Lexical Processes in the Recognition of Japanese Horizontal and Vertical Compounds

Koji Miwa

Seminar für Sprachwissenschaft

Eberhard Karls Universität Tübingen, Germany

Ton Dijkstra

Donders Institute for Brain, Cognition, and Behaviour

Radboud University Nijmegen

Address for the correspondence:

Koji Miwa

Seminar für Sprachwissenschaft

Wilhelmstraße 19, Raum 3.22, Eberhard Karls Universität Tübingen

72074 Tübingen, Germany

E-mail: kojimiwa4@gmail.com

Phone: +49 7071 29-73952

Abstract

This lexical decision eye-tracking study investigated whether horizontal and vertical readings elicit comparable behavioral patterns and whether reading directions modulate lexical processes. Response times and eye movements were recorded during a lexical decision task with Japanese bimorphemic compound words presented vertically. The data were then analyzed together with those obtained in a horizontal lexical decision experiment of Miwa, Libben, Dijkstra, and Baayen (2014). Linear mixed-effects analyses of response times and eye movements revealed that, although response times and first fixation durations were notably shorter in horizontal reading than vertical reading, the vertical reading elicited fewer fixations. Furthermore, while compounds were recognized largely in comparable ways regardless of reading direction, several lexical processes were found to be reading-direction-dependent. Particularly, processing of the first morpheme was modulated by reading direction in a late time frame, such that a horizontal reading advantage was observed for words with a high frequency first morpheme. All in all, the two reading directions do not only differ quantitatively in processing speed, but also qualitatively in terms of underlying processing mechanisms.

Keywords: visual word recognition, reading direction, morphological processing, Japanese, lexical decision, eye-tracking

Lexical Processes in the Recognition of Japanese Horizontal and Vertical Compounds

Horizontal vs. Vertical Readings

In a Google query as of March 9, 2016, the phrase “reading left to right” resulted in 98,800 document counts while the phrase “reading top to bottom” resulted in mere 8,090 counts. This is no surprise. Reading is generally considered to be done horizontally. In the field of vision research, however, differences between horizontal and vertical readings have been studied since the 19th century. In several experiments conducted by Huey (1898), participants read English texts both silently and aloud in both horizontal and vertical alignments. A processing advantage was witnessed for horizontal reading. Half a century later, Tinker (1955) tested an effect of systematic practice in vertical reading. Although vertical reading was 50% slower than horizontal reading before the practice, this disadvantage was minimized to 21.8% after the practice. More recently, Yu, Park, Gerold, and Legge (2010) confirmed that, using alphabetic letter stimuli, vertical reading is slower than horizontal reading and that the processing cost arises due to a perceptual factor (i.e., a smaller visual-span size in vertical reading) rather than an experiential factor (i.e., oculomotor training).

While alphabetic languages are predominantly written horizontally, except in special contexts (e.g., road signs and book titles), Japanese texts are written both horizontally and vertically, depending on genres. While online texts and technical articles are written horizontally, novels (Figure 1a) and newspapers are written vertically. Moreover, Japanese readers grow up in an environment full of horizontal and vertical words (Figure 1b). Because both horizontal and vertical alignments are ubiquitous in Japanese, this language provides a fair testing ground to investigate potential effects of reading directions. In past studies on the effects of reading direction in Japanese, Osaka (1989) reported that saccade length was 5.8 characters on average in horizontal reading while Osaka and Oda (1991) reported that average saccade length in vertical reading was 5.5 characters. The authors concluded that these figures were comparable. Note, however, that there was no direct statistical comparison. More recently, Igawa, Nakayama, Maeda, and Tabuchi (2006) compared horizontal and vertical text readings in Japanese, using eye-tracking. They reported that vertical reading elicited more fixations, slower saccade velocity, and higher subjective difficulty in reading. Apparently, the accumulated knowledge indicates that horizontal reading is, albeit to varying degrees, always faster and less costly than vertical reading.

Morphological Processes

When readers perceive a compound word, such as *blackboard* consisting of the morphemes *black* and *board*, there are three possible ways to recognize it: the morphological parts are processed before the whole (sublexical view: e.g., Taft & Forster, 1975), the whole is processed before the parts (supralexical view: e.g., Giraudo & Grainger, 2001), or the two processes proceed in parallel (dual-route view: e.g., Baayen, Dijkstra, & Schreuder, 1997). If one attempts to construct a language-general model of morphological processing, however, linguistic variations should be taken into account. In Japanese and Chinese, each orthographic symbol (hereafter *character*) represents a morpheme (e.g., 銃 ‘gun’) and is therefore considered to be a morphogram (Rogers, 2005). Although *morphemes* are typically defined as minimal meaningful units, this is not a language-general definition because morphograms in these languages contain a semantic radical, which often represents a general meaning of the morphograms (e.g., 銃 ‘rifle’, 銅 ‘bronze, and 鐘 ‘bell’ all share the semantic radical 釒 ‘steel’ and belong to the same semantic category), as shown in Figure 1c. An issue is how these hierarchically structured morphological units contribute during recognition of morphographic compound words (e.g., 猟銃 'hunting rifle'). Although evidence has been accumulated that all of these morphological components contribute to compound recognition (Miwa, Libben, Dijkstra, & Baayen, 2014; Taft, Zhu, & Peng, 1999), the order of morphological processing is still under debate. Taft et al. (1999) proposed that compound recognition starts from radicals while Miwa et al. (2014) proposed that character units lead the recognition process.

In morphological research, cross-language orthographic variations have four important consequences for processing. First, because characters represent a morpheme, rather than a phoneme, morphographic compounds are typically shorter in length than alphabetic compounds, making supralexical processing more likely. Second, a morphological boundary becomes perceptually obvious in Japanese, which may motivate bottom-up processing from morphemes. Third, compounds with the same number of morphemes always fit into the same rectangle space, which makes sequential left-to-right processing less obligatory (i.e., the head morpheme can be identified without sequential scanning). Fourth, because all morphograms contain a smaller meaningful unit, morphemes may or may not be the minimal meaningful processing unit (Miwa et al., 2014; Taft et al., 1999 respectively). Orthography is therefore not a negligible factor if one attempts to understand language-general morphological processing.

< Insert Figure 1 around here >

The Present Research

Up until now, studies on horizontal vs. vertical reading and on lexical processes have developed as distinct streams of research. In this study, as shown in Figure 1 (c), we linked these issues and investigated how different the two reading directions are and whether the two reading directions affect lexical processes differentially. If we assume that lexical effects (e.g., compound frequency effects) reflect language-general central processes, the effects should be observed regardless of task variations. The present eye-tracking lexical decision study with Japanese bimorphemic compound words examined the following issues.

First, we studied the longstanding issue in vision research: whether there is an overall processing cost for one alignment over the other. This was investigated with Japanese two-character *kanji* words. Among the three orthographic scripts used in Japanese (i.e., *kanji*, *hiragana*, *katakana*), *kanji* is used to represent lexical categories and tend to receive more eye fixations (Kajii, Nazir, & Osaka, 2001). Furthermore, in the BCCWJ word list (National Institute for Japanese Language and Linguistics, 2013), we found that 60% of the kanji words consist of two-characters.

Use of Japanese isolated words for testing reading direction effects is new and effective (for use of isolated words in English, see Yu et al., 2010). Given the prevalence of online texts, which are almost always presented horizontally, we expect some processing cost for vertical reading even in a context of isolated word lexical decision. If the processing cost in vertical reading is perceptual in nature (Yu et al., 2010), a main effect of reading direction should be observed at the earliest fixation. This method is also important for the purpose of understanding word recognition in general and eye movements in a lexical decision task in particular; only response times have received researchers' attention for this most popular psycholinguistic task (Libben & Jarema, 2002), although there are usually multiple eye fixations when words are perceived in isolation (Hyönä, Laine, & Niemi, 1995; Kuperman, Schreuder, Bertram, & Baayen, 2009; Miwa et al., 2014).

Second, we studied whether reading directions modulate lexical processes. Depending on the assumption of attentional units, vertical reading may or may not be different from horizontal reading. If attention is directed to small orthographic units, reading direction has a drastic consequence on lexical processes, because their relative order changes. Inhoff, Pollatsek, Posner, and Rayner (1989) suggested that, in English, attention does not shift on a letter-by-letter basis. For Japanese word reading, Miwa et al.'s (2014) character-driven model similarly predicts that attention does not shift on a radical-by-radical basis. They studied eye fixation durations recorded during lexical decision experiments for horizontally presented Japanese compounds and proposed that morphological processing in Japanese is initially guided by characters, rather than radicals or the whole word unit.

However, it is not yet guaranteed that character-driven processing is the rule invariant across all task conditions. Vertical reading may motivate more analytic processing if readers have more experiential or perceptual constraints in this alignment. Bertram and Hyönä's (2003) visual acuity hypothesis holds that longer words are more likely to be processed from parts. Given the visual acuity constraint in vertical reading (Yu et al., 2010), it is possible that morphological components (i.e., radical and characters) contribute more to lexical processing in a vertical alignment. The lateralized word recognition model of Lavidor, Babkoff, and Faust (2001), which claims that words presented in a non-standard visual format are not processed holistically in the left hemisphere, also supports this prediction, because vertical reading is familiarity-wise less standard than horizontal reading even in Japanese.

Alternatively, vertical reading might motivate holistic processing if vertically presented words are perceived in a qualitatively different manner due to a lack of horizontal eye movements. Furthermore, if expertise in morphographic word recognition is marked by left-side bias, as proposed by Hsiao and Cottrell (2009), vertical alignment may facilitate yet another kind of holistic processing, because a left-side-bias motivates attention to a bit of information from the first and second characters simultaneously (e.g., 犭 'animal' and 釒 'steel' in the example shown in Figure 1). Such qualitative differences between the two alignments were already speculated to exist by Huey (1898): "it is probable that the subject may have a greatest possible speed by one method and a greatest normal speed or speed of comfortable and intelligent reading by quite another" (p. 579).

Using Eye-Tracking to Study Morphological Processing

In this study, we investigated this horizontal vs. vertical issue from a psycholinguistic perspective in an experiment combining lexical decision and eye-tracking. Because participants often make multiple fixations in an isolated word recognition task, this technique allows us to tap into the time course of word recognition (Hyöna et al., 1995; Kuperman et al., 2009; Miwa et al., 2014). The following sections describe the lexical predictors considered in this study, and our predictions regarding their contribution.

Variables Considered in This Study

Table 1 summarizes descriptive statistics for the linguistic and non-linguistic predictors considered in this study.

WholeWordFrequency is a log-transformed compound frequency count collected from the Balanced Corpus of Contemporary Written Japanese (BCCWJ, Maekawa et al., 2014). Because this balanced corpus includes lexical statistics in 11 different genres (e.g., blogs, newspapers, and novels), *WholeWordFrequency* is relatively reading-direction-independent. Previous eye-tracking lexical decision studies (Kuperman et al., 2009; Miwa et al., 2014) reported a small yet significant whole word frequency effect already at the first fixation, not to mention its greater contribution at a later fixation. However, it is not clear whether the effect is perceptual or lexical in nature. If the word frequency effect is largely due to perceptual familiarity, the effect may be modulated by reading directions, with a larger magnitude of the effect for vertical reading. This logic is based on previous reports that the word frequency effect is larger for non-native speakers than for native speakers (Gollan, Montoya, Cera, & Sandoval, 2008).

To gauge lexical activation at the character and radical levels, character token frequencies (*First/SecondCharacterFrequency*) were obtained from the BCCWJ corpus, and radical type frequencies (*First/SecondCharacterRadicalCombinability*), indicating how many characters share a given radical, were obtained from Tamaoka and Makioka’s (2004) lexical database. Morphographic word recognition speed is known to be co-determined by character frequency (Miwa et al., 2014; Tamaoka & Hatsuzuka, 1998) and radical combinability (Feldman & Siok, 1997; Miwa, Libben, & Baayen, 2012; Taft & Zhu, 1997).

Visual complexity of words was also taken into consideration by means of character stroke counts (e.g., 年 ‘year’ is written with 6 strokes). It has been reported that visual complexity gauged by stroke counts co-determines word recognition processes (Tamaoka & Kiyama, 2013; Tamaoka & Takahashi, 1999). We considered stroke counts for the left and right characters separately (*First/SecondCharacterStrokes*), because visual complexities of individual characters are known to contribute in an asymmetrical manner, depending on the location of the eye. Miwa et al. (2014) reported that, although stroke count effects are inhibitory in a fixated region, they are facilitatory in a to-be-fixated region. A question yet to be clarified is whether the same decoding mechanism operates also in vertical reading.

< Insert Table 1 around here >

A Vertical Lexical Decision with Eye-Tracking Experiment

We conducted a lexical decision experiment in which Japanese compound words were presented vertically. The obtained data were then combined with those obtained in a horizontal reading experiment of Miwa et al. (2014). The two levels of *Horizontal* and

Vertical of the factor *ReadingDirection* denoted the former and the latter data sets respectively.

Method

Participants. Twenty native Japanese readers (16 females, median age = 22 years old, *interquartile range* = 3.75) participated at the University of Alberta, Canada. At the time of the experiment, the majority of the participants had stayed in Canada for less than half a year (*median* = 4.5 months, *IQR* = 14 months). For the one-month period preceding the experiment, only a few participants had been exposed to vertical texts more than horizontal texts (*median* = 10% of the time, *IQR* = 22.5%). These statistics provides an idea about how prevalent horizontal and vertical texts are, although it is difficult to estimate how often texts are presented in each direction, because reading direction is largely genre-dependent.

In the analyses that follow, the data from the above 20 participants, who read vertically presented words, were compared to those of the 21 participants tested in Miwa et al. (2014), who read horizontally presented words. There was no significant difference between the two groups of participants in terms of their age and length of stay in Canada. To further safeguard against potential effects of individual differences on lexical processes, the log-transformed participants' length of stay in Canada (*LengthOfStayCanada*) was included in the regression analyses. Because *LengthOfStayCanada* correlated with log-transformed age of the participants ($r = .59$, $p < .0001$), only the former was considered as a measure of individual differences. Idiosyncratic individual variations were accounted for by means of random-intercepts for participants and by-participant random slopes in mixed-effects models.

Materials. The 200 Japanese test words from Experiment 2 of the Miwa et al. (2014) study were used. These are two-character words sampled pseudo-randomly from a large lexical database of Amano and Kondo (2003) so that the words covered the entire frequency range (see Miwa et al., 2014, for more details on the stimulus selection procedure). This sampling procedure minimized researchers' degree of freedom and avoided unnatural sample distributions (see Forster, 2000, for experimenter bias associated with construction of items). As nonword stimuli, the 200 nonwords of Miwa et al. (2014) were used here as well. These were created by changing either the left or right character of an existing word to a non-character. This was done by replacing a sub-character component with another existing component.

Apparatus. The lexical decision experiment was designed and controlled by Experiment Builder (SR Research). During the experiment, movements of participants' right eye were recorded by an Eyelink II head-mounted eye-tracker (SR Research) with a sampling rate of 250 Hz in the pupil-only mode.

Procedure. In each trial, participants were asked to make a two-choice decision regarding whether a presented word was an existing Japanese word or a nonword, by pressing an appropriate button on a Microsoft Sidewinder game pad. A fixation point was always presented first at the location corresponding to the centre of the first character. It also served as a drift correction point, i.e., a target word was not presented until participants fixated on the fixation point. Words were presented in 40-size white font on a black background; the visual angle for each character was approximately 1.64º. This procedure was identical to that in Experiment 2 of Miwa et al. (2014) study, except that words were presented vertically, as shown in Figure 1 (c).

Results and Discussion

Similarities and differences between horizontal and vertical readings. Throughout this paper, R version 3.2.2 (R Development Core Team, 2015) was used for statistical analyses. We initially observed similarities and differences between horizontal and vertical reading in general. The panels (a) to (e) in Figure 2 show item-wise correlation between horizontal and vertical readings for response time, first fixation duration, second fixation duration, fixation count, and average saccade velocity respectively (with the `ggplot2` package, Wickham, 2009). The black dots represent individual words, and the solid lines are a linear smooth with a 95% confidence interval. The dotted lines show hypothetical patterns in which horizontal and vertical readings show identical response and eye movement behaviours.

Response time and eye movement measures were transformed based on visual inspection of Q-Q normality plots and a Box-Cox transformation with the `MASS` package (Box & Cox, 1964; Venables & Ripley, 2002). An inverse transformation was applied to RTs (-1000/RT); an inverse square-root transformation was applied to first fixation durations (-100/√first fixation duration); a square-root transformation was applied to second fixations (√second fixation duration). The plotting was done for correctly responded trials after data trimming (see below for data trimming procedures).

Horizontal and vertical readings can be studied in terms of similarities and differences. First, it is possible to state that horizontal and vertical lexical decision tasks are similar. In all panels, except Panel (e), horizontal and vertical readings show a significant correlation, implying that different reading directions do not alter the underlying cognitive processes drastically and that certain mechanisms are shared between them. Second, some differences between the two reading directions can also be observed. With respect to response times (Panel a), the most data points are located above the hypothetical dotted line, meaning that responses were slower in vertical reading. The observed pattern (the solid line) deviates from the expected line particularly for words that elicited faster responses. The same pattern can be seen in Panel (b). These processing costs associated with vertical reading were supported by significant effects of *ReadingDirection* (reference level = *Horizontal*) as the sole fixed-effect in mixed-effects models: $\beta = 0.127$, $SE = 0.06$, $p = .042$ in a RT analysis, and $\beta = 3.06$, $SE = 0.76$, $p < .001$ in a first fixation duration analysis. In Panel (d), in contrast, more data points are located below the dotted line, indicating that horizontal reading elicited more fixations. A mixed-effects analysis again supported this visual inspection ($\beta = -0.08$, $SE = 0.03$, $p = .012$ for the fixed effect factor *ReadingDirection* in a trial count analysis). This result is not in line with Igawa et al's (2006) report that vertical text reading elicits more eye fixations and implies that the slower recognition speed in the vertical reading is not due to inflated fixation counts. An effect of *ReadingDirection* was not significant in mixed-effects analyses of second fixation durations (Panel c) and saccade velocity (Panel e).

Figure 2 (f) summarizes the approximate time course of horizontal and vertical readings based on median RTs, median first fixation durations, median second fixation durations, and mean fixation counts. The figure shows that vertical reading elicited longer responses and that the cost associated with vertical reading was primarily generated during the first fixation. The observations of Figure 2 and the mixed-effects modeling with *ReadingDirection* as the sole predictor indicate that, even in the context of isolated word recognition, vertical reading entailed significant processing costs. This is in line with previous studies on vertical reading. In the following sections, we reevaluate the effects of *ReadingDirection* with lexical distributional properties as covariates in mixed-effects models and test whether *ReadingDirection* modulates any lexical predictors. Reporting effects of a particular factor both with and without covariates is important, because the two results may not be identical (see Simmons, Nelson, & Simonsohn, 2011, for their suggestions for preventing a false-positive report).

< Insert Figure 2 around here >

Response time analysis. In the sections that follow, (generalized) linear mixed-effects analyses (Baayen, Davidson, & Bates, 2008) were performed using the `lme4` package (Bates, Maechler, & Bolker, 2011) and the `languageR` package (Baayen, 2008, 2011) with *p*-values computed with Satterthwaite approximations to degrees of freedom with the `lmerTest` package (Kuznetsova, Brockhoff, & Christensen, 2015). Because there is an on-going debate with respect to its appropriate setting (Barr, Levy, Scheepers, & Tily, 2013; Bates, Kliegl, Vasishth, & Baayen, 2015; Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2015), the random-effects structures of the models were set only as complex as was justified by a likelihood ratio test using a forward-fitting procedure. Fixed-effects were selected backward, while consulting *p*-values ($p < .05$) and AIC values, from the most complex model with all predictors and interactions between *ReadingDirection* and lexical predictors and *Trial*. The same approach was taken for all mixed-effects analyses reported below unless otherwise noted. Centering was done for all the above numerical predictors by subtracting their mean value from individual values.

The 4,000 data points obtained in this study were combined with the 4,200 data points obtained in Miwa et al., (2014). RTs shorter than 300 ms and longer than 3000 ms were excluded (2 data points). After excluding items that elicited a high error rate (> .40) and incorrect responses, 191 words were kept for analyses (7,139 data points).

A linear mixed-effects model was fitted to the inverse transformed RTs (-1000/RT) initially with all predictors, and potential interactions between *ReadingDirection* and lexical predictors were checked. The random-effect structure of the final model includes intercepts for subjects (SD = 0.18) and items (SD = 0.12), by-subject random slopes for *Trial* (SD = 0.05), *WholeWordFrequency* (SD = 0.01), *SecondCharacterFrequency* (SD = 0.01), *FirstCharacterStrokes* (SD = 0.003), as well as by-item random contrast for *ReadingDirection* (SD = 0.06, indicating greater variance for vertical reading). More complex random-effects structures were not justified by a likelihood ratio test. The standard deviation of the residual error was 0.24, and the R-squared calculated for the correlation between the fitted and observed values was .54.

Table 2 summarizes the fixed-effect structure of the final model. Not surprisingly, there were significant facilitatory contributions of *WholeWordFrequency* and an experimental task effect of *Trial*. Important for the purpose of the present study, there was a main effect of *ReadingDirection*, indicating that vertically presented words were read more slowly than horizontally presented words. *ReadingDirection* also modulated the effect of *FirstCharacterFrequency* (Figure 3a), such that *Horizontal* advantage was more notable for words with high *FirstCharacterFrequency*. The effect of reading direction remained consistent throughout the experiment (i.e., no *ReadingDirection* x *Trial* interaction). In the theoretical framework of Lavidor et al. (2001), if the standard visual format is determined by an individual's experience with a particular alignment, readers with more experience in vertical reading show holistic processing in vertical reading. We considered this possibility by testing the interaction between *WholeWordFrequency* and participants' exposure to vertical reading relative to horizontal reading for the vertical reading data. The prediction was not statistically supported (p = .342).

Although response speed has been a major dependent variable in word recognition research, it captures only the end product of processing. A particular lexical effect might appear and disappear before a response is executed, and the direction of its effect could change from inhibition to facilitation as well. For these reasons, the absence of effects for particular variables in the response time analysis does not imply that these variables are negligible. In order to tap into the time-course lexical processes themselves, the eye movement data were analyzed next.

< Insert Table 2 and Figure 3 around here >

First fixation duration analyses. Overall, the experiment was performed mostly with two fixations on an item (64% and 58% for horizontal and vertical readings respectively). Single fixations and three or more fixations were relatively rare (15.2% for one fixation, 19.1% for three fixations, 3.3% for four fixations, and 0.7% for five fixations). In order to dissect response times to early and late time frames, only correct trials with exactly two fixations were analyzed (5,023 data points). We excluded trials with one fixation, because in these trials the response times are equal to the first fixation durations. In the trials with exactly two fixations, the first fixation durations are perhaps not contaminated by response execution process. The trials excluded in the RT analysis and those with a first fixation duration shorter than 100 ms were excluded here as well.

In the final model of inverse square-root transformed first fixation durations (4,289 data points), the random-effect structure comprises an intercept for subjects (*SD* = 0.23) and items (*SD* = 0.19) and by-subject random slopes for *Trial* (*SD* = 0.05), *FirstCharacterFrequency* (*SD* = 0.03), *FirstCharacterStrokes* (*SD* = 0.02), and by-item random contrasts for *ReadingDirection* (*SD* = 0.08). The standard deviation of the residual error was 0.42, and the R-squared calculated for the correlation between the fitted and observed values was .49.

The fixed-effects are listed in Table 3. There was a main effect of *ReadingDirection*, indicating that advantage of horizontal reading over vertical reading appeared early. *ReadingDirection* also modulated *Trial* (Figure 3b). This indicates that Japanese speakers in this study benefitted from intensive exposure to vertically aligned words, like the English speakers tested by Tinker (1955). Perhaps because participants had less exposure to vertical reading than horizontal reading in their daily life, the former had more room for performance improvement.

The lexical effects at the first fixation were more or less comparable to the early lexical effects reported in Experiment 2 of the Miwa et al. (2014) study, indicating that, at least with respect to first fixation durations, lexical processing is not affected by changing reading direction. Regardless of reading direction, there was an early effect of *WholeWordFrequency* before the eye scanned the entire word (recall that these are trials with two fixations), and character frequency contributed more than radical frequency, with only the former showing a standard facilitatory frequency effect. The asymmetrical contributions of stroke counts imply a foveal feature decoding cost and a parafoveal magnetic attraction (Hyönä & Bertram, 2004), regardless of reading direction.

< Insert Table 3 around here >

Second fixation duration analyses. For the analysis of second fixation durations, the trials excluded in the RT analysis and those with a second fixation duration longer than 800 ms were excluded. In the final model for square-root-transformed second fixations (4,385 data points), the random-effect structure of this model comprises intercepts for subjects (*SD* = 1.16) and items (*SD* = 0.77). By-subject random slopes for *Trial* (*SD* = 0.28), *PreviousRT* (*SD* = 0.79), and *PreviousFixationDuration* (*SD* = 0.96) were also included. The standard deviation of the residual error was 2.37, and the R-squared calculated for the correlation between the fitted and observed values was .54.

The fixed-effects are summarized in Table 4. The main effect of *ReadingDirection* was not significant when it was a sole predictor but significant when the covariates were included. Interestingly, *ReadingDirection* also interacted with *FirstCharacterFrequency*, such that *Horizontal* advantage was greater for words with higher *FirstCharacterFrequency* (Figure 3c). Recall that the interaction involving the same variables was observed in the response time analysis (Figure 3, Panel a). The super-imposed eye-tracking technique successfully identified the locus of the *FirstCharacterFrequency* x *ReadingDirection* interaction to be in a late time frame. The facilitatory contribution of *SecondCharacerFrequency* and the inhibitory contribution of *SecondCharacterStrokes* are in line with the shift of attention from the first character (i.e., the initial fixation point location) to the second character. In this later time frame, *WholeWordFrequency* facilitates processing with the largest magnitude of effect among all the lexical predictors.

< Insert Table 4 around here >

Fixation count analysis. Fixation counts ranged from 1 to 10 fixations (15.2% for 1 fixation, 61.3% for two fixations, 19.1% for three fixations, 3.3% for four fixations, 0.7% for five fixations, and 0.3% for six fixations). For the analysis, trials that were excluded in the RT analysis and those with more than 5 fixations were excluded (13.8% of the data), and 7,068 data points remained.

A generalized linear mixed-effects model with a Poisson distribution was fitted to fixation counts. A final model consisted of the random-effects of participants (*SD* = 0.07) and fixed-effects summarized in Table 5. The R-squared calculated for the correlation between the fitted and observed values was 0.20. It is within our expectation that highly frequent words and later trials elicited fewer fixations. It is, however, contrary to our expectation that horizontal reading elicited more fixations than vertical reading. Because participants tested in the vertical reading task were exposed to horizontal reading 81% of the time and vertical reading 19% of the time on the average, horizontal reading was expected to elicit fewer fixations. Yet, this was not the case. *ReadingDirection* also interacted with *SecondCharacterStrokes*; the magnitude of the *SecondCharacterStrokes* effect was larger in vertical reading (Figure 3d).

< Insert Table 5 around here >

Saccade velocity analysis. Because saccade velocity was reported to be slower in vertical text reading (Igawa et al., 2006), average saccade velocity (visual degrees per second) was analyzed. For saccade velocity analyses, only initial saccades recorded in correctly responded trials were considered (6,228 data points). Because exceedingly large saccades indicate that the eye's fixation location jumped out of the whole word, which was not likely in the present lexical decision experiments, data points with a saccade amplitude larger than 2.5 degrees were excluded. Similarly, because exceptionally short saccades indicate involuntary microsaccades, data points with a saccade amplitude less than 0.2 degrees were excluded. This cut-off point was based on the conservative idea of microssacades being 12 arc min (0.2 degrees) because voluntary miniature saccades may become as short as involuntary microsaccades (Steinman, Haddad, Skavenski, & Wyman, 1973, but see also Martinez-Conde, Otero-Millan, & Macknik, 2014, for a more recently proposed threshold of 1 degree and a review of microsaccade research). After the above trimming procedure, 5,896 data points were kept for analyses. Under the general assumption that there is no information uptake during saccades, which are fast and ballistic (i.e., the destination is pre-determined before a saccade is launched), saccade velocity is expected to reflect relatively early processes that were completed before the end of first fixations.

The random-effect structure of the final model comprises intercepts for participants (*SD* = 7.04) and items (*SD* = 2.01), and by-subject random slopes for *Trial* (*SD* = 1.34), *FirstCharacterFrequency* (SD = 0.85), *FirstCharacterStrokes* (*SD* = 0.28), and *SecondCharacterStrokes* (*SD* = 0.19). By-item random contrasts for *ReadingDirection* was also included (*SD* = 2.62). The standard deviation of the residual error was 9.90, and the R-squared calculated for the correlation between the fitted and observed values was .33.

The fixed effects are summarized in Table 6. Although saccade velocity was slower in the vertical reading than the horizontal reading, as reported by Igawa et al. (2006), the main effect of *ReadingDirection* was not significant. However, its interaction with *WholeWordFrequency* deserves attention (Figure 3, Panel e), because *WholeWordFrequency* is a marker of holistic processing. Apparently, saccades were slower for high frequency compounds recognized in horizontal reading. Although the interpretation of this interaction is not straightforward, the effect of the second character complexity (Panel i) can perhaps be interpreted together. If an upcoming difficulty of processing in a parafoveal region affects foveal processing like a magnet, as suggested by Hyönä and Bertram (2004), a faster saccade velocity for low frequency words with a complex second character would not be unexpected. Saccades were faster also for words with higher *FirstCharacterFrequency* and *SecondCharacterRadicalCombinability*.

< Insert Table 6 around here >

General Discussion

Vertical vs. Horizontal Readings

The present study compared lexical decision performance with Japanese compounds in horizontal and vertical alignments. It was found that, in line with past studies, vertically presented words were read more slowly than horizontally presented words even in the context of an isolated word recognition task. This was largely because first fixations were notably longer in vertical reading, providing support for Yu et al.'s (2010) observation for English words that relative processing delay in vertical reading is due to visual acuity constraint. This, however, does not imply that vertical reading is slower and more costly in all measurements. Interestingly, although eye fixation durations showed a processing advantage for horizontally presented words, fixation counts showed a processing advantage for vertically presented words: Vertical reading elicited a significantly smaller number of fixations. Because more careful processing in the vertical direction reduces the number of fixations, horizontal and vertical readings are considered to differ qualitatively with respect to some underlying psycholinguistic processes. Possibly, this was instigated by a greater motivation in the reader to avoid bottom-up regressive saccades in vertical reading. This interpretation is partially supported by the fact that right-to-left decoding is fully functional (e.g., in Arabic and Hebrew), while bottom-to-top decoding is not encouraged in any existing writing systems. Such variations, however, may be due to historical reasons, anatomical constraints of the eye, or consideration of cognitive efficiency. We leave the empirical comparison of bottom-to-top and right-to-left saccades to future research.

Note that this vertical reading advantage in fixation counts is not in line with the larger fixation counts for vertical reading reported by Igawa et al. (2006). A possible source of this discrepancy lies in the task demands required in the two experiments (i.e., lexical decision in this study and text reading in Igawa et al., 2006). If readers are required to read as fast as possible, the number of bottom-up saccades might be reduced even in text reading.

Reading Directions and Compound Processing

The present study was the first to test whether different text alignments modulate lexical effects. This is important for psycholinguistic research, because models of Japanese word recognition must be formulated in such a way that they can account for different text alignments.

First of all, functional overlaps between horizontal and vertical readings deserve attention, because lexical effects observed to be similar across different reading directions are considered to be *psycholinguistic* in nature. Although an early whole word frequency effect was also reported in the previous lexical decision with eye-tracking studies (Kuperman et al., 2009; Miwa et al., 2014), the present study adds to this finding by demonstrating that the compound frequency effect was not modulated by reading direction in either response time and eye fixation analyses. This result shows that an effect of whole word frequency is, rather than perceptual, largely lexical in nature. Furthermore, in both horizontal and vertical readings, at least in a foveal region, characters, instead of compounds and radicals, were the primary processing units, as proposed by Miwa et al. (2014). This indicates that, with respect to the level of morphological processing, no drastic change (e.g., one route for horizontal reading and another route for vertical reading) is necessary to account for a general pattern of compound processing.

Importantly, however, some lexical effects were modulated by reading direction, indicating that reading direction has consequences for cognitive processes. Bertram and Hyönä’s (2003) visual acuity hypothesis holds that morphological constituents are activated more during the processing of long words due to a limitation in visual acuity. Because visual acuity in a vertical direction is more limited than in a horizontal direction (Yu et al., 2010), it is likely that vertical reading elicits a stronger constituent frequency effect. The first morpheme was indeed processed differently in the two reading directions, as gauged by the *FirstCharacterFrequency* by *ReadingDirection* interaction. For words with a high frequency first morpheme, a horizontal reading advantage (or a vertical reading disadvantage) was observed. If the visual acuity constraint was the source of this interaction, the direction of reading would have modulated the lexical effect during an early timeframe. This was, however, not the case. This *FirstCharacterFrequency* by *ReadingDirection* interaction was observed in RT and second fixation duration analyses, indicating that it is a late effect. One interpretation is that, just as a modifier constituent is initially analyzed as a head in sentential reading (Staub, Rayner, Pollatsek, & Hyönä, 2007), the first character of a word was not recognized as a modifier, or less so, in vertical reading, leading to semantic competition in a late time frame. Importantly, it can be concluded that reading direction has a lasting consequence on the cognitive operations performed later in processing.

Lavidor et al.'s (2001) lateralized word recognition model holds that only words presented in a standard visual format are processed holistically. If the definition of standard visual format is extended to a horizontal alignment, whole compound frequency effects should be modulated by reading direction, most likely during an early time frame. In our study, this was not the case in response time, eye fixation duration, and fixation count analyses. However, in the average saccade velocity analysis, a whole word frequency effect was indeed modulated by reading direction. Therefore, a possibility remains, in line with Lavidor et al.'s (2001) view, that certain holistic perceptual or cognitive processes were modulated by reading direction.

Hsiao and Cottrell (2009) proposed that expert processing is marked by a leftist bias. If the expert-specific left-side bias arises also in reading vertically presented compound words, the left halves of the first and the second character may receive simultaneous attention, triggering more holistic processing in vertical reading. The present results show, however, no sign of greater holistic processing for vertical reading.

Limitations of the Study and Future Directions

Possibilities for future study become clear when we consider the methodological aspects of the present study with respect to statistical analysis and the experimental paradigm.

Multivariate statistical analyses. As Simmons et al. (2011) suggested, reporting a result with and without covariates should be encouraged, because the two results may not be identical. Indeed, in the second fixation analysis of this study, *ReadingDirection* was significant only when covariates were included in the model while the *ReadingDirection* effect observed at the first fixation was always significant with and without covariates. This is not surprising, because non-linguistic effects usually account for a large amount of variance in psycholinguistic experiments (Baayen, 2008). In the response time analysis, for example, when we compared our final model and a model without item-related parameters, the latter linguistically uninformative model could explain as much as 66% of the variance explained by our final model.

With respect to predictors in the mixed-effects models, given the size of the materials, we kept track of effects of only 13 predictors. It is obvious that many other potentially influential predictors can be tested (e.g., familiarity, imageability, neighbourdhood size, semantic transparency, a number of homophones, age of acquisition, vocabulary size). However, consideration of many variables comes with interpretational and analytical challenges associated with multicollinearity (e.g., see Wurm & Fisicaro, 2014, for arguments against the practice of residualizing). Because we limited the number of variables to consider, there was no need for decorrelation in this study.

Between-participants vs. Within-participants designs. In this study, *ReadingDirection* was tested as a between-participants factor, instead of a within-participants factor. There are pros and cons for both approaches. The former design is susceptible to individual differences, but the latter design, as used in Igawa et al. (2006), has limitations as well. When participants are exposed to the two reading directions in a single experiment, the difference between the two conditions become highlighted, which may consequently inform participants of the aim of the study. In this study, across two task conditions, participants were comparable with respect to their age and length of stay in Canada, and idiosyncratic individual differences were minimally accounted for by testing random-intercepts and random-slopes in the mixed-effects models.

Lexical decision with eye-tracking paradigm. In the present study, overall, we obtained consistency between eye-tracking and response time analyses, which successfully allowed us to infer the locus of particular effects. For example, the compound frequency effect observed in the RT analysis is largely a late effect, while the locus of the first character stroke effect is early. Although such consistency provides partial support to the mental chronometry of Donders (1969/1868), it should be noted that end-point responses do not necessarily reflect all cognitive processes. The second character stroke counts, for example, co-determined first and second fixation durations but did not affect RTs. The fact that the contribution of a particular lexical variable is not reflected in a particular dependent measure does not imply that the variable has no contribution. We therefore recommend the recording of multiple measurements.

Conclusion

The present study tested the Japanese language and its script, in which readings in both horizontal and vertical alignments are ecologically valid. The main take home messages are that vertical reading is slower than horizontal reading even in the context of isolated word recognition and that the observed cost was primarily generated during an early time frame, as reflected in the longer first fixation durations in vertical reading than horizontal reading. It is notable, however, that vertical reading also showed a processing advantage: There were fewer fixations in vertical reading than in horizontal reading.

The two alignments did not overly change the pattern of morphological processing at a foveal region; the contributions of character units were higher than those of compound and radical units in an earliest processing stage regardless of reading directions. However, the horizontal advantage (or vertical disadvantage) was greater for words with a high frequency first morpheme, particularly during a late time frame. One interpretation is that it was more difficult to process the first morpheme as a modifier in vertical reading and that the non-modifier (i.e., head) interpretation had an inhibiting consequence for lexicality judgment.

The trade-off between fixation durations and fixation counts, as well as reading directions' modulation of first character frequency and compound frequency effects, provide support for the 118-year-old prediction of Huey (1898): It is not a matter of one alignment being superior to the other, but the two types of readings have some slight yet significant qualitative differences. To fully account for reading across different scripts and reading directions, models of Japanese word recognition should consider which underlying mechanisms, central or peripheral, are responsible for these differences.

References

Amano, S., & Kondo, T. (2003). NTT database series: Lexical properties of Japanese, 2nd release [CD-ROM]. Tokyo, Japan: Sanseido.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. New York, NY: Cambridge University Press.

Baayen, R. H. (2013). *languageR: Data sets and functions with "Analyzing Linguistic Data: A practical introduction to statistics"*. R package version 1.4.1. Retrieved from http://CRAN.R-project.org/package=languageR

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412. doi:10.1016/j.jml.2007.12.005

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language, 37*, 94-117. doi:10.1006/jmla.1997.2509

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. doi:10.1016/j.jml.2012.11.001

Bates, D., Kliegl, R., Vasishth, S., & Baayen, R. H. (2015). Parsimonious mixed models. Manuscript submitted for publication. Retrieved from http://arxiv.org/abs/1506.04967

Bates, D., Maechler, M., & Bolker, B. (2011). *lme4: Linear mixed-effects models using S4 classes*. R package version 0.999375-42. Retrieved from http://CRAN.R-project.org/package=lme4

Bertram, R., & Hyönä. J. (2003). The length of a complex word modifies the role of morphological structure: Evidence from eye movements when reading short and long Finnish compounds. *Journal of Memory and Language, 48*, 615-634. doi:10.1016/S0749-596X(02)00539-9

Box, G. E. P. & Cox, D. R. (1964). An analysis of transformations (with discussion). *Journal of the Royal Statistical Society B, 26*, 211–252.

Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica, 30*, 412-431 (Original work published in 1868). doi:10.1016/0001-6918(69)90065-1

Feldman, L. B., & Siok, W. W. T. (1997). The role of component function in visual recognition of Chinese characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 776-781. doi:10.1037/0278-7393.23.3.776

Forster, K. I. (2000). The potential for experimenter bias effects in word recognition experiments. *Memory & Cognition, 28*, 1109-1115. doi:10.3758/BF03211812

Giraudo, H., & Grainger, J. (2001). Priming complex words: Evidence for supralexical representation of morphology. *Psychonomic Bulletin & Review, 8*, 127-131. doi:10.3758/BF03196148

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory & Language, 58*, 787-814. doi:10.1016/j.jml.2007.07.001

Hsiao, J. H., & Cottrell, G. W. (2009). Not all visual expertise is holistic, but it may be leftist. *Psychological Science, 20*, 455-463. doi:10.1111/j.1467-9280.2009.02315.x

Huey, E. B. (1898). Preliminary experiments in the physiology and psychology of reading. *The American Journal of Psychology, 9*, 575-586. doi:10.2307/1412192

Hyönä, J., & Bertram, R. (2004). Do frequency characteristics of nonfixated words influence the processing of fixated words during reading? *European Journal of Cognitive Psychology, 16*, 104-127. doi:10.1080/09541440340000132

Hyöna, J., Laine, M., & Niemi, J. (1995). Effects of a word's morphological complexity on readers' eye fixation patterns. *Studies in Visual Information Processing, 6*, 445-452. doi:10.1016/S0926-907X(05)80037-6

Igawa, M., Nakayama, N., Maeda, F., & Tabuchi, A. (2006). *Tategaki・yokogaki bunsho ni okeru dokushoji no gankyuundo no hikaku* [Differences in eye movements during reading vertically and horizontally printed sentences]. *Rinsho Ganka, 60*, 1251-1255.

Inhoff, A. W., Pollatsek, A., Posner, M. I., & Rayner, K. (1989). Covert attention and eye movements during reading. *The Quarterly Journal of Experimental Psychology Section A, 41*, 63-89. doi:10.1080/14640748908402353

Kajii, N., Nazir, T. A., & Osaka, N. (2001). Eye movement control in reading unspaced text: the case of the Japanese script. *Vision Research, 41*, 2503-2510. doi:10.1016/S0042-6989(01)00132-8

Kuperman, V., Schreuder, R., Bertram, R., & Baayen, R. H. (2009). Reading of polymorphemic Dutch compounds: Towards a multiple route model of lexical processing. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 876-895. doi:10.1037/a0013484

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015*). lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-25*. Retrieved from http://CRAN.R-project.org/package=lmerTest

Lavidor, M., Babkoff, H., & Faust, M. (2001). Analysis of standard and non-standard visual word format in the two hemispheres. *Neuropsychologia, 39*, 430–439. doi:10.1016/S0028-3932(00)00125-1

Libben, G. & Jarema, G. (2002). Mental lexicon research in the new millennium. *Brain and Language, 81*, 1-10. doi:10.1006/brln.2002.2654

Maekawa, K., Yamazaki, M., Ogiso, T., Maruyama, T., Ogura, H., Kashino, W., … Den, Y. (2014). Balanced corpus of contemporary written Japanese. *Language Resources & Evaluation, 48*, 345-371. doi:10.1007/s10579-013-9261-0

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2015). *Balancing Type I Error and Power in Linear Mixed Models*. Manuscript submitted for publication. Retrieved from http://arxiv.org/abs/1511.01864

Martinez-Conde, S., Otero-Millan, J., & Macknik, S. L. (2014). The impact of micro saccades on vision: towards a unified theory of saccadic function. *Nature Reviews Neuroscience, 14*, 83-96. doi:10.1038/nrn3405

Miwa, K., Libben, G., & Baayen, R. H. (2012). Semantic radicals in Japanese two-character word recognition. *Language and Cognitive Processes, 27*, 142-158. doi:10.1080/01690965.2011.552339

Miwa, K., Libben, G., Dijkstra, T., & Baayen, R. H. (2014). The time-course of lexical activation in Japanese morphographic word recognition: Evidence for a character-driven processing model. *Quarterly Journal of Experimental Psychology, 67*, 79-113. doi:10.1080/17470218.2013.790910

National Institute for Japanese Language and Linguistics. (2012). *Balanced Corpus of Contemporary Written Japanese: Shonagon.* Retrieved from http://www.kotonoha.gr.jp/shonagon/

National Institute for Japanese Language and Linguistics. (2013). *BCCWJ Word List*. Retrieved from http://pj.ninjal.ac.jp/corpus_center/bccwj/freq-list.html

Osaka, N. (1989). Eye fixation and saccade during kana and kanji text reading: Comparison of English and Japanese text processing. *Bulletin of the Psychonomic Society, 27*, 548-550. doi:10.3758/BF03334665

Osaka, N., & Oda, K. (1991). Effective visual field size necessary for vertical reading during Japanese text processing. *Bulletin of Psychonomic Society, 29*, 345-347. doi:10.3758/BF03333939

R Development Core Team (2015). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna, Austria. Retrieved from http://www.R-project.org.

Rogers, H. (2005). *Writing systems: A linguistic approach*. Malden, MA: Blackwell Publishing.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*, 1359-1366. doi:10.1177/0956797611417632

Staub, A., Rayner, K., Pollatsek, A., & Hyönä, J. (2007). The time course of plausibility effects on eye movements in reading: Evidence from noun-noun compounds. *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, 33, 1162-1169. doi:10.1037/0278-7393.33.6.1162

Steinman, R. M., Haddad, G. M., Skavenski, A. A., & Wyman, D. (1973). Miniature eye movement. *Science, 181*, 810-819. doi:10.1126/science.181.4102.810

Taft, M., & Forster, K. (1975). Lexical storage and retrieval of prefixed words. *Journal of verbal learning and verbal behavior, 14*, 638-647. doi:10.1016/S0022-5371(75)80051-X

Taft, M., & Zhu, X. (1997). Submorphemic processing in reading Chinese. *Journal of Experimental Psychology: Learning Memory, and Cognition,* 23, 761-775. doi:10.1037/0278-7393.23.3.761

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

Tamaoka, K. & Hatsuzuka, M. (1998). The effects of morphological semantics on the processing of Japanese two-kanji compound words. *Reading and Writing: An Interdisciplinary Journal, 10*, 293-322. doi:10.1007/978-94-015-9161-4_8

Tamaoka, K., & Kiyama, S. (2013). The effects of visual complexity for Japanese kanji processing with high and low frequencies. *Reading and Writing: An Interdisciplinary Journal, 26*, 205-223. doi:10.1007/s11145-012-9363-x

Tamaoka, K., & Takahashi, N. (1999). The effects of word frequency and orthographic complexity on the writing process of Japanese two-morpheme compound words. *The Japanese Journal of Psychology, 70*, 45-50. doi:10.4992/jjpsy.70.45

Tamaoka, K., & Makioka, S. (2004). New figures for a Web-accessible database of the 1,945 basic Japanese kanji, fourth edition. *Behavior Research Methods, Instruments, & Computers, 36*, 548-558. doi:10.3758/BF03195601

Tinker, M. (1955). Perceptual and oculomotor efficiency in reading materials in vertical and horizontal arrangements. *The American Journal of Psychology, 68*, 444-449. doi:10.2307/1418529

Venables, W. N. & Ripley, B. D. (2002). *Modern applied statistics with S. Fourth edition.* New York: Springer.

Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. New York: Springer.

Wurm, L. H., & Fisicaro, S. A. (2014). What residualizing predictors in regression analyses does (and what it does not do). *Journal of Memory and Language, 72,* 37-48. doi:10.1016/j.jml.2013.12.003

Yu, D. Park, H. Gerold, D., & Legge, G. E. (2010). Comparing reading speed for horizontal and vertical English text. *Journal of Vision, 10*, 1-17. doi:10.1167/10.2.21

Author note

This research was supported by the Izaak Walton Killam scholarship from the Killam Trusts to the first author. Response time and eye movement data after data trimming, accompanying participant and item properties, will be available on the first author's website (http://www.kojimiwa.com/publication.html). The frequency data collected from the Balanced Corpus of Contemporary Written Japanese (BCCWJ) are published with permission of the National Institute for Japanese Language and Linguistics. Authors thank Victor Kuperman and anonymous reviewers for their comments on an earlier version of this paper.

Table 1

Descriptive Statistics for the Predictors Considered in This Study.

Type	Predictor	Range (Mean, SD) / Levels
Compound	WholeWordFrequency $^{(log)}$	4.6: 11.4 (*M* = 6.4, *SD* = 1.4)
Character	FirstCharacterFrequency $^{(log)}$	4.8: 13.6 (*M* = 10.6, *SD* = 1.6)
Character	SecondCharacterFrequency $^{(log)}$	5.1: 13.6 (*M* = 11.0, *SD* = 1.5)
Radical	FirstCharacterRadicalCombinability $^{(log)}$	0.7: 4.7 (*M* = 3.1, *SD* = 1.1)
Radical	SecondCharacterRadicalCombinability $^{(log)}$	0.7: 4.7 (*M* = 3.0, *SD* = 1.1)
Feature	FirstCharacterStrokes	3: 21 (*M* = 9.8, *SD* = 4.1)
Feature	SecondCharacterStrokes	2: 21 (*M* = 9.6, *SD* = 3.8)
Task	ReadingDirection	Levels: Horizontal, Vertical
Task	Trial $^{(/100)}$	0.13: 4.13 (*M* = 2.1, *SD* = 1.1)
Task	PreviousRT $^{(-1000/)}$	-3.2: -0.4 (*M* = -1.6, *SD* = 0.4)
Task	PreviousTrialCorrect	Levels: Correct, Incorrect
Task	PreviousFixationDuration $^{(-100/\sqrt{})}$	-15.8: -2.6 (*M* = -5.2, *SD* = 0.8)
Individual	LengthOfStayCanada $^{(log)}$	0: 4.7 (*M* = 2.0, *SD* = 1.3)

Note. The superscripts represent the method of data transformation. The mean and range are presented for the transformed values before a centering procedure.

Table 2

Fixed-Effects in the Response Time Analysis.

Response time	Type	Estimate	*SE*	*t*-value	*p*-value
(Intercept)		-1.700	0.040	-42.360	< .001
PreviousRT	Task	0.000	0.000	10.142	< .001
Trial	Task	-0.064	0.009	-7.338	< .001
PreviousTrialCorrect (Incorrect)	Task	0.057	0.010	6.111	< .001
ReadingDirection (Vertical)	**Task**	**0.115**	**0.056**	**2.036**	**.049**
WholeWordFrequency	Compound	-0.059	0.006	-8.837	< .001
FirstCharacterFrequency	Character	-0.012	0.009	-1.541	.125
SecondCharacterFrequency	Character	-0.024	0.007	-3.571	< .001
FirstCharacterStrokes	Feature	0.007	0.002	3.073	.002
ReadingDirection (Vertical) x FirstCharacterFrequency	**Task x Character**	**0.015**	**0.006**	**2.674**	**.008**

Note. The partial effects related to reading directions are bolded.

Table 3

Fixed-Effects in the First Fixation Duration Analyses.

First fixation duration	Type	Estimate	*SE*	*t*-value	*p*-value
(Intercept)		-5.114	0.088	-60.723	< .001
Trial	Task	-0.009	0.014	-0.913	.367
PreviousTrialCorrect (Incorrect)	Task	-0.005	0.024	0.215	.830
ReadingDirection (Vertical)	**Task**	**0.249**	**0.077**	**3.534**	**.001**
WholeWordFrequency	Compound	-0.028	0.010	-3.404	.001
FirstCharacterFrequency	Character	-0.091	0.014	-6.212	< .001
SecondCharacterFrequency	Character	0.031	0.012	2.926	.004
FirstCharacterRadicalCombinability	Radical	0.041	0.013	3.086	.002
FirstCharacterStrokes	Feature	0.047	0.005	8.991	< .001
SecondCharacterStrokes	Feature	-0.027	0.004	-5.833	< .001
LengthOfStayCanada	Individual	-0.069	0.030	-2.364	.023
ReadingDirection (Vertical) x Trial	**Task x Task**	**-0.042**	**0.020**	**-2.115**	**.041**

Note. The partial effects related to reading directions are bolded.

Table 4

Fixed-Effects in the Second Fixation Duration Analyses.

Second fixation duration	Type	Estimate	*SE*	*t*-value	*p*-value
(Intercept)		14.620	0.267	55.660	< .001
PreviousFixationDuration	Task	-3.555	0.166	-19.693	< .001
PreviousRT	Task	0.901	0.173	5.750	< .001
Trial	Task	-0.400	0.056	-6.615	< .001
PreviousTrialCorrect (Incorrect)	Task	0.602	0.133	3.844	< .001
ReadingDirection (Vertical)	**Task**	**1.294**	**0.375**	**3.036**	**.004**
WholeWordFrequency	Compound	-0.397	0.049	-7.082	< .001
FirstCharacterFrequency	Character	-0.049	0.067	-0.154	.877
SecondCharacterFrequency	Character	-0.306	0.058	-5.056	< .001
SecondCharacterStrokes	Feature	0.036	0.020	2.116	.036
ReadingDirection (Vertical) x FirstCharacterFrequency	**Task x Character**	**0.169**	**0.067**	**3.149**	**.002**

Note. The partial effects related to reading directions are bolded.

Table 5

Fixed-Effects in the Mixed-Effects Models for Fixation Counts.

Fixation count	Type	Estimate	*SE*	*z*-value	*p*-value
(Intercept)		0.773	0.020	39.180	< .001
PreviousRT	Task	0.127	0.026	4.860	< .001
Trial	Task	-0.028	0.007	-3.790	< .001
ReadingDirection (Vertical)	**Task**	**-0.102**	**0.029**	**-3.570**	**< .001**
WholeWordFrequency	Compound	-0.036	0.006	-6.160	< .001
SecondCharacterStrokes	Feature	0.013	0.003	4.120	< .001
ReadingDirection (Vertical) x SecondCharacterStrokes	**Task x Feature**	**0.009**	**0.004**	**2.010**	**.045**

Note. The partial effects related to reading directions are bolded.

Table 6

Fixed-Effects in the Mixed-Effects Models for Saccade Velocity.

Average saccade velocity	Type	Estimate	SE	t-value	p-value
(Intercept)		50.215	1.555	32.220	< .001
Trial	Task	0.243	0.241	1.730	.092
ReadingDirection (Vertical)	**Task**	**-2.762**	**2.227**	**-0.982**	**.332**
WholeWordFrequency	Compound	-0.401	0.162	-1.876	.062
FirstCharacterFrequency	Character	0.475	0.208	2.321	.024
SecondCharacterRadicalCombinability	Radical	0.367	0.150	2.529	.012
FirstCharacterStrokes	Feature	0.021	0.064	0.353	.725
SecondCharacterStrokes	Feature	0.414	0.056	7.139	< .001
ReadingDirection (Vertical) x WholeWordFrequency	**Task x Compound**	**0.549**	**0.229**	**2.103**	**.037**

Note. The partial effects related to reading directions are bolded.

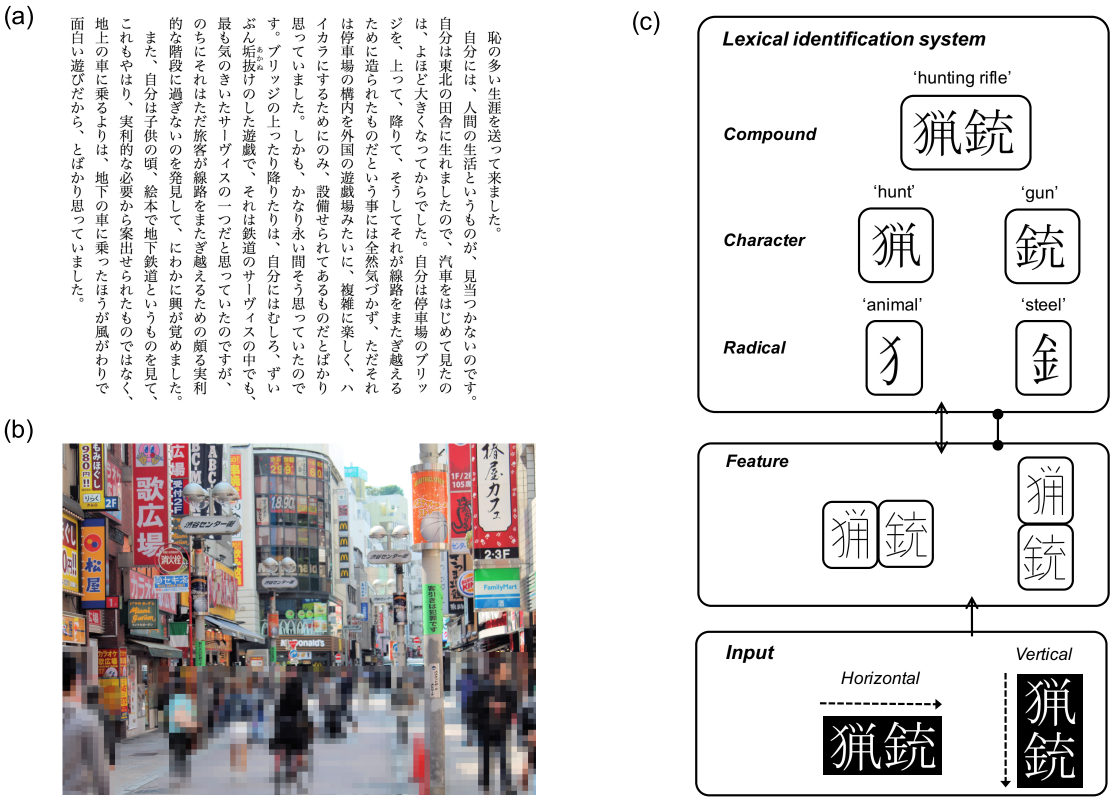

Figure 1. Horizontal and vertical writing in Japanese (Panels a and b) and a hierarchical morphological structure of bimorphemic compounds in Japanese and two reading directions (Panel c).

Note: The dotted arrows indicate the direction of reading. Panel (a) and Panel (b) are adapted with permission of Aozora in Browsers and Åsa Gilbertson respectively.

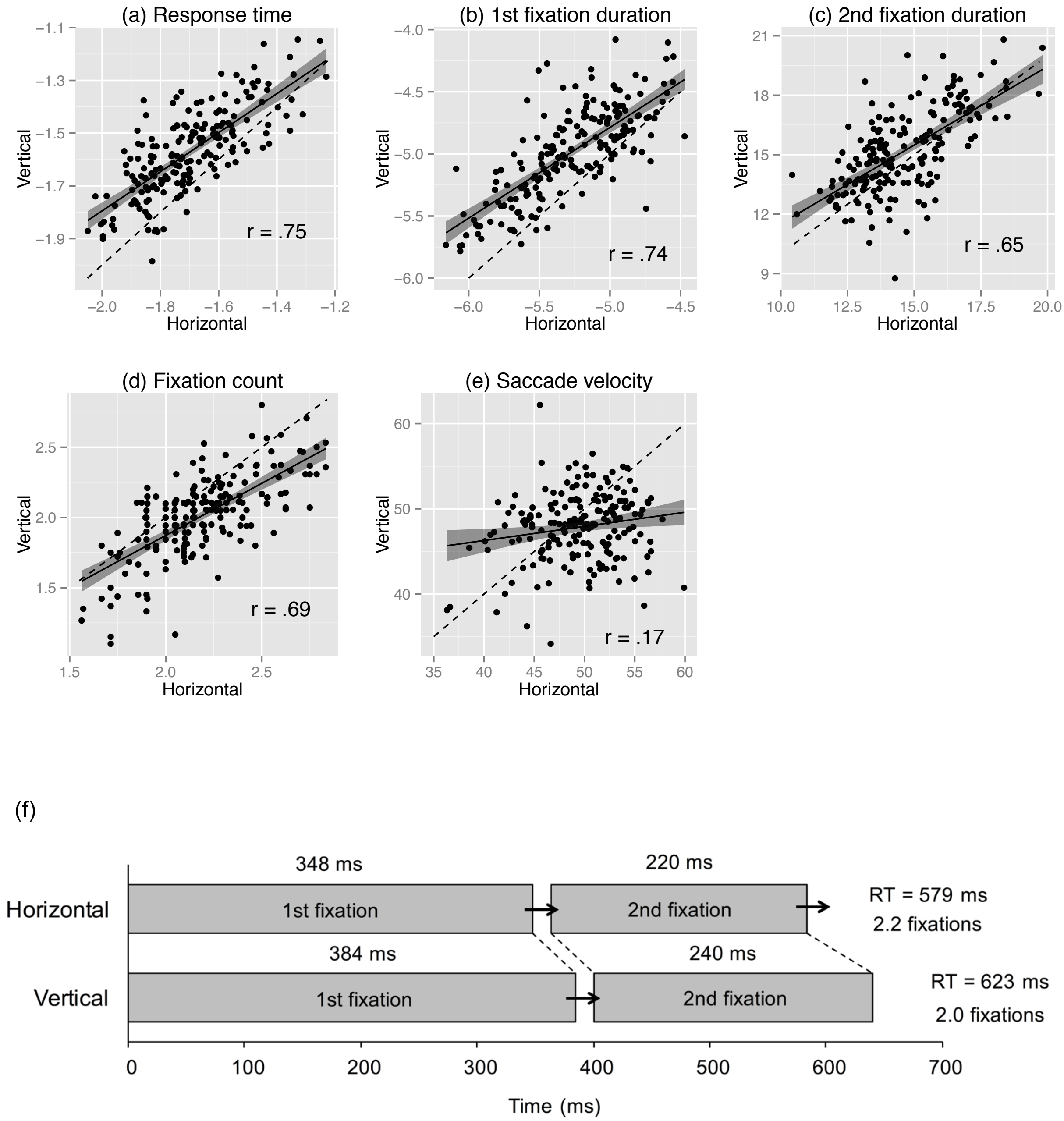

Figure 2. Scatter plots showing correlations between horizontal and vertical readings and a graphical summary of the time-course of horizontal and vertical readings.

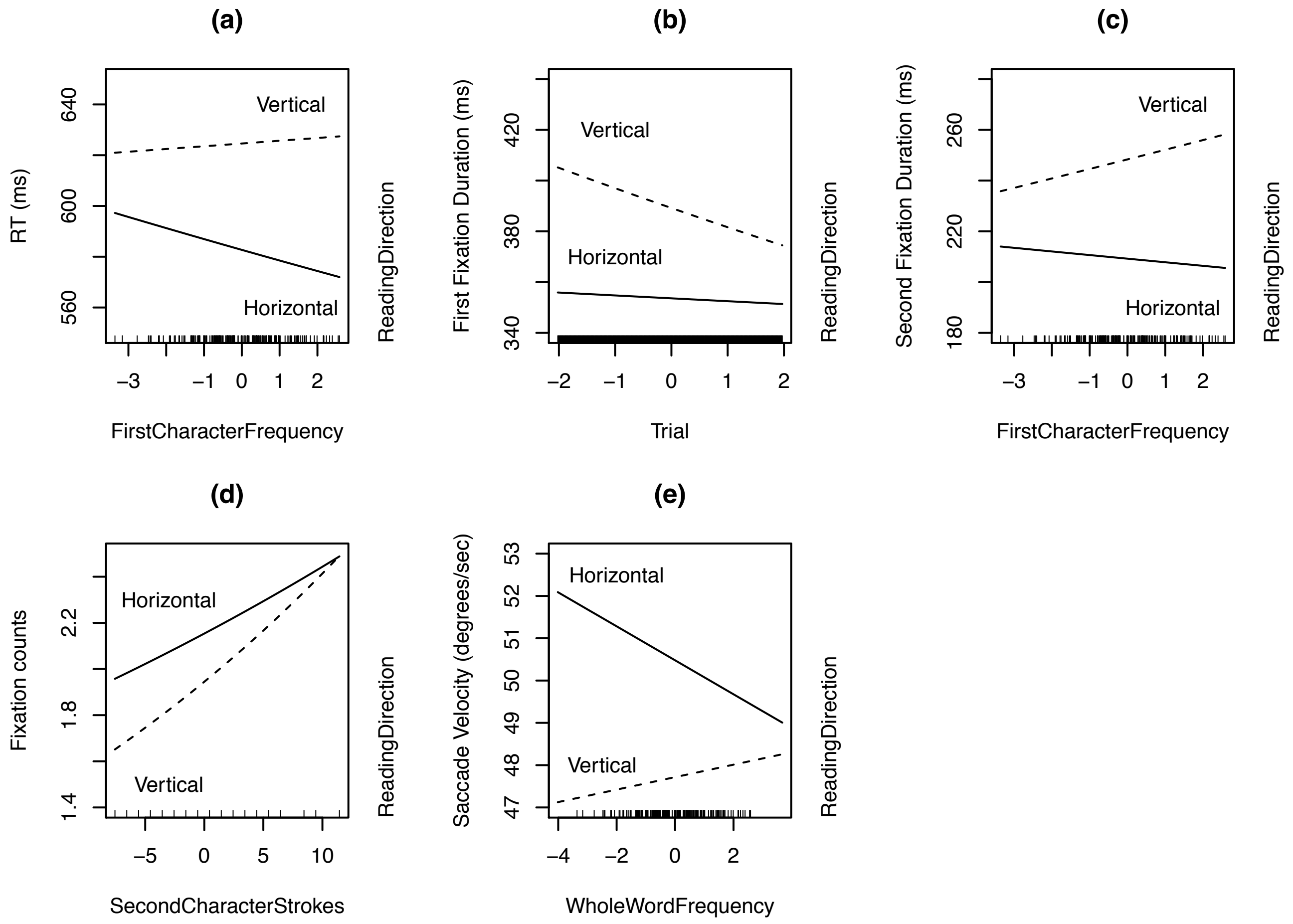

Figure 3. Interactions between lexical predictors and reading direction observed in the mixed-effects model for response times, first fixation durations, second fixation durations, fixation count, and saccade velocity.

Note: The rugs reflect a distribution of the predictor values.