

Article

Word $n+2$ preview effects in three-character chinese idioms and phrases

Yu, Lili, Cutter, Michael, Yan, Guoli, Bai, Xuejun, Fu, Yu, Drieghe, Denis and Liversedge, Simon Paul

Available at <http://clock.uclan.ac.uk/22349/>

Yu, Lili, Cutter, Michael, Yan, Guoli, Bai, Xuejun, Fu, Yu, Drieghe, Denis and Liversedge, Simon Paul ORCID: 0000-0002-8579-8546 (2016) Word $n+2$ preview effects in three-character chinese idioms and phrases. Language, Cognition and Neuroscience, 31 (9). pp. 1130-1149. ISSN 2327-3798

It is advisable to refer to the publisher's version if you intend to cite from the work.
<http://dx.doi.org/10.1080/23273798.2016.1197954>

For more information about UCLan's research in this area go to <http://www.uclan.ac.uk/researchgroups/> and search for <name of research Group>.

For information about Research generally at UCLan please go to <http://www.uclan.ac.uk/research/>

All outputs in CLoK are protected by Intellectual Property Rights law, including Copyright law. Copyright, IPR and Moral Rights for the works on this site are retained by the individual authors and/or other copyright owners. Terms and conditions for use of this material are defined in the [policies](#) page.

Word $n+2$ Preview Effects in Three-Character Chinese Idioms and Phrases

Lili Yu ^{1,2}, Michael G. Cutter ², Guoli Yan ¹, Xuejun Bai ¹, Yu Fu ¹, Denis Drieghe ², and
Simon P. Liversedge ²

¹ Tianjin Normal University

² University of Southampton

Author Note

Lili Yu, Guoli Yan, Xuejun Bai, and Yu Fu, Academy of Psychology and Behaviour, Tianjin Normal University, 241 Weijing Road, Tianjin, China. Michael G. Cutter, Denis Drieghe, and Simon P. Liversedge, School of Psychology, University of Southampton, Southampton, England, SO17 1BJ.

Correspondence regarding this article should be addressed to Simon P. Liversedge, School of Psychology, Shackleton Building, University of Southampton, Highfield Campus, Southampton SO17 1BJ. E-mail: S.P.Liversedge@soton.ac.uk.

Word Count: 13,593

Abstract

Prior research using the boundary paradigm suggests that Chinese readers only process word $n+2$ in the parafovea when word $n+1$ is a single character, high-frequency word. We attempted to replicate these findings (Experiment 1), and investigated whether greater $n+2$ preview effects are observed when word $n+1$ and $n+2$ form an idiom rather than a phrase (Experiment 2). Experiment 1 replicated prior findings, although additional analyses of word $n+1$ and $n+2$ as a single region revealed significant preview effects regardless of word $n+1$ frequency. In Experiment 2 there was a main effect of phrase type, such that idioms were read more quickly than phrases, and significant $n+2$ preview effects. There was no interaction between these variables, suggesting that idioms are not parafoveally processed to a greater extent than phrases. These results suggest that $n+2$ preview effects in Chinese occur under several circumstances. Factors influencing the observation of these effects are discussed.

Keywords: Parafoveal processing, eye movements, word $n+2$ preview effect, multi-word units, Chinese reading

1
2
3 During a single fixation in reading, processing is not confined to the fixated word.
4
5 Rather, information is also extracted from words in the parafovea. In an alphabetic language
6
7 such as English readers extract up to 3–4 characters to the left of fixation, and 14–15
8
9 characters to the right of fixation (McConkie & Rayner 1975; 1976). In a logographic,
10
11 character based language such as Chinese readers extract up to 1 character to the left of
12
13 fixation, and 3 characters to the right of fixation (Inhoff & Liu, 1998). This is known as the
14
15 perceptual span. In both languages the rightward span is roughly equivalent to three words,
16
17 including the currently fixated word. However, this represents the maximum amount of
18
19 information that can be processed during a fixation, and does not account for the factors that
20
21 influence the extent of parafoveal processing from one fixation to the next. The current article
22
23 investigates some of these factors. Specifically, we investigate whether readers are more
24
25 likely to process a word further into the parafovea when it and a preceding word together
26
27 form a multi-word unit, which contains more than one word, but may have a unified meaning
28
29 and therefore a single lexical entry.
30
31
32
33
34

35 Before proceeding it is important to outline two theoretical approaches to eye-
36
37 movement control, and how they account for parafoveal processing. The first approach takes
38
39 a serial view to lexical processing, and is embodied in the E-Z Reader model (Reichle,
40
41 Rayner, & Pollatsek, 2003). In this model lexical processing is confined to one word, and
42
43 attention does not shift to subsequent words until prior words have been identified. The
44
45 amount of time it takes to identify a word is determined by its processing difficulty. Once a
46
47 word is identified, attention shifts to the next word. Typically, there is a lag between attention
48
49 shifting and a saccade being made, such that attention tends to arrive at the next word before
50
51 the eye, resulting in parafoveal processing of this word. When this parafoveal word can be
52
53 identified quickly, processing may also begin on the next word before a saccade is made. The
54
55 second approach takes the view that multiple words are processed simultaneously, and is
56
57
58
59
60

1
2
3 embodied in the SWIFT model (Schad & Engbert, 2012). All words within the perceptual
4 span are processed in parallel, with the extent of the perceptual span (in characters) being
5 determined by foveal processing difficulty. Thus, the number of words that are processed in a
6 single fixation is determined by the length of each word due to visual acuity limitations, and
7 the processing load of the foveal word. Essentially, the main controversies between the two
8 approaches to eye movement control are whether lexical processing is carried out one word at
9 a time, and whether the factors that influence the spatial extent of parafoveal processing are
10 lexical or perceptual. Thus, investigating parafoveal processing of multi-word units—which
11 technically consist of more than one word, but may have a unified lexical representation—
12 has important implications for these contrasting theoretical approaches. To be clear, if the
13 processing of parafoveal information was found to be influenced by whether upcoming words
14 form a unit (i.e. have a corresponding lexical representation), then this might imply that
15 multiple words may be processed simultaneously in the parafovea even though lexical
16 processing progresses in a serial and sequential manner.
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33

34
35 One way in which parafoveal processing has been investigated is using the boundary
36 paradigm (Rayner, 1975). In this paradigm an invisible boundary precedes a target word in a
37 sentence. Prior to the point of fixation crossing this boundary either a correct or incorrect
38 parafoveal preview of the target word will be available. As the eye crosses the boundary this
39 preview is replaced by the target word. Reduced fixation times on this word after receiving a
40 correct, as opposed to incorrect preview, is referred to as *preview benefit*. Prior research has
41 demonstrated that when the preview is of the word directly after the boundary (i.e. word $n+1$)
42 then a preview benefit is reliably obtained (see Cutter, Drieghe, & Liversedge, 2015 for a
43 review).
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Studies have also investigated whether a preview benefit is obtained from a word two words to the right of an invisible boundary (i.e. word $n+2$). Typically, in alphabetic languages effects of a word $n+2$ preview are only observed when word $n+1$ is three letters long (Angele & Rayner, 2011; Kliegl, Risse, & Laubrock, 2007; Radach, Inhoff, Glover, Vorstius, 2013; Risse & Kliegl, 2011; 2012), but not when it is longer (Angele, Slattery, Yang, Kliegl, & Rayner, 2008; Rayner, Juhasz, & Brown, 2007). This is in line with both E-Z Reader and SWIFT. In E-Z Reader it is easier to identify a short rather than long word $n+1$, allowing processing to begin on word $n+2$ prior to a saccade being made across the boundary. In SWIFT a short word $n+1$ will allow word $n+2$ to remain in the perceptual span, allowing it to be processed from word n . Thus, one factor determining the spatial extent of parafoveal pre-processing is the length of word $n+1$.

More recently Cutter, Drieghe, and Liversedge (2014) demonstrated that word $n+2$ can be parafoveally processed, despite being preceded by a long word $n+1$, when these two words form a single lexical unit. In this study words $n+1$ and $n+2$ after the boundary formed a specific type of multi-word unit—a spaced compound (e.g. *teddy bear*)—and each word had their previews manipulated to be either correct or a non-word. Crucially, word $n+1$ was on average 5.65 characters long, meaning that no $n+2$ preview effects were expected on the basis of prior literature. Nevertheless, when there was a correct preview of the first constituent of the spaced compound readers obtained a substantial $n+2$ preview benefit. Cutter et al. concluded that this was due to the first constituent licensing the processing of the second constituent as part of a lexically represented multi-word unit. Thus, a second variable that determines whether word $n+2$ is parafoveally processed in alphabetic languages is whether it forms a single unit with word $n+1$. In the current study we investigated whether a similar phenomenon occurs in Chinese reading; namely, that when word $n+2$ forms a multi-word

1
2
3 unit with word $n+1$, parafoveal processing of word $n+2$ may occur to a larger extent than
4
5 previously observed.
6

7
8 There are a number of differences between alphabetic languages and Chinese that are
9 relevant to the study of parafoveal processing. Chinese is a logographic script, in that it
10 consists of several thousand visually complex characters, rather than strings of letters. Each
11 character occupies the same amount of space. A single character can be a word by itself
12 approximately 20% of the time, combine with a second character to form a longer word about
13 70% of the time, or form an even longer word with more characters 10% of the time (Modern
14 Chinese Frequency Dictionary, 1986). Thus, the space that a word takes up in Chinese is
15 typically smaller than in an alphabetic language. Another difference is that there are no word
16 spaces in Chinese, which further increases the density of Chinese text (Bai, Yan, Liversedge,
17 Zang, & Rayner, 2008). Due to the greater spatial density of Chinese, words to the right of
18 fixation are typically closer to the fovea, and are thus more available for parafoveal
19 processing.
20
21
22
23
24
25
26
27
28
29
30
31
32
33

34
35 Despite the factors discussed above, prior research suggests that word $n+2$ is only
36 rarely parafoveally processed in Chinese. Yang, Wang, Xu, and Rayner (2009) manipulated
37 the previews of the second and third character after a boundary, which together formed a
38 single word. This word was preceded by a single character high-frequency word, and
39 significant preview benefits were observed. However, in a follow-up study in which word
40 $n+1$ was controlled to be low-frequency, Yang, Rayner, Li, and Wang (2012) did not observe
41 an $n+2$ preview effect. Yan, Kliegl, Shu, Pan, and Zhou (2010) observed a similar pattern of
42 effects in a single experiment. These studies suggest that in Chinese, word $n+2$ is only
43 parafoveally processed when word $n+1$ is high frequency. However, it may be that in Chinese,
44 much like in English, words $n+1$ and $n+2$ can form a multi-word unit together. In this case
45 word $n+2$ may be parafoveally processed regardless of the frequency of word $n+1$. One such
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 possible multi-word unit that occurs in Chinese is an idiom, the processing of which is
4
5 investigated in the current study.
6

7
8 In Chinese, an idiom is defined as a fixed phrase or short sentence that consists of
9
10 several words, and has a unified meaning (Huang & Liao, 2007). Similar to the spaced
11
12 compounds used in Cutter et al.'s study, idioms are made up of more than one word in a
13
14 linguistic sense; however, all of the constituent words together express a single concept, and
15
16 as such may have a single lexical entry. Specifically, an idiom has a figurative meaning,
17
18 which is different from the literal meaning of its constituent words. For example, a literal
19
20 reading of *kick the bucket* refers to somebody forcefully hitting a bucket with their foot.
21
22 However, when used figuratively this idiom refers to somebody dying. It is unlikely that this
23
24 figurative meaning could be computed by identifying and combining the individual words.
25
26 Consequently, an idiom may also be stored and processed as a multi-word unit, and the
27
28 meaning retrieved rather than computed. Several theories have addressed this possibility.
29
30
31

32
33 According to the lexical representation hypothesis (Swinney & Cutler, 1979) idioms
34
35 are stored in the mental lexicon as long words. Upon encountering the first word of an idiom
36
37 a reader will begin activating the whole unit's lexical entry. The configuration hypothesis
38
39 (Cacciari & Tabossi, 1988) proposes that rather than being stored as single lexical units,
40
41 idioms are stored as a configuration of their constituent words at a higher level of memory.
42
43 As an idiom is encountered each word is lexically identified, and feeds activation up to this
44
45 configuration. Once sufficient activation has built up, an idiomatic key is reached, leading to
46
47 the retrieval of the idiom's figurative meaning. In this theory idioms are treated as larger units
48
49 once the idiomatic key has been reached. Hybrid theories (Libben & Titone, 2008; Titone &
50
51 Connine, 1999) propose that when idioms are encountered the figurative meaning is retrieved
52
53 as a lexicalised unit, while at the same time each constituent word is also identified. The
54
55 immediacy with which the whole idiom's lexical entry is retrieved is influenced by a number
56
57
58
59
60

1
2
3 of factors, such as the frequency of the idiom's initial constituent word (Libben & Titone,
4
5 2008). Despite their differing assumptions, all three theories predict that, at least at some
6
7 point during the identification of an idiom, it will be processed as a larger unit. Consequently,
8
9 an idiom's constituent words may be processed in parallel to a greater extent than individual
10
11 words that do not form an idiom.
12
13

14
15 A large amount of research in alphabetic languages has investigated the processing of
16
17 idioms (e.g., *break the ice*) relative to novel phrases (e.g., *break the cup*). Findings suggest
18
19 that idioms are processed faster than phrases, across several measures including response
20
21 times in a phrase classification task (Swinney & Cutler, 1979), self-paced reading times
22
23 (Conklin & Schmitt, 2008), and fixation times on final words (Underwood, Schmitt, &
24
25 Galpin, 2004) and across the whole idioms (Siyanova-Chanturia, Conklin, & Schmitt, 2011).
26
27 This processing advantage has been viewed as evidence that idioms are lexically represented.
28
29 The logic behind this is that idioms are processed faster due to their meaning being directly
30
31 retrieved, whereas a phrase's meaning has to be computed.
32
33

34
35 There are several reasons why idioms may be processed as a single unit in Chinese.
36
37 As mentioned above, there are no word boundaries in Chinese text, and consequently readers
38
39 must segment text into appropriate processing units online. However, the rules of how
40
41 Chinese words are constructed are complex; a single Chinese character may be a word by
42
43 itself or often combine with other characters to form a longer word, and thus ambiguity
44
45 regularly exists amongst Chinese readers as to where word boundaries lie (e.g., Liu, Li, Lin &
46
47 Li, 2013). Therefore, it is possible that since idioms convey a figurative meaning they might
48
49 be segmented into and processed as a single word, despite being comprised of more than one
50
51 word.
52
53
54
55
56
57
58
59
60

1
2
3 One prior study has used eye-tracking methodology to examine the processing of
4 Chinese idioms. Carrol and Conklin (2015) presented four-character Chinese idioms, which
5 were translated into English, to bilingual native Chinese speakers. It was found that the final
6 word of these translated idioms received shorter fixations than the final word of a matched
7 phrase. While this study suggests a processing advantage for the final parts of a Chinese
8 idiom, it remains unclear whether such effects occur during the reading of non-translated
9 Chinese idioms, and how the characteristics of Chinese text discussed earlier may affect this
10 process. Also, the focus of the current study was on the processing of three- rather than four-
11 character Chinese idioms. Furthermore, prior studies using tasks other than sentence reading
12 tasks have also found that three-character idioms are processed more quickly than normal
13 phrases. Gu and Liao (1995) asked participants to judge whether a linguistic unit was
14 meaningful, and found that response times were faster for three-character idioms than for
15 matched phrases. Similar results were observed in a naming task, such that idioms were
16 named more quickly than matched phrases (Zhang & Shi, 2009). If we assume that this
17 processing advantage is due to idioms being retrieved as multi-word units from the lexicon
18 we may, arguably, expect their constituent words to be processed in parallel during reading.
19 As such, the final word of an idiom may be processed to a greater extent in the parafovea
20 than the same word in a novel phrase. If this were the case, we may observe similar preview
21 effects in Chinese idioms as those observed by Cutter et al. (2014) in English spaced
22 compounds.

23
24
25 However, it is worth noting that many of the aforementioned studies reporting these
26 effects, in both English and Chinese, have not controlled the relative predictability of the final
27 words of the idioms and the phrases. This is important since the models of eye-movement
28 control discussed earlier attribute a large proportion of variance in reading times to a word's
29 predictability in context. As such, the differences between the idioms and phrases observed in

1
2
3 previous studies may be due to the high degree of predictability in idioms, rather than the
4
5 idioms having a special representational status. Thus, a further aim of our study was to
6
7 establish the idiom processing advantage in natural Chinese reading using eye-tracking
8
9 methodology, while taking care to control the predictability of the later parts of our target
10
11 items. To be clear, if we still observe an idiom processing advantage with the predictability of
12
13 the idiom from the prior context controlled, the processing advantage has to be due to the
14
15 special representational status of idioms; if, on the other hand, the idiom processing
16
17 advantage is missing from the current study, it may indicate that within idiom predictability
18
19 plays an important role in idiom processing.
20
21
22

23
24 Before proceeding, we feel it is necessary to explain why we chose to investigate
25
26 idioms in the current study, in contrast to the spaced compounds used by Cutter et al. While
27
28 we consider both idioms and spaced compounds to fall under the broader category of multi-
29
30 word units, they are not without their differences. We will save an in-depth discussion of
31
32 these differences, and the way in which they may have affected the results of the current
33
34 study, until the General Discussion. However, for now it is important to note that the idioms
35
36 used in the current study may differ from spaced compounds in several respects, such as the
37
38 level of similarity between the meaning constructed from their constituent words and whole
39
40 unit meaning (i.e. the level of decomposability¹) and the amount of ambiguity this may cause
41
42 during processing. Given that these differences exist, the question may arise as to why we
43
44 used idioms rather than spaced compounds in the current study. At a practical level the
45
46 answer is obvious; there is no such thing as a spaced compound in Chinese on account of the
47
48 language being entirely unspaced. While the most similar Chinese item we could have chosen
49
50 to test our suggestions about multi-word processing of would have been unspaced compounds,
51
52 this would have been fairly uninteresting; compounds are generally considered as ‘words’
53
54 within Chinese rather than multi-word units (Huang & Liao, 2007), and as such would not be
55
56
57
58
59
60

1
2
3 considered a compelling test of our hypothesis. Thus, we settled on idioms as the most
4
5 appropriate linguistic unit in Chinese. In addition, at a theoretical level, we are of the opinion
6
7 that the way in which different types of multi-word units have often been considered quite
8
9 separately in the literature may well be to the detriment of overall theoretical understanding
10
11 of the way in which the human cognitive system processes them. For the most part, separate
12
13 literatures and theories have developed around the processing of idioms and compound words,
14
15 despite similar issues being relevant to the processing of both. For example, the issue of
16
17 decomposability of meaning has been investigated for both compounds and idioms, in
18
19 addition to models being developed for both which allow for meaning retrieval by both a
20
21 word-by-word and whole unit route (see Pollatsek, Hyönä, & Bertram, 2000 and Kuperman,
22
23 Schreuder, Bertram & Baayen, 2010 for compounds, and Libben & Titone, 2008 for idioms).
24
25 It does not seem overly bold to advance a position whereby we should expect to see these
26
27 different units being processed in a similar manner, and in cases where this does not occur,
28
29 consider differences in relation to overarching accounts of multi-word unit processing, rather
30
31 than specific idiosyncratic theoretical accounts. As such, while we acknowledge that idioms
32
33 are different to spaced compounds, we consider it worthwhile to examine whether they are
34
35 processed in a similar manner, and, if not, to consider the reasons why.
36
37
38
39
40

41
42 To summarise, the primary aim of the current study was to investigate whether an $n+2$
43
44 preview benefit can be observed in Chinese when words $n+1$ and $n+2$ form an idiom together,
45
46 despite the low-frequency of the first word. Furthermore, we were interested in examining
47
48 whether there was a more general processing advantage for idioms over matched phrases
49
50 when predictability is controlled. We placed three-character idioms and matched phrases
51
52 directly after an invisible boundary. In both item types the first character formed a low-
53
54 frequency word (and hence was word $n+1$ after the boundary) and the second and third
55
56 characters together formed a word (and hence together were word $n+2$ after the boundary).
57
58
59
60

1
2
3 The preview of word $n+2$ was manipulated. Given that prior studies have not observed $n+2$
4
5 preview effects in Chinese when word $n+1$ was low-frequency, no main effects of our
6
7 preview manipulation were expected. However, we did hypothesise an interaction, such that
8
9 an $n+2$ preview effect would be observed in the idioms, but not the phrases. We made this
10
11 hypothesis on the basis that idioms may have multi-word unit representational status, such
12
13 that the first character of the idiom may licence the processing of the entire unit (idiom), thus
14
15 leading to the early extraction of information from word $n+2$. In contrast the first character of
16
17 a phrase should not licence the processing of a larger lexical unit (the phrase) and therefore
18
19 the early extraction of information from word $n+2$. In this way the second constituent of an
20
21 idiom would be processed in the parafovea despite the processing difficulty of the first
22
23 constituent, in a similar manner to the second constituent of a spaced compound in Cutter et
24
25 al.'s (2014) study. If this did occur, it would be strong evidence for the idioms being
26
27 parafoveally processed as a multi-word unit. Finally, we hypothesised that the idioms should
28
29 be read faster than the matched phrases, even with the predictability of word $n+2$ being
30
31 controlled. In addition to investigating whether idioms are processed as single lexical units
32
33 during reading, we also attempted to replicate the findings of Yang et al. (2009) and Yang et
34
35 al. (2012) (i.e. word $n+2$ preview effects are only observed in Chinese text when word $n+1$ is
36
37 highly frequent) in the same testing session as our main study. The primary reason for doing
38
39 this was to establish that our participants were showing similar preview effects to those
40
41 observed in prior studies.
42
43
44
45
46
47

48 Experiment 1

49
50
51 Experiment 1 was an attempt to replicate findings from previous studies on $n+2$
52
53 preview benefit in Chinese reading, using the same stimuli as Yang et al. (2009) and Yang et
54
55 al. (2012). The reason for including this replication of “benchmark” $n+2$ preview effects was
56
57 to be certain that parafoveal processing was occurring in the way that we would anticipate
58
59
60

1
2
3 based on previous findings. For example, if we obtained null $n+2$ preview effects in both the
4
5 idioms and matched phrases we wanted to be sure that it was due to the particular way the
6
7 idioms were being processed rather than our participants simply not extracting information
8
9 from word $n+2$ regardless of how favourable the circumstances were. Replicating Yang et
10
11 al.'s (2009) effect would allow us to reject the latter explanation. Conversely, if we observed
12
13 $n+2$ preview effects in both the idioms and phrases we wanted to ensure that it was due to
14
15 some characteristic of our stimuli that differed from those used by Yang et al. (2012), rather
16
17 than our participants extracting a greater deal of parafoveal information than those tested in
18
19 prior studies. Replicating Yang et al.'s (2012) null effect would allow us to reject the latter
20
21 explanation.
22
23

24 25 26 **Method**

27
28
29 **Participants.** 40 (20 males and 20 females) native Chinese speakers who were
30
31 graduate students studying in England with an average age of 25 and 40 (10 males and 30
32
33 females) native Chinese speakers who were undergraduate students studying in China with an
34
35 average age of 21 years with normal or corrected to normal vision participated.² Eighteen
36
37 additional participants were tested, but excluded from the analysis due to noticing more than
38
39 three display changes.
40
41

42
43 **Apparatus.** Participants from England were tested with an SR Research Eyelink 1000
44
45 Desktop mounted eye tracker with a sample rate of 1000 Hz. Head rests were used to reduce
46
47 head movements. Participants from China were tested with an SR Research Eyelink 2000
48
49 Tower mounted eye tracker with a sample rate of 1000 Hz. The display monitors were 21"
50
51 CRT monitors with a refresh rate of 120 Hz. The distance between the monitor and
52
53 participant was 75cm. Sentences were presented on a black background, and each character
54
55 occupied 0.9 degrees of visual angle. Participants' right eyes were tracked, although viewing
56
57
58
59
60

1
2
3 was binocular. The setting of the monitors, the distance between the monitor and participant,
4
5 and the stimuli display were exactly the same in both laboratories.
6
7

8 **Materials and Design.** 20 items with a high-frequency single-character first word and
9
10 two-character second word from Yang et al.'s (2009) Experiment 2, and 20 items with a low-
11
12 frequency single-character first word and two-character second word from Yang et al.'s
13
14 (2012) experiment were used in the current experiment. These were interspersed with the 40
15
16 items from Experiment 2 and 20 filler items, and all items were preceded by 6 practice
17
18 sentences.
19
20

21
22 The boundary paradigm (Rayner, 1975) was used. The boundary was located before
23
24 the high- or low-frequency single-character word, and either a correct or incorrect preview of
25
26 the two-character word $n+2$ was presented. The incorrect previews were pseudo-words,
27
28 consisting of two real characters, with both characters different to the characters in the correct
29
30 previews (see Figure 1). Although the stimuli in this experiment included both items with a
31
32 high-frequency and low-frequency word $n+1$, these two subsets of stimuli were treated as two
33
34 independent experiments, and were not manipulated as an independent variable. Therefore,
35
36 analyses were carried out separately for items with high- and low-frequency word $n+1$. The
37
38 characteristics of the stimuli are listed in Table 1.
39
40
41

42
43 **Procedure.** An information sheet and a consent form were presented to the
44
45 participants prior to beginning the experiment. A three-point horizontal calibration grid was
46
47 used, with an acceptance criterion of a maximum error below 0.40 degrees and an average
48
49 error below 0.20 degrees. Prior to each sentence, there was a drift check in the center of the
50
51 screen, followed by a gaze-contingent target that occupied half of a character on the first
52
53 character of each sentence. The participant needed to gaze at the target for 20ms for the trial
54
55 to begin. After the target was triggered, the sentence would appear on the screen, with the
56
57
58
59
60

1
2
3 first character beginning in the same position as the gaze contingent target. If the participants
4 failed to trigger the gaze contingent target (0.45 degrees of visual angle), or the participants'
5 level of error on the drift correct increased beyond 0.9 degrees, they were recalibrated.
6
7 Participants were recalibrated regularly, with each participant being recalibrated a minimum
8
9 of twice.
10
11
12

13
14
15 For sentences presented with the boundary paradigm, before participants' eyes
16 crossed the boundary, either a correct or incorrect word $n+2$ was displayed as a preview, and
17 as soon as participants' eyes crossed the boundary, the incorrect word $n+2$ was replaced by
18 the correct target in the text. The boundary change was programmed using SR Research
19 Experiment Builder. The boundary triggered after a single sample was detected beyond it,
20 and the delay for the display change to occur was on average 5.8 and 6.4ms in England and
21 China, respectively. The difference between the two laboratories in terms of the time needed
22 to implement the display change after the boundary was crossed was negligible (0.6ms) even
23 though it was statistically significant ($t(8478) = 11.07, p < .001$).
24
25
26
27
28
29
30
31
32
33
34

35
36 Participants were instructed to read for comprehension. After reading each sentence
37 the participant pressed a button to move on. One third of the sentences were followed by a
38 yes/no comprehension question. The experiment took approximately 30 minutes.
39
40
41
42

43 **Results**

44
45
46 Across all participants an average of 94% of the questions were answered correctly.
47
48 Linear mixed-effect models constructed using the lme4 package (Bates, Maechler, Bolker, &
49 Walker, 2014) in R (R Development Core Team, 2013) were used to analyse the data.
50
51 Preview type was a fixed factor in the model, and contrasts were set up using the `contr.sdif`
52 function in the MASS library (Venables & Ripley, 2002). This function sets up contrasts so
53 that the output from the models represents main effects and the intercept corresponds to the
54
55
56
57
58
59
60

1
2
3 grand mean. Items with a high-frequency word $n+1$ and items with a low-frequency word
4
5 $n+1$ were analysed separately. Participants and items were treated as random factors, with
6
7 both random intercepts and random slopes included in the models. Since two different groups
8
9 of participants took part in the experiment, we also constructed models in which participant
10
11 group was included as (i) a main effect, or (ii) both as a main effect, and allowed to interact
12
13 with $n+2$ preview type. This was done to ensure no confounding effects of this factor. The fit
14
15 of each of these models to the data was compared with the simplest model (i.e. the one
16
17 including the main effect of word $n+2$ preview type alone). If one of the models including
18
19 participant group had a significantly better fit to the data, this model is reported. Otherwise,
20
21 the simplest model is reported.
22
23
24

25
26 Trials in which participants blinked in any of the critical regions (14%), or in which
27
28 the display change triggered early (9%) or completed more than 10ms after fixation onset
29
30 (4%) were excluded. Altogether, 27% of the data were excluded.
31
32

33
34 Analyses were carried out on four regions: word n (the word prior to the boundary);
35
36 word $n+1$; word $n+2$; and a composite region consisting of both word $n+1$ and word $n+2$. The
37
38 composite region was used due to the high skipping rates in the current study. Word $n+1$ was
39
40 skipped 51% of the time in the current study, and on 49% of trials the composite region
41
42 received a single, first-pass fixation. It has been proposed (e.g. Li, Liu, & Rayner, 2011) that
43
44 Chinese readers may target saccades beyond the end of the currently fixated word, and upon
45
46 landing attempt to identify as many characters as possible. If this is correct then it is likely
47
48 that the effect of a preview manipulation on word $n+2$ would affect fixations across the entire
49
50 three character composite region, rather than just fixations on word $n+2$.³ Furthermore, the
51
52 analysis across this region will retain greater statistical power than either individual region,
53
54 due to less data being lost due to skipping. Due to these issues we consider this composite
55
56 region to provide the most informative measure of our manipulation. It is worth noting that
57
58
59
60

1
2
3 Yang et al. did not analyse a composite region, and so any effects observed in this region
4
5 represent an extension, rather than replication, of their work.
6
7

8 For each region, five measures were computed. First fixation duration (FFD; the
9 duration of the first fixation in an interest area); single fixation duration (SFD; the duration of
10 a fixation when it was the only fixation in the interest area); gaze duration (GD; the sum of all
11 fixations in a region from the first fixation until a saccade leaving the region); go-past time
12 (GP; the sum of all fixations from the first fixation in a region until a saccade was made to
13 the right of the region); skipping probability (the proportion of times that a region was
14 skipped). The means and standard errors are shown in Table 2. The output from the models is
15 displayed in Table 3. To increase the normality of the data, the fixation time measures were
16 log-transformed. The skipping probability data were analyzed using logistic models due to
17 the variable's binary nature.
18
19
20
21
22
23
24
25
26
27
28
29
30

31 **Word n , and Word $n+1$.** There were no significant effects of the preview
32 manipulation on fixation times on the pre-target word or word $n+1$, for either type of stimuli.
33
34
35

36 **Word $n+2$.** There were no significant effects of the $n+2$ preview when word $n+1$ was
37 low-frequency. However, when word $n+1$ was high-frequency, a 22ms $n+2$ preview effect
38 was observed in gaze durations. Also, when word $n+1$ was low-frequency, there was a
39 significant effect of whether participants were tested in England or China in FFD, SFD and
40 GD (Participants tested in China's means = 307, 311, 394ms for each measure; participants
41 tested in England's means = 290, 284, 342ms). This effect did not interact with the preview
42 type, which means that regardless of whether fixation times were longer or shorter for a
43 particular group of participants, word $n+2$ preview effects were observed when $n+1$ was high
44 frequency, but not when $n+1$ was low frequency. Given the lack of interaction, we did not
45 further analyse these two data sets separately.
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 **Composite region.** On the composite region significant effects of the preview type
4 were observed regardless of word $n+1$ frequency. In the case of a high-frequency word $n+1$
5 this effect was reliable across all fixation time measures. When word $n+1$ was low-frequency
6 this effect was reliable across all fixation time measures. When word $n+1$ was low-frequency
7 this effect was reliable across all fixation time measures. When word $n+1$ was low-frequency
8 this effect was only reliable in gaze duration. Furthermore, a main effect of participant group
9 appeared in go-past time. Again, there was no interaction between group and word $n+2$
10 preview benefit.
11
12
13
14
15

16 17 **Discussion**

18
19
20 Experiment 1 was an attempt to replicate Yang et al.'s findings, and the analyses on
21 the regions examined by Yang et al. (i.e., words n , $n+1$, and $n+2$) provided a full replication.
22 Consistent with their findings, we observed $n+2$ preview effects of 22ms on word $n+2$ in gaze
23 duration when word $n+1$ was high-frequency. Yang et al. (2009) found a marginal ($p=0.051$)
24 effect in gaze duration of 22ms, and no effect in other measures. Furthermore, Yang et al.
25 (2012) did not find any effect of the word $n+2$ preview when a low-frequency word $n+1$ was
26 used. Thus, the pattern of effects was replicated in the current experiment, and our results are
27 consistent with Yang et al.'s findings.
28
29
30
31
32
33
34
35
36
37

38
39 However, we also performed analyses on a composite region consisting of both word
40 $n+1$ and word $n+2$, which Yang et al. did not examine. In this region we observed an $n+2$
41 preview benefit regardless of the frequency of word $n+1$. When word $n+1$ was high-
42 frequency, there was an $n+2$ preview effect of 17ms, 31ms, 30ms, and 25ms in FFD, SFD,
43 GD, and GP. When word $n+1$ was a low-frequency word, word $n+2$ preview benefit was only
44 observed in gaze duration, the size of which was 24ms. The finding that $n+2$ preview effects
45 are obtained even when word $n+1$ is a low-frequency word might be considered to be
46 inconsistent with previous findings. However, it should perhaps not come as a surprise that
47 we observed an effect of our preview manipulation in our composite region, and that these
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 effects were more stable than in the individual word regions. As discussed above it is likely
4
5 that the effect of a preview manipulation on word $n+2$ would affect fixations across the entire
6
7 three character composite region, rather than just fixations on word $n+2$. Furthermore, the
8
9 analysis across this region retained greater statistical power than either individual region.
10

11
12 In summary, the results of the current experiment replicated Yang et al.'s studies.
13
14 Word $n+2$ preview effects were observed during fixations on word $n+2$ only when word $n+1$
15
16 was high-frequency. However, in a further analysis which Yang et al. did not conduct $n+2$
17
18 preview effects were observed for both types of stimuli. This suggests that while word $n+2$
19
20 preview effects were not entirely prevented by a low-frequency word $n+1$, they were
21
22 certainly delayed and decreased relative to when word $n+1$ was high frequency. The $n+2$
23
24 effect only appeared in gaze duration in the analysis of the composite region when word $n+1$
25
26 was a low-frequency word, whereas it appeared across multiple measures and regions when
27
28 word $n+1$ was high frequency. As such, the issue of whether word $n+2$ is processed
29
30 differently when it is part of an idiom relative to novel phrase remains interesting. To be clear,
31
32 it may be that in novel phrases similar $n+2$ preview effects will be observed as in the low-
33
34 frequency items, while in idioms we may expect to see a similar pattern of effects to the high-
35
36 frequency items.
37
38
39
40
41

42 Experiment 2

43
44
45 Experiment 2 was carried out to investigate parafoveal processing of Chinese idioms
46
47 relative to matched phrases. We hypothesised that the reading times in idioms would be
48
49 shorter than phrases. We also predicted that the parafoveal processing of an idiom's
50
51 constituent words (word $n+1$ and $n+2$ after an invisible boundary) may be increased relative
52
53 to parafoveal processing of a matched phrase's constituent words. To recapitulate, these
54
55 idioms and phrases were selected in order to be as similar as possible to the items with a low-
56
57
58
59
60

1
2
3 frequency word $n+1$ from Experiment 1 (254 per million in Experiment 1, 188 per million for
4
5 idioms and 162 per million for phrases in Experiment 2), such that $n+2$ preview effects
6
7 should be relatively modest in the phrases. However, due to the lack of visual word
8
9 boundaries and the necessity of online word segmentation in Chinese, as well as the
10
11 characteristics of Chinese idioms, it may be that such multi-word units are segmented and
12
13 processed as single words. Was this the case, even with a low-frequency word $n+1$, $n+2$
14
15 preview effects should be observed in the idioms, but not the phrases.
16
17

18 19 Method

20
21
22 **Participants.** The same participants were tested as in Experiment 1 in the same
23
24 testing session.
25

26
27
28 **Apparatus.** The same apparatus was used as in Experiment 1.
29

30
31 **Materials and design.** 40 verb-noun structure idioms and matched phrases with a
32
33 low-frequency word $n+1$ were used in the present experiment. The syntactic structure, and the
34
35 frequencies of word $n+1$ and word $n+2$ of the idioms and phrases were matched with the
36
37 phrases used in Yang et al.'s (2012) study. The idioms were selected from the *Chinese Idiom*
38
39 *Dictionary* (2009). 16 participants who did not take part in the eye-tracking study were asked
40
41 to define the figurative meaning of the idiom and rate their familiarity on a 6-point scale.
42
43 Idioms which were correctly defined by over 88% of the participants were selected, and the
44
45 average familiarity of the selected idioms was 4.95. According to the *Modern Chinese Words*
46
47 *Dictionary* (5th edition, 2005), word $n+1$ was always a single-character verb by itself, and the
48
49 second and the third character within the idiom or phrase formed a two-character noun, in
50
51 which the second character after the boundary was not typically processed as an individual
52
53 single-character word.⁴ Thus, word $n+1$ and $n+2$ after the boundary should have been
54
55 processed as such.
56
57
58
59
60

1
2
3 As shown in Figure 2 and Table 4, the idioms and phrases only differed in their first
4 character, with this character being controlled for stroke complexity and frequency. The
5 preceding sentential context was the same for both idioms and phrases, with the predictability
6 of the whole idioms and phrases from the sentence context being controlled. Moreover, the
7 predictability of word $n+2$ between the idioms and phrases from the context including word
8 $n+1$ was matched, as was the whole sentence naturalness. Each of these factors was rated by a
9 separate group of 16 participants, none of who participated in the eye tracking experiment.
10 Finally, the stroke complexity and frequency of the characters of word $n+2$ was matched
11 between the correct and incorrect condition. The characteristics of these stimuli are listed in
12 Table 4.
13
14
15
16
17
18
19
20
21
22
23
24
25

26 The invisible boundary was directly before the first character of an idiom or phrase,
27 and either a correct or incorrect preview of word $n+2$ was presented prior to participants
28 crossing the boundary. The incorrect previews were of the same kind as those used in
29 Experiment 1.
30
31
32
33
34
35

36 **Procedure.** The procedure was identical to Experiment 1.
37

38 39 Results

40
41 Linear mixed-effect models constructed using the lme4 package in R were used to
42 analyse the data. Word $n+2$ preview type and phrase type were treated as fixed factors, and
43 were allowed to interact. Contrasts were set up using the `contr.sdif` function from the MASS
44 library. These successive differences contrasts were set up for both the two-level variables
45 such that the intercept corresponds to the grand mean and the fixed factor estimates to the
46 main effects. The interaction between the two factors was also included in the model.
47
48
49
50
51
52
53

54 Participants and items were treated as random factors in the analysis model, with random
55 slopes for each fixed factor. We also constructed models in which the participant group was
56
57
58
59
60

1
2
3 included as only a fixed factor, or was allowed to interact with other factors. Again, the
4
5 **contr.sdif** function was used to set up the successive differences contrasts on this variable.
6
7 When a more complex model significantly improved the fit of the model, this model is
8
9 reported. If the more complex models did not significantly improve the fit of the data the
10
11 simplest model is reported. In the case of any significant interactions between our variables
12
13 **we set up custom contrasts to test the underlying simple effects, and constructed a new linear**
14
15 **mixed-effects model with these contrasts.**
16
17
18

19
20 The same data exclusion criteria were used as in Experiment 1. 19% of the trials were
21
22 lost due to participants blinking in any of the critical regions, 11% and 6% of the trials were
23
24 lost because the display change triggered early or completed more than 10ms after fixation
25
26 onset, respectively. In total, 35% of the data were excluded. Analyses were carried out on the
27
28 same regions and for the same measures as in Experiment 1. The means and standard errors
29
30 are shown in Table 5, and the model outputs in Table 6.
31
32

33
34 **Word n .** There were no significant effects of the phrase type, the preview
35
36 manipulation, or the interaction between the phrase type and preview. There was however a
37
38 significant interaction between the $n+2$ preview and participant group in the skipping of the
39
40 pre-target word. **Simple effects** contrasts revealed that this was due to a significant effect of
41
42 the $n+2$ preview for participants tested in England ($b = -0.42$, $z = -2.953$, $p = 0.003$), but not
43
44 for those tested in China ($b = 0.23$, $z = 1.541$, $p = 0.120$). The participants tested in England
45
46 skipped the pre-target word 37% of the time given a valid $n+2$ preview, but this increased to
47
48 45% when there was an invalid $n+2$ preview. This suggests that the non-word previews
49
50 attracted the point of fixation towards them, in a manner similar to low-frequency words in
51
52 alphabetic languages (e.g. Hyönä & Bertram, 2004).
53
54
55
56
57
58
59
60

1
2
3 **Word $n+1$.** There was a main effect of preview type in gaze durations of 23ms, in
4
5 addition to an interaction between preview type and phrase type in first fixation durations.
6
7 Simple effects contrasts were performed in order to further investigate this interaction. There
8
9 were significant $n+2$ preview effects in the phrases of 36ms, while no $n+2$ preview effects
10
11 were significant in the idioms ($b = -0.02, t = -0.61$).
12
13

14
15 Furthermore, in FFD, SFD, and GD, there was a significant interaction between the
16
17 word $n+2$ preview effect and participant group. The participants tested in England showed a
18
19 28ms, 32ms, and 37ms $n+2$ preview effect in FFD ($b = -0.10, t = -3.11$), SFD ($b = -0.22, t = -$
20
21 2.99) and GD ($b = -0.12, t = -3.36$), respectively. Participants tested in China did not show
22
23 any preview effects ($b = -0.01, t = -0.22$; $b = -0.20, t = -0.45$; $b = -0.00, t = -0.02$).
24
25

26
27 **Word $n+2$.** There was a significant main effect of the $n+2$ preview type in FFD
28
29 (15ms), SFD (21ms) and GD (24ms). While this effect was not significant in GP, there was a
30
31 trend of a 15ms preview benefit. There was a significant effect of phrase type in FFD (18ms)
32
33 and GD (30ms), such that a target word was fixated for less time when it was the final word
34
35 of an idiom rather than phrase. This trend was also present in SFD (12ms) and GP (20ms).
36
37 No interaction between the two factors was observed. There was a main effect of participant
38
39 group in gaze duration. However, this did not interact with any of our variables of interest.
40
41

42
43 **Composite region.** In the composite region there were significant word $n+2$ preview
44
45 effects across all fixation time measures. This effect was 22ms, 53ms, 45ms, and 38ms in
46
47 FFD, GD, GP, and SFD, respectively. There was a 48ms significant effect of phrase type in
48
49 GD, with shorter gaze durations for idioms than for phrases. There was no significant
50
51 interaction between the two factors.⁵
52
53

54 55 **Discussion**

56
57
58
59
60

1
2
3 A main effect of our preview manipulation on word $n+2$ was observed in first, single,
4 and gaze durations on word $n+2$, and in all fixation time measures on the composite region.
5
6 Furthermore, there were significant effects of phrase type on first fixation and gaze durations
7 on word $n+2$, and gaze durations across the entire unit, such that idioms were processed faster
8 than the matched phrases. Finally, there was no interaction between our two variables on
9 word $n+2$ or the composite region, and an interaction in the opposite direction to what was
10 predicted during fixations on word $n+1$.
11
12
13
14
15
16
17
18

19 This pattern of results is somewhat surprising. We hypothesised that we would
20 observe an interaction between our two variables, such that participants would obtain an $n+2$
21 preview effect only when word $n+2$ was the second constituent of an idiom. We did not
22 observe this interaction. Had this lack of interaction occurred alongside a null effect of our
23 preview manipulation it would not have been surprising. This would simply have suggested
24 that the idioms were not processed as multi-word units in the parafovea, meaning that word
25 $n+2$ was not processed to a greater extent in idioms than in other contexts where null effects
26 have been observed. However, instead we observed a main effect of our preview
27 manipulation, with substantial effects being observed in idioms and matched phrases. These
28 effects were larger and more stable than those observed in prior studies (Yan et al., 2010;
29 Yang et al., 2009, 2012) and Experiment 1 of the current paper. Clearly, something else about
30 our stimuli drove these effects. We will return to this issue in the General Discussion.
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

47 The direction of the interaction between our two variables on word $n+1$ was also
48 surprising. In this region we observed significant effects of the $n+2$ preview in phrases, but
49 not idioms. It is not unusual for $n+2$ preview effects to appear on word $n+1$ (see Risse &
50 Kliegl, 2012). However, what was puzzling was the fact that the effect was only observed
51 within the phrases. This suggests an earlier integration of information extracted from word
52 $n+2$ in these items. We have no explanation as to why this effect may have occurred, and will
53
54
55
56
57
58
59
60

1
2
3 not discuss it further. However, it is worth noting that this effect was not present in our
4
5 analysis of a composite region. We consider this to be a more representative measure of the
6
7 processing of our items, on the basis that it consists of all instances in which the target region
8
9 was fixated. In contrast, the fixation time data on word $n+1$ represents a sub-set of about 50%
10
11 of trials.
12

13
14 While we did not observe the expected interaction, we did observe main effects of
15
16 phrase type—even though the predictability of the final word of the idiom was controlled—
17
18 with idioms being read more quickly than phrases, which indicated an idiom processing
19
20 advantage in Chinese over phrases. These results are in line with the findings from English
21
22 idioms. For example, Siyanova-Chanturia et al. (2011) found that English idioms were
23
24 fixated for a shorter period of time than matched phrases. The results of the current
25
26 experiment are also in agreement with findings from decision and naming tasks in three-
27
28 character Chinese idioms, in which idioms were responded to faster than matched phrases
29
30 (Gu & Liao, 1995; Zhang & Shi, 2009). Furthermore, our results are in line with findings by
31
32 Carrol and Conklin (2015), showing that the final word of a translated four-character Chinese
33
34 idiom received shorter fixations than the final word of a matched phrase. Our findings show
35
36 that this advantage is also present for three-character idioms in their non-translated form.
37
38
39
40
41

42 Interestingly there was an interaction between participant group and preview type on
43
44 word $n+1$. This interaction took the shape of participants in England gaining an $n+2$ preview
45
46 benefit, and those in China not. This interaction was not observed during fixations on word
47
48 $n+2$ or the composite target region. Thus, it seems that readers from both groups extracted
49
50 information from word $n+2$, but that participants from England integrated this on
51
52 intermediate fixations (i.e. on word $n+1$), whereas participants from China did not integrate it
53
54 until fixating word $n+2$. This could be due to the fact that participants tested in England were
55
56 primarily graduate students studying abroad with more experience of reading, whereas
57
58
59
60

1
2
3 participants tested in China were all freshmen in the university. Veldre and Andrews (2014)
4
5 investigated how reading skill was associated with readers' eye movement behaviour by
6
7 using the moving window paradigm. Higher skilled readers were more disrupted when
8
9 reading text presented in a small window size (i.e. 3 characters), indicating a greater
10
11 sensitivity to parafoveal information for these highly skilled readers. In the current study
12
13 participants tested in England may have had a greater efficiency in reading, and as such more
14
15 substantial preview effects (However, see Whitford & Titone, 2015).
16
17
18

19 **General Discussion**

20
21
22 The current study aimed to investigate the parafoveal processing of idioms and
23
24 matched phrases in Chinese reading, alongside replicating prior research in this area.
25
26 Specifically, we were interested in the extent to which word $n+2$ was processed parafoveally
27
28 when it was the second constituent of an idiom, rather than part of a phrase. We investigated
29
30 this across two experiments. In Experiment 1 we tested the effect of word $n+1$ frequency on
31
32 word $n+2$ preview effects, with the expectation of replicating prior research. We undertook
33
34 this replication in order to ensure that the participants tested in our study were processing
35
36 parafoveal information to an equivalent extent to those tested by Yang et al. (2009; 2012).
37
38 This replication was successful, with us observing near identical effects to both Yang et al.
39
40 studies when restricting our analysis to the same interest areas used in these studies. Thus, we
41
42 can be sure that any additional effects we observed in either experiment were due to
43
44 theoretically interesting reasons, as opposed to any extraneous factors to do with our
45
46 participants or experimental setup. Through the analysis of a further interest area we also
47
48 observed additional effects in this experiment, suggesting that the modulation of $n+2$ preview
49
50 effects by word $n+1$ frequency is not as simple as previously thought. In Experiment 2 we
51
52 investigated word $n+2$ preview effects in three-character Chinese idioms and matched
53
54 phrases with a low-frequency word $n+1$ in the same population of participants. We
55
56
57
58
59
60

1
2
3 hypothesised that if idioms are processed as multi-word units then word $n+2$ preview effects
4 would be observed in these items, but not the matched phrases. We also set out to investigate
5 how idioms are foveally processed, with prior research suggesting that they would be fixated
6 for less time than matched phrases. We will first address our findings in relation to $n+2$
7 preview effects, and then discuss our results concerning idiom processing.
8
9
10
11
12
13

14 **$N+2$ Preview Effects**

15
16
17
18 While Experiment 1 showed a modulating effect of $n+1$ frequency on $n+2$ preview
19 effects, this was less clear than in prior studies. When word $n+1$ was high-frequency, $n+2$
20 preview effects were observed across several measures and in both a composite region and on
21 word $n+2$ itself. In contrast, when word $n+1$ was low-frequency this effect only appeared in
22 gaze durations across the composite region. While these effects of an $n+2$ preview given a
23 low-frequency word $n+1$ were not stable across measures, they are more than has been
24 observed in prior studies. Furthermore, the results of Experiment 2 showed that word $n+2$
25 preview effects were surprisingly stable and large in both idioms and phrases compared to
26 either type of item in Experiment 1, despite word $n+1$ always being low frequency in this
27 experiment. Thus, it seems that Chinese readers extract information from word $n+2$ more
28 reliably than has previously been observed.
29
30
31
32
33
34
35
36
37
38
39
40
41
42

43 The fact that the $n+2$ preview effects in Experiment 2 were more stable than in
44 Experiment 1 when word $n+1$ was a low-frequency single-character word is surprising, since
45 the two subsets of stimuli were tested during the same testing session using the same
46 participants, and were well matched on many factors (see Tables 1 and 4), such as the
47 frequency of word $n+1$ ($t(27) = 1.23, p = .229$) and the frequency of word $n+2$ ($t(56) = 0.23,$
48 $p = .814$). The predictability of word $n+2$ was also controlled to a similar level, with the
49 predictability being 4.0% for the phrases used in Yang et al.'s study and 0.3% in the current
50
51
52
53
54
55
56
57
58
59
60

1
2
3 study.⁶ It seems that factors other than word $n+1$ frequency and those we controlled for may
4
5 modulate the extent to which parafoveal information is processed during Chinese reading.
6
7 Potential factors that might explain the differences in $n+2$ preview benefits between the two
8
9 subsets of phrases with a low-frequency word $n+1$ were carefully examined using
10
11 independent t-tests. The only difference that was both significant and substantial between the
12
13 phrases with a low frequency word $n+1$ in Experiment 1 and the phrases in Experiment 2 was
14
15 the frequency of the characters within word $n+2$ (first character of word $n+2$, $t(46) = 2.52$, p
16
17 $= .007$; second character of word $n+2$, $t(42) = 4.22$, $p < .001$). The frequency of the first and
18
19 second character in word $n+2$ was 965 and 712 per million in Experiment 2, and only 295
20
21 and 130 per million in Experiment 1. Thus, it seems that the most likely factor driving our
22
23 effect was the character frequency of word $n+2$.
24
25
26
27

28
29 Prior studies have shown evidence of character frequency within words influencing
30
31 word recognition in Chinese. Zhang and Peng (1992) found effects in a lexical decision task,
32
33 by varying character frequency while controlling word frequency. Words with high-
34
35 frequency characters were identified more quickly and accurately than words with low-
36
37 frequency characters. Yan, Tian, Bai, and Rayner (2006) independently varied the frequency
38
39 of a target word and the characters composing it during a sentence reading task. Results
40
41 showed that both word frequency and character frequency influenced the processing of the
42
43 fixated word in first-pass reading times. Crucially, fixations were longer on words with low-
44
45 frequency characters than words with high-frequency characters. In a similar manner,
46
47 character frequency might also influence the processing of words located in the parafovea.
48
49 Ma, Li, & Rayner (2015) investigated how the properties of a parafoveal word influence the
50
51 processing of the currently fixated word (a parafoveal-on-foveal effect); they manipulated
52
53 both the frequency of the first character of and the whole of a two-character word in the
54
55 parafovea. They found that the first character frequency—but not the whole word
56
57
58
59
60

1
2
3 frequency—of the parafoveal word affected fixation times on the currently fixated word.
4
5 Thus, their results suggest that the characteristics of Chinese characters might be processed in
6
7 parallel in the parafovea independently of the words they form.
8
9

10 Our findings of word $n+2$ preview benefit in Chinese are compatible with parallel
11
12 models of eye movement control due to the assumption that all words within the perceptual
13
14 span are processed in parallel. Since all of the words within the foveally determined
15
16 perceptual span are processed in parallel, a high-frequency character will simply be processed
17
18 faster than a low-frequency character, leading to a larger preview benefit from a high-
19
20 frequency parafoveal character than a low-frequency parafoveal character (Inhoff & Rayner,
21
22 1986). However, how the activation of the characters (within words) and the word's
23
24 activation interact in Chinese reading, is not yet clear in this kind of model.
25
26
27
28

29 The $n+2$ preview benefit observed in this study may not be entirely compatible with a
30
31 serial model of eye movement control, such as E-Z Reader. In this model, for word $n+2$ to be
32
33 parafoveally processed, word $n+1$ must be identified whilst fixation remains on word n . In
34
35 other words, $n+2$ preview effects should only be obtained when word $n+1$ can be easily
36
37 identified. In our experiment, word $n+2$ preview effects were still observed in gaze durations
38
39 within the composite region despite the low-frequency of word $n+1$. However, it is not
40
41 necessarily impossible that word $n+1$ was often identified from word n despite its low-
42
43 frequency. Indeed, with a skipping rate of almost 52% it seems reasonable to assume that this
44
45 word was often identified parafoveally. In cases when word $n+1$ can be identified quickly,
46
47 and an attention shift occurs to word $n+2$, it is reasonable to assume that character frequency
48
49 influences the processing of word $n+2$, leading to greater activation of word $n+2$ in the
50
51 parafovea when its characters are high frequency. Furthermore, it is also possible that the
52
53 effect occurred due to the necessity of online word segmentation in Chinese, as opposed to
54
55 being the result of attention being deployed to word $n+2$.
56
57
58
59
60

1
2
3 Current models of Chinese word segmentation (e.g. Li, Rayner, & Cave, 2009)
4
5 propose that characters from multiple upcoming words may be processed in parallel, with
6
7 these characters feeding activation up to congruent word representations. As such, it may be
8
9 that due to the ambiguity of word boundaries in Chinese, information is extracted from
10
11 characters from several words in the parafovea, and that the frequency of these characters
12
13 will influence the extent of this activation. Note that if this is true, our results would not
14
15 necessarily violate an assumption of serial lexical processing, due to the fact that the effects
16
17 were driven by the properties of sub-lexical units (i.e. the characters) rather than the lexical
18
19 unit (i.e. the words). Further research is required to establish concretely that $n+2$ preview
20
21 effects in Chinese are modulated by character frequency.
22
23

24
25
26 Before moving on, we also feel it is necessary to briefly discuss whether the lack of
27
28 an enhanced $n+2$ preview effect in idioms relative to phrases is problematic for a serial
29
30 processing account. Recall, our original hypothesis was that if parafoveal processing is
31
32 limited by lexical as well as perceptual factors, as is the case in a serial approach, then idioms
33
34 should be processed to a greater extent in the parafovea than phrases. As such, our results
35
36 may seem problematic for a model such as E-Z Reader. However, we believe that our
37
38 findings with regard to the processing of idioms and phrases may tell us more about the
39
40 conditions required for multi-word unit processing than serial versus parallel lexical
41
42 processing in general. This will become clear in the following section.
43
44
45

46 47 **Chinese Idiom Processing**

48
49
50 Our hypothesis that idioms may be stored and processed as larger multi-word units
51
52 was only partially supported by the current results. In terms of parafoveal processing, our
53
54 results showed no obvious effects of the phrase type, with our only effect being an interaction
55
56 on word $n+1$ in the opposite direction to that which was hypothesised. Thus, it does not seem
57
58
59
60

1
2
3 as though Chinese idioms, at least those with a low predictability of the final words, are
4
5 immediately segmented into larger units in the parafovea and processed as a single word.⁷
6
7

8 Importantly, however, there were significant main effects of phrase type for foveal
9
10 inspection times of idioms. First of all, the terminal word of our items received shorter
11
12 fixations when they were part of an idiom as opposed to part of a phrase. This replicates an
13
14 effect that has previously been observed in English idioms (see Underwood, Schmitt, &
15
16 Galpin, 2004) and translated Chinese idioms (see Carrol & Conklin, 2015). Furthermore,
17
18 idioms as a whole received shorter gaze durations than phrases, an effect that has also
19
20 previously been observed in English (Siyanova-Chanturia, Conklin, & Schmitt, 2011). More
21
22 generally, this pattern of effects supports findings from a range of experimental paradigms
23
24 showing that idioms are processed more quickly than phrases. It is worth noting that we
25
26 observed these effects despite controlling the relative predictability of the later parts of our
27
28 idioms and phrases, which suggests idioms have a special representational status, and this is
29
30 not determined by the within idiom predictability. To the best of our knowledge this is the
31
32 first study that has shown the idiom processing advantage in natural Chinese reading.
33
34 Furthermore, it is also the first to show such effects whilst controlling for potential
35
36 predictability effects.
37
38
39
40
41

42 Our results provided little support for the lexical representation hypothesis (Swinney
43
44 & Cutler, 1979), which proposes that upon encountering the initial word of an idiom, readers
45
46 will activate lexical representations of the individual constituent words and whole idiom.
47
48 Upon beginning to parafoveally process the first constituent of an idiom (i.e. word $n+1$), the
49
50 representation of the entire idiom should have become activated, and thus word $n+2$ should
51
52 have been pre-processed to a greater extent than in the matched phrases regardless of the
53
54 predictability of the final parts of the idioms. This did not occur, with the pre-processing of
55
56 word $n+2$ in an idiom not differing from the pre-processing of word $n+2$ in a matched phrase.
57
58
59
60

1
2
3 The configuration hypothesis (Cacciari & Tabossi, 1988) and hybrid theories (e.g. Libben &
4
5 Titone, 2008) provide better explanations for our findings. In both of these theories the
6
7 processing of an idiom begins with the identification of each individual word, and the idiom
8
9 is only processed as a whole once a certain level of activation has been reached. This level of
10
11 activation may not have been reached prior to readers crossing the invisible boundary in our
12
13 study, hence explaining why the idioms were not pre-processed in the parafovea to a greater
14
15 extent than the phrases. However, sufficient activation would have been reached once the
16
17 idioms were directly fixated, leading to the overall processing advantage over phrases.
18
19

20
21 It is worth considering the characteristics of our idioms which may have led to them
22
23 not being processed as multi-word units in the parafovea, in contrast to the spaced
24
25 compounds investigated by Cutter et al. (2014). They proposed a mechanism that would
26
27 allow for the increased parafoveal processing of multi-word units in the context of an
28
29 interactive activation approach to word identification. Essentially, their mechanism worked
30
31 on the assumption that if a multi-word unit is stored in the lexicon, then the first constituent
32
33 would activate the lexical entry for the whole multi-word unit, and that this activation would
34
35 lead to the extension of processing further into the parafovea. The findings of the current
36
37 study suggested that while the idiom was processed as a single lexical unit once foveal
38
39 processing commenced—indicated by the standard idiom processing advantage—the
40
41 activation of the whole unit by the first character was not adequate enough to lead to greater
42
43 parafoveal processing. One obvious difference between the spaced compounds used in Cutter
44
45 et al.'s study and the idioms used in the current study as multi-word units, which may
46
47 influence the speed at which the first constituent activates a larger unit, is how likely this first
48
49 constituent is to appear as part of the multi-word unit, rather than in other contexts. In Cutter
50
51 et al.'s study the first constituent could be followed by very few words, with a predictability
52
53 norming study showing that the second constituent was 97% predictable given the first, and a
54
55
56
57
58
59
60

1
2
3 corpus search showing that the first constituent would be followed by the second constituent
4
5 42% of the time. In contrast, the second constituent of the idioms in the current study was
6
7 only 1.3% predictable given the first constituent, with a search of the CCL (Center for
8
9 Chinese Linguistics PKU) corpus showing that the first constituent would be followed by the
10
11 second a mere 0.42% of the time. Within Cutter et al.'s processing mechanism, it can be
12
13 assumed that when the first constituent of a multi-word unit is first encountered in the
14
15 parafovea, the lexical entry of this word alone will be activated, as well as any multi-word
16
17 units starting with this constituent. The level of activation of these various competing lexical
18
19 representations will be strongly influenced by their relative frequency. In cases when the first
20
21 constituent does not tend to appear outside of the context of the multi-word unit (e.g. the
22
23 spaced compounds in Cutter et al.'s study), then multi-word unit activation will be high,
24
25 causing the rapid expansion of processing across this whole unit. In cases when the first
26
27 constituent regularly appears outside of the context of the multi-word unit (e.g. the idioms in
28
29 the current study), then activation of the multi-word unit will be relatively modest compared
30
31 to activation of the word alone, leading to the continued processing of this constituent as an
32
33 isolated word. In this case, sufficient activation of the multi-word unit would not occur until
34
35 it is directly fixated, and further excitation of this multi-word unit is caused by the latter
36
37 constituents. Future studies on this topic are required to test this possibility, perhaps by
38
39 manipulating the predictability of the later part of multi-word units, or the transitional
40
41 probability between the first constituent and the rest of the unit.
42
43
44
45
46
47

48
49 There may, of course, be other differences between the idioms used in the current
50
51 study and the spaced compounds used by Cutter et al. As discussed above, idioms have both a
52
53 literal compositional meaning (i.e. *kick the bucket* referring to hitting a bucket forcefully with
54
55 a foot) and a figurative idiomatic meaning (i.e. *kick the bucket* referring to dying). One could
56
57 argue that the extra processing resources required to disambiguate the dual meanings of an
58
59
60

1
2
3 idiom somehow led to the suppression of an $n+2$ preview effect in our idioms⁸, perhaps
4
5 through a constriction of the perceptual span. In contrast the spaced compounds used by
6
7 Cutter et al. may have had fairly decomposable meanings, allowing for a larger perceptual
8
9 span and the processing of word $n+2$. However, we do not consider this position to be
10
11 supported by either our data, or the larger literature. It is worth reiterating that we observed
12
13 very little difference between our idioms and phrases in terms of preview benefit, while also
14
15 finding that the idioms were fixated for less time, regardless of preview type. This pattern of
16
17 effects would seem to be contrary to an interpretation reliant on processing problems caused
18
19 by meaning ambiguity. Presumably, if participants had struggled with resolving the conflict
20
21 between the literal and figurative meaning of the idioms, then we should have seen evidence
22
23 of this such that participants would remain fixated on the ambiguous idioms for longer than
24
25 the unambiguous phrases. This is the opposite of what we found. Furthermore, it seems
26
27 reasonable to argue that even if there was an effect of meaning ambiguity, this should not
28
29 have adversely affected the parafoveal lexical activation of the idioms (or the individual
30
31 words of which they were comprised), since at this point in processing there should not yet be
32
33 two conflicting meanings. A further point of note is that there is not a dichotomous
34
35 relationship in terms of spaced compounds always having a decomposable meaning, and
36
37 idioms always having an entirely non-decomposable meaning unrelated to their constituent
38
39 words. Rather, both of these classes of multi-word units vary on their level of
40
41 decomposability. For example, while some items from Cutter et al.'s study were highly
42
43 decomposable (e.g. *olive oil* being oil extracted from olives), it is hard to see how somebody
44
45 could derive the meaning of terms such as *tikka masala*, *helter skelter*, *ferris wheel*, *hula*
46
47 *hoop*, and *cocker spaniel* by simply combining the meaning of the two constituent words.
48
49 Idioms also vary in terms of their decomposability. For example, while there is little relation
50
51 between literally kicking a bucket and dying (i.e. a non-decomposable idiom), in the idiom
52
53
54
55
56
57
58
59
60

1
2
3 *pop the question*, meaning to ask someone to marry you, there is a relatively transparent link
4
5 between the idea of *the question* and a marriage proposal. Thus, it is not entirely clear that
6
7 idioms and spaced compounds do fundamentally differ in terms of the level of ambiguity
8
9 between their constituent-based meaning and whole phrase meaning. In the current study
10
11 there was a mixture of decomposable and non-decomposable idioms; for example, the non-
12
13 decomposable idiom “炒鱿鱼” literally reads as *stir-fry squid* with an idiomatic meaning of
14
15 *fire someone*, while the decomposable idiom “敲警钟” literally reads as *toll the alarm bell*
16
17 with an idiomatic meaning of *give a warning*. Presumably if the results of the current study
18
19 were influenced by a conflict between the literal and figurative interpretation of idioms, then
20
21 $n+2$ preview effects would vary between these two different types. Prior research has actually
22
23 investigated the differential processing of these two types of idioms, with results being
24
25 inconclusive. While a comprehensive review of this literature is beyond the scope of the
26
27 current article, there is one paper worth highlighting. Libben and Titone (2008) investigated a
28
29 large number of factors influencing the processing of idioms using regression analysis.
30
31 Critically, in the task most resembling natural reading, no effect of decomposability was
32
33 observed on reading times. In contrast, the frequency of the initial verb of the idiom did
34
35 significantly influence reading times; idioms with infrequent initial verbs were read faster
36
37 than idioms with highly frequent initial verbs. Furthermore, in offline predictability ratings
38
39 they found that a low frequency initial verb led to a higher predictability rating for the rest of
40
41 the idiom than a high frequency initial verb. Taken together, these findings suggest that the
42
43 frequency of the initial verb—and the related issue of how well this predicts the rest of the
44
45 idiom—is more important to the lexical retrieval of idioms than decomposability. For this
46
47 reason, we consider that the way in which the initial constituents of our idioms may activate
48
49 (or rather, fail to activate) the latter parts at a lexical level to be the more likely explanation of
50
51 the discrepancy between the current study and the results of Cutter et al.
52
53
54
55
56
57
58
59
60

1
2
3 To close, the current study investigated factors influencing the parafoveal processing
4 of word $n+2$ in Chinese. In Experiment 1 we observed that while $n+2$ preview effects are
5 modulated by the frequency of word $n+1$, this effect is not as strong as has previously been
6 proposed. Our findings also seem to suggest that $n+2$ preview effects in Chinese are
7 modulated by the character frequency of word $n+2$. This latter finding seems quite novel, and
8 may warrant further investigation under other controlled conditions. While idioms were not
9 processed to a greater extent than phrases in the parafovea, there were clear differences in
10 fixation times on the terminal word and across the whole unit. Thus, it seems that idioms are
11 processed differently to compositional phrases, though the extent to which this occurs during
12 parafoveal as well as foveal processing remains unclear. Further investigation of factors
13 influencing the time course of idiom processing should shed light on this aspect of Chinese
14 reading.
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Acknowledgements

This project was initiated while the first author was visiting the University of Southampton on State Scholarship Fund 201208120023 from the China Scholarship Council. We are grateful for support from the Recruitment Program of Global Experts (1000 Talents Award from Tianjin). This project was also supported by the third author's Grant: Ministry of Education Project of Key Research Institute of Humanities and Social Sciences in Universities 15JJD190003. We would like to thank Jinmian Yang and Keith Rayner for supplying the stimuli used in Experiment 1.

For Peer Review Only

References

- Angele, B., & Rayner, K. (2011). Parafoveal processing of word $n+2$ during reading: Do the preceding words matter? *Journal of Experimental Psychology: Human Perception and Performance*, 37, 1210-1220. doi: 10.1037/a0023096
- Angele, B., Slattery, T. J., Yang, J., Kliegl, R., & Rayner, K. (2008). Parafoveal processing in reading: Manipulating $n+1$ and $n+2$ previews simultaneously. *Visual Cognition*, 16, 697-707. doi:10.1080/13506280802009704
- Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. New York, NY: Cambridge University Press. doi:10.1017/CBO9780511801686
- Bai, X., Yan, G., Liversedge, S. P., Zang, C., & Rayner, K. (2008). Reading spaced and unspaced Chinese text: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, 34, 1277-1287. doi: 10.1037/0096-1523.34.5.1277
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4 classes* (R Package Version 1.1-7). Retrieved from <http://CRAN.R-project.org/package=lme4>
- Carrol, G., & Conklin, K. (2015). Cross language lexical priming extends to formulaic units: Evidence from eye-tracking suggests that this idea 'has legs'. *Bilingualism: Language and Cognition*. Advance online publication. doi:10.1017/S1366728915000103
- Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language*, 27, 668-683. doi:10.1016/0749-596X(88)90014-9

- 1
2
3 *Chinese Character Information Dictionary (Hanzi Xinxi Zidian)*. (1988). Beijing, China:
4
5 Science Press.
6
7
8 *Chinese Idiom Dictionary [Hanyu Guanyongyu Cidian]*. (2009). Beijing, China: The
9
10 Commercial Press.
11
12
13 Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly
14
15 than nonformulaic language by native and nonnative speakers? *Applied Linguistics*,
16
17 29, 72-89. doi: 10.1093/applin/amm022
18
19
20
21 Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2014). Preview benefit in English spaced
22
23 compounds. *Journal of Experimental Psychology: Learning, Memory, and Cognition*,
24
25 40, 1778-1786. doi: 10.1037/xlm0000013
26
27
28 Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2015). How is information integrated across
29
30 fixations in reading? In A. Pollatsek & R. Treiman (Eds.), *The Oxford handbook of*
31
32 *reading*. (pp. 245-260). NY: Oxford University Press.
33
34
35 doi:10.1093/oxfordhb/9780199324576.013.32
36
37
38 Gu, B., & Miao, X. (1995). A study on the comprehension of Chinese idioms [in Chinese].
39
40 *Acta Psychologica Sinica*, 27, 15-19.
41
42
43 Huang, B., & Liao, X. (Eds.). (2007). *Modern Chinese Language (Updated 3rd edition)*
44
45 *[Xiandai Hanyu (Zengding di san ban)]*. Beijing, China: Advanced Education press.
46
47
48
49 Hyönä, J., & Bertram, R. (2004). Do frequency characteristics of nonfixated words influence
50
51 the processing of fixated words during reading? *European Journal of Cognitive*
52
53 *Psychology*, 16, 104-127. doi: 10.1080/09541440340000132
54
55
56
57
58
59
60

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- Inhoff, A. W., & Liu, W. (1998). The perceptual span and oculomotor activity during the reading of Chinese sentences. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 20-34. doi: 10.1037/0096-1523.24.1.20
- Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics*, *40*, 431-439. doi: 10.3758/BF03208203
- Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773-795. doi: 10.1080/01621459.1995.10476572
- Kliegl, R., Risse, S., & Laubrock, J. (2007). Preview benefit and parafoveal-on-foveal effects from word $n+2$. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 1250-1255. doi: 10.1037/0096-1523.33.5.1250
- Kuperman, V., Schreuder, R., Bertram, R., & Baayen, R. H. (2009). Reading polymorphemic Dutch compounds: Toward a multiple route model of lexical processing. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 876-895. doi:10.1037/a0013484
- Li, X., Liu, P., & Rayner, K. (2011). Eye movement guidance in Chinese reading: Is there a preferred viewing location? *Vision Research*, *51*, 1146-1156. doi: 10.1016/j.visres.2011.03.004
- Li, X., Rayner, K., & Cave, K. R. (2009). On the segmentation of Chinese words during reading. *Cognitive Psychology*, *58*, 525-552. doi: 10.1016/j.cogpsych.2009.02.003
- Libben, M. R., & Titone, D. A. (2008). The multidetermined nature of idiom processing. *Memory & Cognition*, *36*, 1103-1121. doi: 10.3758/MC.36.6.1103

- 1
2
3 Liu, P., Li, W., Lin, N., & Li, X. (2013). Do Chinese readers follow the national standard
4 rules for word segmentation during reading? *PloS one*, 8, e55440. doi:
5 10.1371/journal.pone.0055440
6
7
8
9
10 Ma, G., Li, X., & Rayner, K. (2015). Readers extract character frequency information from
11 nonfixated-target word at long pretarget fixations during Chinese reading. *Journal of*
12 *Experimental Psychology: Human Perception and Performance*. Advance online
13 publication. doi:10.1037/xhp0000072
14
15
16
17
18
19
20 McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation
21 in reading. *Perception & Psychophysics*, 17, 578-586. doi: 10.3758/BF03203972
22
23
24
25
26 McConkie, G. W., & Rayner, K. (1976). Asymmetry of the perceptual span in reading.
27 *Bulletin of the Psychonomic Society*, 8, 365-368. doi: 10.3758/BF03335168
28
29
30
31 McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye movement control
32 during reading: The location of initial eye fixations on words. *Vision Research*, 28,
33 1107-1118. doi: 10.1016/0042-6989(88)90137-X
34
35
36
37
38 *Modern Chinese Frequency Dictionary (Xiandai Hanyu Pinly Cidian)*. (1986). Beijing:
39 Beijing Language and Culture University Press.
40
41
42
43 *Modern Chinese Words Dictionary (5th edition) [Xiandai Hanyu Cidian (Di wu ban)]*.
44 (2005). Beijing, China: The Commercial Press.
45
46
47
48
49 Morey, R. D., Rouder, J. N., & Jamil, X. (2015). *BayesFactors: Computation of Bayes*
50 *factors for common designs. Version 0.9.12*. Retrieved from <http://bayesfactorppl.r->
51 [forge.r-project.org/](http://bayesfactorppl.r-forge.r-project.org/)
52
53
54
55
56
57
58
59
60

- 1
2
3 Pollatsek, A., Hyönä, J., & Bertram, R. (2000). The role of morphological constituents in
4
5 reading Finnish compound words. *Journal of Experimental Psychology: Human*
6
7 *Perception and Performance*, 26, 820-833. doi: t0.1037//0096-1523.26.2.820
8
9
- 10 R Development Core Team. (2013). *R: A language and environment for statistical computing*.
11
12 Vienna, Austria: R Foundation for Statistical Computing. Retrieved from
13
14 <http://www.Rproject.org/>
15
16
- 17 Radach, R., Inhoff, A. W., Glover, L., & Vorstius, C. (2013). Contextual constraint and $n+2$
18
19 preview effects in reading. *The Quarterly Journal of Experimental Psychology*, 66,
20
21 619-633. doi: 10.1080/17470218.2012.761256
22
23
- 24 Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*,
25
26 7, 65-81. doi: 10.1016/0010-0285(75)90005-5
27
28
- 29 Rayner, K., Juhasz, B. J., & Brown, S. J. (2007). Do readers obtain preview benefit from
30
31 word $n+2$? A test of serial attention shift versus distributed lexical processing models
32
33 of eye movement control in reading. *Journal of Experimental Psychology: Human*
34
35 *Perception and Performance*, 33, 230-245. doi: 10.1037/0096-1523.33.1.230
36
37
- 38 Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement
39
40 control in reading: Comparisons to other models. *Behavioral and Brain Sciences*, 26,
41
42 445-476. doi: 10.1017/S0140525X03000104
43
44
- 45 Risse, S., & Kliegl, R. (2011). Adult age differences in the perceptual span during reading.
46
47 *Psychology and Aging*, 26, 451-460. doi: 10.1037/a0021616
48
49
- 50 Risse, S., & Kliegl, R. (2012). Evidence for delayed parafoveal-on-foveal effects from word
51
52 $n+2$ in reading. *Journal of experimental psychology: Human perception and*
53
54 *performance*, 38, 1026-1042. doi: 10.1037/a0027735
55
56
57
58
59
60

- 1
2
3 Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes
4 factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*, 365-374.
5
6 doi:10.1016/j.jmp.2012.08.001
7
8
9
10 Schad, D. J., & Engbert, R. (2012). The zoom lens of attention: Simulating shuffled versus
11 normal text reading using the SWIFT model. *Visual Cognition*, *20*, 391-421. doi:
12
13 10.1080/13506285.2012.670143
14
15
16
17 Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An
18 eye-tracking study of idiom processing by native and non-native speakers. *Second*
19
20 *Language Research*, *27*, 251-272. doi: 10.1177/0267658310382068
21
22
23
24 Swinney, D. A., & Cutler, A. (1979). The access and processing of idiomatic expressions.
25
26 *Journal of Verbal Learning and Verbal Behavior*, *18*, 523-534. doi: 10.1016/S0022-
27
28 5371(79)90284-6
29
30
31
32 Titone, D. A., & Connine, C. M. (1999). On the compositional and noncompositional nature
33 of idiomatic expressions. *Journal of Pragmatics*, *31*, 1655-1674. doi: 10.1016/S0378-
34
35 2166(99)00008-9
36
37
38
39 Underwood, G., Schmitt, N., & Galpin, A. (2004). The eyes have it: An eye-movement study
40 into the processing of formulaic sequences. In, N. Schmitt (Eds.) *Formulaic*
41
42 *Sequences* (pp. 155-172). Amsterdam, John Benjamins.
43
44
45
46
47 Veldre, A., & Andrews, S. (2014). Lexical quality and eye movements: Individual differences
48 in the perceptual span of skilled adult readers. *The Quarterly Journal of Experimental*
49
50 *Psychology*, *67*, 703-727. doi: 10.1080/17470218.2013.826258
51
52
53
54 Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S*. New York:
55
56 Springer
57
58
59
60

- 1
2
3 Whitford, V., & Titone, D. (2014). Second-language experience modulates eye movements
4 during first- and second-language sentence reading: Evidence from a gaze-contingent
5 moving window paradigm. *Journal of Experimental Psychology: Learning, Memory,*
6 *and Cognition, 41*, 1118-1129. doi: 10.1037/xlm0000093
7
8
9
10
11
12 Yan, G., Tian, H., Bai, X., & Rayner, K. (2006). The effect of word and character frequency
13 on the eye movements of Chinese readers. *British Journal of Psychology, 97*, 259-268.
14 doi: 10.1348/000712605X70066
15
16
17
18
19 Yan, M., Kliegl, R., Shu, H., Pan, J., & Zhou, X. (2010). Parafoveal load of word $n+1$
20 modulates preprocessing effectiveness of word $n+2$ in Chinese reading. *Journal of*
21 *Experimental Psychology: Human Perception and Performance, 36*, 1669-1676. doi:
22 10.1037/a0019329
23
24
25
26
27
28
29 Yang, J., Rayner, K., Li., & Wang, S. (2012). Is preview benefit from word $n+2$ a common
30 effect in reading Chinese? Evidence from eye movements. *Read Writ, 25*, 1079-1091.
31 doi: 10.1007/s11145-010-9282-7
32
33
34
35
36
37 Yang, J., Wang, S., Xu, Y., & Rayner, K. (2009). Do chinese readers obtain preview benefit
38 from word $n+2$? Evidence from eye movements. *Journal of Experimental Psychology:*
39 *Human Perception and Performance, 35*, 1192-1204. doi: 10.1037/a0013554
40
41
42
43
44 Zang, C., Liversedge, S. P., Bai, X., & Yan, G. (2011). Eye movements during Chinese
45 reading. In, Gilchrist, Iain and Everling, Stefan (Eds.) *The Oxford Handbook of Eye*
46 *Movements* (pp. 961-978). Oxford, GB, Oxford University Press.
47
48
49
50
51
52 Zhang, B., & Peng, D. (1992). Decomposed storage in the Chinese lexicon. In H. C. Chen &
53 O. J. L. Tzeng. (Eds.), *Language Processing in Chinese* (pp. 131-149), North-Holland.
54
55
56
57
58
59
60

Zhang, J., & Shi, Y. (2009). The production mechanism of Chinese idioms [in Chinese]. *Acta Psychologica Sinica*, 41, 659-675.

For Peer Review Only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Footnotes

¹ Terminology differs between the idiom and compound literature for this factor. Given the focus of this paper we adopt the terminology traditionally used for idioms (i.e. decomposable vs. non-decomposable) rather than compounds (i.e. transparent vs. opaque).

² We were forced to include different participant groups due to where the first author was based at the different times of testing, rather than for any theoretical reason. It is worth noting that all participants were exposed to all experimental conditions, making this a within subject design.

³ It should be noted that effects appearing on word $n+1$ and the composite region could, of course, also be due to delayed parafoveal-on-foveal effects (Risse & Kliegl, 2012), as well as a preview benefit appearing in mislocated fixations. Regardless of the exact reason why effects may be distributed over a larger region, at a theoretical level this analysis still demonstrates an effect of upcoming parafoveal information having been processed prior to the eyes crossing the invisible boundary.

⁴ The probability of the second character after the boundary forming a word with other characters, as opposed to a single character word, was 0.78. This was calculated by dividing the frequency that this character appeared as part of a multi-character word by the total frequency of the character.

⁵ While no significant interactions were observed on word $n+2$ or in our composite region, there were clearly numerical differences in the size of preview benefit effects for idioms and phrases, albeit of inconsistent directions across measures. In order to affirm that we had obtained a null effect, rather than failed to obtain an effect due to a lack of statistical power, we undertook Bayes factor analysis. Bayes factors provide a value indicating the extent to which a dataset provides support for one of two competing hypotheses. Decreasing

1
2
3 Bayes values (below 1) provide increasing support for one hypothesis, while increasing
4
5 Bayes values (above 1) provide support for the alternative hypothesis (see Kass & Raftery,
6
7 1995 and Rouder, Morey, Speckman, & Province, 2012 for a more in-depth discussion). We
8
9 calculated Bayes factors using the BayesFactor package (Morey, Rouder, & Jamil, 2015) in R
10
11 to ascertain whether our results for word $n+2$, and for our composite region were more
12
13 adequately reflected by a model including an additive, or interactive relationship between
14
15 phrase type and $n+2$ preview type. For all reading time measures on word $n+2$ Bayes factors
16
17 were between 0.1 and 0.15, representing substantial support for an additive model. For first
18
19 fixation, gaze duration, go past time, and single fixation duration on the composite region,
20
21 Bayes factors were 0.07, 0.07, 0.12, and 0.38. Three of these values also represent at least
22
23 substantial support for the additive effect over the interactive effect, while one represents
24
25 anecdotal support. These values suggest that the lack of a significant interaction in our main
26
27 analyses were due to the null effect being supported, rather than a lack of statistical power.
28
29
30
31

32
33 ⁶ Although the predictability from the text prior to word $n+2$ was slightly higher in the
34
35 phrases used in Yang et al.'s (2012) study than the phrases in the current Experiment 2 ($t(19)$
36
37 = -1.79, $p = 0.089$), the difference favored Yang et al.'s stimuli producing larger word $n+2$
38
39 preview effects, which was obviously not the case in the current results. Therefore, in the
40
41 current study the predictability of word $n+2$ could not be the cause of larger word $n+2$
42
43 preview effect.
44
45

46
47 ⁷ It could be argued that a ceiling effect might exist for the word $n+2$ preview effect
48
49 observed in the current study, and this ceiling effect might prevent a larger word $n+2$ preview
50
51 effect being observed in idioms. Although we do not reject this possibility, we do consider it
52
53 unlikely. Indeed, our initial question was whether there would be any $n+2$ preview effects at
54
55 all given the frequency of word $n+1$.
56
57
58
59
60

⁸We thank an anonymous reviewer for this insight.

For Peer Review Only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure Captions

Figure 1. An Example of Stimuli in Experiment 1. The two characters in bold comprise word $n+2$. They were not in bold and in same font as the other characters in the experiments. The character after the boundary “|” and prior to word $n+2$ is word $n+1$.

Figure 2. An Example of Stimuli in Experiment 2. The two characters in bold comprise word $n+2$. They were not in bold and in the same font as the other characters in the experiment. The character after the boundary “|” and prior to word $n+2$ is word $n+1$. Sentence frames were the same for idioms and phrases prior to the boundary.

PARAFOVEAL PROCESSING OF CHINESE IDIOMS

Table 1

Characteristics of Stimuli from Two Subsets of Stimuli in Experiment 1

	Type	Low-frequency word $n+1$	High-frequency word $n+1$
Word $n+1$	Number of strokes	8.80 (3.14)	6.65 (3.82)
	Word frequency	254 (306)	6070 (8149)
Word $n+2$ Correct preview	Word frequency	6.23 (5.80)	8.19 (8.31)
	Number of strokes of the first character	9.65 (2.70)	9.40 (3.20)
	Frequency of the first character	295 (357)	262 (314)
	Number of strokes of the second character	9.50 (2.46)	9.70 (2.42)
Word $n+2$ Incorrect preview	Frequency of the second character	130 (121)	194 (267)
	Number of strokes of the first character	9.65 (2.70)	9.25 (3.08)
	Frequency of the first character	294 (353)	265 (320)
	Number of strokes of the second character	9.50 (2.46)	9.65 (3.13)
	Frequency of the second character	130 (123)	165 (248)

Note. Word frequency was taken from the People's Daily Corpus (1998), and character frequency was taken from the Chinese Character Information Dictionary (1988). The frequencies are per million.

PARAFOVEAL PROCESSING OF CHINESE IDIOMS

Table 2

*Means (Standard Errors) for Fixation Time and Fixation Probability Measure for All**Target Regions in Experiment 1*

Frequency of word $n+1$	Analysis region	Preview type	Measures					
			FFD	SFD	GD	GP	SP	
Low-frequency	Word n	Correct	290 (8)	296 (10)	296 (9)	305 (9)	0.53 (0.02)	
		Incorrect	283 (6)	273 (8)	301 (9)	313 (10)	0.55 (0.02)	
	Word $n+1$	Correct	319 (8)	327 (10)	329 (9)	443 (10)	0.47 (0.02)	
		Incorrect	316 (7)	308 (8)	339 (9)	444 (10)	0.43 (0.02)	
	Word $n+2$	Correct	296 (5)	292 (7)	375 (10)	479 (10)	0.14 (0.02)	
		Incorrect	301 (5)	301 (8)	360 (9)	472 (9)	0.11 (0.01)	
	Composite Region	Correct	315 (6)	318 (12)	531 (14)	616 (16)	0.03 (0.01)	
		Incorrect	320 (6)	341 (15)	555 (15)	632 (16)	0.03 (0.01)	
	High-frequency	Word n	Correct	270 (5)	273 (7)	276 (6)	287 (6)	0.47 (0.02)
			Incorrect	273 (5)	278 (6)	280 (6)	294 (6)	0.51 (0.02)
Word $n+1$		Correct	257 (5)	258 (6)	261 (6)	369 (8)	0.55 (0.02)	
		Incorrect	266 (5)	274 (7)	270 (6)	371 (7)	0.54 (0.02)	
Word $n+2$		Correct	268 (4)	275 (6)	319 (6)	434 (7)	0.14 (0.01)	
		Incorrect	280 (4)	286 (6)	340 (7)	456 (7)	0.12 (0.01)	
Composite region		Correct	262 (4)	269 (8)	404 (9)	449 (10)	0.04 (0.01)	
		Incorrect	279 (4)	300 (8)	434 (8)	474 (9)	0.03 (0.01)	

Note. FFD = first fixation duration; SFD = single fixation duration; GD = gaze

duration; GP = go past time; SP = skipping probability

PARAFOVEAL PROCESSING OF CHINESE IDIOMS

Table 3

Fixed Effects Estimates from the LME Models for All Measures Across All Target Regions in Experiment 1

Region	Effect	Measure																		
		FFD			SFD			GD			GP			SP						
		b	SE	t	b	SE	t	b	SE	t	b	SE	t	b	SE	t	b	SE	t	z
Word <i>n</i> +1 High Frequency																				
Word <i>n</i>	Preview	-0.02	0.03	-0.75	-0.02	0.03	-0.65	-0.02	0.03	-0.62	-0.02	0.03	-0.7	-0.19	0.12	-1.63				
Word <i>n</i> + 1	Preview	-0.03	0.03	-1.34	-0.04	0.04	-0.99	-0.03	0.03	-0.95	-0.01	0.03	-0.21	0.06	0.11	0.57				
Word <i>n</i> + 2	Preview	-0.04	0.02	-1.80	-0.05	0.03	-1.53	-0.06	0.03	-2.07	-0.04	0.02	-1.70	0.15	0.17	0.85				
Composite region	Preview	-0.06	0.02	-2.80	-0.12	0.04	-2.81	-0.09	0.03	-3.37	-0.09	0.03	-3.41	0.37	0.32	1.16				
Word <i>n</i> +2 Low Frequency																				
Word <i>n</i>	Preview	-0.02	0.04	-0.45	0.04	0.05	0.87	-0.03	0.04	-0.69	-0.03	0.04	-0.73	-0.09	0.13	-0.71				
Word <i>n</i> + 1	Preview	-0.01	0.03	-0.28	0.02	0.04	0.46	-0.04	0.03	-1.13	-0.02	0.03	-0.45	0.20	0.13	1.47				
Word <i>n</i> + 2	Preview	-0.02	0.02	-0.77	-0.02	0.04	-0.59	0.02	0.03	0.57	-0.01	0.05	-0.17	0.30	0.20	0.13				
Group	Group	-0.07	0.03	-2.03	-0.12	-0.12	-2.71	-0.14	0.04	-3.38	-	-	-	-	-	-				
Composite region	Preview	-0.02	0.03	-0.98	-0.09	0.07	-1.38	-0.07	0.03	-2.19	-0.06	0.03	-1.94	0.15	0.39	0.38				
Group	Group	-	-	-	-	-	-	-	-	-	-0.15	0.06	-2.55	-	-	-				

Note. Significant terms are presented in bold. FFD = first fixation duration; SFD = single fixation duration; GD = gaze duration; GP = go past time; SP = skipping probability

PARAFOVEAL PROCESSING OF CHINESE IDIOMS

Table 4

Mean (Standard Deviation) for Characteristics of both Idioms and Phrases

	Type	Idioms	Phrases	
Word <i>n</i> +1	Number of strokes	9.32 (3.96)	9.15 (3.03)	
	Word frequency	188 (230)	162 (191)	
Word <i>n</i> +2	Word frequency	6.86 (10.90)	6.86 (10.90)	
	Number of strokes of the first character	7.22 (3.90)	7.22 (3.90)	
	Correct preview	Frequency of the first character	965 (1603)	965 (1603)
		Number of strokes of the second character	8.15 (3.68)	8.15 (3.68)
	Incorrect preview	Frequency of the second character	712 (854)	712 (854)
		Number of strokes of the first character	7.25 (3.88)	7.25 (3.88)
		Frequency of the first character	878 (1353)	878 (1353)
		Number of strokes of the second character	8.12 (3.65)	8.12 (3.65)
	Sentence	Frequency of the second character	687 (830)	687 (830)
		Naturalness	4.75 (0.48)	4.56 (0.55)
Predictability from text prior word <i>n</i> +1		0.000 (0.000)	0.000 (0.000)	
Predictability from text prior word <i>n</i> +2		0.013 (0.051)	0.003 (0.014)	

PARAFOVEAL PROCESSING OF CHINESE IDIOMS

Table 5

Mean (Standard Error) for Fixation Time and Fixation Probability Measure for All Target Regions in Experiment 2

Analysis region	Phrase type	Preview type	Measures					
			FFD	SFD	GD	GP	SP	
Word n	Idiom	Correct	249 (5)	246 (5)	265 (7)	278 (8)	0.41 (0.02)	
		Incorrect	255 (5)	254 (6)	281 (8)	295 (9)	0.42 (0.02)	
	Phrase	Correct	252 (5)	254 (6)	269 (7)	285 (8)	0.42 (0.02)	
		Incorrect	255 (5)	257 (7)	284 (9)	294 (9)	0.43 (0.02)	
	Word $n + 1$	Idiom	Correct	294 (7)	301 (9)	301 (7)	417 (9)	0.49 (0.02)
			Incorrect	294 (6)	295 (8)	306 (8)	408 (8)	0.52 (0.02)
Phrase		Correct	270 (6)	272 (8)	282 (8)	394 (10)	0.54 (0.02)	
		Incorrect	306 (8)	308 (11)	322 (9)	426 (10)	0.51 (0.02)	
Idiom		Correct	276 (5)	276 (7)	330 (9)	442 (9)	0.18 (0.02)	
		Incorrect	289 (5)	299 (8)	358 (9)	458 (9)	0.16 (0.02)	
Word $n + 2$	Phrase	Correct	292 (5)	289 (8)	364 (11)	452 (10)	0.16 (0.02)	
		Incorrect	308 (6)	309 (10)	384 (11)	487 (10)	0.13 (0.02)	
	Idiom	Correct	284 (5)	284 (9)	435 (11)	500 (13)	0.05 (0.01)	
		Incorrect	302 (5)	333 (12)	484 (12)	528 (12)	0.04 (0.01)	
	Composite region	Correct	283 (4)	294 (10)	478 (15)	527 (16)	0.04 (0.01)	
		Incorrect	310 (5)	318 (15)	537 (15)	590 (16)	0.03 (0.01)	

Note. FFD = first fixation duration; SFD = single fixation duration; GD = gaze

duration; GP = go past time; SP = skipping probability

PARAFOVEAL PROCESSING OF CHINESE IDIOMS

Table 6

Fixed Effects Estimates from the LME Models for All Measures Across All Target Regions of Experiment 2

Region	Effect	Measure														
		FFD			SFD			GD			GP			SP		
		<i>b</i>	<i>SE</i>	<i>t</i>	<i>b</i>	<i>SE</i>	<i>t</i>	<i>b</i>	<i>SE</i>	<i>t</i>	<i>b</i>	<i>SE</i>	<i>t</i>	<i>b</i>	<i>SE</i>	<i>z</i>
Word <i>n</i>	Phrase	-0.01	0.02	-0.35	-0.00	0.02	-0.11	-0.00	0.02	0.02	0.00	0.02	0.08	0.03	0.10	0.31
	Preview	-0.01	0.02	-0.40	-0.02	0.02	-0.71	-0.03	0.02	-1.39	-0.02	0.02	-1.03	-0.10	0.10	-0.94
	Phrase*Preview	0.01	0.03	0.46	0.05	0.04	1.19	0.02	0.04	0.56	0.04	0.04	1.05	-0.01	0.20	-0.06
Word <i>n</i> + 1	Group	-	-	-	-	-	-	-	-	-	-	-	-	-0.04	0.21	-0.17
	Preview*Group	-	-	-	-	-	-	-	-	-	-	-	-	-0.65	0.21	-3.12
	Phrase	-0.03	0.02	-1.27	-0.04	0.03	-1.34	-0.02	0.03	-0.89	-0.03	0.03	-1.18	0.06	0.09	0.69
Word <i>n</i> + 2	Preview	0.06	0.02	-2.55	-0.05	0.03	-1.55	-0.06	0.02	-2.50	-0.04	0.02	-1.57	-0.01	0.09	-0.08
	Phrase*Preview	-0.09	0.04	-2.06	-0.08	0.06	-1.39	-0.09	0.05	-1.90	-0.09	0.05	-1.92	0.27	0.19	1.44
	Group	-0.04	0.04	-1.20	-0.01	0.04	-0.34	-0.05	0.04	-1.23	-	-	-	-	-	-
Composite region	Preview*Group	-0.09	0.04	-2.06	-0.16	0.06	-2.75	-0.11	0.05	-2.33	-	-	-	-	-	-
	Phrase	-0.05	0.02	2.8	0.02	0.03	0.81	0.05	0.02	2.21	0.02	0.03	0.70	-0.25	0.13	-1.97
	Preview	-0.04	0.02	-2.3	-0.06	0.03	-2.45	-0.07	0.02	-2.88	-0.05	0.03	-1.90	0.14	0.13	1.12
Word <i>n</i> + 2	Phrase*Preview	0.01	0.04	0.2	0.02	0.05	0.40	0.03	0.04	0.80	-0.05	0.05	-1.10	0.20	0.25	0.77
	Group	-	-	-	-	-	-	0.10	0.04	-2.53	-	-	-	-	-	-
	Phrase	0.01	0.02	0.50	-0.00	0.04	-0.06	0.06	0.02	2.37	0.04	0.03	1.56	-0.25	0.24	-1.03
Word <i>n</i> + 2	Preview	-0.07	0.02	-4.30	-0.12	0.04	-2.76	-0.12	0.03	-4.69	-0.11	0.02	-4.54	0.15	0.24	0.62
	Phrase*Preview	-0.01	0.03	-0.40	0.08	0.07	1.19	-0.01	0.04	-0.04	-0.05	0.04	-1.18	0.33	0.48	0.68

Note: Significant terms are presented in bold. FFD = first fixation duration; SFD = single fixation duration; GD = gaze duration; GP = go past time; SP = skipping probability

1
2
3
4
5
6
7 钟伯伯拿刀切**西瓜**给客人吃。(Correct preview with a low-frequency word $n+1$)
8
9 钟伯伯拿刀切**手铐**给客人吃。(Incorrect preview with a low-frequency word $n+1$)
10
11 English translation: Uncle Zhong uses a knife to cut **watermelon** for the guests.
12
13 女主人煎了**鲫鱼**给客人吃。(Correct preview with a high-frequency word $n+1$)
14
15 女主人煎了**枪河**给客人吃。(Incorrect preview with a high-frequency word $n+1$)
16
17 English translation: The hostess fried a **carp** for the guests.
18
19

20 Figure 1. An Example of Stimuli in Experiment 1. The two characters in bold comprise word $n+2$. They were
21 not in bold and in same font as the other characters in the experiments. The character after the boundary
22 “|” and prior to word $n+2$ is word $n+1$.

23 114x47mm (96 x 96 DPI)