. (1995) in particular provide insight to the sociolinguistic implications discussed above since the incongruous radical semantics are activated when the semantic domain of the character (i.e., the meaning of the overall character) is engaged. According to the Script Relativity Hypothesis (Pae, 2020), which links scriptal characteristics to differences in cognition, Chinese readers should be influenced when judging the semantic category of a combinatorial character when its radical is semantically incongruous—a phenomenon attributed to the writing system. In order to conclusively demonstrate that incongruous radical semantics are activated during character semantic judgment, however, it remains necessary to implement a better control of semantic relatedness in a task such as the one used by Leck et al.

In the present study, we used characters such as those illustrated in Fig. 1 to examine whether skilled readers activate the semantics of radicals even when they are unrelated to the meaning of the character and thus introduce irrelevant information. We employ a modified version of the Van Orden (1987) task, which requires deeper semantic processing than other common tasks such as lexical decision. The participant is given a categorization cue such as “is it a part of the human body” followed by a target character and must then decide if it is a member of the category. In a critical contrast, an incongruous character (e.g., 职 “job”) is pitted against an unrelated character (e.g., 网 “net”). Though both characters do not belong to the category, the Incongruous character contains a semantic radical that does (耳 “ear”). If the semantics of the radical is computed, it should be more difficult to respond “no” to the Incongruous character compared to the unrelated character, thus resulting in longer reaction time or lower accuracy rate for the incongruous characters. In contrast, if sublexical semantics from the radicals do not impact character-level semantic judgment, accuracy and reaction time of these two types of characters should

Table 1 Summary of conditions and examples of a set of stimuli

Reference category example	Target example	Condition	Correct response	Average frequency (SD)	Average stroke number (SD)
在厨房里 “In the Kitchen”	券 “ticket”	Incongruous	No	23,735 (32,808.5)	10 (3.2)
	宪 “statue”	Unrelated	No	24,273 (31,948.1)	9 (2.4)
	盐 “salt”	Overall Related	Yes	16,723 (20,352.1)	11 (2.5)
	刀 “knife”	Radical Only	Yes	118,726 (235,088.0)	5 (2.5)

be comparable. Essential to the design, we collected auditory semantic relatedness ratings on the critical items in order to draw more reliable interpretations of the data.

Method

Participants

Fifty-one speakers of Mandarin (33 females, $M_{\text{age}} = 23.4$, range = 18.4–34.8) at a midwestern university in the US participated in the study in person. Participants were recruited through the department participation-for-credit program and either received course credit or a small payment for their participation. All participants reported that they had completed at least 9 years of formal schooling in Mainland China and were native speakers of Mandarin and proficient readers of Chinese. A separate group of 35 college-aged participants (18 females, $M_{\text{age}} = 20.8$, range = 19.2–22.6) were recruited from a Mandarin-speaking region in northeastern China to provide independent, auditory semantic relatedness ratings on the stimuli during online sessions. These raters were all native speakers of Mandarin without hearing or language impairments.

Design

The study employed a categorization task based on one introduced by Van Orden (1987). On each trial, a semantic category (e.g., “in the kitchen”) was first displayed on the screen. A target character (e.g., 券 “ticket”) then appeared at the center of the screen and participants were asked to indicate whether the target character designated a member of the semantic category that appeared just before it.

Four stimulus conditions were employed (Table 1). In the Overall Related condition, the target character’s meaning as well as that of its radical were both related to the reference category (e.g., reference category: “in the kitchen”; target character: 盐 “salt,” with radical 皿 “dish”). In the Incongruous condition, the character was not a member of the category, but contained a radical related to it. For example, the target 券 “ticket” is unrelated to the category “in the kitchen”, but contains the radical 刀 “knife”. In the Unrelated condition, the character was unrelated to the category and

contained no related radicals (e.g., “in the kitchen”, 宪 “statue”). Finally, in the Radical Only condition, the target character was the radical from the corresponding item in the Overall Related condition. Similar to items from the Overall Related condition, items in the Radical Only condition were included so that there were approximately equal numbers of “yes” and “no” trials. Reference categories (e.g., “in the kitchen”) were created to match the Incongruous condition.

The specific conditions of critical interest are the Incongruous and Unrelated conditions. Reference categories in the Incongruous condition were designed to align with the target character’s radical but not that character’s overall meaning. In the Unrelated condition, the category was semantically unrelated to all aspects of the target character’s meaning including sublexical components of that character. If radical semantics impact character processing, responses in the Incongruous condition should differ from those of the Unrelated condition, given that the targets in the Incongruous condition contain sublexical components that are semantically related to their corresponding reference category.

Materials

Ninety-two unique combinatorial characters that contain both a semantic radical and a phonetic element were used as targets. Table 1 provides additional details (see data repository online at <https://osf.io/hqxdk/> for all stimuli).

Character frequencies were determined using the Modern Chinese Character Frequency List (Da, 2004), which lists character frequencies based on a corpus of 248 million characters. Stimuli were selected so as to control for grammaticality, target complexity (i.e., number of strokes) and frequency across the conditions. The only exception to this principle was in the case of the stimuli selected for the Radical Only condition, given that the radicals are, by nature, visually simpler and almost always more frequently occurring than complex characters. Across the targets, there were 23 radicals embedded in the complex characters. Characters in the Overall Related condition and the Unrelated condition were selected based on their congruence with the reference category and were adjusted to match the character selected for the semantic condition in complexity (number of strokes) and frequency. The Radical Only condition was the radical that appeared in the Incongruous and Overall Related condition character’s sublexical unit. The characters selected for the Unrelated condition were completely unrelated to the reference category. To ensure stimuli for this task were appropriate linguistically and culturally, all characters selected were checked by two native speaking research assistants as well as a linguist who is a native speaker of the language.

In addition to the target characters, 23 unique category prompts were developed based on the radicals in the targets. Each category prompt was a 3- to 4-character long phrase and was matched with a range of 4–24 targets (note that a category always appeared with at least four targets given that there are four conditions in the experiment). Category prompts were used with more than one target because constructing stimuli of this kind presents natural constraints on stimuli development. We made efforts to create categories that were ideally suited to elicit semantic

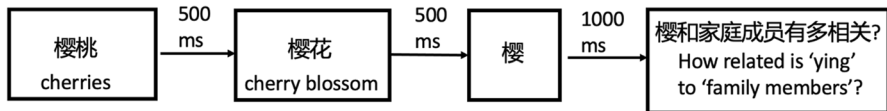


Fig. 2 An example of the auditory semantic relatedness rating task

activation of the orthographic elements of interest, among other constraints across conditions, rather than create unique categories across stimuli (therefore reducing the quality of the stimuli, the number of stimuli or both).

Procedure

On each trial, the reference category appeared at the center of the screen for 1500 ms. This was followed by a fixation cross shown for 200 ms, indicating the location of the upcoming target character. The target character was then shown in the same position as the fixation cross for 200 ms to encourage rapid responses. Participants were instructed to respond as quickly and as accurately as possible. After a response was provided, a new trial began. Participants were tested individually in a quiet room and began with 4 practice trials. They were instructed to indicate whether the target character belonged to the reference category by pressing “yes” or “no” keys (the “z” and “n” on the keyboard, also marked green and red, respectively). The experiment was run in PsychoPy 2 (Pierce, 2007) in Python 2.7. Latency and accuracy data were collected for every trial and were recorded from the onset of the target to the onset of the participant’s response. Response accuracy was also recorded. Each participant completed a total of 136 category-target pairs over a period of approximately 10 min.

The experiment employed a within-subject design. Stimuli were presented in four blocks, with 34 trials in each block. Each target was assigned to one and only one block. While a category could appear more than once in a block, each target within its set of four would never be presented in the same block (see Table 1 for an example set). The presentation of stimuli within each block as well as the order of blocks were randomized. In order to minimize the number occasions where an Incongruous trial occurs adjacent to its Radical Only trial (the same radical appears twice consecutively), the randomization also ensured that two items (e.g., Incongruous and Radical Only) that shared the same radical did not occur consecutively.

Auditory semantic relatedness ratings

To ensure that the variability in the degree to which the meaning of a target character is comparably related to its reference category across the critical conditions (i.e., Incongruous and Unrelated), an auditory semantic relatedness task was carried out with an independent group of raters. Participants rated the relatedness of each target-reference pair on a 1–7 scale with 1 labeled “not at all related” and 7 labeled “highly related”. In order to minimize homophone-related processing difficulties,

Table 2 Descriptive statistics of the four conditions in an auditory semantic relatedness task

Condition	M	SD	Min	Max
Overall Related	6.29	1.35	1	7
Radical only	6.08	1.47	1	7
Incongruous	1.58	1.15	1	7
Unrelated	1.35	.87	1	7

on each trial, participants first heard two disyllabic or trisyllabic context words that contained the target character before hearing the target character pronounced in isolation; they were then asked how related the target character was to the reference category. Figure 2 illustrates the paradigm.

Additional considerations were given to the selection of the context words to ensure that the two context words for each target elicited similar semantic meaning from the target character. High frequency context words were prioritized² over low frequency items in the event that multiple possible context words with similar semantic meanings existed. A full list of context words, their glossary, as well as all trials and auditory recordings can be found in the online repository.

Since our goal is to show that how a concept is *written* impacts the degree to which the meaning of a target character is related to its reference category across the Incongruous versus Unrelated conditions, we expect that the ratings on the Incongruous and Unrelated conditions should be comparable in this auditory task since the overall meaning of the target characters are intended to be unrelated to their reference categories in both of these conditions. A full list of context words, their glossary, as well as auditory recordings of the trials can be found online.

Results

We first report the rating results from the auditory semantic relatedness task. The pattern of means and other descriptive data for the ratings are summarized in Table 2 below.

The ratings on the two conditions of interest (Incongruous and Unrelated) exhibited a 0.12 numeric difference based on their means, however, this difference was not statistically significant, $F(1, 32)=0.44$, $p=0.51$, 95% CI=[− 0.46, 0.23] (see Table 3). These results indicate that target characters in the Incongruous condition were not on average more semantically similar to their corresponding reference categories than targets in the Unrelated condition, and that any effects observed in our task should be due to script-induced differences (Fig. 3).

We then trimmed the data by identifying outliers by participant accuracy, item accuracy, and reaction time (RT). Using mean accuracy scores for each participant

² There were two exceptions, both in the Incongruous condition, where the context words were low frequency since the target word itself was of very low frequency and no context words (in which the target word is embedded) could fit the criterion.

Table 3 Summary statistics for the ratings between the incongruous and unrelated conditions

	<i>b</i>	<i>SE</i>	<i>F</i>	<i>p</i>	95% CI
Intercept	1.47	.10	229.45	< .001	[1.28, 1.66]
Condition	− .12	.17	.44	= .51	[− .46, .23]

Coefficient estimates based on $n=2039$ with by subject and by item random intercepts included for random effects using the lmer() method in the R package lme4 (Bates et al., 2015) in R version 4.04 (R Core Team, 2020). Coefficients represent units of the Rating DV, with the negative estimate associated with moving from the *Incongruous* to the *Unrelated* condition. The *p*- and *F* values reported were calculated using the Anova() method (type III Wald F tests with Kenward–Roger dfs) in the car package version 3.0 (Fox & Weisberg, 2019). 95% confidence intervals were calculated using confint() in the stats package (Stats Core Team, 2020). These same notes apply to other statistical models reported here and will not be repeated in subsequent tables except where alternative methods apply

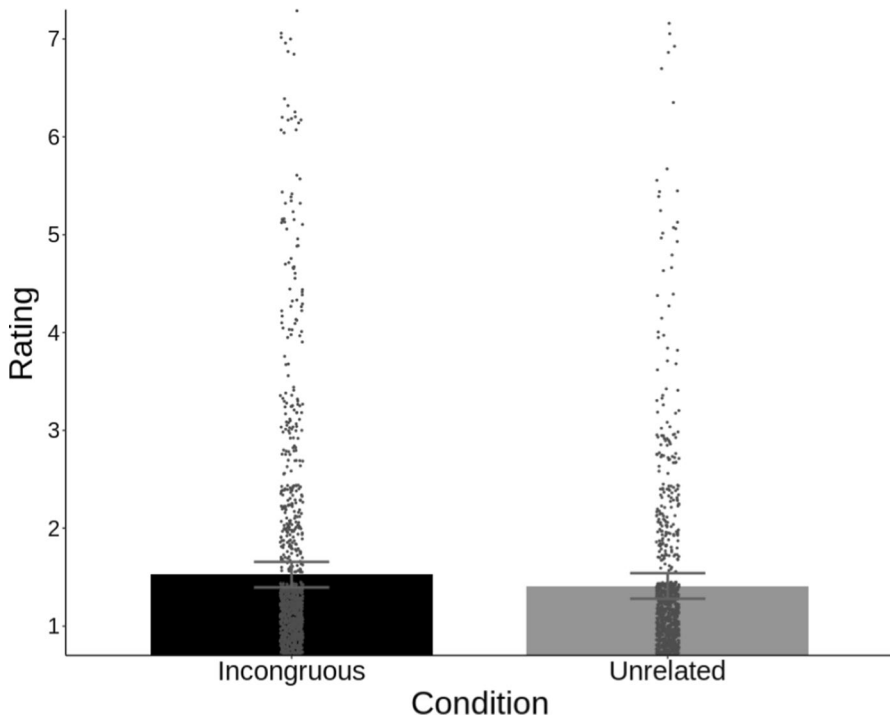


Fig. 3 Model predictions and raw data for semantic ratings in the incongruous and unrelated conditions are shown. Bars represent predicted means for two conditions and points are raw data. Error bars are standard error of the point estimate. See Table 3 for information about the statistical model

Table 4 Descriptives of response accuracy

Condition	<i>M</i>	<i>SD</i>
Overall related	.88	.32
Radical only	.94	.23
Incongruous	.93	.26
Unrelated	.97	.17

by collapsing across all trials, two participants with mean accuracy below 55% were excluded from analyses. Four items with a mean accuracy below 63% in the Incongruous and Radical Only conditions were excluded as well. Trials showing response times 3 SD and greater (2742 ms) and less than 300 ms were excluded given that such processing times likely indicate these responses were not reliable. The criteria for RT resulted in comparable numbers of excluded observations across the four conditions (see supplement for more description regarding outliers). Removing observations for these targets resulted in the elimination of 532 responses (7.7% of all observations), although trimming the data in this way did not change the results in any meaningful way. Analyses reported here involve only the Incongruous and Unrelated conditions given that these two conditions elicited “no” answers and were the key contrast in this study.

For accuracy, the outcome variable was the accuracy of the participant on the trial. Since a given reference category-target pair was either correct or incorrect, the accuracy scores on each item were dichotomous (0 or 1). Descriptive statistics for accuracy are shown in Table 4.

A statistical model was estimated using the `glmer` function in `lme4` (Bates, et al., 2015). For the purposes of modeling our data, we focused only on the Incongruous and Unrelated conditions where accurate responses are “no.” Data for the “yes” responses in the other two conditions (the Radical Only and Overall Related) could not be directly compared to conditions requiring “no” responses and so were excluded from these analyses. In this logistic model, the participant’s accuracy was predicted by the condition of the target (Incongruous or Unrelated), which was entered as a fixed factor. The semantic similarity rating for each target (to its corresponding reference category) was included as a Covariate. Results show that although participants were very accurate at the task (Table 5), they were more likely to make incorrect responses in the Incongruous condition compared to the Unrelated condition ($b = -0.82$, $SE = 0.27$, $p < 0.01$, $z = -3.07$, 95% CI = $[-1.36, -0.30]$). In terms of probabilities, the logits reported here translate to a decrease in probability of 0.02 (between the Unrelated to Incongruous conditions), resulting in a probability of accuracy in the Incongruous condition of 0.96 compared to 0.98 in the Unrelated condition (these values correspond to the predicted means shown in Fig. 4) (Table 6).

A similar statistical model was constructed predicting response times from the trial condition with the semantic rating for each target included as a covariate, focusing again on our two conditions of interest. Descriptive data for all four conditions are provided in Table 6.

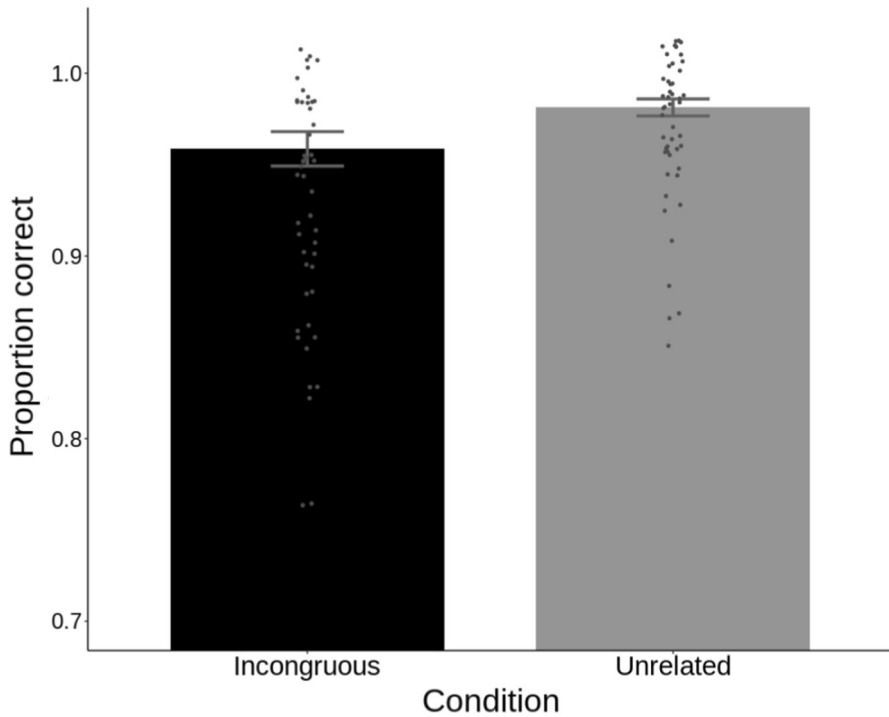


Fig. 4 Model predictions of accuracy in focal conditions is shown along with raw data. The proportion correct and error bars for the two focal conditions are shown as predictions from a linear mixed effects model of response accuracy. Points correspond to participants and are generated from raw data. Error bars indicate standard error of the point estimate

Table 5 Model outputs for the effect on accuracy

	<i>b</i>	<i>SE</i>	<i>z</i>	<i>p</i>
Intercept	4.79	.40	12.00	< .001
Condition	– .82	.27	– 3.07	< .01
Rating	– .52	.19	– 2.74	< .01

Coefficient estimates based on $n=3199$ with by subject and by item random intercepts included for estimates of random effects. “Rating” refers to the semantic rating for each target. The linear mixed effects model was estimated using the binomial family given the dichotomous outcome variable (accuracy), centered on the *Unrelated* condition. Coefficients are provided as logits, where the estimate of condition is associated with moving from the *Unrelated* to the *Incongruous* condition

Only the correct trials were included for statistical models for RT and outliers were excluded using the criteria described previously. In the model of RTs, participants in the *Incongruous* condition were associated with responses that were 52.25 ms slower than those in the *Unrelated* condition ($b=52.25$, $SE=16.71$, $F(1$,

Table 6 Descriptive statistics for response times (ms) across the four conditions

Condition	<i>M</i>	<i>SD</i>	max	min
Overall related	826	323	2739	350
Radical only	788	312	2720	352
Incongruous	870	330	2739	421
Unrelated	834	305	2694	402

74.36)=12.46, $p < 0.001$, 95% CI=[18.98, 86.76]). Model outputs are shown in Table 7. See Fig. 5 for a visual depiction of the effect.

These effects suggest that when the radical provides semantically incongruent information, the semantic processing of the overall character is impacted – an indication that readers activated irrelevant semantics from the radicals even when the semantic information interfered with the processing of the character. Analyses of reaction time provide additional support for this hypothesis; the reaction time analyses are perhaps even stronger than those provided by the accuracy data given that the overall accuracy across conditions was very high.

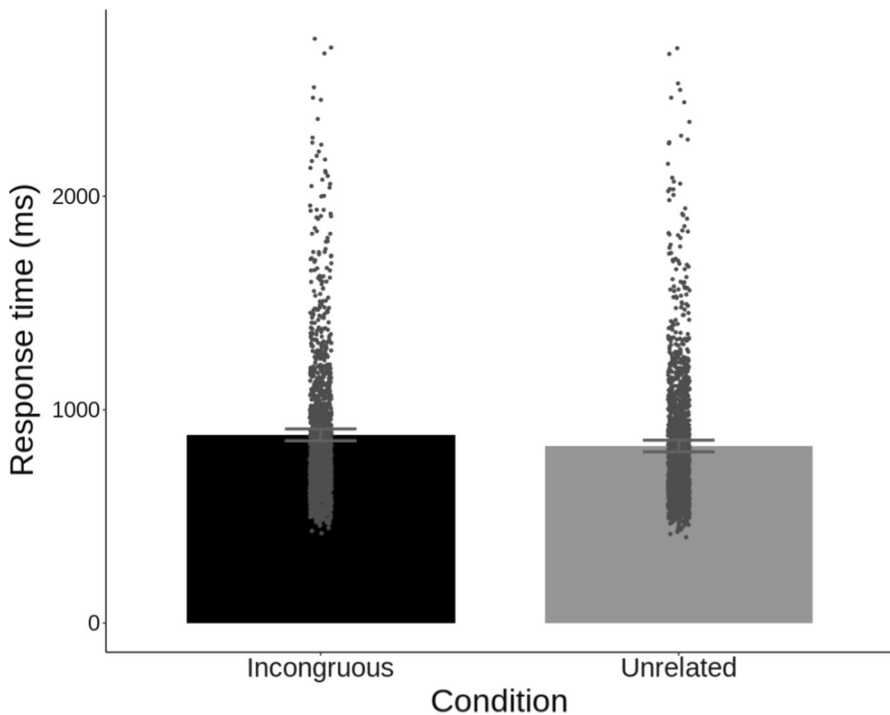


Fig. 5 Model predictions for RT in focal conditions are shown along with raw data. Data points correspond to observations taken from raw data. Error bars indicate standard error of the point estimate

Table 7 Model outputs for the effect on response times

	<i>b</i>	<i>SE</i>	<i>F</i>	<i>p</i>	95% <i>CI</i>
Intercept	782.19	33.34	550.23	< .001	[717.08, 847.36]
Condition	52.25	16.71	9.72	< .01	[18.98, 86.76]
Rating	30.31	12.89	5.52	< .05	[4.92, 55.49]

Coefficients estimated with $n=3038$, where focal effect shown is centered on the Unrelated condition. Degrees of freedom were approximated using Kenward–Roger approach with $F(1, 74.36)$ for the effect of condition and $F(1, 57.72)$ for the effect of the semantic rating. This method can result in noninteger values for degrees of freedom (Judd et al., 2012)

Discussion

The present study employed a semantic categorization task to examine the effects of radical semantics at the sub-lexical level. As expected, when the semantics of the radical were potentially misleading in relation to the meaning of the overall character, participants took longer and were less accurate at judging the semantic category of the character (e.g., deciding that 券 “ticket” does not belong to the “in the kitchen” category yielded longer RT and lower accuracy rate due to the visual presence of the radical 刀 “knife”) compared to when no misleading radical was involved (e.g., 宪 “statue” yielded faster RT and higher accuracy rate since neither the character itself nor its radical suggest any relation to the “in the kitchen” category).

By collecting auditory semantic relatedness ratings on the stimuli and incorporating them in the analysis, the present study offers improved control measures on the design compared to Leck et al. (1995). Our findings also verify that when radicals contain semantic information that is incongruent to the meaning of the combinatorial character, the overall, character-level, semantic activation is slowed, and readers experience more difficulty at judging the semantic category of the overall character.

As previously shown by Leck et al. (1995), radical semantics impact character-level semantic category judgement. Our results replicate this pattern with added control measures. These findings suggest that rather than simply impeding the processing of a character and arriving at a meaning, the semantics of a character’s radical may subtly condition processing, which could include aspects of cognition outside of reading individual characters in controlled experimental tasks like the one employed here. Since the majority of Chinese characters are combinatorial characters like the ones tested in the present experiment, it would therefore be of great interest to both psycholinguistic and sociolinguistic research to determine the extent to which such activation influences other forms of language processing. As suggested by our results, a *ticket* in the Chinese writing system may evoke concepts related to the *kitchen* because it contains a *knife* in its orthographic form. These data provide *prima facie* evidence for the presence of sublexical effects of this kind, but how such effects impact other forms of processing, such as at the word or phrase levels, requires further investigation. For instance, as each combinatorial character has

at least two units of semantic meaning embedded within it, the possible semantic impact of incongruent radicals and the extent to which they may condition the overall semantic processing of the whole character, and how that may in turn condition the processing at the phrase or sentence level, offers a potentially rich field of further investigation. Carefully controlling for the influence of sub-character semantic activation will be critical for lexical analysis studies of this kind.

Turning to the possible socio-linguistic significance of this phenomenon, we may consider more culturally salient instances of radical-character incongruence than the kitchen-knife example provided above. For example, researchers in sociology and gender studies have long suspected that the compositional nature of certain Chinese characters may predispose readers to internalize sexist and misogynist assumptions. Tan (1990) identified many combinatorial characters (e.g., 嫌 “dislike”, 奸 “evil”, 妒 “jealous”, 婪 “greedy”) with the radical 女 “female”, and suggested that “derogatory meanings” as well as “patriarchal standards for femininity” were inherent in the Chinese writing system (Tan, 1990, p. 637). Hall & Ames, (2000, p. 84) further divided these characters into two groups. They showed that in the first group, the relationships and roles of women are defined by the very characters through which they are represented, e.g., 妈 “mother” and 姑 “aunt”. The second group consists of characters that are “a rather damning collection of negative character traits and attitudes which are directly associated with the female gender,” for example, 妄 “absurdity” and 婷 “graceful” (of course not a negative trait in and of itself but one delineating desired female characteristics from the point of view of a dominant male gaze). Future psycholinguistic research on lexical access building on the findings reported here may be able to address the potential effects engendered by the culturally salient semantics of what would otherwise be considered incongruent radicals such as 女 “female” when used as a constituent component of a character denoting negative qualities or emotions. Indeed, these examples call into question the very categorization employed in this study: it is assumed that the radical 女 “female” should be considered irrelevant or incongruous in the host of words in Chinese that employ it (奴 “slave,” 奸 “wicked,” 妒 “envy,” 妖 “seductive, evil.”).³ One must at least allow for the possibility that the prominence of the radical indicating femininity in so many lexical items of this class performs a consistent function in assigning negative characteristics to the female gender (Lewis & Lupyan, 2020; Lewis et al., 2021). One could explore this same concept further through analysis of characters used to write names for historically non-Chinese peoples which exhibit a tendency to employ incongruent radicals with animalistic associations (Ford, 2020).⁴ While

³ This is not to say that all characters in Chinese with the 女 *female* radical have negative connotations. Many are simply neutral, denoting social roles or labels associated with women, whereas others represent desirable qualities, albeit often from a male perspective.

⁴ For example, there is a widely-discussed tendency for ethnonyms of foreign peoples to contain what might otherwise be considered the semantically incongruent radical 犭 “dog.” In such contexts, the animalistic associations conferred by this radical (or others such as 虫 “worm/insect,” also employed in foreign ethnonyms) do not seem entirely coincidental to a reader’s processing of the characters denoting such ethnic groups.

these patterns may be historical accidents, such diachronic language characteristics nonetheless may impact modern day readers of the script. It is important to note, however, the results from the present study employed a character-level semantic judgement task, and it will be important for future studies to assess the extent to which, if at all, these sub-lexical activations impact semantic activation beyond the single-character level.

Previous studies have attempted to resolve the debate surrounding radical's status vis-a-vis the organization of the adult reading system (Ding et al., 2004; Taft & Zhu, 1997; Taft et al., 2000). While Ding et al. (2004) concluded that radicals are a basic unit of Chinese orthographic representation, behavioral and modeling work by Yang et al. (2009) further specified that radicals' influence on lexical access may be restricted to early stage of character identification because they are more consistently related to semantics. Our work takes these conclusions further in relation to radicals' status as functional units, insofar as their semantics are activated in relation to the categorization of the overall character. These results also highlight the importance of investigating the extent to which radical semantics are activated during reading beyond the single character level and incorporating such findings into modeling works. Though the present study was not intended to examine whether sublexical activation follows a hierarchical model (see model by Taft, 2006), nor does it attempt to solve the time course of activation, this type of research nonetheless challenges notions in which characters are the unit of processing in reading Chinese by raising the question of whether the radical semantics are activated in the system. It also suggests ways in which aspects of the Chinese writing system may semantically predispose readers to associate words and concepts with categories that seem incongruous at the surface level.

In sum, our study examined the impact of an idiosyncratic characteristic of written Chinese on skilled reading. Our results point to the possibility that the properties of a given writing system may potentially influence concepts and word meanings, a conclusion that would significantly support the Script Relativity Hypothesis, according to which the script we read can shape cognition (Pae, 2020). These results thus serve as a point of departure for further lines of inquiry not only into the complex processes involved in interpreting units of meaning in reading Chinese but also the socio-cultural implications that such processes may have.

Acknowledgement We thank Dr. Randolph B. Ford for his contributions on the sociocultural significance of ethnonym naming conventions.

Authors' contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by MCB and MB. The first draft of the manuscript was written by TW and MCB and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This work has been supported by the Vilas Trust at UW-Madison and by the Institute of Education Sciences, US Department of Education, through Award #R305B150003 to UW-Madison. The opinions expressed are those of the authors and do not represent views of the US Department of Education.

Availability of data and material Data and experimental materials can be found on OSF at <https://osf.io/hqxdk/>.

Code availability Code can be found in the same repository as data and other materials: <https://osf.io/hqxdk/>.

Declarations

Conflict of interest The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. [arXiv:1406.5823](https://arxiv.org/abs/1406.5823)
- Boltz, W. (1999). Language and writing. In M. Loewe & E. Shaughnessy (Eds.), *The Cambridge History of Ancient China: From the origins of civilization to 221 BC* (pp. 74–123). Cambridge University Press. <https://doi.org/10.1017/CHOL9780521470308.004>.
- Cheng, C. M. (1981). Perception of Chinese characters. *Acta Psychologica Taiwanica*, 23(2), 137–153.
- Da, J. (2004). *Chinese text computing*. <http://lingua.mtsu.edu/chinese-computing>.
- DeFrancis, J. (1984). *The Chinese language: Fact and fantasy*. University of Hawaii Press.
- Ding, G., Peng, D., & Taft, M. (2004). The nature of the mental representation of radicals in Chinese: A priming study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(2), 530–539.
- Feldman, L. B., & Siok, W. W. (1999). Semantic radicals contribute to the visual identification of Chinese characters. *Journal of Memory and Language*, 40(4), 559–576. <https://doi.org/10.1006/jmla.1998.2629>
- Ford, R. B. (2020). *Rome, China, and the Barbarians: Ethnographic traditions and the transformation of empires*. Cambridge University Press.
- Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (3rd edn.). SAGE Publications. <https://socialsciences.mcmaster.ca/jfox/Books/Companion/>.
- Frost, R. (2012). Towards a universal model of reading. *Behavioral and Brain Sciences*, 35(5), 263–279. <https://doi.org/10.1017/S0140525X11001841>
- Hall, D. L., & Ames, R. T. (2000). Sexism, with Chinese characteristics. In C. Li (Ed.), *The sage and the second sex: Confucianism, ethics, and gender* (pp. 75–96). Open Court.
- Ho, C. S. H., Ng, T. T., & Ng, W. K. (2003). A “radical” approach to reading development in Chinese: The role of semantic radicals and phonetic radicals. *Journal of Literacy Research*, 35(3), 849–878. https://doi.org/10.1207/s15548430jlr3503_3
- Hoosain, R. (1991). *Psycholinguistic implications for linguistic relativity: A case study of Chinese*. Lawrence Erlbaum Associates Inc.
- Hsu, C. H., Tsai, J. L., Lee, C. Y., & Tzeng, O. J. L. (2009). Orthographic combinability and phonological consistency effects in reading Chinese phonograms: an event-related potential study. *Brain and Language*, 108(1), 56–66. <https://doi.org/10.1016/j.bandl.2008.09.002>.
- Ju, D., & Jackson, N. E. (1995). Graphic and phonological processing in Chinese character identification. *Journal of Reading Behavior*, 27(3), 299–313. <https://doi.org/10.1080/10862969509547885>
- Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, 103(1), 54–69. <https://doi.org/10.1037/a0028347>
- Leck, K. J., Weekes, B. S., & Chen, M. J. (1995). Visual and phonological pathways to the lexicon: Evidence from Chinese readers. *Memory & Cognition*, 23(4), 468–476. <https://doi.org/10.3758/BF03197248>
- Lewis, M., Cooper Borkenhagen, M., Converse, E., Lupyan, G., & Seidenberg, M. (2021). What might books be teaching young children about gender? *Psychological Science*. <https://doi.org/10.1177/09567976211024643>

- Lewis, M., & Lupyan, G. (2020). Gender stereotypes are reflected in the distributional structure of 25 languages. *Nature Human Behaviour*, 4(10), 1021–1028. <https://doi.org/10.1038/s41562-020-0918-6>
- McBride-Chang, C., Bialystok, E., Chong, K. K., & Li, Y. (2004). Levels of phonological awareness in three cultures. *Journal of Experimental Child Psychology*, 89(2), 93–111.
- Norman, J. (1988). *Chinese*. Cambridge University Press.
- Packard, J. L., Chen, X., Li, W., Wu, X., Gaffney, J. S., Li, H., & Anderson, R. C. (2006). Explicit instruction in orthographic structure and word morphology helps Chinese children learn to write characters. *Reading and Writing*, 19(5), 457–487. <https://doi.org/10.1007/s11145-006-9003-4>
- Pae, H. K. (2020). From linguistic relativity to script relativity. In *Script effects as the hidden drive of the mind, cognition, and culture* (pp. 37–58). Springer. https://doi.org/10.1007/978-3-030-55152-0_3
- Pearce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1–2), 8–13. <https://doi.org/10.1016/j.jneumeth.2006.11.017>
- Perfetti, C. A. (2020). Foreword. In H. K. Pae, *Script effects as the hidden drive of the mind, cognition, and culture* (pp. vii–viii). Springer.
- Perfetti, C. A., & Zhang, S. (1991). Phonological processes in reading Chinese characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(4), 633–643. <https://doi.org/10.1037/0278-7393.17.4.633>
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Seidenberg, M. S. (2011). Reading in different writing systems: One architecture, multiple solutions. In P. McCardle, B. Miller, J. R. Lee, & O. J. L. Tzeng (Eds.), *The extraordinary brain series. Dyslexia across languages: Orthography and the brain–gene–behavior link* (pp. 146–168). Paul H Brookes Publishing.
- Shu, H. (2003). Chinese writing system and learning to read. *International Journal of Psychology*, 38(5), 274–285. <https://doi.org/10.1080/00207590344000060>
- Shu, H., & Anderson, R. C. (1997). Role of radical awareness in the character and word acquisition of Chinese children. *Reading Research Quarterly*, 32(1), 78–89. <https://doi.org/10.1598/RRQ.32.1.5>
- Taft, M., Zhu, X., & Ding, G. (2000). The relationship between character and radical representation in Chinese. *Acta Psychologica Sinica*, 32(Suppl.), 3–12
- Taft, M. (2006). Processing of characters by native Chinese readers. In P. Li, L.-H. Tan, E. Bates, & O. Tzeng. (Eds.), *Handbook of East Asian psycholinguistics* (Vol. 1). Cambridge University Press. <https://doi.org/10.1017/CBO9780511550751.023>.
- Taft, M., & Zhu, X. (1997). Submorphemic processing in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(3), 761–775. <https://doi.org/10.1037/0278-7393.23.3.761>
- 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>
- Tan, D. (1990). Sexism in the Chinese language. *The National Women's Studies Association Journal*, 2(4), 635–639. <http://www.jstor.org/stable/4316075>.
- Tan, L. H., Liu, H. L., Perfetti, C. A., Spinks, J. A., Fox, P. T., & Gao, J. H. (2001). The neural system underlying Chinese logograph reading. *NeuroImage*, 13(5), 836–846. <https://doi.org/10.1006/nimg.2001.0749>
- Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition*, 15(3), 181–198. <https://doi.org/10.3758/BF03197716>
- Wang, T. (2018). How to define linguistic units from a psycholinguistic approach: A case study of morpheme. In *Proceedings of the 26th annual meeting of the international association of Chinese linguistics* (pp. 72–88). Harvard University Press.
- Wu, X., Li, W., & Anderson, R. C. (1999). Reading instruction in China. *Journal of Curriculum Studies*, 31(5), 571–586. <https://doi.org/10.1080/002202799183016>
- Yang, J., McCandliss, B. D., Shu, H., & Zevin, J. D. (2009). Simulating language-specific and language-general effects in a statistical learning model of Chinese reading. *Journal of Memory and Language*, 61(2), 238–257. <https://doi.org/10.1016/j.jml.2009.05.001>
- Yu, L., & Reichle, E. D. (2017). Chinese versus English: Insights on cognition during reading. *Trends in Cognitive Sciences*, 21(10), 721–724. <https://doi.org/10.1016/j.tics.2017.06.004>

Zhou, X., & Marslen-Wilson, W. (1999). The nature of sublexical processing in reading Chinese characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(4), 819–837. <https://doi.org/10.1037/0278-7393.25.4.819>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.