Manuscript

Eye-Movement Evidence for the Mental Representation of Strokes in Chinese Characters

Lili Yu[1], Jianping Xiong[2], Qiaoming Zhang[3], Denis Drieghe[4], & Erik D. Reichle [5]

1. Department of Cognitive Science, Macquarie University, Australia
2. Institute of Education, Henan Normal University, China
3. School of Educational Science, Ludong University, China
4. Department of Psychology, University of Southampton, UK
5. Department of Psychology, Macquarie University, Australia

Working draft; please do not cite without permission.

Address correspondence to: Lili Yu, Department of Cognitive Science, Macquarie University, NSW 2109, Australia; or via e-mail to: lili.yu@mq.edu.au. This work was supported a National Institute of Health grant RO1HD075800 awarded to the last author.

Word count = 3,861

Although strokes are the smallest identifiable units in Chinese words, the fact that they are often embedded within larger units (i.e., radicals and/or characters that comprise Chinese words) raises questions about *how* and even *if* strokes are separately represented in lexical memory. The present experiment examined these questions using a gaze-contingent boundary paradigm (Rayner, 1975) to manipulate the parafoveal preview of the first of two-character target words. Relative to a normal preview, the removal of whole strokes was more disruptive (i.e., resulting in longer looking times on targets) than the removal of an equivalent amount of visual information (i.e., number of pixels) from strokes located either in similar locations or throughout the entire character. These findings suggest that strokes are represented as discrete functional units rather than visual features or integral parts of the radicals/characters in which they are embedded. We discuss the theoretical implications of this conclusion for models of Chinese word identification.

The compositional role of individual letters in the mental representation of printed words in English and other alphabetic writing systems is irrefutable, as demonstrated by the fact that *all* existing computational models of word identification make explicit theoretical assumptions about how letters are combined to access other lexical information from memory (e.g., Davis, 2010; Gomez, Ratcliff, & Perea, 2008; Whitney, 2001). What remains less clear, however, is whether or not, in written Chinese, the individual *strokes* (so named because Chinese was historically written using a brush and ink) that comprise words play a similar role and thus are similarly represented in the lexicon. This question arises because, in contrast to alphabetic writing systems, where words are composed from linear arrays of letters, Chinese words are composed of 1-4 *characters* (see Fig. 1) that are themselves composed of 1-36 strokes. These strokes are often also arranged into clusters called *radicals* that as a whole can convey semantic or phonological information, and it is because of this 2-dimensional, hierarchical structuring of Chinese words that empirical efforts to understand how they are represented in memory have yielded mixed results about both the relative importance of strokes (e.g., as compared to radicals) and how they might be represented in the lexicon (for empirical and theoretical reviews, see respectively Yu & Reichle, 2017, and Reichle & Yu, 2017).

Evidence for the mental representation of strokes comes from an eye-movement experiment reported by Yan et al. (2012; see also Tseng, Chang, & Wang, 1965). In this experiment, 15%, 30%, or 50% of the strokes were removed from target characters, with the removed strokes being those that: (1) are normally written early when a character is written; (2) are normally written later in a character; and (3) do not contribute to the overall “envelop” or shape of a character. The key finding was that, with 30% or more of the strokes being removed, the removal of the early strokes was more disruptive than the removal of the later strokes, with the removal of shape-preserving strokes being the least disruptive. (Flores d'Arcais, 1994, observed a similar processing advantage for early compared to later strokes using both character naming and character-identity judgements.) This suggests that strokes are represented in the lexicon, but perhaps to varying degrees, with earlier-written strokes being more important than internal or later-written strokes—similar to what has been shown with English, where the initial letters of a word are more important than internal or ending letters (Rayner, White, Johnson & Liversedge, 2006; White, Johnson, Liversedge, & Rayner, 2008).

Other studies have also examined the possible differential weighting of information in the representation of Chinese characters by removing fragments from characters rather than individual strokes (e.g., removing " " or " " from the character "象"; Liu, 1983; Peng, 1982; Tsao & Wang, 1983). These studies collectively show that removing fragments from the left and/or upper parts of characters is more disruptive to their identification than removing fragments from the right and/or lower parts. Expanding upon this finding using an eye-movement study, Wang et al. (2013) used a visual-redundancy metric to remove character segments of varying levels of informativeness. Perhaps not surprisingly, reading was more disrupted by the removal of informative (non-redundant) segments, and these informative segments tended to be located on the left side of the characters. Possible explanations for the latter finding include the fact that, when reading from left to right, the left sides of characters are closer to the high-acuity center of vision and subject to less lateral interference from proximal strokes, and/or the fact that the meanings of characters are often conveyed by semantic radicals that tend to be located more often on the left sides of characters.

These possible interpretations are further complicated by Peng's (1982) analyses of the *composability* or number of characters that are consist with a given character fragment. These analyses were based on 3,000 characters and indicated that stroke-fragment information in the upper left quadrants of characters tended to have great composability (i.e., are consistent with more possible characters) than other strokes, underscoring the fact that any location-related differences in stroke processing may reflect one or more of several variables, including differences related to visual acuity, lateral interference, the order in which the strokes are normally written, the locations of semantic radicals, and possible differences in visual redundancy and/or composability of the strokes themselves.

Given this unresolved complexity, the main objective of the current experiment is not to disentangle the different accounts of stroke processing, but is instead more modest—to demonstrate the psychological reality of strokes as functional units in the identification of Chinese characters. In other words, the current experiment is intended to show that strokes are processed and represented as discrete, separable units rather than as sets of visual features or integral components of the radicals or characters in which they are embedded. To do this, we used a gaze-contingent *boundary paradigm* (Rayner, 1975) in which some type of preview of a target word is immediately replaced by the target as reader's eyes move across an invisible "boundary" located to the left of the target. In our variant of this paradigm, we manipulated the preview of the first character in 2-character target words (see Fig. 1). As Figure 2 shows, the preview was: (1) the actual target character, providing a baseline against which to measure the possible effects of the other three previews; (2) the target character with some number of the earliest-written strokes removed; (3) the target character with some proportion of pixels removed from strokes in the left side or upper-left quadrant of the character; or (4) the target character with the same proportion of pixels removed, but from across all of the strokes in the character. These previews were intended to dissociate the disruptive effects of removing strokes per se from the effects of removing visual information from informative parts of the character, and from visually degrading the character as a whole.

Figures 1 & 2

Method

Participants. 68 undergraduates with normal or corrected-to-normal vision from Henan Normal University took part in the experiment. (Data from an additional 11 participants were collected but excluded from our analyses because they noticed the display changes four or more times.)

Materials and Design. 68 single-radical target characters were selected to avoid any potential difficulty with interpreting our results. As illustrated in Figure 1, there were four preview conditions: (1) *identical preview*; (2) *stroke-removal preview*, wherein early-written strokes were removed from a character; (3) *fragment-removal preview*, wherein pixels from the left side or upper-left quadrant were removed; and (4) *segment-removal preview*, wherein the initial segment of each stroke in a character were removed. The overall proportion of pixels removed across conditions 2-4 was ~30%, to equate visual density across the conditions and to maximize the disruption caused by the manipulation (as per Yan et al., 2012). To render the stroke- and fragment-removal conditions maximally dissimilar, the strokes that were removed in the latter condition were selected from the left side or upper-left quadrant of the character, subject to the constraint that they were not the same (early-written) strokes removed from the former condition. (This was verified using the *RadicalLocator* software[1]; see Yu, Reichle, Jones, & Liversedge, 2015). 16 extra participants who did not participate in the eye-tracking experiment indicated how many characters could be generated from each of the previews in conditions 2-4; on average, the previews are congruent with 1.1 characters and did not differ across conditions [F (2, 203) = 0.30, p = 0.74].

All target characters were the initial characters of 2-character target words, which were embedded near the middle of their sentence frames. Target plausibility and sentence naturalness ratings were collected from 15 subjects who did not participate in the eye-tracking experiment using a 5-point scale; targets were rated as plausible (M = 4.1 out of 5) and sentences as natural (M = 4.0 out of 5). 15 additional subjects completed cloze norms on the target words; targets were unpredictable, only being guessed on average 1.4% of the time.

The preview conditions were counterbalanced across participants and items using a Latin-square design so that each participant encountered all conditions as often but read each sentence only once.

Apparatus. Participants' eye movements were recorded by an SR Research EyeLink-1000 desktop eye-tracker with a sampling rate of 1,000 Hz. A chin rest was used to minimize participants' head movements. The display monitor was a NESO

FS210A CRT monitor with a refresh rate of 120 Hz and a screen resolution of 1,024 × 768 pixels. Sentences were displayed as Song font in black (RGB: 0, 0, 0) on a grey background (RGB: 150, 150, 150), with 36 pixels per character and a 1-pixel gap between characters. The distance between the monitor and participant was approximately 75 cm, so that each character occupied ~1° of visual angle. Although viewing was binocular, only the participants' right eyes were tracked.

Procedure. Participants gave informed consent prior to their participation and were instructed to read sentences for comprehension. A three-point horizontal calibration was used at the beginning of the experiment to ensure eye-tracking accuracy, with a maximum acceptable error of 0.4° (*Mean* < 0.2°). Prior to each trial, a drift-calibration dot appearing at the location of first character in the sentences, which was followed by a gaze-contingent trigger that participants needed to fixate for 20 ms to display the sentence. The participant then read the sentence at his/her own pace and pressed the space-key when done.

Each participant read 138 sentences consisting of 10 practice sentences, followed by 68 experimental sentences randomly interspersed with 60 filler sentences. 35% of the sentences were followed by yes/no comprehensive questions. Participants answered 89% of these questions correctly, indicating that they comprehended the sentences.

Results

Linear Mixed Models (*LMMs*) using the *lme4* package (version 1.1-12, Bates et al., 2015) in *R* (R Core Team 2016) were used to analyze the data. A treatment contrast was used in the analysis, with the identical-preview being treated as a baseline and three planned contrasts compared it with each of the preview conditions. One extra contrast between the stroke-removal and fragment-removal previews was made to assess the disruption caused by removing strokes versus left-side/upper-left quadrant pixels. Both participants and items were entered in the models as random effects, with the maximum random effects structure specifying intercepts and slopes for preview effects across participants and items (Barr, Levy, Scheepers, & Tily, 2013). Absolute *t*- and *z*-values equal to or greater than 1.96 indicates significance using $\alpha = 0.05$.

Four regions of interest were included in our analyses: the pre-target word, the target character (i.e., the first character of the target word), the target word, and the post-target word. Five dependent measures were calculated for each region: (1) *first-fixation duration* (*FFD*), or the duration of the initial first-pass fixation in a region; (2) *single-fixation duration* (*SFD*), or the duration of the first-pass fixation in a region that is fixated exactly once; (3) *gaze duration* (GD), or the sum of all first-pass fixations in a region; (4) *go-past time* (*GP*), or the sum of all fixations from the first fixation into a region until the eyes exit the region to the right (i.e., this measure includes any regressive fixations that exit a region to the left); and (5) *skipping probability* (*PrS*), the probability of a region being skipped.

Prior to completing the analyses, all fixations shorter than 60 ms or longer than 600 ms were removed (4.0% of the data), as were fixations more than three standard deviations above the mean per participant for each measure, resulting in the loss of an additional 0.1%-2.1% of the data across measures. We also removed trials in which participants blinked while fixating a critical region (4.0% of trials), the display change was triggered prematurely (5.6% of trials) or required more than 10 ms to complete

(5.3% of trials), or a saccade "hooked back" to the pre-target word after triggering the display change (3.4% trials). These exclusions left 3,779 trails.

Because the log-transformed data yielded a similar pattern of results as the untransformed data, the latter are reported for transparency. The means and standard deviations are shown in Table 1, and the statistical models for target characters and words are shown in Table 2.

Tables 1 & 2

Pre-target Words. On the pre-target words, first fixations (b = -8.29, SE = 4.00, t = 2.08) and gaze durations (b = -16.66, SE = 6.19, t = 2.69) were shorter with the segment-removal than identical preview, producing a *parafovea-on-fovea* effect in which orthographic properties of a parafoveal word modulated fixations on a foveal word[2].

Target Character. On the target characters, none of the three preview conditions differed from the identical preview for first-fixation or single-fixation durations (all ts < 1.40). However, the stroke-removal preview was more disruptive than the fragment-removal preview, resulting in longer first- (b = -11.46, SE = 5.95, t = 1.93) and single-fixation durations (b = -11.66, SE = 5.86, t = 1.99). This finding is consistent with the hypothesis that strokes are important functional units, and that their removal is more problematic for lexical processing than the removal of comparable amounts of visual information from strokes in similar locations (i.e., in the left side and/or upper-left quadrant of characters). Also, consistent with this hypothesis is that both gaze durations (marginally significant, b = 12.08, SE = 6.31, t = 1.91), and go-past times (b = 49.67, SE

= 17.36, t = 2.86) were longer for the stroke-removal than identical preview, with neither the fragment- nor segment-removal conditions exhibiting this pattern (all ts < 1.56). However, the observed differences between stroke-removal and fragment-removal previews for early measures were absent for both gaze durations and go-past times. Finally, the skipping rates for target characters were equally high across the preview conditions (58% on average; all ts < 1.76).

Target Word. The fixation-duration measures exhibited a similar pattern on the target words as on the target characters, with marginally inflated first and single-fixations durations in the stroke-removal relative to fragment-removal previews (FFD: b = -7.62, SE = 4.38, t = 1.74; SFD: b = -8.73, SE = 4.57, t = 1.91), and longer go-past times in the stroke-removal relative to identical previews (b = 26.53, SE = 11.91, t = 2.23). However, the skipping rate was higher for identical previews than for stroke-removal previews (b = -0.38, SE = 0.13, t = 3.05), segment-removal previews (b = -0.40, SE = 0.13, t = 3.22), and marginally for fragment-removal previews (b = -0.22, SE = 0.12, t = 1.81)

Post-Target Word. No effects were observed on the post-target words expect an elevated skipping rate following the segment-removal than identical preview (b = 0.22, SE = 0.11, t = 2.01). We suspect that this finding may reflect the slight reduction of target-character skipping in this condition—that it may have afforded more preview and thus more skipping of the post-target word.

Discussion

In the present experiment, a boundary paradigm was used to examine the processing costs associated with removing individual strokes from characters, fragments of strokes from specific within-character locations, and segments of all within-character strokes. The first hypothesis addressed by this experiment concerns the basic representation of strokes in lexical memory: If strokes are represented in the lexicon, then the processing costs associated with their removal from preview should be greater than the processing costs associated with the removal of the same amount of information from fragments in similar locations. Consistent with the hypothesis, the stroke-removal preview caused more disruption to target characters/words processing than did the fragment-removal preview, as indexed by both longer first- and single-fixations durations in the former than latter condition, and by longer gaze duration and go-past times for the stroke-removal preview compared to the identical preview condition.

Prior analyses (see Wang et al., 2013; Yan et al., 2012) have suggested that the contour of the incomplete characters is highly correlated to how readily they can be recognized because characters' missing informative segments also have the lowest proportion of overlapping perimeters and vertices to their original characters. Therefore, to ensure that the observed disruption in the stroke-removal condition was actually due to the cost associated with the missing strokes per se, and not differences in the contours of the stroke-removal versus fragment-removal previews, we compared the convex perimeters of the stroke-removal versus fragment-removal previews; a paired *t*-test indicated there was no difference between the two conditions [$t\ (134) = 1.19$, $p = 0.23$].

We also examined whether the removal of information from two character locations (i.e., left sides vs. upper-left quadrants) in the fragment-removal condition differentially affected our findings. Although the amount of data from the target characters is insufficient to perform separate analyses for each location (for fixation-duration measures, left side $N = 564$-591, upper-left quadrant $N = 530$-561), such analyses were possible (see Table 3) using the target-word data (left side N = 910-1,140, upper-left quadrant N = 859-1,116). Contrasts revealed, with fragments removed from the upper-left quadrants of characters, removing strokes was more disruptive for both first-fixation (marginally, b = -10.90, SE = 5.87, t = 1.86) and single-fixation durations (b = -16.88, SE = 6.99, t = 2.41); when fragments being removed from left sides of characters, removing strokes resulting in longer go-past times (b = -39.95, SE = 19.92, t = 2.01). Thus, although removing fragments from different parts of characters may have affected different processing stages (i.e., removing left/upper-left fragments affected early/late processing), the removal of strokes was more disruptive to overall processing than the removal of either type of fragment.

Table 3

Finally, it is worth noting that the current findings have theoretical implications for existing and future accounts of Chinese word identification (see Reichle & Yu, 2017). Foremost, our results are more consistent with models that postulate the representation of strokes as discrete functional units (e.g., Li, Rayner, & Cave, 2009) than models that do not (e.g., Perfetti, Liu, & Tan, 2005). Similarly, our results show that models require assumptions about the strokes being weighted differentially during processing, although whether this is due to visual-acuity limits (e.g., Li et al., 2009), the serial allocation of attention (e.g., Taft, Zhu, & Peng, 1999), and/or other factors (e.g., the order in which strokes are normally written) remains unclear. Finally, considerable effort has been directed towards explaining how, in alphabetic language, the order of letters is encoded and represented in the lexicon (e.g., Davis, 2010). We suspect that written Chinese will provide an informative arena for evaluating these competing hypotheses because of the inherent complexity of Chinese words (e.g., the fact that strokes are arranged along two spatial dimensions as compared to one for letters in alphabetic writing systems).

References

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

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1–48. http://dx.doi.org/10.18637/jss.v067.i01

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

Flores d'Arcais, G. B. (1994). Order of strokes writing as a cue for retrieval in reading Chinese characters. *European Journal of Cognitive Psychology, 6,* 337–355. http://dx.doi.org/10.1080/09541449408406519

Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: a model of letter position coding. *Psychological review, 115*, 577–601. http://dx.doi.org/10.1037/a0012667

Hyönä, J. (1995). Do irregular letter combinations attract readers' attention? Evidence from fixation locations in words. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 1365–1373. http://dx.doi.org/10.1037/0096-1523.21.1.68

Li, X., Rayner, K., & Cave, K. R. (2009). On the segmentation of Chinese words during reading. *Cognitive Psychology*, *58*, 525-552. https://doi.org/10.1016/j.cogpsych.2009.02.003

Liu, I. M. (1983). Cueing function of fragments of Chinese characters in reading. *Acta Psychologica Taiwainica, 25*, 85–90. (in Chinese)

Peng, R. X. (1982). A preliminary report on statistical analysis of the structure of Chinese characters. *Acta Psychologica Sinica, 14*, 385–390. (in Chinese)

Perfetti, C. A., Liu, Y., & Tan, L. H. (2005). The Lexical Constituency model: Some implications of research on Chinese for general theories of reading. *Psychological Review*, *112*, 43-59. http://dx.doi.org/10.1037/0033-295X.112.1.43

Plummer, P., & Rayner, K. (2012). Effects of parafoveal word length and orthographic features on initial fixation landing positions in reading. *Attention, Perception, & Psychophysics, 74*, 950–963. https://doi.org/10.3758/s13414-012-0286-z

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7,* 65–81. https://doi.org/10.1016/0010-0285(75)90005-5

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

Reichle, E. D., & Yu, L. (2017). Models of Chinese Reading: Review and Analysis. *Cognitive Science*. Manuscript in Press.

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

Tsao, Y. C., & Wang, T. G. (1983). Information distribution in Chinese characters. *Visible Language, 17*, 357–364.

Tseng, S. C., Chang, L. H., & Wang, C. C. (1965). An informational analysis of the Chinese language: I. The reconstruction of the removed strokes of the ideograms in printed sentence-texts. *Acta Psychologica Sinica, 10*, 299–306. (in Chinese)

Wang, H. C., Schotter, E. R., Angele, B., Yang, J., Simovici, D., Pomplun, M., & Rayner, K. (2013). Using singular value decomposition to investigate degraded Chinese character recognition: evidence from eye movements during reading. *Journal of research in reading, 36,* 35–50. http://dx.doi.org/10.1111/j.1467-9817.2013.01558.x

White, S. J., Johnson, R. L., Liversedge, S. P., & Rayner, K. (2008). Eye movements when reading transposed text: The importance of word beginning letters. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 1261–1276. http://dx.doi.org/10.1037/0096-1523.34.5.1261

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

Yan, G., Bai, X., Zang, C., Bian, Q., Cui, L., Qi, W., Rayner, K., & Liversedge, S. P. (2012). Using stroke removal to investigate Chinese character identification during reading: Evidence from eye movements. *Reading and Writing, 25*, 951–979. https://doi.org/10.1007/s11145-011-9295-x

Yu, L., & Reichle, E. D. (2017). Chinese versus English: Insights on Cognition during Reading. *Trends in Cognitive Sciences, 21*, 721–724. https://doi.org/10.1016/j.tics.2017.06.004

Yu, L., Reichle, E. D., Jones, M., & Liversedge, S. P. (2015). Radicallocator: A software tool for identifying the radicals in Chinese characters. *Behavior Research Methods, 47,* 826–836. https://doi.org/10.3758/s13428-014-0505-8

Endnotes

1. Radical Locator was used to calculate visual-similarity scores between the stroke-removal previews and each of the two types of fragment-removal previews (i.e., previews generated by removing information from the left side vs. upper-left quadrant of target characters). The fragment-removal preview that was most dissimilar to the stroke-removal preview was then selected for use in the experiment.

2. Although not predicted, this result is consistent with the hypothesis that irregular orthographic patterns “pop out”, rapidly drawing the eyes away from the pre-target region (Hyönä, 1995; Plummer & Rayner, 2012), and suggests that the segment-removal condition was slightly anomalous.

Table Captions

Table 1. Mean fixation times and skipping probabilities, as a function of preview types for all regions of interest. (Standard deviations are in parentheses.)

Table 2. LMM fixed-effect estimates for all measures and preview conditions on target characters and words.

Table 3. Mean target word fixation times and skipping probabilities for previews involving the removal of information from the left side vs. upper-left quadrant of characters. (Standard deviations are in parentheses.)

Figure Captions

Figure 1. An example of the boundary paradigm, the experimental sentences, and the regions of interest. Panel A shows the four possible previews of the first character in the target words prior to a reader's eyes crossing an invisible boundary. Panel B shows display change resulting in the target being made visible after a reader's eyes have crossed the boundary. (The English translation of the example sentence is: *His favorite way to relax is to go into the study and look at his stamp collection,* with the underlined word being the two-character target word in the example sentence.)

Figure 2. Examples of the four preview conditions: (1) identical preview; (2) stroke-removal preview; (3) fragment-removal preview; and (4) segment-removal preview. The gray portions of the character are for illustrative purposes and correspond to the pixels that were removed to create each of the previews.

Table 1. Mean fixation times and skipping probabilities, as a function of preview types for all regions of interest. (Standard deviations are in parentheses.)

Region	Dependent Measure	Identical Preview	Stroke-Removal Preview	Fragment-Removal Preview	Segment-Removal Preview
Pre-Target Word	FFD	241 (80)	237 (74)	243 (78)	233 (76)
	SFD	238 (79)	235 (74)	240 (77)	231 (77)
	GD	273 (119)	267 (110)	271 (117)	256 (107)
	GP	309 (175)	306 (175)	305 (174)	295 (167)
	PrS	0.23 (0.42)	0.22 (0.42)	0.22 (0.42)	0.22 (0.41)
Target Character	FFD	244 (82)	254 (87)	242 (82)	247 (81)
	SFD	243 (82)	255 (89)	243 (83)	250 (81)
	GD	249 (91)	264 (96)	256 (98)	258 (91)
	GP	292 (180)	348 (220)	329 (221)	326 (224)
	PrS	0.59 (0.49)	0.59 (0.49)	0.59 (0.49)	0.55 (0.50)
Target Word	FFD	249 (86)	254 (87)	247 (85)	249 (84)
	SFD	250 (87)	256 (89)	247 (83)	251 (84)
	GD	295 (142)	308 (143)	302 (145)	299 (141)
	GP	356 (235)	386 (240)	376 (252)	371 (254)
	PrS	0.23 (0.42)	0.18 (0.38)	0.20 (0.40)	0.18 (0.38)
Post-Target Word	FFD	235 (82)	239 (81)	238 (82)	234 (78)
	SFD	234 (81)	238 (79)	235 (81)	232 (76)
	GD	273 (132)	273 (124)	275 (133)	265 (131)
	GP	353 (276)	344 (315)	369 (308)	365 (331)
	PrS	0.26 (0.44)	0.26 (0.44)	0.26 (0.44)	0.30 (0.46)

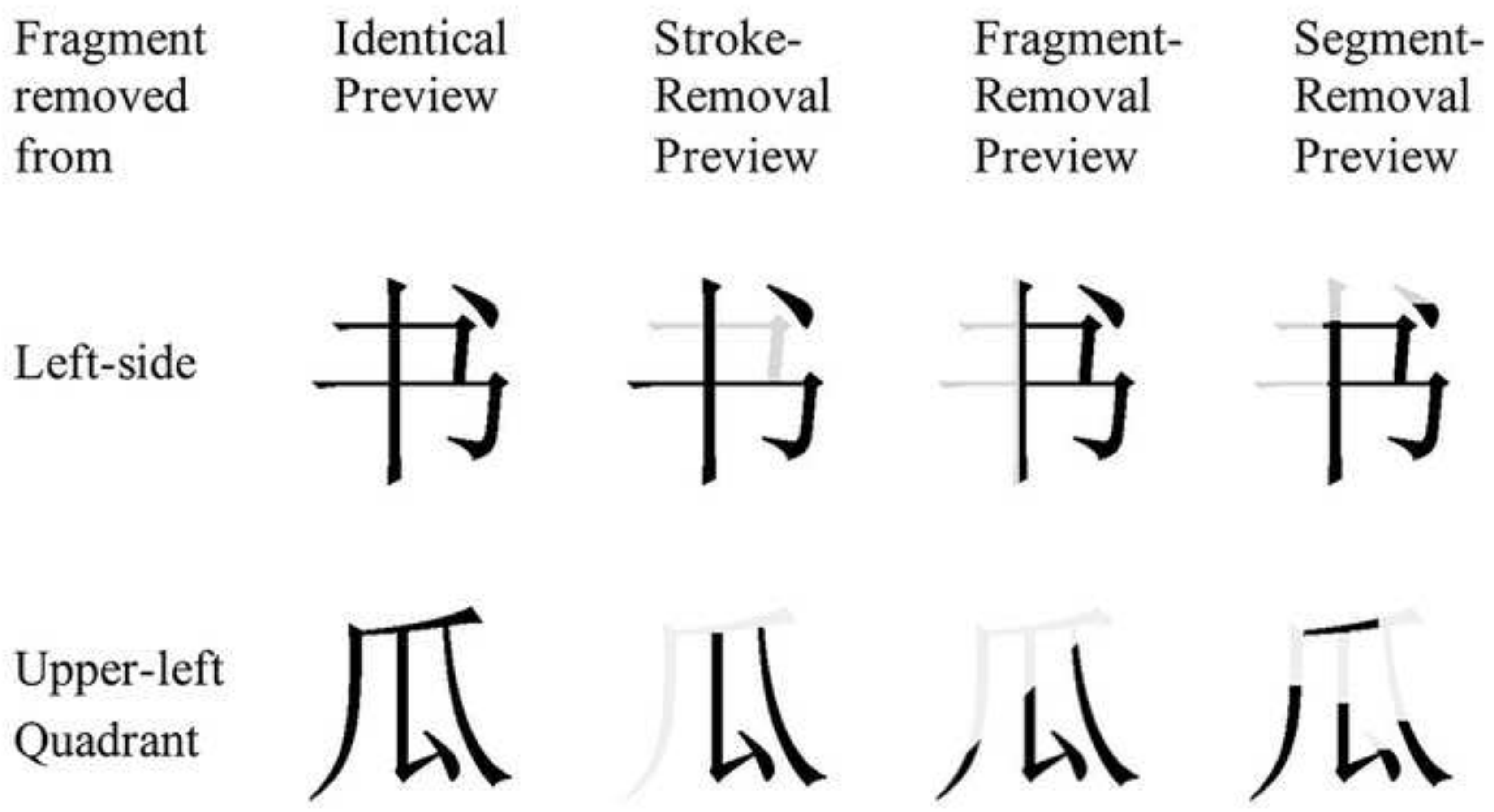
Table 2. LMM fixed-effect estimates for all measures and preview conditions on target characters and words.

Dependent Measure	Contrast	Target Character			Target Word		
		b	*SE*	*t*	*b*	*SE*	*t*
FFD	Intercept	243.22	5.11	47.61	246.97	5.59	44.15
	Stroke-Removal	6.97	6.06	1.15	3.68	4.47	0.82
	Fragment-Removal	-4.34	5.67	-0.77	-3.77	4.63	-0.81
	Segment-Removal	0.13	5.89	0.02	-0.50	4.56	-0.11
	Intercept	243.99	4.77	51.19	246.85	4.82	51.24
	Stroke-Removal vs. Fragment-Removal	**-11.46**	**5.95**	**-1.93**	-7.62	4.38	-1.74
SFD	Intercept	242.56	5.15	47.08	248.58	5.90	42.13
	Stroke-Removal	8.73	6.27	1.39	5.61	5.01	1.12
	Fragment-Removal	-2.68	5.90	-0.45	-2.96	5.21	-0.57
	Segment-Removal	3.43	6.13	0.56	0.51	5.11	0.10
	Intercept	245.10	4.95	49.52	249.45	5.12	48.70
	Stroke-Removal vs. Fragment-Removal	**-11.66**	**5.86**	**-1.99**	**-8.73**	**4.57**	**-1.91**
GD	Intercept	246.67	6.69	36.85	288.99	9.71	29.75
	Stroke-Removal	**12.08**	**6.31**	**1.91**	11.00	8.44	1.30
	Fragment-Removal	3.59	6.28	0.57	4.24	7.30	0.58
	Segment-Removal	6.14	6.19	0.99	3.82	7.85	0.49
	Intercept	252.11	5.44	46.39	293.98	9.16	32.11
	Stroke-Removal vs. Fragment-Removal	-8.88	6.69	-1.33	-6.49	7.72	-0.84
GP	Intercept	292.03	11.32	25.79	353.75	14.32	24.70
	Stroke-Removal	**49.67**	**17.36**	**2.86**	**26.53**	**11.91**	**2.23**
	Fragment-Removal	25.76	16.64	1.55	11.80	12.06	0.98
	Segment-Removal	23.50	17.48	1.35	11.57	13.68	0.85
	Intercept	318.12	11.00	28.92	366.72	13.78	26.62
	Stroke-Removal vs. Fragment-Removal	-22.87	14.73	-1.55	-14.12	11.95	-1.18
PrS	Intercept	0.39	0.11	3.73	-1.53	0.16	-9.40
	Stroke-Removal	0.01	0.10	0.06	**-0.38**	**0.13**	**-3.05**
	Fragment-Removal	0.02	0.10	0.18	-0.22	0.12	-1.81
	Segment-Removal	-0.17	0.10	-1.75	**-0.40**	**0.13**	**-3.22**
	Intercept	0.36	0.09	4.14	-1.78	0.15	-12.13
	Stroke-Removal vs. Fragment-Removal	0.02	0.12	0.19	0.24	0.16	1.54

Table 3. Mean target word fixation times and skipping probabilities for previews involving the removal of information from the left side vs. upper-left quadrant of characters. (Standard deviations are in parentheses.)

Dependent Measure	Fragment-Removal Position	#Trails in Analysis	Identical Preview	Stroke-Removal Preview	Fragment-Removal Preview	Segment-Removal Preview
FFD	Left	1,140	247 (87)	253 (86)	252 (92)	250 (87)
	Upper-Left	1,116	252 (86)	255 (88)	242 (78)	249 (81)
SFD	Left	910	246 (88)	250 (87)	251 (89)	249 (87)
	Upper-Left	859	254 (86)	262 (91)	243 (77)	252 (81)
GD	Left	1,124	288 (138)	308 (156)	300 (144)	302 (149)
	Upper-Left	1,109	304 (146)	308 (130)	303 (147)	296 (134)
GP	Left	1,116	343 (229)	404 (266)	359 (231)	383 (262)
	Upper-Left	1,094	369 (242)	368 (210)	392 (269)	359 (245)
PrS	Left	1,446	0.24 (0.43)	0.16 (0.37)	0.22 (0.41)	0.20 (0.40)
	Upper-Left	1,396	0.22 (0.42)	0.19 (0.39)	0.17 (0.38)	0.15 (0.36)


(A)
Removal manipulation
他最喜欢的放松方式就是钻进书房研究他收集的邮票。
Identical Preview
Stroke-Removal Preview
Fragment-Removal Preview
Segment-Removal Preview
(B)
他最喜欢的放松方式就是钻进书房研究他收集的邮票。
Pre-target Word
Target Word
Post-target Word


Fragment removed from
Identical Preview
Stroke-Removal Preview
Fragment-Removal Preview
Segment-Removal Preview
Left-side
书
Upper-left Quadrant
瓜