Do alternating-color words facilitate reading aloud text in Chinese? Evidence with developing and adult readers

Manuel Perea1,2 · Xiaoyun Wang3

Abstract Prior research has shown that colors induce perceptual grouping and, hence, colors can be used as word dividers during reading (Pinna & Deiana, 2014). This issue is particularly important for those writing systems that do not employ interword spaces (e.g., Chinese). The rationale is that alternating colors across words in these scripts may facilitate the process of word identification without altering the spatial distribution of text. Here, we tested whether color alternation across words produces a benefit in a reading-aloud task in native speakers of Chinese. Participants had to read two texts: one with color alternation across words and the other with monochrome words. Experiments 1 and 2 were conducted with adult readers, whereas Experiment 3 was conducted with developing readers (Grade 2 children). Results showed that color information facilitated reading aloud a text for adult readers—restricted to texts containing technical, unfamiliar words (Experiment 2)—and developing readers. We examined the implications of these findings in the context of literacy and vocabulary training.

Keywords Word recognition · Literacy training · Reading · Color information

To delimit where each word begins and ends in a text, most writing systems employ interword spaces. Unsurprisingly, the removal of interword spaces in Western languages (e.g., apoetcan surviveeverythingbutamisprint) has a large reading cost relative to the spaced format (e.g., see Pollatsek & Rayner, 1982; Rayner, Fischer, & Pollatsek, 1998). Importantly, Chinese and other writing systems (Japanese, Thai, among others) do not employ interword spaces (e.g., 大 象 打 算 在 森 林 开 一 家 商 店 in Chinese [The elephant plans to open a shop in the forest]), and skilled readers can read the resulting text without apparent difficulty—note that around 72% of Chinese words are formed by two characters, but words composed of one, three, and four characters are also common (6%, 12%, and 10%, respectively; see Li, Bicknell, Liu, Wei, & Rayner, 2014).

Inserting interword spaces in unspaced scripts does not typically produce faster reading times than the standard unspaced format (e.g., see Bai, Yan, Liversedge, Zang, & Rayner, 2008; Sainio, Hyönä, Bingushi, & Bertram, 2007; Winkel, Radach, & Lukzaneeyanawin, 2009, for evidence in Chinese, Japanese Kanji, and Thai, respectively). As argued by Bai et al. (2008), this outcome can be the net result of two opposing effects. On the one hand, word-spaced text may facilitate the process of word identification (i.e., the inclusion of spaces between words helps delineate word boundaries and it may also reduce lateral inhibition), but, on the other hand, spaced text is visually unfamiliar. Evidence supporting this interpretation comes from an eye-tracking experiment conducted by Zang, Liang, Bai, Yan, and Liversedge (2013) in Chinese. Zang et al. (2013) found briefer first-pass fixation durations on words in spaced than in unspaced sentences with adults and children—this would be consistent with the idea that word identification is facilitated by the presence of interword spaces. However, the advantage of the sentences with interword spaces vanished in the total reading times, which is consistent with the idea of a cost due to the visual unfamiliarity of the stimuli.
What we should also note here is that inserting interword spaces in unspaced scripts such as Chinese has a beneficial effect under some circumstances (e.g., when reading unfamiliar words). Blythe et al. (2012) found that native speakers of Chinese (both children and adults) benefited from the presence of interword spaces around a target stimulus when learning new words. Likewise, Shen et al. (2012) found that inserting interword spaces had a beneficial effect on nonnative readers that were learning Chinese relative to the standard unspaced format. This was so for both nonnative readers from another unspaced script (e.g., Japanese, Thai) and nonnative readers from spaced scripts (e.g., Korean, English; see also Bai et al., 2013, for similar evidence with native English speakers learning Chinese).

Thus, inserting interword spaces to delineate word boundaries may help word learning and word identification in unspaced scripts, which has important implications for reading instruction methods (see Shen et al., 2012). However, an unintended consequence of this manipulation is that it necessarily alters the spatial distribution of the sentences (e.g., the sentence 大象 打算在森林开一家商店 would be presented as 大象 打算 在 森林 开 一家 商店; see also Häikiö, Hyönä, & Bertram, 2015, for a similar observation regarding the insertion of hyphens in Finnish books for beginning readers).

In the present series of experiments, we examine the efficiency of other, less intrusive, visual cues that delimit word boundaries without modifying the spatial distribution of text. One such option is the use of bold emphasis across alternate words, as in the sentence this sentence is easy to read. Perea and Acha (2009) found that alternating-bold sentences in Spanish produced faster reading times than did unspaced sentences. However, they also found a large reading cost of alternating-bold sentences relative to the standard-spaced sentences (see also Leyland, Kirkby, Juhasz, Pollatsek, & Liversedge, 2013, for a similar manipulation using shading as a visual cue). Here, we examined the role of a more powerful visual cue (i.e., the use of alternating-color words) in an unspaced writing system (Chinese) by comparing reading aloud times in mono-color versus multicolor text (e.g., 大象 打算 在 森林 开 一家 商店 vs. 大象 打算 在 森林 开 一家 商店 [The elephant /plans to/ in /the forest /open /a /shop]).

The rationale behind the present experiments is that color information may induce “visual grouping and wholeness” (Pinna & Deiana, 2014, p. 319; see also Goldfarb & Treisman, 2011; Häikiö et al., 2015; Livingstone & Hubel, 1988; Perea, Tejero, & Winskel, 2015; Pinna, Uccula, & Tanca, 2010; Treisman & Gelade, 1980). If the unit of processing is the word, and color demarcation operates at that level of granularity, then it should produce a facilitative effect during reading. Indeed, a number of recent experiments in spaced orthographies with skilled adult readers have shown that color information helps delineate word boundaries during reading (e.g., Perea et al., 2015; Pinna & Deiana, 2014; see also Häikiö et al., 2015, for a similar manipulation at the syllable level). Specifically, Pinna and Deiana (2014) compared reading speed, reading easiness (as rated by readers on a scale from 0 to 100), and comprehension scores in a reading-aloud task in Italian using unspaced mono-color text and alternating-color text (e.g., tobeornottobe vs. tobeornottobe, respectively). Pinna and Deiana found faster reading times and increased reading easiness for the alternating-color unspaced text over the mono-color unspaced text; on the other hand, there were no differences in comprehension scores. In another experiment, Pinna and Deiana (2014) found higher values in reading easiness for alternating-color text than for mono-color text in the presence of interword spaces (e.g., to be or not to be vs. to be or not to be; see also Pinna et al., 2010, for converging evidence using a reading-aloud task). Pinna and Deiana concluded: “The results demonstrated that the chromatic similarity can influence the process of segmentation of words and, therefore, the phenomenal grouping and shape formation” (p. 348). Likewise, Perea et al. (2015) examined the role of color information in a sentence-reading experiment in which participants silently read three types of sentences in Spanish while their eye movements were registered: (1) unspaced sentences in which each alternate word was displayed in different color (e.g., tobeornottobe); (2) spaced sentences in which each alternate word was displayed in different color (e.g., to be or not to be); and (3) unspaced mono-color sentences (e.g., tobeornottobe). Perea et al. found a small reading cost in the unspaced, alternating-color sentences relative to the spaced sentences (around 5% of the total reading time), whereas the unspaced mono-color sentences produced a dramatically larger reading cost (around 31% of the total reading time; see also Reingold, Sheridan, Meadmore, Dreige, & Liversedge, 2016, for the use of color as a selective cue during sentence reading). While color information appears to be a powerful visual cue to segment words, Perea et al. (2015) pondered, “An important question for further research is whether children’s reading benefits from the use of unspaced sentences with alternating colour words (e.g., when learning unspaced languages: Chinese, Thai)” (p. 13). The present study aims to provide an answer to this question.

The main goal of the present experiments is to examine whether adult and developing readers of an unspaced writing system (Chinese) can benefit from the use of color information regarding word boundaries when reading aloud a text. Specifically, in the present experiments, we compared the reading times of multicolor texts (e.g., 大象 打算 在 森林 开 一家 商店) and standard mono-color texts (e.g., 大象 打算在 森林 开 一家 商店)—comprehension scores were also obtained for each text. As shown by Blythe et al. (2012) and Shen et al. (2012), there is a facilitative effect of demarcation in Chinese during silent sentence reading when measuring on-line eye-movement data—via inserting interword spaces. Here, we examined whether alternating-color words (i.e., a visual cue that
does not alter the spatial distribution of text) facilitate reading aloud in Chinese. It is important to stress here that, for developing readers (Grade 2 children), reading aloud may be more prevalent than silent reading. For each experiment, we compiled two texts of similar length and difficulty for participants to read aloud. One of the texts was presented in a multicolored format (i.e., using an alternating-color manipulation), and the other text was presented in a mono-colored format. The order/condition of the texts was fully counterbalanced. We employed four colors rather than two colors to avoid potential perceptual grouping between the \( n - 1 \) word and the \( n + 1 \) word (see also Pinna & Deiana, 2014, for a similar procedure). To segment the words for the multicolor texts, we followed the guidelines from the *Contemporary Chinese Language Word Segmentation Specification for Information Processing* (State Administration of Quality Supervision, and Standardization Administration of China, 1993). As word segmentation among Chinese readers may not necessarily follow all these recommendations (see Liu, Li, Lin, & Li, 2013), we verified that their agreement on word segmentation was at ceiling level (see Method sections).

Experiments 1 and 2 were conducted with adult skilled readers, whereas Experiment 3 was conducted with developing readers (Grade 2 children). The difference between Experiments 1 and 2 was the difficulty of the texts (the difficulty ratings in a 1 to 7 Likert scale, where 1 = *very easy* and 7 = *very difficult*, were 2.0–2.2 in Experiment 1 and 5.8–5.9 in Experiment 2). The rationale behind this manipulation is that texts that include technical, relatively unfamiliar words may induce more reliance on word segmentation cues such as color alternation than standard easy-to-read texts (see Blythe et al., 2012, for evidence using interword spacing as visual cues when learning new words). Finally, the materials in Experiment 3 were adapted to the reading proficiency of Grade 2 children.

As indicated earlier, there is some evidence showing that, when measuring total reading times in relatively easy-to-read texts, adult readers of unspaced writing systems do not obtain much benefit from the use of visual cues such as interword spaces to delimitate word boundaries. For instance, Bai et al. (2008) failed to find differences in total reading times between unspaced and spaced sentences in Chinese with adult readers (see also Sainio et al., 2007, for similar evidence in Japanese Kanji, and Winskel et al., 2009, for similar evidence in Thai). Bai et al. (2008) also conducted an experiment in which each alternating word in Chinese was highlighted (e.g., as in 大象打仗在森林开一家商店) and failed to find any differences in the total reading times between highlighted and standard sentences. More related to the present study, Perea et al. (2015) conducted a sentence reading experiment with adult skilled readers of Thai in which each alternating word was in different color (green/red) or in the same color (either green or red). They found no differences in reading times between alternating-color and mono-color sentences. Taken together, these findings suggest that skilled readers of unspaced writing systems learn to develop some nonvisual cues to help delimitate word boundaries during reading (e.g., see Kasisopa, Reilly, Luksaneeyanawin, & Burnham, 2013, for discussion). Thus, we do not expect (large) differences in reading-aloud times between multicolored and mono-colored texts for standard, easy-to-read texts in Chinese (Experiment 1). Critically, the scenario may be different when adult readers read technical, difficult texts that contain a number of unfamiliar words (Experiment 2) or in an immature word recognition system (Experiment 3). In this latter scenario, the potential benefit from the visual cues that delineate word boundaries may override the cost due to lack of visual familiarity (see Blythe et al., 2012; Shen et al., 2012; Zang et al., 2013, for a similar reasoning with respect to inserting interword spaces). Clearly, if color information does benefit reading aloud a text in Chinese, this would have practical implications for reading instruction and vocabulary acquisition methods in Chinese and presumably other unspaced scripts.

**Experiment 1 (adult skilled readers with standard texts)**

**Method**

**Participants**

Twenty-four students (18 females, \( M_{\text{age}} = 25.7 \) years, range: 24–28 years) from Zhejiang University (China) participated voluntarily in the experiment. They were native Mandarin Chinese speakers with normal (or corrected-to-normal) vision, and none of them was color-blind. Informed consent was obtained from all participants.

**Materials**

We selected two text passages from the *Baidu Encyclopedia*. Text A was about the evolution of wolves, and Text B was about the historical review of galaxies. In Text A, there were 384 characters (or 214 words) and 37 punctuation marks; in Text B, there were 384 characters (or 212 words) and 31 punctuation marks. Fifteen university students who did not participate in the reading-aloud experiment were asked to evaluate the difficulty of the two texts separately on a 1 to 7 Likert scale (1 = *very easy*, 7 = *very difficult*). The data showed similar scores for Text A (mean = 2.0, \( SD = 0.7 \)) and Text B (mean = 2.2, \( SD = 0.7 \)), \( t(14) = -1.146, p = .271 \). As there may be some disagreement on word segmentation among Chinese readers, the same raters who evaluated the degree of text difficulty were also asked to rate their agreement on the text segmentation by color on a 1 to 7 Likert scale (1 = *strongly disagree*, 7 = *strongly agree*). The mean score was 6.9...
We conducted a paired samples t test on the participants’ responses. Participants were instructed to read aloud two texts (either in mono-color or multicolor format) while trying to comprehend them. We computed the total reading time in seconds, and this was transformed into words per minute (i.e., reading speed). Each text was followed by two yes/no comprehension questions on the texts. Participants answered these questions orally.

Results failed to show any differences in reading speed between multicolor and mono-color texts, \( t(23) = -2.31, p = .02 \) (137.4 vs. 137.9 wpm, respectively). In fact, there was moderate evidence for invariance, as this corresponds to a Bayes factor of BF\(_{01} = 4.629\) (i.e., the present data is 4.6 times more likely under the null model than the alternative model–Cauchy \( r = 0.707\)).

The present null effect is consistent with previous research in Chinese (e.g., Bai et al., 2008; Zang et al., 2013) and other unspaced scripts (Japanese Kanji: Sainio et al., 2007; Thai: Perea et al., 2015). Before accepting the conclusion that alternating-color words do not facilitate reading aloud text in Chinese with adult skilled readers, it is important to consider an alternative explanation. Color information may be more effective at signaling word boundaries when the text contains technical, infrequent words—in the same way that inserting spaces between words in Chinese is particularly effective when learning new words (Blythe et al., 2012). The texts employed in the current experiment were easy to read: the level of difficulty was rated as 2.0–2.2 out of 7 (1 = very easy, 7 = very difficult). A stronger test of the role of color information during reading aloud in Chinese would require adult skilled readers to read more difficult texts containing technical, unfamiliar words. This was the goal of Experiment 2. The experimental procedure was parallel to that in Experiment 1, except that we selected two texts extracted from academic papers: one on medical principles of phlegm stagnation and the other on the immune responses of invertebrates. The difficulty level of these two texts was rated around 5.8–5.9 out of 7. As we may be dealing with a small-sized effect, sample size was increased to 40 participants.

Experiment 2 (adult skilled readers with difficult texts)

Method

Participants

Forty college students (27 females, \( M_{\text{age}} = 21.9\) years, range: 19–29 years) from Zhejiang University (China) took part in the experiment. They were native Mandarin Chinese speakers with normal (or corrected-to-normal) vision, and none of them was color-blind. Informed consent was obtained from all participants.

Materials

Two texts were extracted from two published academic papers. Text A was about the medical principles of phlegm stagnation, and Text B was about the immune responses of...
invertebrates. These texts are presented in Appendix B. To evaluate the difficulty of each text, we recruited the same 15 raters as in Experiment 1. The average level of difficulty of two texts on a 1 to 7 Likert scale was very similar: 5.9 for Text A ($SD = 0.6$) and 5.8 for Text B ($SD = 0.7$), $t(14) = 0.292, p = .774$—note that the texts in Experiment 2 (mean = 5.8, $SD = 0.5$) were more difficult than those in Experiment 1 (mean = 2.1, $SD = 0.6$), $t(14) = 19.843, p < .001$. These 15 raters also agreed on how the two texts were segmented into words (mean = 6.8, $SD = 0.4$). In Text A, there were 394 characters (or 205 words) and 23 punctuation marks; in Text B, there were 394 characters (or 205 words) and 26 punctuation marks. The arrangement of the two texts was the same as in Experiment 1.

**Results and discussion**

The reading speed (in words per minute) and the comprehension scores (in percentage) are presented in Table 1. Comprehension scores were 85.6% and 83.1% for the multi-color and mono-color texts, respectively (i.e., these scores were quite high despite the difficulty of the texts). As in Experiment 1, we conducted a paired t test on the participants’ reading times as a function of color: mono-color versus multicolor.

Results showed faster reading times for the multicolor than for the mono-color texts, $t(39) = 3.228, p = .003$ (110.2 vs. 104.5 wpm, respectively). This corresponded to a Bayes factor of BF$_{10} = 13.43$—this index represents $p(data|H_1)/p(data|H_0)$. Thus, given the present data, the alternate hypothesis is 13.43 times more likely than the null hypothesis, using the standard prior distribution for the alternative model --Cauchy $r = 0.707$. This can be considered as strong support for the alternate hypothesis (see Jeffreys, 1961, for some guidelines on how to interpret Bayes factors).

In sum, color information may facilitate reading aloud in adult Chinese readers when the texts are difficult to read—note that, unsurprisingly, reading speed was lower in the current experiment than in Experiment 1 (107.3 vs. 137.7 wpm, respectively). The question now is whether developing readers (Grade 2 children) also benefit from alternating-color text when reading aloud in Chinese.

**Experiment 3 (developing readers)**

**Method**

**Participants**

Twenty-four second-grade children (13 females, $M_{age} = 7.5$ years, age range: 7–9 years) from a primary school in Taizhou (Zhejiang province, China) took part voluntarily in the experiment. All of them were native speakers of Mandarin Chinese and had normal (or corrected-to-normal) vision. None of them was color-blind or had been diagnosed with learning (or reading) problems. Parental informed consent was obtained for all participants.

**Materials**

We selected two text passages that were adapted from primary students’ compositions from the Baidu Library. Text A was a story about a tiger, a leopard, and a lion; Text B was a story about shoes and socks. In Text A, there were 377 characters (or 219 words) and 57 punctuation marks; in Text B, there were 377 characters (or 220 words) and 55 punctuation marks. The two texts are presented in Appendix C. In an initial step, two second-grade students who did not participate in the experiment assessed the familiarity of the words. All of them were native speakers of Mandarin Chinese and had normal (or corrected-to-normal) vision. None of them was color-blind or had been diagnosed with learning (or reading) problems. Parental informed consent was obtained for all participants.

## Table 1

<table>
<thead>
<tr>
<th></th>
<th>Words per minute</th>
<th>Comprehension scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adults (standard) (Exp. 1)</td>
<td>Multicolored 137.4 (4.1)</td>
<td>95.8 (2.8)</td>
</tr>
<tr>
<td>Adults (standard) (Exp. 1)</td>
<td>Mono-colored 139.3 (3.1)</td>
<td>93.7 (3.4)</td>
</tr>
<tr>
<td>Adults (difficult) (Exp. 2)</td>
<td>Multicolored 110.2 (2.4)</td>
<td>85.6 (2.5)</td>
</tr>
<tr>
<td>Adults (difficult) (Exp. 2)</td>
<td>Mono-colored 104.5 (2.2)</td>
<td>83.1 (2.6)</td>
</tr>
<tr>
<td>Children (Exp. 3)</td>
<td>Multicolored 81.5 (2.1)</td>
<td>97.9 (2.1)</td>
</tr>
<tr>
<td>Children (Exp. 3)</td>
<td>Mono-colored 77.2 (2.0)</td>
<td>95.8 (2.9)</td>
</tr>
</tbody>
</table>
appropriate in difficulty for second-grade children, and (2) the two texts were equally reading friendly. In a third step, 14 Grade 2 primary school students who did not take part in the reading aloud experiment assessed the difficulty of the two texts on a 1 to 7 Likert scale (1 = very difficult, 7 = very easy). Results showed similar scores for Text A (mean = 5.9, SD = 1.1) and Text B (mean = 5.8, SD = 0.9), t(13) = 0.458, p = 0.655. These individuals also rated their agreement on the text segmentation by color on a 1 to 7 Likert scale (1 = strongly disagree, 7 = strongly agree), and the mean score was 6.9 (SD = 0.4). The arrangement of the two texts was the same as in Experiment 1. After reading each text, participants had to answer two comprehension questions.

Design and procedure

Design and procedure were the same as in Experiment 1.

Results and discussion

The reading speed (in words per minute) and the comprehension scores (in percentage) are presented in Table 1. Comprehension scores were at ceiling level: 97.9% and 95.8% for the multicolor and mono-color texts, respectively. The statistical analyses were analogous to those in Experiments 1 and 2.

Results showed faster reading times for the multicolor than for the mono-color texts, t(23) = 3.55, p = .002 (81.53 vs. 77.19 wpm, respectively). This corresponded to a Bayes factor of BF10 = 22.66, thus reflecting strong evidence toward the alternate hypothesis (Jeffreys, 1961).

Therefore, this experiment supports the view that, with developing readers, color information provides an effective visual cue to delineate words in Chinese.

General discussion

In the current experiments, we asked whether color information regarding word boundaries facilitates reading aloud a text in an unspaced script (Chinese). Results showed that color information did help text reading aloud not only for developing readers (Experiment 3) but also for adult skilled readers when the texts contained technical, unfamiliar words (Experiment 2). Taken together, these results generalize the findings reported in previous studies (Bai et al., 2013; Blythe et al., 2012) in which inserting interword spaces (i.e., another visual cue) facilitated silent reading in Chinese—in the current experiments, the effect was obtained when reading aloud a text. Importantly, this facilitative effect was obtained using a visual cue (i.e., color information) that does not alter the spatial distribution of the text, thus providing a less intrusive visual cue than inserting interword spaces (see Häkiö, Bertram, & Hyönnä, 2016, for a similar point).

The findings of Experiments 2 and 3 are consistent with the claim that color enhances perceptual grouping during reading (see Häkiö et al., 2015; Perea et al., 2015; Pinna & Deiana, 2014). Color information can serve as a visual cue to delimit word boundaries, and this facilitates the process of word identification when reading aloud in Chinese. These results may have implications for the reading instruction methods employed in unspaced writing systems (see also Bai et al., 2013; Shen et al., 2012, for discussion). At the early stages of learning to read—or when encountering texts with unfamiliar words—color information may provide an extra visual cue to help to segment the words, eliminating the readers’ burden to group the characters to form words. A fair question is to what extent the current findings can be generalized to a silent-reading scenario. Previous research has shown that colors are a powerful visual cue to segment words during silent reading in spaced orthographies (e.g., see Perea et al., 2015), which suggests that color information may also have a facilitative effect during silent reading in Chinese. Therefore, while we acknowledge that more data should be gathered, the present findings can be taken to suggest that publishing companies could, in the future, use color information for texts with unfamiliar words and in beginning texts for children (or for learners of Chinese [or other unspaced writing systems] as an L2).

Future research should examine the effect of multicolored text on poor readers compared to typically developing readers during silent reading in Chinese. For letter spacing in the Roman script, the effect of interletter spacing distance (usually interpreted as crowding effects) is larger in poor readers than in typical readers (e.g., see Moll & Jones, 2013; Perea, Panadero, Moret-Tatay, & Gómez, 2012; Spinelli, De Luca, Judica, & Zoccolotti, 2002; Zorzi et al., 2012). Similarly, multicolored text might have differential effects on poor readers compared to good readers. Furthermore, in an applied scenario, it may be desirable to examine both reading speed and comprehension by using texts of various difficulty levels—these texts would require answering both surface and deep comprehension questions (see Perea et al., 2012, for discussion). We acknowledge that when readers are highly proficient in the reading process—or when the text is easy to read, the effect of color information as a visual cue may diminish. Indeed, as shown in Experiment 1, skilled adult readers of Chinese seem to be capable of effectively segmenting and processing text in the absence of explicit visual cues corresponding to word boundaries; for instance, several adult readers indicated after the experiments that color information was distracting to them.

In sum, the present experiments demonstrated that color information is a powerful visual cue to segment words of an unspaced writing system (Chinese) not only in developing readers but also in adult skilled readers—at least for texts
containing technical, difficult words. Despite the fact that participants had not been trained to read multicolored text, their reading speed was higher than with mono-colored text. Importantly, alternating color across words keeps exactly the same the spatial distribution of the text as the standard monocolored text in unspaced scripts, thus providing a clearer scenario to normal reading than inserting spaces between words—we acknowledge that further research should compare the effectiveness of these two visual cues: insertion of interword spaces and color alternation. The present findings open a window of opportunities for further research. Two research lines seem particularly promising: (a) to examine the developmental trajectory of the effect of color information in Chinese across the different grades of elementary school (or across learning levels, in the case of adult learners of Chinese as L2)—note that a standardized set of texts varying in difficulty would allow a more in-depth evaluation of both reading speed and comprehension; and (b) to register the participants’ eye movements when reading multicolored versus mono-colored text in Chinese (or any other unspaced writing system), as this may help determine in detail the time course of the effect of color information during word identification and eye-movement guidance.

Author note The research reported in this article has been partially supported by Grant PSI2014-53444-P from the Spanish Ministry of Economy and Competitiveness. We thank Raymond Bertram, Tuomo Häikiö, and Simon Liversedge for their helpful comments on an earlier version of this article.

Appendix A

Texts used in Experiment 1

Text A

Two million years ago, there were three top killers in the food chain on the earth. They were the titan bird, the saber-toothed tiger, and the wolf, with the tiny tit birds the most powerful and the wolf the least powerful among the three. But now we can only see the existence of the wolf. Why did the tiny tit birds and the saber-toothed tigers that were gigantic, ran quickly, and attacked rapidly die out? These two kinds of animals were larger in size and more powerful than the wolf, but they lacked a social system, often living and hunting alone. The biggest disadvantage of living alone was that they were unable to protect their young effectively when they went out hunting, leading to the sharp decrease of the offspring’s survival. Second, although the titan bird and the saber-toothed tiger were powerful in hunting, they were relatively slow in digesting and eating, often failing to keep their own prey. On the contrary, the wolves operate as a team, thus they can protect the offspring from other predators more effectively. Moreover, the wolves are quick on the trigger and good at attacking the prey’s weakness, which is to bite their thighs and make them paralyzed. Then, the wolves can eat up a deer within 20 minutes, turning it into a pile of bones and leaving no chance for other animals to plunder their wealth. So as time goes on, in the long process of evolution, the other two top killers were extinct, but the wolves survive because they lived and hunt in groups. The wolves stood at the top of the food chain, but 1 million years later, humans who were more capable of cooperation occupied that position.

Questions:
1. The wolves have a higher offspring survival rate than that of the tiny tit birds and the saber-toothed tigers.
2. The wolves are slow in digesting and eating.

Text B


Translation Two million years ago, there were three top killers in the food chain on the earth. They were the titan bird, the saber-toothed tiger, and the wolf, with the tiny tit bird the most powerful and the wolf the least powerful among the three. But now we can only see the existence of the wolf. Why did the tiny tit birds and the saber-toothed tigers that were gigantic, run quickly, and attacked rapidly die out? These two kinds of animals were larger in size and more powerful than the wolf, but they lacked a social system, often living and hunting alone. The biggest disadvantage of living alone was that they were unable to protect their young effectively when they went out hunting, leading to the sharp decrease of the offspring’s survival. Second, although the titan bird and the saber-toothed tiger were powerful in hunting, they were relatively slow in digesting and eating, often failing to keep their own prey. On the contrary, the wolves operate as a team, thus they can protect the offspring from other predators more effectively. Moreover, the wolves are quick on the trigger and good at attacking the prey’s weakness, which is to bite their thighs and make them paralyzed. Then, the wolves can eat up a deer within 20 minutes, turning it into a pile of bones and leaving no chance for other animals to plunder their wealth. So as time goes on, in the long process of evolution, the other two top killers were extinct, but the wolves survive because they lived and hunt in groups. The wolves stood at the top of the food chain, but 1 million years later, humans who were more capable of cooperation occupied that position.

Questions:
1. The wolves have a higher offspring survival rate than that of the tiny tit birds and the saber-toothed tigers.
2. The wolves are slow in digesting and eating.

Text B
When you look at the sky on a clear night, you can see a light gauze-like band across the sky, looking like a great river in the sky. That is the galaxy. The galaxy is in north–south direction in the summer and east–west direction in the winter. In the past, with the underdevelopment of science, the ancients had no idea about what the galaxy was, so they imagined it as a river in the sky. The famous fairy tale in China that the cowherd (Altair) and the weaving maid (Vega) meet once a year when magpies fly together to form a bridge over the sky. That is the galaxy. The galaxy is in north–south direction in the summer and east–west direction in the winter. In the past, with the underdevelopment of science, the ancients had no idea about what the galaxy was, so they imagined it as a river in the sky. The famous fairy tale in China that the cowherd (Altair) and the weaving maid (Vega) meet once a year when magpies fly together to form a bridge over the sky. That is the galaxy. The galaxy is in north–south direction in the summer and east–west direction in the winter. The famous fairy tale in China that the cowherd (Altair) and the weaving maid (Vega) meet once a year when magpies fly together to form a bridge over the sky. Altair, which is on the east side of the galaxy, is the brightest star in Aquila; Vega, which is on the west side of the galaxy, is the brightest star in Lyra. Westerners imagine the galaxy as the milk from Hera when she was feeding an infant. So they call it the milky way. The beautiful fairy tale is no substitute for a satisfactory scientific explanation. What is the galaxy? In the early 17th century, the great Italian scientist Galileo invented the telescope and found that the galaxy is made up of hundreds of millions of stars. Actually, the milky way is just a part of the galaxy. It is the light band formed by the projection of the main body of the galaxy on the celestial sphere, and it is the round surface full of stars that people see from the side when they are placed in. The numerous stars are very close to us, and they are mixed with the interstellar dust and gas, so it is difficult for people to tell them apart with their naked eyes, making them look like a smog-covered light band. Questions: 1. The galaxy that we see is the whole part of the galaxy. 2. It is Galileo who invented the telescope.

Appendix B

Texts used in Experiment 2

Text A

The galaxy. The galaxy is in north–south direction in the summer and east–west direction in the winter. In the past, with the underdevelopment of science, the ancients had no idea about what the galaxy was, so they imagined it as a river in the sky. The famous fairy tale in China that the cowherd (Altair) and the weaving maid (Vega) meet once a year when magpies fly together to form a bridge over the sky. That is the galaxy. The galaxy is in north–south direction in the summer and east–west direction in the winter. In the past, with the underdevelopment of science, the ancients had no idea about what the galaxy was, so they imagined it as a river in the sky. The famous fairy tale in China that the cowherd (Altair) and the weaving maid (Vega) meet once a year when magpies fly together to form a bridge over the sky. Altair, which is on the east side of the galaxy, is the brightest star in Aquila; Vega, which is on the west side of the galaxy, is the brightest star in Lyra. Westerners imagine the galaxy as the milk from Hera when she was feeding an infant. So they call it the milky way. The beautiful fairy tale is no substitute for a satisfactory scientific explanation. What is the galaxy? In the early 17th century, the great Italian scientist Galileo invented the telescope and found that the galaxy is made up of hundreds of millions of stars. Actually, the milky way is just a part of the galaxy. It is the light band formed by the projection of the main body of the galaxy on the celestial sphere, and it is the round surface full of stars that people see from the side when they are placed in. The numerous stars are very close to us, and they are mixed with the interstellar dust and gas, so it is difficult for people to tell them apart with their naked eyes, making them look like a smog-covered light band. Questions: 1. The galaxy that we see is the whole part of the galaxy. 2. It is Galileo who invented the telescope.
Questions:
1. In traditional Chinese medicine, the phlegm has nothing to do with the blood stasis and they have different symptoms.
2. Traditional Chinese medicine holds that the phlegm stagnation is concerned with qi, while Western medicine believes that it correlates with the oxidized modification products of the oxidized low density lipoprotein.
3. Traditional Chinese medicine believes that phlegm and blood stasis are homologous, and they share the symptom atherosclerosis.
4. Lishui Tiaozi capsule is able to increase oxidized modification products caused by the oxidized low density lipoprotein.

Text B

Vertebrates which recognize microorganisms and cells of the immune system. When infected with microorganisms, invertebrates have innate immune defense system. An important step of the immune responses in invertebrates is the recognition of the nonself. The key to recognizing the nonself is the existence of the pattern recognition receptors and different pattern recognition receptors will lead to the same immune defense reaction.

Appendix C

Texts used in Experiment 3

Text A

A large number of vertebrates are infected with microorganisms, invertebrates usually activate various cellular and humoral immune responses. There are several steps as follows to clean the pathogens when the invertebrates suffer the infection. First, they need to recognize the microorganism as the "nonself." Then, the extracellular cascade reaction that activates serine proteases and relieves the serine protease inhibitor is triggered to amplify the signal of infection to a stronger danger signal, or to remove the false alarm. This is followed by the activation of the target open gene through the signal transduction pathway. At last, the effector response system will be activated to remove or destroy the invaders, including the phagocytosis, the encapsidation, the activation of protease's cascade reaction and melanization, the synthesis of antimicrobial peptide, and so on. Studies have shown that the recognition of the nonself is attributed to the existence of some pattern recognition receptors that are either specific and soluble or able to combine with the cell membrane. They can recognize or combine with the conservative pathogen-associated molecular patterns at the surface of the microorganisms, which does not exist in the host cells. It can be seen that the pattern recognition receptors activate innate immune responses through the recognition of pathogen-associated molecules. But the immune responses vary among different pattern recognition receptors. For example, the peptidoglycan recognition protein could lead to a melanization cascade reaction, phagocytosis, and signal transduction, while the gram-negative bacteria binding protein plays a part in innate immune signal transduction in bacterial lipopolysaccharide reaction.

Translation

After being infected with microorganisms, invertebrates usually activate various cellular and humoral immune responses. There are several steps as follows to clean the pathogens when the invertebrates suffer the infection. First, they need to recognize the microorganism as the "nonself." Then, the extracellular cascade reaction that activates serine proteases and relieves the serine protease inhibitor is triggered to amplify the signal of infection to a stronger danger signal, or to remove the false alarm. This is followed by the activation of the target open gene through the signal transduction pathway. At last, the effector response system will be activated to remove or destroy the invaders, including the phagocytosis, the encapsidation, the activation of protease's cascade reaction and melanization, the synthesis of antimicrobial peptide, and so on. Studies have shown that the recognition of the nonself is attributed to the existence of some pattern recognition receptors that are either specific and soluble or able to combine with the cell membrane. They can recognize or combine with the conservative pathogen-associated molecular patterns at the surface of the microorganisms, which does not exist in the host cells. It can be seen that the pattern recognition receptors activate innate immune responses through the recognition of pathogen-associated molecules. But the immune responses vary among different pattern recognition receptors. For example, the peptidoglycan recognition protein could lead to a melanization cascade reaction, phagocytosis, and signal transduction, while the gram-negative bacteria binding protein plays a part in innate immune signal transduction in bacterial lipopolysaccharide reaction.

Translation

After being infected with microorganisms, invertebrates usually activate various cellular and humoral immune responses. There are several steps as follows to clean the pathogens when the invertebrates suffer the infection. First, they need to recognize the microorganism as the "nonself." Then, the extracellular cascade reaction that activates serine proteases and relieves the serine protease inhibitor is triggered to amplify the signal of infection to a stronger danger signal, or to remove the false alarm. This is followed by the activation of the target open gene through the signal transduction pathway. At last, the effector response system will be activated to remove or destroy the invaders, including the phagocytosis, the encapsidation, the activation of protease's cascade reaction and melanization, the synthesis of antimicrobial peptide, and so on. Studies have shown that the recognition of the nonself is attributed to the existence of some pattern recognition
the intoxicating fragrance of meat. The tiger’s eyes lighted up, and it approached the buffalo gradually. At the same time, the king of running, the leopard, just polished its claws and came out to look for food. He also found the small buffalo and ran straight to the green grass. The tiger and the leopard met. The leopard asked the tiger quietly, “Tiger, elder brother, both of us found the small buffalo. How about a contest? The winner gets the buffalo and the loser leaves.” The tiger nodded and agreed. They went to a clearing and also invited the lion to be the referee. The lion said solemnly, “You two are both powerful in the forest! Look, there is a mountain, the first one who arrive at the top of the mountain will win.” Just the moment the lion finished talking, the tiger and leopard immediately ran desperately toward the mountain. The lion blinked its eyes and smiled, satisfied: “What two stupid guys! Now the small buffalo is my dish!” The lion came to the grass, held the small buffalo in its mouth, and disappeared into the dense forest. Having not realized that they were cheated by the lion, the tiger and the leopard were still running desperately toward the top of the mountain.

Questions:
1. The lion, the tiger, and the leopard all wanted to get the buffalo.
2. At last, the tiger won the buffalo.

Text B

Once upon a time, a pair of socks and a pair of shoes who can speak were good friends. Every day, the owner wore this pair of socks and shoes. But one day, the socks and shoes quarreled with each other. Socks said: “You are dirty! I am clean!” Shoes said: “You can’t keep up with me!” After that, the owner had to wear only socks to walk on the road. With such thin socks to stand on the floor, the feet turned red with cold. After a while, the socks were worn into holes. Then the socks were ill, so the owner let them rest. Then the shoes were on duty. The feet were very uncomfortable when they were bare in the cold shoes. Others who noticed this kept talking about the owner behind his back, and laughed at him secretly! The owner lost face and was very unhappy when he went back home. The feet thought it was their fault that made their owner lose face as they didn’t work on socks and shoes. So the feet got socks and shoes together and said sincerely and sternly to them, “Don’t blame each other for the stink. Actually, it’s my fault. I make the stink out of the sweat. I will take shower regularly in the future.” Hearing that the feet elder brother was criticizing himself, socks and shoes felt embarrassed and bowed their heads. From then on, socks and shoes became inseparable friends. They spent every day happily with their owner.

Questions:
1. The socks and the shoes never quarreled with each other.
2. It is the sweat from the feet that makes the shoes and socks stink.

References
Häkiö, T., Bertram, R., & Hyöni, J. (2016). The hyphen as a syllabification cue in reading bisyllabic and multisyllabic words among