Visual trimorphemic compound recognition in a morphographic script

[1] Koji Miwa, [2] Gary Libben, [3] Yu Ikemoto

[1] Seminar für Sprachwissenschaft, Eberhard Karls Universität Tübingen, Germany

[2] Department of Applied Linguistics and Department of Psychology, Brock University, Canada

[3] School of Human Welfare Studies, Kwansei Gakuin University, Japan

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

Acknowledgment:

This study was supported by the Izaak Walton Killam predoctoral grant from the Killam trusts to the first author.

Abstract

This lexical decision with eye-tracking study investigated how Japanese trimorphemic compounds (e.g., 体温計 'clinical thermometer') are recognized. The questions answered were, in the course of decomposing and composing Japanese trimorphemic compounds, (1) whether recognition processes are tuned for a specific branching direction, (2) whether the morphological processing proceeds in a bottom-up combinatorial manner, and (3) whether the three constituents of trimorphemic compounds are equally important and processed serially. Mixed-effects regression analyses of response times and fixation durations revealed that a left-branching advantage appears in a late time frame and that, although there was early processing of the whole compound from the first fixation, a character frequency effect was also observed. Furthermore, the first and the third, but not the second, constituent frequencies contributed to compound recognition. This bathtub-like effect was further supported by corpus-based evidence: the conditional probability for the second constituent is incomparably high.

Keywords: morphological processing, trimorphemic compound, lexical decision with eye-tracking, Japanese

Word count: Abstract (147) + Main text (10072) + References (1874)

Visual trimorphemic compound recognition in a morphographic script

Bimorphemic compound recognition

There has been a sizable number of studies on processing of bimorphemic compounds (e.g., Kuperman, Bertram, & Baayen, 2008; Libben, 2006; Marelli & Luzzatti, 2012; Pollatsek, Hyönä, & Bertram, 2000; see also Frost, Grainger, & Carreiras, 2008 for a review). This is not surprising because complex words provide an opportunity to study how processing efficiency is achieved in human mind and what efficiency means in the first place (McClelland & Patterson, 2002a, 2002b; Pinker & Ullman, 2002a, 2002b). Accumulated evidence suggests, at least in the context of visual processing, that processing efficiency is achieved not by decreasing the mount of processing nor by minimizing the number of representations but instead by maximizing opportunities for processing (Libben, 2006). For instance, in recognizing a compound word, such as *baseball*, frequencies of occurrences of all morphological constituents (i.e., the compound *baseball*, the first constituent *base*, and the second constituent *ball*) co-determine response speed. Current morphological processing models differ with respect to the order of processing: whole-then-part (Giraudo & Grainger, 2001), part-then-whole (Taft, 2004; Taft & Nguyen-Hoan, 2010), or part-and-whole (Diependaele, Duñabeitia, Morris, & Keuleers, 2011; Kuperman, Schreuder, Bertram, & Baayen, 2009; Pollatsek et al., 2000).

For languages with different writing systems, such as Chinese and Japanese, it has similarly been reported that lexical distributional properties of morphological constituents contribute to compound word recognition processes. Note that, because morphographic scripts used in these languages represent morphemes (Rogers, 2005), bimorphemic compounds are written with two orthographic symbols, such as 体温 *taion* ‘temperature’ (hereafter, an individual morphological constituent represented by a morphographic symbol is called *character*). Evidence of character component processing in compound recognition came from a large number of studies (e.g., Taft, Zhu, & Peng, 1999; Tamaoka, 2005, 2007; Tamaoka & Hatsuzuka, 1998; Miwa, Libben, Dijkstra, & Baayen, 2014; Nakayama, Sears, Hino, & Lupker, 2014; Zhou, Marslen-Wilson, Taft, & Shu, 1999). Although these studies all agree on the point that constituent characters are activated during morphographic compound recognition, consensus has not been reached regarding to what extent compound words are decomposed at the very initial stage of lexical access. In the multilevel model of Taft et al. (1999), compound words and complex characters are initially decomposed not into characters but into even smaller components called radicals, and word recognition proceeds in a bottom-up manner. Indeed, in many cases, it is possible to decompose characters into smaller sub-character components, one of which indicates a basic category meaning (e.g., 体 ‘body’ consists of components イ + 本, where the semantic radical イ implies that the whole character meaning relates to a human body). In a character-driven processing model of Miwa, Libben et al. (2014), on the other hand, compound words are recognized predominantly from characters, instead of radicals. The evidence came from lexical decision with eye-tracking experiments, showing that the early fixation durations were co-determined by large character frequency effects and small radical frequency effects. Given that this trend remained unchanged across different task demands and different font sizes, they concluded that characters serve as the primary processing units in the earliest time frame, although radical-level representations are assumed in this model as well.

Trimorphemic compound recognition

Given the accumulated results on bimorphemic compound processing, it is reasonable to question to what extent trimorphemic compound recognition processes can be predicted from our knowledge of bimorphemic compound recognition. On the surface, extension from bimophemic to trimorphemic compounds appears to be simple (i.e., only one additional morpheme). However, trimorphemic compounds provide a unique opportunity to disentangle the problem of composition from the problem of decomposition.

It is expected that two processes are at work in recognizing compound words — a process to identify constituents (decomposition) and a process to understand the whole word meaning (composition) — regardless of whether these processes work in sequence or in parallel. For a bimorphemic compound AB, once the word is decomposed into A and B, there is only one possible structural choice regarding how to compose these morphemes for proper interpretation: [AB]. For this reason, it is difficult to investigate a composition process using bimorphemic compounds. On the contrary, for trimorphemic compounds ABC, upon decomposition of words into morphemes A, B, and C, there are multiple possibilities for composition: [[AB][C]] or [[A][BC]] or [[A][B][C]]. For example, after a trimorphemic compound *pork back rib* is decomposed into three morphemes, these morphemes may be composed as [[pork back][rib]] or [[pork][back rib]] or [[pork][back][rib]]. Depending on how constituent morphemes are composed, the resulting whole word carry different meanings: ‘a rib of a pork back’ or ‘a back rib of a pork’ or ‘a pork back rib.’ Thus, trimorphemic compounds are qualitatively different from bimorphemic compounds and possess characteristics of both words and sentences. This morphological problem has been conceptualized in the form of hierarchical tree structure in theoretical linguistics (Spencer, 1997). A question yet to be answered is whether such hierarchical structures are actually used online during several hundred milliseconds of visual word recognition process.

Compared to the large number of bimorphemic compound processing research, trimorphemic compound research is rather limited, and the limited number of past studies do not point to a single conclusion with respect to effects of branching directions. An early study of Libben (1993) reported no significant difference between left- and right-branching words in a naming experiment with nonsense multimorphemic words, but a nonsense word naming task with a Broca's patient showed a right-branching advantage (Libben, 1994). On the contrary, Pollatsek, Drieghe, Stockall, and de Almeida’s (2010) study on ambiguous trimorphemic compound recognition in sentential reading reported a left-branching advantage. Krott et al. (2004), too, reported left-branching advantage for German and Dutch languages and stated that the result is consistent with lexical statistics in these languages, indicating that there are more left-branching words than right-branching words. de Almeida and Libben (2005) tested ambiguous trimorphemic compound processing in sentential reading and pointed out that morphological structures may not be constructed from the beginning; it is possible that early processing is morphologically blind and a morphological structure is chosen later. These inconsistencies may come from various sources: languages, participants, or tasks.

Although the above studies were done with languages with alphabetic scripts, cross-linguistic orthographic differences should be taken into account in investigating the issue of morphological decomposition and composition. That is, morphological decomposition is clearly more challenging in alphabetic scripts than that in morphographic scripts. If the readers are not familiar with German or related languages, for example, it must be nearly impossible to decompose multimorphemic German compounds (e.g., *Fußballweltmeisterschaft* 'soccer world cup') into constituents (e.g., *Fuß-ball-welt-meister-schaft).* However, without any knowledge of the Japanese language, readers can decompose multimorphemic words (e.g., 非課税証明書 'a certificate of tax exemption') into constituents properly (e.g., 非-課-税-証-明-書) due to the inter-character spacing and the orthographical scripts representing a morpheme. Effects of a visual cue in segmenting compounds were tested by Inhoff, Radach and Heller (2000). In their study, spacing inserted into German compounds facilitated morphological segmentation in an early phase. Similarly, Bertram, Kuperman, Baayen, and Hyönä (2011) inserted a hyphen in trimorphemic compounds. Although this is not conventional in the tested languages Dutch and Finnish, hyphenation was received favorably by readers, particularly towards the end of the experiment, as long as it did not motivate incorrect interpretation of words. Given these findings, in processing Japanese trimorphemic compounds, readers are expected to take advantage of inter-character spacing inherent in the morphographic script, which may motivate qualitatively different processes. For example, if the third character happens to be crucial for its status as a morphological head (e.g., 体温計), then there is no need to process from left to right in locating the third character, as it always occurs in the same physical location separated from the preceding character by space. To the best of our knowledge, there is no chronometric experimental study that investigated how and when lexical properties at different morphological levels contribute to trimorphemic compound recognition in a morphographic script. Using multivariate analyses for response time and eye movement data, we tested multiple hypotheses and aimed to provide a bird's-eye view of the contributions of 12 lexical predictors.

Possible processing architectures for Japanese trimorphemic compounds

A great number of studies investigated processing of bimorphemic compounds in morphographic scripts. It is not clear, however, whether Japanese trimorphemic compound processing can be predicted from our knowledge of bimorphemic compound processing. On one hand, the two are comparable because it is efficient to process the two types of compounds using the same cognitive architecture. On the other hand, a composition process, which is crucial in three-character compound processing, has not been investigated. Japanese is considered to be a left-branching language with respect to its syntactic structure. Because linguistic structures tend to pattern in a systematic manner, apparently motivated by cognitive reasons, such as processing and representational efficiency (Whaley, 1997), it is possible to expect that the morphological structure of Japanese is similarly left-branching. That is, a left-branching morphological structure embedded in a left-branching syntax is intuitively less costly than right-branching morphology in left-branching syntax. An inspection of the Balanced Corpus of Contemporary Written Japanese (BCCWJ, Maekawa et al., 2014) confirmed that there are more left-branching compounds (61.8%) in Japanese than right-branching compounds

(24.5%) and those without any branching structure (13.7%). If a hierarchical architecture is psychologically real, and if the word recognition system is tuned for a left-branching structure, processing three-character compounds from left to right on a character-by-character basis may fail for some types of words (e.g., right-branching words). Yet, native speakers of Japanese recognize familiar three-character compounds effortlessly. To conceptualize the problem more clearly, several possible processing architectures are presented in Figure 1. Throughout the sections that follow, individual morphographic constituents are called *characters,* a two-character compound is called a *bimorphemic compound*, and a three-character compound is called a *trimorphemic compound.*

(Figure 1 about here)

Hypothetical processing architectures

Left-branching architecture. Miwa, Libben et al. (2014) manipulated the initial fixation point in lexical decision and concluded that, for bimorphemic compounds, it is optimal to process from the left character to the right character. If trimorphemic compound recognition similarly proceeds from left to right, then one possible architecture is the one tuned for a left-branching structure, in which a hierarchical structure can be established by connecting the incoming constituent to the existing constituent (Figure 1 Panel a). This approach is, for example, applicable to a trimorphemic complex words *science fiction writer*, which can be correctly recognized as *science*, *science fiction*, and then *science fiction writer.* Similarly, a Japanese trimorphemic compound 体温計 *taionkei* ‘thermometer,’ which can be correctly recognized sequentially as 体 ‘body,’ 体温 ‘body temperature,’ and then 体温計 ‘clinical thermometer.’ This architecture is analogous to the late closure and minimal attachment principles in sentence processing (Frazier, 1978 as cited in Frazier & Rayner, 1982), although the working-memory-related cost saving is expected to be negligible in compound processing, relative to sentence processing.

It is also possible to expect a left-branching-tuned processing, because Japanese is typologically considered to be a left-branching language and because there are more left-branching words, as mentioned above. Experimental investigation of branching directions of languages dates back to Forster’s (1966, 1968) studies on sentence completion. In these studies, participants’ performance to complete sentences, in which either the beginning or the end were missing, was affected by the branching direction of the participants’ language. A drawback of such a left-branching-specific processing is a likely misanalysis of right-branching words. Compound words such as *Toronto drug store* cannot be properly recognized (i.e., **Toronto drug* + *store*). For the same reason, Japanese compounds with a right-branching structure, such as 北半球 *kitahankyu* ‘northern hemisphere (literally, north-half-sphere)’ cannot be decomposed into *北半 and 球. According to this left-branching architecture, assuming that words are processed from left to right, we expect a large character frequency effect in an early time frame, together with a moderate compound frequency effect and a lesser or null effect of whole word (trimorphemic compound) frequencies. Processing advantage for words with the left-branching structure is also expected.

Right-branching architecture. The architecture tuned for right-branching words is visualized in Figure 1 Panel b. Although this is intuitively not appropriate for a left-branching language, the architecture is in line with the prefix stripping hypothesis (Taft & Forster, 1975), in which a complex word's initial morphological constituent to be automatically stripped off. In this architecture, the right-branching word 北半球 'northern hemisphere' can be correctly recognized as a combination of 北 'north' and 半球 'hemisphere.' However, if such initial constituent-stripping were the language-general rule, trimorphemic compound recognition would fail for the majority of Japanese words at an early processing stage. According to this right-branching architecture, we can expect processing advantage for words with the right-branching structure.

Architecture without any intermediate compound level. If the processing architecture is strictly tuned for a particular branching structure as in Panels (a) and (b), extra processing cost has to be expected for words with an incompatible branching structure. If we assume the same processing mechanism operating for trimorphemic compounds of all kinds, then one solution is not to assume any mechanisms to distinguish different branching directions (Panel c). Although this architecture may not be appealing from the perspective of theoretical linguistics, there is cognitive motivation: there is no apparent cost arising from a misanalysis. According to this architecture, a bimorphemic compound frequency effect is not expected to appear at any point in time, not to mention an effect of branching directions.

Both-branching architecture. It is also possible to construct all possible morphological structures from the beginning, and an appropriate interpretation is chosen at a later stage (Figure 1, Panel d). This appears to create redundancy. However, redundancy maximizes opportunities for further processing and reduces cost of later misanalysis (Libben, 2006). Like the previous architecture (Panel c), processing advantage is not assumed for a specific branching structure. Unlike the previous architecture, however, a bimorphemic compound processing is expected. In an early time frame, the magnitude of a bimorphemic compound frequency effect is expected to be bigger than that of a trimorphemic compound frequency effect but smaller than that of a character frequency effect.

Supralexical and sublexical processing. We have so far provided several possible processing architectures assuming that morphological word recognition starts from the character level. However, it should be noted that there are other possibilities not depicted in Figure 1: supralexical and sublexical processing. If trimorphemic compounds are processed in a top-down manner, as predicted by a supralexical model (Giraudo & Grainger, 2001), then we should observe a large whole word frequency effect before morphological constituent frequency effects. In contrast, if trimorphemic compounds are processed in a strictly bottom-up manner from sub-character radical constituents, as predicated by sublexical model (Taft et al., 1999), then we should observe large radical frequency effects before character and whole word frequency effects arise. Although we did not ignore these possibilities during the forthcoming data analyses, it should be note here that a supra-lexical model has never been proposed in morphographic word recognition research and that character effects were incomparably stronger than radical and whole compound effects in the early time frame of two-character word recognition process (Miwa, Libben et al., 2014). Therefore, it was predicted that strictly bottom-up and top-down processes were less likely than the above mentioned character-driven processes in three-character compound recognition as well.

Methodological consideration and aims of this study

Use of eye-tracking in a lexical decision experiment

Response time has been a commonly used dependent variable in word recognition research (Libben & Jarema, 2002). Whenever response times are measured, an implicit assumption is that response times reflect all processes from the moment of stimulus perception to the moment of a button-press, following the mental chronometry of Donders (1969/1868). However, this should be kept in mind in interpreting results because response times do not necessarily capture the entire processing from perception to recognition to response. One solution to this problem is recording of eye movements during a word recognition experiment, which provides more dependent measures and allows researchers to tap into a time-course of word recognition process (Hyönä, Laine, & Niemi, 1995; Kuperman et al., 2009; Miwa, Libben et al., 2014; Miwa, Dijkstra, Bolger, & Baayen, 2014; see Bertram, 2011 for a review of eye-tracking studies on morphological processing). In previous studies, the results in a response time analysis were qualitatively more similar to those in a later time frame than those in an early time frame, although early and late measures both showed a decent correlation with response times (Miwa, Libben et al., 2014, Miwa, Dijkstra et al., 2014).

In this lexical decision with eye-tracking study, participants' eye movements (i.e., fixation duration, fixation count, and saccade amplitude) were measured, as well response times (RT), and analyzed using linear mixed-effects regression modeling (Baayen, 2008;

Baayen, Davidson, & Bates, 2008). Emphasis was placed on response times and fixation durations because they can be interpreted in a like manner, given the shared unit of measurement in milliseconds. Analyses of early and late fixation durations allow us to see what variables co-determine lexical processing in early and late time frames respectively. Saccade amplitudes were analyzed to safeguard against the possibility that different eye movement measures point to overly different conclusions. Although analyses of fixation durations motivated Miwa, Libben et al. (2014) to propose character-driven processing, where to fixate next may depend more on the properties of the whole compound (i.e., trimorphemic compound frequency and branching direction).

Research questions

Using a regression design, we tested multiple hypotheses. First, we investigated whether recognition processes are tuned for a specific branching direction and, if so, when a branching-direction effect manifests itself. If there is a language-specific morphological tree or something equivalent entrenched in memory (Figure 1 Panels a and b), then an effect of branching directions should be observed. We expected left-branching advantage because Japanese is a left-branching language with the majority of three-character words being left-branching, as well as because processing advantage was found in a reading task for left-branching words even in a right-branching language (Pollatsek et al., 2010). If Japanese word recognition system is strictly tuned to optimize left-branching processing, then the branching effect should co-determine the first fixation durations. However, following the maximization of opportunity principle (Libben, 2006), it is possible to predict that the early processing is blind to branching structures (Panels c and d).

Second, we investigated whether the morphological processing proceeds in a bottom-up combinatorial manner. Because bimorphemic compound recognition starts with activation of constituent characters, accompanying a small yet significant contribution of a whole compound unit (Miwa, Libben et al., 2014), it is similarly expected that, in trimorphemic compound recognition, the first fixation durations are co-determined by large character frequency effects and, if there is any, a small whole trimorphemic compound frequency effect. However, if the lexical processing proceeds in a bottom-up manner, we should also expect a frequency effect of a bimorphemic compound inside a trimorphemic compound (e.g., 体温 in 体温計). This second question relates to the first question because a morphology-sensitive architecture implies bottom-up processing.

Third, we investigated whether the three characters of trimorphemic compounds are equally important and processed serially. Note that there may be different answers to this question at different levels. At a lexical level, morphological headedness (Libben, Gibson, Yoon, & Sandra, 2003) and left-to-right preferential processing in Japanese (Miwa, Libben et al., 2014) deserve attention. Because Japanese compounds are predominantly right-headed (e.g., 体温計 ‘thermometer, literally body-temperature-measure’ is a kind of a ‘measure,’ as opposed to a kind of a ‘body.’ see also Kageyama, 2010 for Japanese morphology), the third character frequency may contribute more than frequencies of the first and the second characters. Indeed, in multivariate studies on Japanese bimorphemic compound recognition (Miwa, Libben, & Baayen, 2012; Miwa, Libben et al., 2014), the contribution of the second character frequency to response times was always more than or equal to that of the first character frequency. With respect to processing direction, although a morphological head is located at the right-hand side, it is expected that Japanese trimorphemic compound words are read from left to right. Miwa, Libben et al. (2014) manipulated the initial fixation position (left, right, and centre) in their lexical decision with eye-tracking experiment, but frequency of the left character co-determined early fixation durations more than that of the right character regardless.

At a feature level, we coded visual complexity for each character separately, instead of considering the complexity of a whole trimorphemic compound because visual complexity does not always inhibit processing as is commonly believed. The direction of visual complexity effects depends on the eye position. In Miwa, Libben et al.'s (2014) lexical decision study, during an early time frame, visual complexity of the fixated region inhibited processing and that of unfixated region facilitated processing. Such fixation-dependent asymmetrical effects of visual complexity were in line with a magnetic attraction account proposed by Hyönä and Bertram (2004). The present study more clearly tests the magnetic force account of a parafoveal-on-foveal effect (note, however, that this is a within-word parafoveal-on-foveal effect. For more commonly investigated between-word parafoveal-on-foveal effects, see Drieghe, 2011 for a review). If the asymmetrical effects of character strokes observed in Miwa, Libben et al. (2014) are not two-character-word-specific phenomena, then at the first fixation on a trimorphemic compound, complexity of the first character in the foveal area should lead to inhibition and that of the second character should lead to facilitation, possibly accompanying slightly weaker facilitatory effect of visual complexity of the third character.

Predictors considered in this study

Table 1 shows descriptive statistics of lexical predictors considered in this study. *Branching* is a factor with two levels of *LeftBranching* and *RightBranching*, referring to the morphological branching direction (e.g., *science fiction writer* is left-branching).

(Table 1 about here)

The log-transformed frequencies of occurrences for various morphological units (*LogTrimorphCompFreq LogBimorphCompFreq*, *LogFirstCharFreq LogSecondCharFreq*, and *LogThirdCharFreq*) were collected from the Balanced Corpus of Contemporary Written Japanese (BCCWJ) listing approximately 104 million words (Maekawa et al., 2014). Although *LogTrimorphCompFreq* and *LogBimorphCompFreq* correlated decently ($r = .46$ $p < .01$), we included the original predictors together without residualizing and observed the consequence, following suggestions by Wurm and Fisicaro (2014) and York (2012).

Although NTT frequencies (Amano & Kondo, 2003), based on newspaper texts published in the years from 1985 to 1998, were popular among many past studies on Japanese processing (e.g., Hino, Kusunose, Lupker, & Jared, 2013; Miwa et al., 2012; Miwa, Libben et al., 2014; Nakayama et al., 2014; Tamaoka, 2007; White, Hirotani, & Liversedge, 2011), we opted for the BCCWJ frequency measures in this study. The

BCCWJ frequency appears to fit lexical decision data better, based on a goodness-of-fit test, possibly because the BCCWJ is more contextually diverse and more up-to-date. When NTT frequency and BCCWJ frequency were compared as a sole fixed-effect in mixed-effects models fitted to response time data for 474 bimorphemic Japanese compounds tested in Miwa, Libben et al. (2014) study, the latter was identified to be a better frequency measure (ΔAIC = -55.2, $p < 0.0001$ in a likelihood ratio test).

Not to neglect the possibility that morphographic words are initially decomposed into radical units (Taft et al, 1999), we considered radical combinability. *LogFirstCharacterRadicalCombinability*, *LogSecondCharacterRadicalCombinability*, and *LogThirdCharacterRadicalCombinability* were log-transformed radical combinability measures taken from Tamaoka and Makioka’s (2004) lexical database for *kanji* characters. These were type frequencies of radicals, indicating how many characters share a given semantic radical (e.g., the radical 氵 ‘water’ is seen in 103 characters, including 海 ‘sea’, 港 ‘port’, and 滝 ‘waterfall’).

Because visual word recognition necessarily starts with decoding of visual features, visual complexity was considered by means of character stroke counts (*FirstCharacterStrokes*, *SecondCharacterStrokes*, and *ThirdCharacterStrokes*). We considered stroke counts for the three characters separately, instead of stroke counts of entire trimorphemic compounds. Because the direction of visual complexity effects can be facilitatory or inhibitory depending on where the eye is fixating (Miwa, Libben et al., 2014), this allows us to study visual complexity effects more precisely.

In addition to the above lexical predictors, we also considered participants’ months of stay in Canada and several task-related predictors (*Trial*, *PreviousResponseCorrect*, *invPreviousRT,* and *PreviousFixationDuration*). *Trial* indicates at what point in the experiment a particular trial took place. *PreviousResponseCorrect* (levels*: Correct, Incorrect*) is a factor coding response accuracy in the previous trial (reference level = *Correct*). *invPreviousRT* refers to inversely transformed response time (-1000/RT) in an immediately preceding trial. In the second subgaze duration reported below, *logPreviousFixationDuration* (i.e., log-transformed first fixation duration) was included as a covariate. All the above numerical predictors were centered by subtracting the mean value from their individual values.

In this study, we opted for mixed-effects regression techniques (Baayen, 2008; Baayen, Davidson, & Bates, 2008) to assess participants, items, and task effects simultaneously in a single statistical model. With this technique, there was no need for extensive pre-experimental matching and dichotomization of numerical variables, which are not optimal for statistical reasons (Baayen & Milin, 2010; MacCallum, Zhang, Preacher, & Rucker, 2002).

A lexical decision with eye-tracking experiment

Methods

Participants. Twenty-one native speakers of Japanese (14 females, 19 right-handed) were tested. They were 23.6 years old on average (median = 22 years, interquartile range = 7). At the time of the experiment, the majority of the participants had stayed in Canada for less than half a year (median = 3 months, *IQR* = 10), and they were exposed to Japanese 42.4% of the time (*SD* = 18.4). The participants' log-transformed months of stay in Canada correlated strongly with log-transformed age (*r* =

.80, $p < .01$). To safeguard against potential effects of their expatriate status on lexical processes, their log-transformed months of stay in Canada was tested in the forthcoming regression models.

Materials. The lexical decision experiment consisted of 200 existing trimorphemic words and 200 non-existing nonwords (see Appendix A). The 200 existing trimorphemic either had a left-branching structure (100 words) or a right-branching structure. Because the whole compound frequency is considered to be the most influential predictor in Japanese compound recognition (see Miwa, Libben, & Baayen, 2012 for a random forest variable importance ranking), the frequencies of occurrence were matched for the left- and the right-branching words, $t(198) = -0.26$, $p = .79$, so that the two groups of words become comparable with respect to ease of processing. For all other lexical properties between the left- and right-branching words, we opted for post-experimental statistical control by including relevant predictors as covariates in regression models. If left- and right-branching words are inherently different with respect to specific lexical distributional properties, this can be statistically controlled by including an interaction the lexical effects and the factor *Branching*.

In order to elicit "no" responses, 200 nonwords were created by randomly combining kanji characters. The existing words and the nonwords were matched with respect to stroke counts, $t(398) = 0.125$, $p = .90$, not to motivate responses purely based on visual complexity.

Apparatus. The experiment was designed with Experiment Builder software (SR Research, Canada). Words were presented on a 24-inch LCD display. An EyeLink II head-mounted eye-tracker (SR Research, Canada) was used to track eye movements, in the pupil-only mode with a sampling rate of 250 Hz. Tracking was done binocularly, but the recording was done only for the right eye for all participants.

Procedure. In this experiment, participants were instructed to decide as quickly and accurately as possible whether a presented word is an existing legitimate word in Japanese (= Yes) or not (= No), by pressing an appropriate button of a Microsoft SideWinder game-pad. The Yes-button was pressed by the index finger of their dominant hand. Participants were instructed, in each trial, to fixate on a fixation point, which also served as a drift-correct point. Given that the first character is more likely to be fixated than the second and the third characters when reading Japanese trimorphemic words in sentential reading (Kajii, Nazir, & Osaka, 2001), the fixation mark was placed at the first character position. Words were presented in white Mincho font on a black background. From a viewing distance of 70 cm, the visual angle for each kanji character was approximately 1.3 degree (the physical width of each character was 16 mm). The experiment started with 12 practice trials and lasted approximately an hour, including two breaks. Participants received feedback about accuracy and speed of their responses after the practice and at each break point.

Results and discussion

In the sections that follow, (generalized) linear mixed-effects analyses were conducted in R version 3.2.2 (R Development Core Team, 2015), using the R packages *lme4* (Bates, Maechler, Bolker, & Walker, 2014) with p-values computed with Satterthwaite approximations to degrees of freedom with *lmerTest* package (Kuznetsova, Brockhoff, & Christensen, 2015). The partial effects were then visualized using *languageR* package (Baayen, 2013). Transformations of outcome variables were based on visual inspection of Q-Q normality plots and Box-Cox transformation with *MASS* package (Box & Cox, 1964; Venables & Ripley, 2002).

Response time analysis. All participants responded with relatively high accuracy (range 89.4 - 99.5%) and were therefore kept for the analyses. Four words were removed from the analyses because average response accuracy to these words was less than 70%. In addition, incorrect responses and responses faster than 300 ms were removed. The above trimming procedure removed 6% of the data, and 3,897 data points were kept.

We fitted linear mixed-effects models to inversely transformed response times (-1000/RT). In the final model, after outliers with absolute standardized residuals exceeding 2.5 standard deviation units were removed (2.4% of the data), random intercepts for subjects (*SD* = 0.15) and item (*SD* = 0.09), together with random slopes for subjects with respect to *PreviousRT* (*SD* = 0.05) and *Trial* (*SD* = 0.03) were included. Random slopes involving other predictors were also tested but not justified by likelihood ratio tests. While there are competing proposals regarding an optimal random-effects structure (Barr, Levy, Scheepers, & Tily, 2013; Bates, Kliegl, Vasishth, & Baayen, 2015; Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2015), we opted for a forward-fitting procedure and kept only those random intercepts and random slopes justified by likelihood ratio tests throughout this study. The standard deviation of the residual error of the model was 0.23. The R^2 calculated for the correlation between the fitted and observed values was .46.

Fixed-effects were selected by backward elimination based on $p < .05$ and AIC values, and the final list of fixed-effects is summarized in Table 2. Figure 2 visualizes partial effects of the lexical predictors. In Figure 2, it is notable that although Panels (b), (c), (d), and (f) all show a downward slope, Panel (e) does not, meaning that all morphological constituents (i.e., trimorphemic compound, bimorphemic compound, and character) contributed to RTs, except the second character. RTs were not significantly co-determined by branching directions (Figure 2, Panel a), stroke counts (Panels g, h, and i), radical combinability measures, nor subjects' months of stay in Canada. However, it is too hasty to conclude that these predictors did not contribute at all in the recognition of trimorphemic compounds. Their effects might have appeared and disappeared well before responses, or their effects were inhibitory at one point and facilitatory at another point. To see more accurate pictures, the participants' eye movements were analyzed next.

(Table 2 and Figure 2 about here)

Fixation count analysis. Fixation counts were analyzed first because only trials with multiple fixations are informative for the purposes of studying the time course of lexical effects: multiple fixations can dissect the recognition process into smaller time frames. The words and trials removed in the above response time analysis were removed here as well. After removing trials with a blink, 3643 data points were at our disposal (12% of the data was removed in this trimming procedure).

Given the font size typical in lexical decision experiments, one may predict that words could be read with a single fixation. This was true. 10% of the trials were read with a single fixation. However, 70% of the trials were read with two fixations (17% with three fixations, and 2% with four fixations), implying that the participants could read with a single fixation but did not most of the time. This is not surprising because even bimorphemic compounds were read with two fixations in the context of a lexical decision experiment (Miwa, Libben et al., 2014).

Generalized linear mixed-effects models with a Poisson distribution were fitted to fixation counts. The final model comprised of the random-effects of subjects (*SD* = 0.12) and, as fixed-effects, significant facilitatory effects of *Trials* and *logThirdCharFreq* (Table 3). This indicates that participants made fewer fixations for words with high *logThirdCharFreq* and for trials in later parts of the experiment (effect sizes = -0.2 and -0.3 fixation counts respectively). In the sections below, only a subset of trials with exactly two or three fixations was used in fixation duration analyses (87% of the data).

First fixation duration analysis. The words and trials removed in the response time analysis were removed here as well, and there were 3,195 data points at our disposal. Linear mixed-effects models were fitted to log-transformed first fixation durations, and outliers with standardized residuals exceeding 2.5 standard deviation units were removed (2.9% of the data) for the final model. The random-effects structure of the final model comprised random intercepts for subjects (*SD* = 0.15) and random slopes per subjects for *Trial* (*SD* = 0.02), *logFirstCharFreq* (*SD* = 0.01), and *FirstCharacterStrokes* (*SD* = 0.01). Random slopes involving other predictors were also tested but not justified by likelihood ratio tests. The standard deviation of the residual error was 0.14. The R^2 calculated for the correlation between the fitted and observed values was 0.44. The fixed-effects were identified through backward elimination based on $p < .05$ and AIC values, and the final list of fixed-effects is summarized in Table 3. Significant partial effects of the lexical predictors are visualized in Figure 3.

(Table 3 and Figure 3 about here)

Because visual word recognition is expected to start with perception of visual features, the first fixation durations, not surprisingly, reflected analyses of stroke features. The visual complexity of the first character had an inhibitory effect (Figure 3, Panel g), but visual complexity of the second and third character had a facilitatory effect (Panels h and i). This asymmetrical pattern of results may appear to be counterintuitive, but the results are in line with Hyönä and Bertram's (2004) magnetic force account that processing difficulty in a parafoveal region attracts an eye movement, with the magnitude of attraction depending on the distance between the current eye position and the source of difficulty. Recall that the fixation mark was placed at the first character position in this experiment. Consequently, participants fixated at the center of the first character without exception (*M* = 1.52, *SD* = 0.17, given that 1.0 and 2.0 refer to the left edges of the first character and the second character respectively). The facilitatory effect of visual complexity in a yet-to-be fixated parafoveal region was also observed in bimorphemic compound recognition (Miwa, Libben et al., 2014), and the present results indicate that even the visual complexity at the third character position facilitates processing. Interestingly, this asymmetrical pattern was almost completely reversed in the saccade amplitude analysis (Appendix B, Panels g, h, i) in which saccade amplitudes were decreased by the visual complexity of the second and third characters but increased by that of the third character. That is, the visual complexity of the third character elicited the eye to travel further. These results indicate that feature-level analyses are done for all three characters simultaneously during the earliest time frame.

Figure 3 also shows a facilitatory *logFirstCharFreq* (Panel d), accompanying an effect of *logTrimorphCompFreq* (Panel b). It is notable that there was no effect of *Branching* (Panel a) and *LogBimorphCompFreq* (Panel c). Because there is an effect of *logTrimorphCompFreq* without an effect of *LogBimorphCompFreq*, the result is not consistent with an architecture that postulates a compound frequency effect (Figure 1 a). However, this is still consistent with the architectures that assumes a direct link from the first character to the whole trimorphemic compound level (Figure 1 b and c).

Although some past studies proposed decomposition of characters into radicals in an early time frame, radical combinability measures did not contribute to the first fixation durations. This is consistent with the character-driven, not radical-driven, processing proposed by Miwa, Libben et al. (2014): early contribution of radicals remain negligible compared to that of characters even in trimorphemic compound recognition.

Second subgaze duration analysis. To study lexical processes in a later time frame, mixed-effects models were fitted to square-root-transformed second subgaze durations. For trials with three fixations, second subgaze durations were the sum of the second and third fixation durations, and second subgaze durations were identical to second fixation durations for trials with exactly two fixations. As in the first fixation duration analysis, with the words and trials removed in the response time analysis excluded, there were 3,195 data points to start with. Note that, in this later time frame, the eye moved to the second character region: *Median* = 2.33 (*IQR* = 0.33) for the second fixations and *Median* = 2.81 (*IQR = 1.22*) for the third fixations, given that positions 2.0 and 3.0 refer to the left edges of the second character and the third character respectively. It is clear that even when three fixations were made, the eye rarely traveled to the third character.

With outliers with standardized residuals exceeding 2.5 standard deviation units removed (2.6% of the data), the final model was fitted. The random structure of the final model comprised random intercepts for subjects (*SD* = 0.15), per-subject random slopes for *Trial* (*SD* = 0.02), per-subject random slopes for *logFirstCharFreq* (*SD* = 0.01), and per-subject random slopes for *FirstCharacterStrokes* (*SD* = 0.01). Random slopes involving other predictors were also tested but not justified by likelihood ratio tests. The standard deviation of the residual error of the model was 0.14. The R^2 calculated for the correlation between the fitted and observed values was 0.54.

The fixed-effects of the model are summarized in Table 3, and the partial effects of the lexical predictors are visualized in Figure 4. In this late timeframe, a main effect of *Branching* was observed, favoring a left-branching structure (Figure 4, Panel a). The second subgaze durations were also co-determined by large facilitatory effects of *logTrimorphCompFreq* and *logBimorphCompFreq* (Panels b and c) and a facilitatory effect of *logThirdCharFreq* (Panel f). The inhibitory effect of *SecondCharacterStrokes* (Panel h) is reasonable because this is where the eye was fixating.

The lack of the first character frequency effect (Panel d) may be because first character processing was more or less completed and because the eye moved away from the first character. The lack of the second character frequency effect (Panel e) is, however, more counter-intuitive. In bimorphemic compound recognition, both first and second character frequency effects were observed (Miwa, Libben et al., 2014). Similarly, in trimorphemic compound recognition, it is natural to expect all characters to become activated for successful recognition of the word. Yet, this was not the case.

Radical combinabilities and participants' months of stay in Canada were not significant co-determinants of the second subgaze durations either.

(Figure 4 about here)

Lexical statistics: predictability of characters at different positions

Thus far, neither the RT analysis nor the eye movement analyses captured a contribution of *logSecondCharFreq*. In other words, recognition of trimorphemic compounds was achieved with activation of the first and the third characters. This then generates questions as to whether the second character is qualitatively special and why successful trimorphemic compound recognition is achievable with the first and the third characters only. One possible scenario is that native speakers of Japanese have lexical distributional information entrenched in memory and make use of it during word recognition, just as native speakers of Dutch and German make use of lexical distributional statistics of interfixes inside trimorphemic compounds as a probabilistic cue in processing compounds (Krott et al., 2004).

In order to understand this bathtub-like effect, we calculated conditional probabilities for characters at different positions to test the hypothesis that the three characters in trimorphemic compounds are not equally predictable. From the frequency list of Balanced Corpus of Contemporary Written Japanese (BCCWJ) containing 2,434,619 words (National Institute for Japanese Language and Linguistics, 2013), we sampled 112,854 Japanese trimophemic compounds categorized as general nouns. These words were then morphologically segmented by MeCab-0.996 (Kudo, 2013) and the majority of the words were identified to have a left-branching structure (61.8% for left-branching, 24.5% for right-branching, and 13.7% for others). This classification is consistent with the linguistic typological understanding of Japanese as a left-branching language and shows that the vast majority of Japanese trimorphemic compounds have a clear branching structure (i.e., not ambiguous).

In Figure 5, the violin plots depicted with *vioplot* R-package (Adler, 2005) visualize, for the 112,854 trimorphemic compounds, distributions of conditional probability for the first character, the second character, and the third character respectively (Panels a, b, and c). The conditional probability measures predictability of each character when the other characters are known. For the trimorphemic word 動物園 ‘zoo,’ for example, the conditional probabilities of individual characters at different positions are as follows:

$$\mathrm{P}(1|\ 2\ 3) = \mathrm{P}(動|物\ 園) = \mathrm{P}(動 \cap 物\ 園)/\mathrm{P}(物\ 園) = 0.82 \tag{1}$$

$$P(2|\,1\,3) = P(\text{物}|\text{動 園}) = P(\text{物} \cap \text{動 園})/P(\text{動 園}) = 1.00 \tag{2}$$

$$P(3|\,1\,2) = P(\text{園}|\text{動 物}) = P(\text{園} \cap \text{動 物})/P(\text{動 物}) = 0.59 \tag{3}$$

As shown in (1), when the second character 物 and the third character 園 are given, the trimorphemic compound is predicted to be 動物園 82% of the time (it can also be 植物園 'botanical garden'). When the first character 動 and the third character 園 are given as in (2), however, the second character is always 物 without an exception. When the first character 動 and the second character 物 are given as in (3), it is more difficult to predict the trimorphemic compound because there are many possibilities at the third character position: 動物学 'zoology,' 動物界 'animal kingdom,' 動物食 'animal diet,' 動物臭 'animal odor,' 動物名 'animal name' and more. All in all, for the word 動物園, the predictability of the second character is incomparably higher than that of the first and third characters

Interestingly, this seems to be a general trend across all trimorphemic words. The panels (a), (b), and (c) in Figure 5 indicate that conditional probability of the second character (median = .83) is generally higher than that of the first and the third characters (medians = .50 and .16) to a statistically significant extent ($p < 0.01$ in a generalized linear mixed-effects model for the binomial family with a logit link function). This means that, if native speakers of Japanese are given the first and the third characters, they can correctly guess the second character most of the time. These lexical statistics help us to understand why second character frequency effects were not observed at all in the present lexical decision experiment: the second character can be predicted once the first and the third characters are recognized. These lexical statistics also help us to re-interpret Kajii et al.'s (2001) report that Japanese speakers tend to fixate on the first character but not the second character of trimorphemic compounds in sentential reading. Perhaps, readers avoid fixating on the second character (i.e., the word center) because the second character is least informative.

(Figure 5 about here)

The panels (d), (e), and (f) of Figure 5 show conditional probabilities for characters at different positions separately for 69,723 left-branching words and 27,699 right-branching words, excluding 15,432 words that MeCab-0.996 identified to be neither. It is clear that predictability of the first and third characters depends on whether a word is left-branching or right-branching. For right-branching words, the first character is difficult to predict (Panel d), but the third character is highly predictable (Panel f). This pattern reverses for left-branching words: the first character is more or less predictable, but the third character is not. One interpretation is that, when trimorphemic compounds (e.g., 北半球 'northern hemisphere') are morphologically interpreted to be combination of a monomorphemic word (e.g., 北 'north') and a bimorphemic compound (e.g., 半球 'hemisphere'), the monomorphemic word is difficult to predict. It might be the case that the second character is most predictable because it is always part of a bimorphemic compound.

Because conditional probabilities differ greatly between left- and right-branching structures, the response time and eye movement data were reanalyzed with, as covariates, conditional probabilities for the first, second, and third characters. Conditional probability measures, however, did not emerge as significant predictors, and the mixed-effects models reported thus far remained unchanged. The lack of significant additive contribution of conditional probability to on-going eye movements is perhaps expected, given the insignificant additive contribution of conditional co-occurrence probability reported by Ong and Kliegl (2008), although their conditional probability measure was for prediction of an upcoming word in sentential reading. This implies that conditional probability was not computed online for individual compounds in each trial. Instead, conditional probability served as a probabilistic cue to set a global processing strategy for the entire experiment session: it is important to process the first and third characters because they are more informative than the second character.

General discussion

This lexical decision with eye-tracking study revealed that, in recognition of Japanese trimorphemic compounds presented in isolation, the first character and the whole trimorphemic compound unit become active in an early time frame. During this early time frame, visual complexity of the second and third characters shortened fixation durations (see also comparable finding in a saccade amplitude analysis in Appendix B). This facilitatory effect of visual complexity in a parafoveal region can be considered something analogous to magnetic force (Hyönä & Bertram, 2004). According to this account of a parafoveal-on-foveal effect, strength to attract eye movements depends on the distance from the current fixation and the source of complexity. At the first fixation, there was no significant effect of branching direction. The early effect of trimorphemic compound frequencies should deserve attention as well, with respect to Bertram and Hyönä's (2003) visual acuity hypothesis, which claims that longer words are more likely to be processed from parts. In this study, holistic processing was still at work for morphographic trimorphemic compounds from the very first fixation, possibly because Japanese trimorphemic compounds are not long enough to exceed the visual acuity limitation. If so, a question that remains for future research is whether four- or five-multimorphemic words are processed from constituent morphemes.

In a late time frame, after a saccade was made rightward at the second subgaze, the frequency of the trimorphemic compound unit co-determined eye fixation durations to a great extent, accompanying activation of the third character and the bimorphemic compound units. It is apparent that weight of processing shifted to the right and also to the higher level of the morphological representations. Processing advantage for left-branching words was also observed in this later time frame. Recall that the majority of trimorphemic compounds have a left-branching structure, and the recognition system, or more specifically a composition stage, is apparently sensitive to such lexical statistics.

The sum of the above lexical effects in these early and later time frames roughly corresponded to those observed in the response time analysis, in which all the morphological units contributed to lexical decision responses except the second character. Successful recognition of compound words without explicit activation of a constituent may appear to be counter-intuitive. However, the result is compatible with any dual- or multiple-routes models of lexical processing. As de Almeida and Libben (2002) stated, “prelexical activation helps but is not necessary [… because] when constituent activation is disrupted, it doesn't *have the opportunity* to make the difference it otherwise would” (p. 113).

Possible motivations behind the bathtub-like effect

The above results overall show that activation of a complex word does not require activation of its all constituent units. In the case of the present experiment, a second character frequency effect was not observed anywhere. There remains a question: why was the second character not actively processed? Here, we present five possibilities.

First, the importance of peripheral components evokes an association with the bathtub effect in word recall (Aitchison, 1987); the first and last units are recalled better than the middle unit. It is, however, not likely that this was the sole motivation because compounds are too short to be memory-demanding, and the underlying motivation is not expected to be found in short-term memory

Second, the focus on peripheral components also evokes an association with a terminal feature effect in letter identification (Fiset et al., 2008). Terminal features are known to be particularly important when alphabetic letters (e.g., A) are identified, and trimorphemic compounds might have been perceived as a coherent unit like an alphabetic letter. However, this is not likely either because second character was processed at the feature level (Figure 3 and 4, Panel h).

Third, because the motivation for peripheral constituent processing cannot be attributed to short-term memory nor visuoperceptual attention, we considered the morphological transcendence hypothesis (Libben, 2015). From this perspective, two morphologically transcended variants of the second characters are assumed, and they might have been both activated in qualitatively different manners. More specifically, in a triconstituent compound structure composed of constituents ABC, A keeps its morphological role as modifier, no matter what the configuration of the compound (in right-headed languages like Japanese and English), and C keeps its role as head. However, the constituent B needs to change its roles. For left branching, it gives up its head role to be part of a modifier; for right branching, it gives up its modifier role to A, which now functions as the modifier of the ABC compound. To check this hypothesis, we considered positional family sizes of the second character so that we could see whether they interact with the factor *Branching* (e.g., perhaps for left-branching compounds, it is the modifier family size of B in the embedded compound AB that plays a role). However, reanalyses with mixed-effects modeling showed that this was not the case.

Fourth, the lack of any second character frequency effect might have been due to optimization of processing based on conditional probability calculated in each trial. In order to test this hypothesis, we reanalyzed with conditional probabilities of characters at different positions. They, however, did not emerged as significant predictors.

Finally, although conditional probability did not affect processing locally in each trial, it might have affected lexical processing globally throughout the experiment. In a corpus analysis, it became evident that the second character is highly predictable given the other two characters. It might have been the case that such uneven conditional probabilities of characters at different positions were known beforehand and that readers utilized them as part of their processing strategy throughout the experiment. With all other motivation above being less likely, this is at present the best explanation. It is important to note that the bathtub-like effect we observed was not a byproduct of a particular response strategy developed during the experiment because there was no interaction between *logSecondCharFreq* and *Trial* ($p = .98$ in a mixed-effects model).

Revisiting the hypothetical morphological architectures

It is now time to revisit the processing architectures presented in Figure 1. The model tuned for left-branching words (Panel a) is not likely because, at the first fixation, there was no effect of *LogBimorphCompFreq*, although an effect of *LogTrimorphCompFreq* was already observed. The architecture tuned for right-branching words (Panel b) can accommodate early effect of *LogTrimorphCompFreq* without *LogBimorphCompFreq,* assuming that characters were processed from left to right. However, this is not likely either because there was a left-branching advantage at the second fixation. The architecture without any hierarchical morphological structure (Panel c) similarly can accommodate the lack of any *LogBimorphCompFreq* effect at the first fixation. However, this is not perfectly correct either because there were effects of *LogBimorphCompFreq* and *Branching* at the second subgaze. Finally, the both-branching architecture (Panel d) is not likely either because an effect of *LogTrimorphCompFreq* was observed before that of *LogBimorphCompFreq*.

All in all, because none of these architectures is perfectly compatible with the results, it is not possible to pick one model and discard the others. Yet, conversely speaking, each of these models is partially compatible with the results. One way to think about this issue is that early pre-lexical decomposition processes follow the branchless architecture (Panel c) and late post-lexical composition processes are sensitive to lexical statistics favoring a left-branching structure (Panel a) more than a right-branching structure (Panel b). The architecture to build all possible morphological trees (Panel d) seems to be least likely, and this cast another vote to de Almeida and Libben's (2005) conclusion that two morphological structures are not simultaneously built for trimorphemic compounds. Yet, another approach to this issue is to question the linguistic structure or any stable and well-defined representations. Libben (2014, 2015) claims, just as an electron can be paradoxically perceived to be both particles and a wave in a quantum physics, a word, too, may possess duality. In other words, there is no word in the mind, not to mention structures, as a stable or well-defined representational unit. If one is to proceed along this line, a research question is, according to Libben (2014), not "how are compound words represented in the mind?" but "what mental representations correspond to compound words?"

Cross-linguistic differences and similarities

The present study investigated trimorphemic compound processing in a morphographic script, leaving us with a question: what are similar and what are different across different writing systems? Recall that the morphographic script used in Japanese has two unique characteristics: individual symbols are considered to encode morphemes, and there are inter-morpheme spaces in trimorphemic compounds. The former characteristics offers physically fixed word length for compounds consisting of the same number of morphemes. The latter characteristics segment compounds into morphemes and makes the morphological constituents immediately identifiable even at the visuoperceptual level. Given these unique characteristics, at least the earliest processing stage is expected to be language-specific. We predict that facilitatory visual complexity effects of characters in parafoveal regions are language-specific because the second and third character are always found in the same regions in a distinguishable manner, thanks to inter-character space and constant word length.

Although the early lower level processing is expected to language-specific, the higher level cognitive processes are expected to be language-general. In this study, we observed a bathtub-like effect (i.e., frequency effects of the 1st and 3rd, but not 2nd, constituents) and a left-branching advantage. Although specific lexical effects may differ across languages, patterns of lexical processing go hand in hand with lexical distributional statistics and gradually tuned to optimize processing (see Krott et al., 2004 for Dutch and German speakers' awareness of conditional probabilities of interfixes).

In addition, multimorphemic compound processing studies of Kuperman (2008, 2009), Miwa, Libben et al. (2014), and the present study provides a functional overlap; the first constituent's contribution is largest in the early time frame, and a whole compound frequency effect appears before the second/third constituent frequency effects, irrespective of the number of morphemes (i.e., bimorphemic or trimorphemic) and the target language (i.e., Dutch, Finnish, or Japanese).

Revisiting use of eye-tracking in a lexical decision experiment

Following previous lexical decision with eye-tracking studies (Kuperman et al., 2009; Miwa, Dijkstra et al., 2014, Miwa, Libben et al., 2014), the present study, too, demonstrated that an eye-tracking technique successfully complements a lexical decision experiment with response time as a sole dependent variable. Lexical decision has been the most popular task in psycholinguistics. An ideal scenario believed in the field is that what we see in a RT analysis is the sum of what we see in an early fixation analysis and in a late fixation analysis. Interestingly, this was more or less true for the character frequency effects (i.e., Panel b to f in Figure 2 ≈ Panel b to f in Figure 3 + Panel b to f in Figure 4; note that all the plots in Figures 2, 3, 4 have comparable y-axes with an 80-millisecond range). However, this is not always the case. In this study, for example, the effects of stroke counts were not observed in the RT analysis, but they clearly co-determined eye movements. It is possible that the effects appeared and disappeared before a response (in the case of the first and third character stroke effects) or that an early inhibitory effect was cancelled out by a late facilitatory effect (in the case of the second character stroke effect). Inconsistencies are sometimes observed between an eye movement analysis and a response time analysis. This is why combining the multiple measures is valuable.

Limitation of the present study

In this study, responses and eye movements were measured in the context of a visual lexical decision task with words presented in isolation. For future studies, we summarize three obvious limitations of the study.

Task-induced effects. We did not study eye movements in a context of natural reading. We studied recognition processes for trimorphemic compounds presented in isolation by means of eye movements. There were some positive aspects of this design –

(1) the results were not colored by sentence contexts, (2) the first fixation position was perfectly controlled, and (3) the results are more or less generalizable to isolated wording reading in real life (e.g., reading signs, advertisements, and simple descriptions). The design becomes problematic only when one attempts to generalize the results to what is called 'natural' reading (see Kuperman, Drieghe, Keuleers, & Brysbaert, 2013 for a comparison of lexical decision and reading). Because of the nature of the isolated reading task with the lexical decision task demand, span of effective vision might have been changed, and there was no parafoveal preview benefit in this study. We predict that readers are not expected to make as many multiple fixations if the same trimorphemic compounds appear in a sentential context.

Legitimacy of stimuli. In this study, we opted for constructing a stimulus set, as opposed to sampling randomly from a large data base. We confirmed that the set of items used in this study were, although the items were chosen from high frequency words, log-normally distributed and qualitatively comparable to randomly selected samples with respect to word frequencies between left- and right-branching types (see Appendix C for details).

Potentially influential other variables. As Simmons, Nelson, & Simonsohn (2011) warned, there is a researcher degree of freedom with respect to what predictors to be included in analyses. The present study considered, for parsimony, only 12 lexical predictors, one participants' trait, and three task-related variables although there are many other potentially influential lexical variables: for example, familiarity, imageability, semantic transparency, orthographic neighbors, and homophone counts. Some of these variables are known to correlate highly with the predictors used in this study, and investigation of their independent contributions requires further statistical care (e.g., principal component analysis) at the expense of parsimony and interpretation ease. We therefore leave it to future research. Interested readers are encouraged to test efficacy of other predictors, using our response time and eye fixation duration data published online. At the moment, because Miwa and Libben et al.'s (2014) Japanese bimorphemic compound study did not observe these unconsidered variables modulating compound and character frequency effects, we expect that inclusion of these variables is not going to change the pattern of results reported in the present study.

Conclusion

This regression study was a modest attempt to understand how multiple lexical distributional predictors contribute to the visual recognition of Japanese trimorphemic compounds. A lexical decision with eye-tracking experiment revealed several notable phenomena: (1) late left-branching advantage, (2) character-driven, as opposed to radical-driven, bottom-up processing, and (3) uneven and asymmetrical contributions of constituent characters reflected in the bathtub-like constituent frequency effect and the early within-word parafoveal-on-foveal effect. While the character-driven processing and the within-word parafoveal-on-foveal effect were also found in a previous bimorphemic Japanese compound processing study, the left-branching advantage and the bathtub-like effect are specific to trimorphemic compounds and in line with lexical distributional characteristics of trimorphemic compounds found in a corpus. Native speakers of Japanese apparently make use of lexical distributional statistics in the course of recognizing compounds in a brief second.

Acknowledgments

The response time and fixation duration data in a long format, accompanying participant and item properties, are downloadable from the first author’s website (http://kojimiwa.com/publication.html). The descriptions of the variables and data trimming procedures are found in this paper. The frequency data obtained from Balanced Corpus of Contemporary Written Japanese (BCCWJ) are published with permission of the National Institute for Japanese Language and Linguistics. We thank Denis Drieghe and an anonymous reviewer for their comments on the earlier version of this paper.

References

Adler, D. (2005). *vioplot: Violin plot. R package version 0.2.* http://wsopuppenkiste.wiso.uni-goettingen.de/~dadler

Aitchison, J. (1987). *Words in the mind: An introduction to the mental lexicon*. Oxford: Basil Blackwell.

Amano, S., & Kondo, T. (2003). *NTT database series: Lexical properties of Japanese, 2nd release* [CD-ROM]. Tokyo: 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.* 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., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research, 3*, 12-28. doi:10.21500/20112084.807.

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. *arXiv preprint arXiv:1506.04967*.

Bates, D., Maechler, M., Bolker, B., &Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7*. http://CRAN.R-project.org/package=lme4.

Bertram, R. (2011). Eye movements and morphological processing in reading. *The Mental Lexicon, 6*, 83-109. doi:10.1075/ml.6.1.04ber

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

Bertram, R., Kuperman, V., Baayen, R. H., & Hyönä, J. (2011). The hyphen as a segmentation cue in triconstituent compound processing: It's getting better all the time. *Scandinavian Journal of Psychology, 52,* 530-544. doi:10.1111/j.1467-9450.2011.00914.x

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

de Almeida, R. & Libben, G. (2002). Compound pre-access decomposition: Effects of constituent disruption. *Folia Linguistica, 36*, 97-115. doi:10.1515/flin.2002.36.1-2.97

de Almeida, R. G., & Libben, G. (2005). Changing morphological structures: The effect of sentence context on the interpretation of structurally ambiguous English trimorphemic words. *Language and Cognitive Processes, 20,* 373-394. doi:10.1080/01690960444000232

Diependaele, K., Duñabeitia, J. A., Morris, J., & Keuleers, E. (2011). Fast morphological effects in first and second language word recognition. *Journal of Memory and Language, 64*, 344-358. doi:10.1016/j.jml.2011.01.003

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

Drieghe, D. (2011). Parafoveal-on-foveal effects in eye movements during reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *Oxford Handbook on Eye Movements* (pp. 839-855). New York: Oxford University Press

Fiset, D., Blais, C., Éthier-Majcher, C., Arguin, M., Bub, D., & Gosselin, F. (2008). Features for identification of uppercase and lowercase letters. *Psychological Science, 19,* 1161-1168. doi:10.1111/j.1467-9280.2008.02218.x

Forster, K. I. (1966). Left-to-right processes in the construction of sentences. *Journal of Verbal Learning and Verbal Behavior, 5,* 285-291. doi:10.1016/S0022-5371(66)80032-4

Forster, K. I. (1968). Sentence completion in left- and right-branching languages. *Journal of Verbal Learning and Verbal Behavior, 7,* 296-299. doi:10.1016/S0022-5371(68)80004-0

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

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambigous sentences. *Cognitive Psychology, 14*, 178-210. doi:10.1016/0010-0285(82)90008-1

Frost, R., Grainger, J., & Carreiras, M. (2008). Advances in morphological processing: An introduction. *Language and Cognitive Processes, 23*, 933-941. doi:10.1080/01690960802321925

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

Hino, Y., Kusunose, Y., Lupker, S. J., & Jared, D. (2013). The processing advantage and disadvantage for homophones in lexical decision task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 529-551. doi:10.1037/a0029122

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

Inhoff, A. W., Radach, R., & Heller, D. (2000). Complex compounds in German: Interword spaces facilitate segmentation but hinder assignment of meaning. *Journal of Memory and Language, 42*, 23-50. doi:10.1006/jmla.1999.2666

Kageyama, T. (2010). Variation between endocentric and exocentric word structures. *Lingua, 120*, 2405–2423. doi:10.1016/j.lingua.2010.04.005

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

Kudo, T. (2013). MeCab: Yet another part-of-speech and morphological analyzer. Retrieved May 30, 2015, from http://taku910.github.io/mecab/

Krott, A., Libben, G., Jarema, G., Dressler, W., Schreuder, R., & Baayen, H. (2004). Probability in the grammar of German and Dutch: Interfixation in triconstituent compounds. *Language and speech, 47*, 83-106. doi:10.1177/00238309040470010401

Kuperman, V., Bertram, R., & Baayen, R. H. (2008). Morphological dynamics in compound processing. *Language and Cognitive Processes, 23*, 1089-1132. doi:10.1080/01690960802193688

Kuperman, V., Drieghe, D., Keuleers, E., & Brysbaert, M. (2013). How strongly do word reading times and lexical decision times correlate? Combining data from eye movement corpora and megastudies. *Quarterly Journal of Experimental Psychology, 66,* 563-580. doi:10.1080/17470218.2012.658820

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*. http://CRAN.R-project.org/package=lmerTest

Libben, G. (1993). Are morphological structure computed during word recognition? *Journal of Psycholinguistic Research, 22*, 535-544. doi:10.1007/BF01068253

Libben, G. (1994). The role of hierarchical morphological structure: A case study. *Journal of Neurolinguistics, 8*, 49-55. doi:10.1016/0911-6044(94)90006-X

Libben, G. (2006). Why study compound processing? An overview of the issues. In G. Libben & G. Jarema (Eds.), *The representation and processing of compound words* (pp. 1–22). New York: Oxford University Press.

Libben, G. (2014). The nature of compounds: A psychocentric perspective. *Cognitive Neuropsychology, 31,* 8-25. doi:10.1080/02643294.2013.874994

Libben, G. (2015). *The quantum metaphor and the organization of words in the mind*. Unpublished manuscript. Department of Applied Linguistics, Brock University, St. Catharines, Canada.

Libben, G., Gibson, M., Yoon, Y. B., & Sandra, D. (2003). Compound fracture: The role of semantic transparency and morphological headedness. *Brain and Language, 84,* 50–64. doi:10.1016/S0093-934X(02)00520-5

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

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7*, 19-40. doi:10.1037/1082-989X.7.1.19

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

Marelli, M., & Luzzatti, C. (2012). Frequency effects in the processing of Italian nominal compounds: modulation of headedness and semantic transparency. *Journal of Memory and Language, 66,* 644-664. doi:10.1016/j.jml.2012.01.003

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2015). Balancing Type I Error and Power in Linear Mixed Models. *arXiv preprint arXiv:1511.01864*.

McClelland, J. L., & Patterson, K., (2002a). 'Words or Rules' cannot exploit the regularity in exceptions. *Trends in Cognitive Sciences, 6*, 464-465. doi:10.1016/S1364-6613(02)02012-0

McClelland, J. L., & Patterson, K., (2002b). Rules or connections in past-tense inflections: what does the evidence rule out? *Trends in Cognitive Science, 6*, 465-472. doi:10.1016/S1364-6613(02)01993-9

Miwa, K., Dijkstra, T., Bolger, P., & Baayen, R. H. (2014). Reading English with Japanese in mind: Effects of frequency, phonology, and meaning in different-script bilinguals. *Bilingualism: Language and Cognition,17,* 445-463. doi:10.1017/S1366728913000576

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

Nakayama, M., Sears, C. R., Hino, Y., & Lupker, S. J. (2014). Do masked orthographic neighbor primes facilitate or inhibit the processing of Kanji compound words. *Journal of Experimental Psychology: Human Perception and Performance, 40*, 813-840. doi:10.1037/a0035112

National Institute for Japanese Language and Linguistics. (2013). *BCCWJ Word List.* Retrieved July, 15, 2015, from http://pj.ninjal.ac.jp/corpus_center/bccwj/freq-list.html

Ong, J. K., & Kliegl, R. (2008). Conditional co-occurrence probability acts like frequency in predicting fixation durations. *Journal of Eye Movement Research, 2*, 1-7. doi:10.16910/jemr.2.1.3

Pinker, S., & Ullman, M. T. (2002a). The past and future of the past tense. Trends in *Cognitive Sciences, 6*, 456-463. doi:10.1016/S1364-6613(02)01990-3

Pinker, S., & Ullman, M. T. (2002b). Combination and structure, not gradedness, is the issue. *Trends in Cognitive Sciences, 6,* 472-474. doi:10.1016/S1364-6613(02)02013-2

Pollatsek, A, Drieghe, D., Stockall, L., & de Almeida, R. G. (2010). The interpretation of ambiguous trimorphemic words in sentence context. *Psychonomic Bulletin & Review, 17*, 88-94. doi:10.3758/PBR.17.1.88

Pollatsek, A., Hyönä, J., & Bertram, R. (2000). The role of morphological constituents in reading Finnish compound words. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 820-833. doi:10.1037/0096-1523.26.2.820

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

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

Spencer, A. (1997). Morphological theory and English. *Links & Letters, 1*, 71-84.

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

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *The Quarterly Journal of Experimental Psychology, 57,* 745-765. doi:10.1080/02724980343000477

Taft, M., & Forster, K. I. (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., & Nguyen-Hoan, M. (2010). A sticky stick? The locus of morphological representation in the lexicon. *Language and Cognitive Processes, 25,* 277-296. doi:10.1080/01690960903043261

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. (2005). The effect of morphemic homophony on the processing of Japanese two-kanji compound words. *Reading and Writing, 18,* 281-302. doi:10.1007/s11145-005-3354-0

Tamaoka, K. (2007). Rebounding activation caused by lexical homophony in the processing of Japanese two-kanji compound words. *Reading and Writing, 20*, 413-439. doi:10.1007/s11145-006-9036-8

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

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

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S. Fourth edition.* Springer, New York. ISBN 0-387-95457-0

Whaley, L. J. (1997). *Introduction to typology: The unity and diversity of language*. Thousand Oaks, CA: SAGE Publications.

White, S., Hirotani, M., & Liversedge, S. P. (2011). Eye movement behaviour during reading of Japanese sentences: Effects of word length and visual complexity. *Reading and Writing, 25*, 981-1006. doi:10.1007/s11145-010-9289-0

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

York, R. (2012). Residualization is not the answer: Rethinking how to address multicollinearity. *Social Science Research, 41*, 1379–1386. doi:10.1016/j.ssresearch.2012.05.014

Zhou, X., Marslen-Wilson, W., Taft, M., & Shu, H. (1999). Morphology, orthography, and phonology in reading Chinese compound words. *Language and Cognitive Processes, 14*, 525-565. doi:10.1080/016909699386185

Table 1. Descriptive statistics of lexical predictors considered in this study. The mean was calculated before the centering procedure. Comp = compound, Char = character, Trimorph = Trimorphemic compound, Bimorph = bimorphemic compound

Lexical predictors	Type	Range	Levels, Mean (SD)
Branching	Branching		LeftBranching, RightBranching
LogTrimorphCompFreq	Trimorph	2.57 : 7.87	5.01 (1.1)
LogBimorphCompFreq	Bimorph	3.22 : 11.12	7.93 (1.5)
LogFirstCharFreq	Character	6.91 : 13.39	10.84 (1.2)
LogSecondCharFreq	Character	6.88 : 13.39	10.96 (1.3)
LogThirdCharFreq	Character	7.81 : 13.14	10.78 (1.2)
LogFirstCharRadicalCombinability	Radical	0 : 4.69	2.77 (1.3)
LogSecondCharRadicalCombinability	Radical	0 : 4.69	2.91 (1.4)
LogThirdCharRadicalCombinability	Radical	0 : 4.69	2.80 (1.2)
FirstCharacterStrokes	Feature	1 : 20	8.59 (3.7)
SecondCharacterStrokes	Feature	1 : 21	8.68 (3.6)
ThirdCharacterStrokes	Feature	3 : 21	9.63 (3.8)

Table 2. The fixed-effects structure of the linear mixed-effects model for response times

	Estimate	Std.Error	t-value	p-value
(Intercept)	-1.874	0.034	-53.682	< 0.001
Trial	-0.045	0.008	-5.103	< 0.001
PreviousResponseCorrect (Incorrect)	0.096	0.017	5.308	< 0.001
invPreviousRT	0.133	0.017	7.100	< 0.001
logTrimorphCompFreq	-0.053	0.008	-6.006	< 0.001
logBimorphCompFreq	-0.015	0.006	-2.726	0.007
logFirstCharFreq	-0.018	0.006	-2.392	0.018
logThirdCharFreq	-0.015	0.006	-2.290	0.023

Table 3. The fixed-effects structure of the (generalized) linear mixed-effects models for fixation counts, first fixation durations, and second subgaze durations.

(a) Fixation counts

	Estimate	Std.Error	t-value	p-value
(Intercept)	0.730	0.029	25.247	< 0.001
Trial	-0.029	0.010	-2.840	0.005
logThirdCharFreq	-0.028	0.009	-2.945	0.003

(b) First fixation durations

	Estimate	Std.Error	t-value	p-value
(Intercept)	5.650	0.033	166.759	< 0.001
Trial	0.012	0.005	1.951	0.066
PreviousResponseCorrect (Incorrect)	-0.035	0.011	-2.847	0.004
logTrimorphCompFreq	-0.005	0.002	-2.603	0.009
logFirstCharFreq	-0.008	0.003	-2.923	0.007
FirstCharacterStrokes	0.011	0.001	7.376	< 0.001
SecondCharacterStrokes	-0.004	0.001	-3.894	< 0.001
ThirdCharacterStrokes	-0.002	0.001	-2.578	0.010

(c) Second subgaze durations

	Estimate	Std.Error	t-value	p-value
(Intercept)	19.021	0.385	46.153	< 0.001
logPreviousFixationDuration	-4.646	0.117	-35.189	< 0.001
Trial	-0.413	0.049	-7.319	< 0.001
PreviousResponseCorrect (Incorrect)	1.205	0.271	3.761	0.001
invPreviousRT	1.025	0.121	7.173	< 0.001
Branching (RightBranching)	0.383	0.146	2.120	0.036
logTrimorphCompFreq	-0.318	0.071	-4.711	< 0.001
logBimorphCompFreq	-0.174	0.052	-3.344	0.001
logThirdCharFreq	-0.157	0.057	-2.428	0.016
FirstCharacterStrokes	-0.025	0.020	-1.601	0.114
SecondCharacterStrokes	0.054	0.018	2.775	0.006

Figure captions

Figure 1. Hypothetical processing architectures for Japanese trimorphemic compounds. Note: To be concise, sub-character radical representations and the response/decision system are assumed but not depicted.

Figure 2. Partial effects of lexical predictors in the mixed-effects model fitted to inverse-transformed response times. Note: The rugs represent a distribution of predictor values. The RTs were back-transformed to the original scale for the visualization. Trimorph = trimorphemic compound, Bimorph = bimorphemic compound.

Figure 3. Partial effects of lexical predictors in the mixed-effects model fitted to log-transformed first fixation durations. Note: The rugs represent a distribution of predictor values. The first fixation durations were back-transformed to the original scale for the visualization.

Figure 4. Partial effects of lexical predictors in the mixed-effects model fitted to square-root-transformed second fixation durations. Note: The rugs represent a distribution of predictor values. The second fixation durations were back-transformed to the original scale for the visualization. A line is drawn in Panel (a) simply to make the effect more visible, and it is does not imply any continuity between the *LeftBranching* and *RightBranching* types.

Figure 5. Violin plots showing distributions of conditional probability for characters at different positions. Note: The white dots and the corresponding values represent a median.

Figure 1

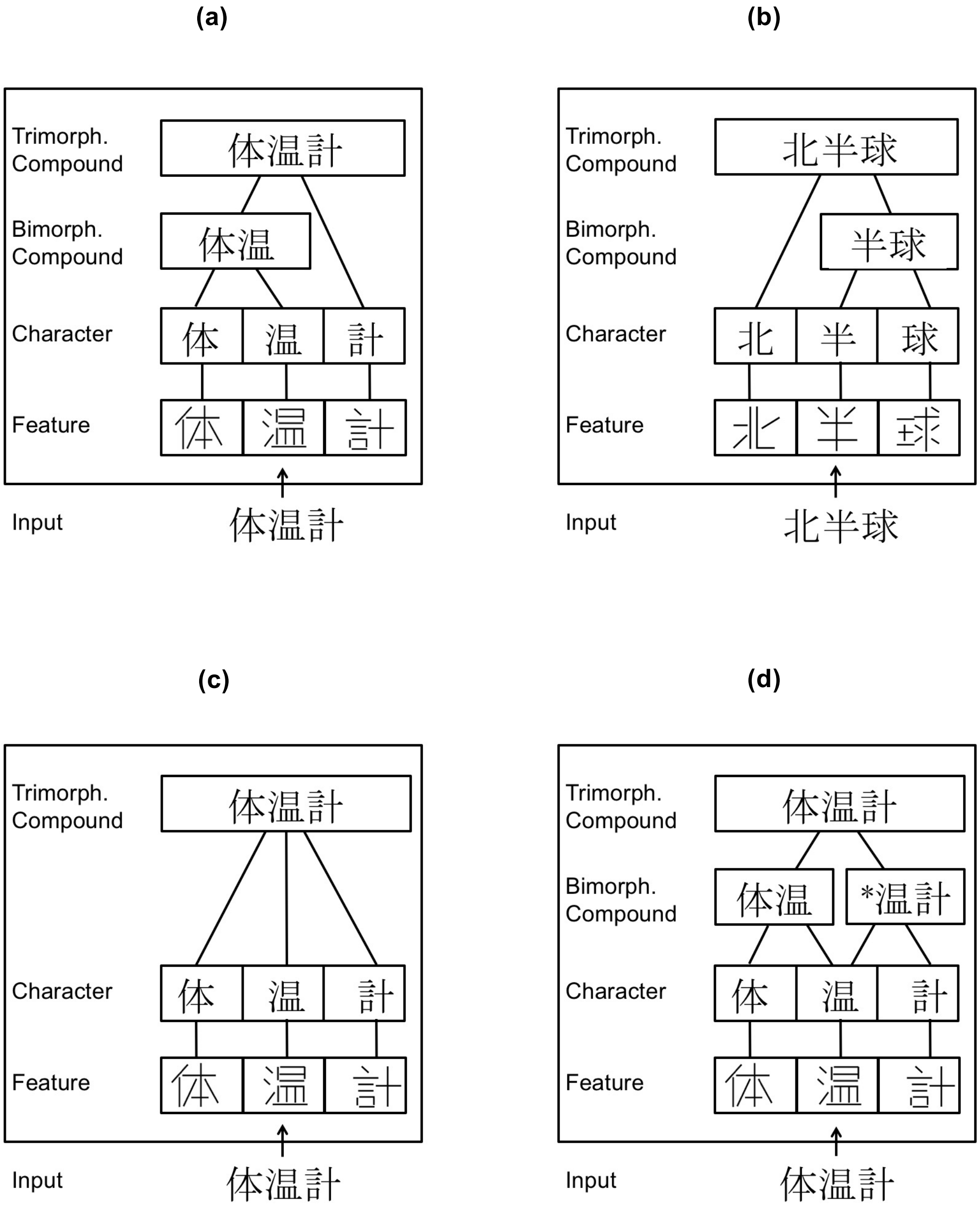

(a)
Trimorph. Compound
体温計
Bimorph. Compound
体温
Character
体
温
計
Feature
体
温
計
Input
体温計
(b)
Trimorph. Compound
北半球
Bimorph. Compound
半球
Character
北
半
球
Feature
北
半
球
Input
北半球
(c)
Trimorph. Compound
体温計
Character
体
温
計
Feature
体
温
計
Input
体温計
(d)
Trimorph. Compound
体温計
Bimorph. Compound
体温
*温計
Character
体
温
計
Feature
体
温
計
Input
体温計

Figure 2

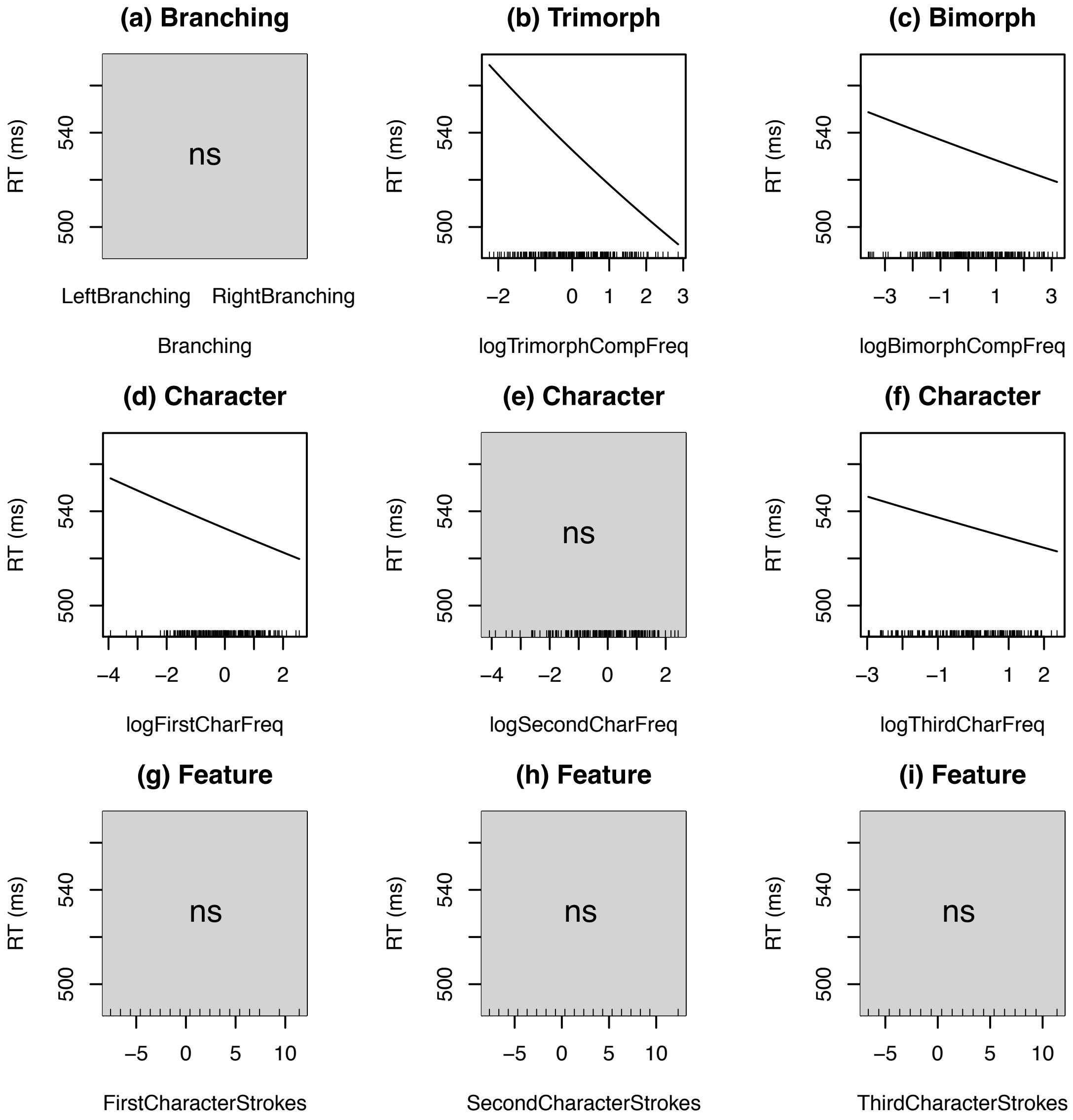

(a) Branching
RT (ms)
500
540
ns
LeftBranching
RightBranching
Branching
(b) Trimorph
−2
0
1
2
3
logTrimorphCompFreq
(c) Bimorph
−3
−1
1
3
logBimorphCompFreq
(d) Character
−4
−2
0
2
logFirstCharFreq
(e) Character
logSecondCharFreq
(f) Character
−3
−1
1
2
logThirdCharFreq
(g) Feature
−5
0
5
10
FirstCharacterStrokes
(h) Feature
SecondCharacterStrokes
(i) Feature
ThirdCharacterStrokes

Figure 3

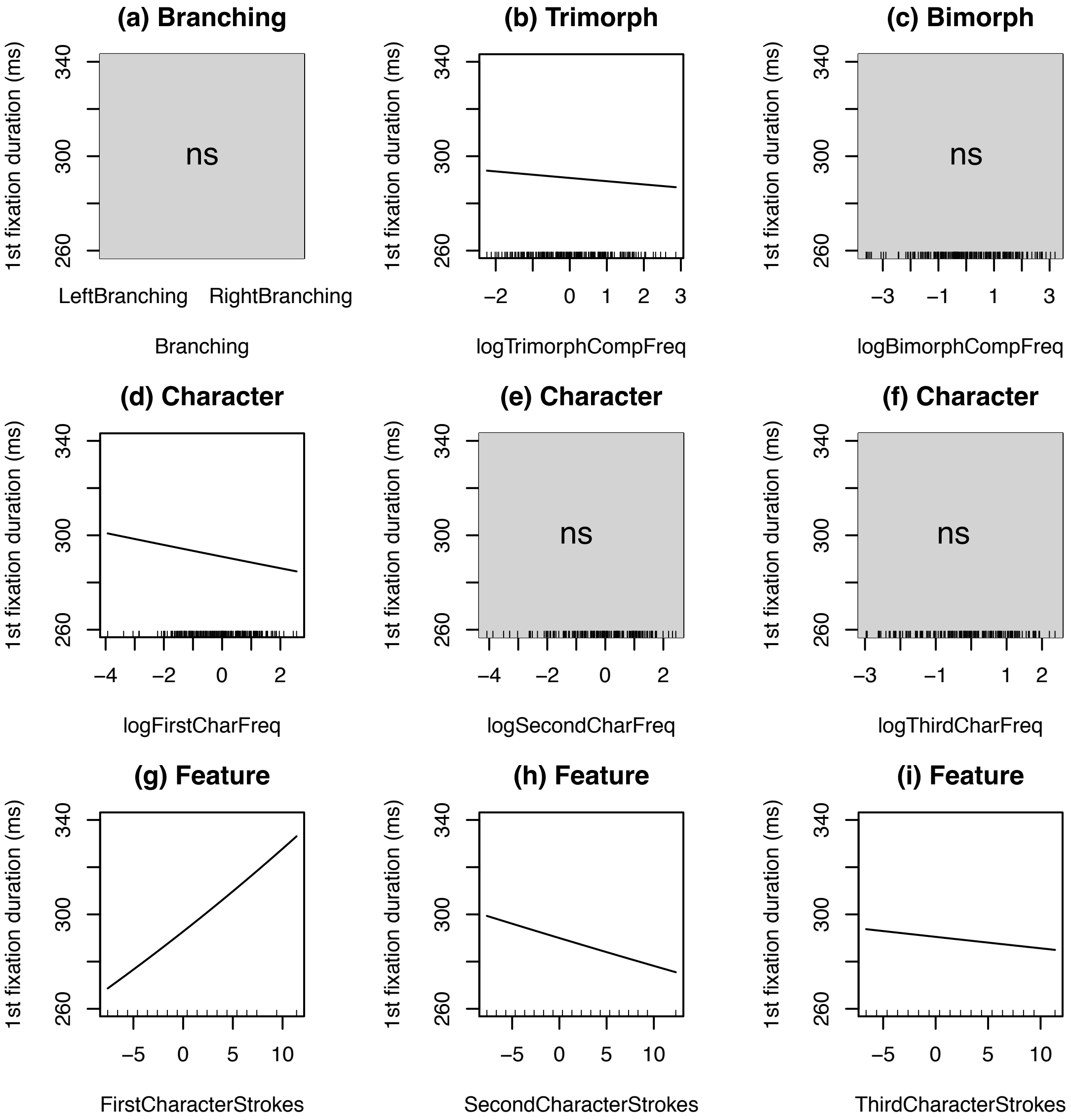

(a) Branching
1st fixation duration (ms)
340
300
260
ns
LeftBranching
RightBranching
Branching
(b) Trimorph
1st fixation duration (ms)
340
300
260
−2
0
1
2
3
logTrimorphCompFreq
(c) Bimorph
1st fixation duration (ms)
340
300
260
ns
−3
−1
1
3
logBimorphCompFreq
(d) Character
1st fixation duration (ms)
340
300
260
−4
−2
0
2
logFirstCharFreq
(e) Character
1st fixation duration (ms)
340
300
260
ns
−4
−2
0
2
logSecondCharFreq
(f) Character
1st fixation duration (ms)
340
300
260
ns
−3
−1
1
2
logThirdCharFreq
(g) Feature
1st fixation duration (ms)
340
300
260
−5
0
5
10
FirstCharacterStrokes
(h) Feature
1st fixation duration (ms)
340
300
260
−5
0
5
10
SecondCharacterStrokes
(i) Feature
1st fixation duration (ms)
340
300
260
−5
0
5
10
ThirdCharacterStrokes

Figure 4

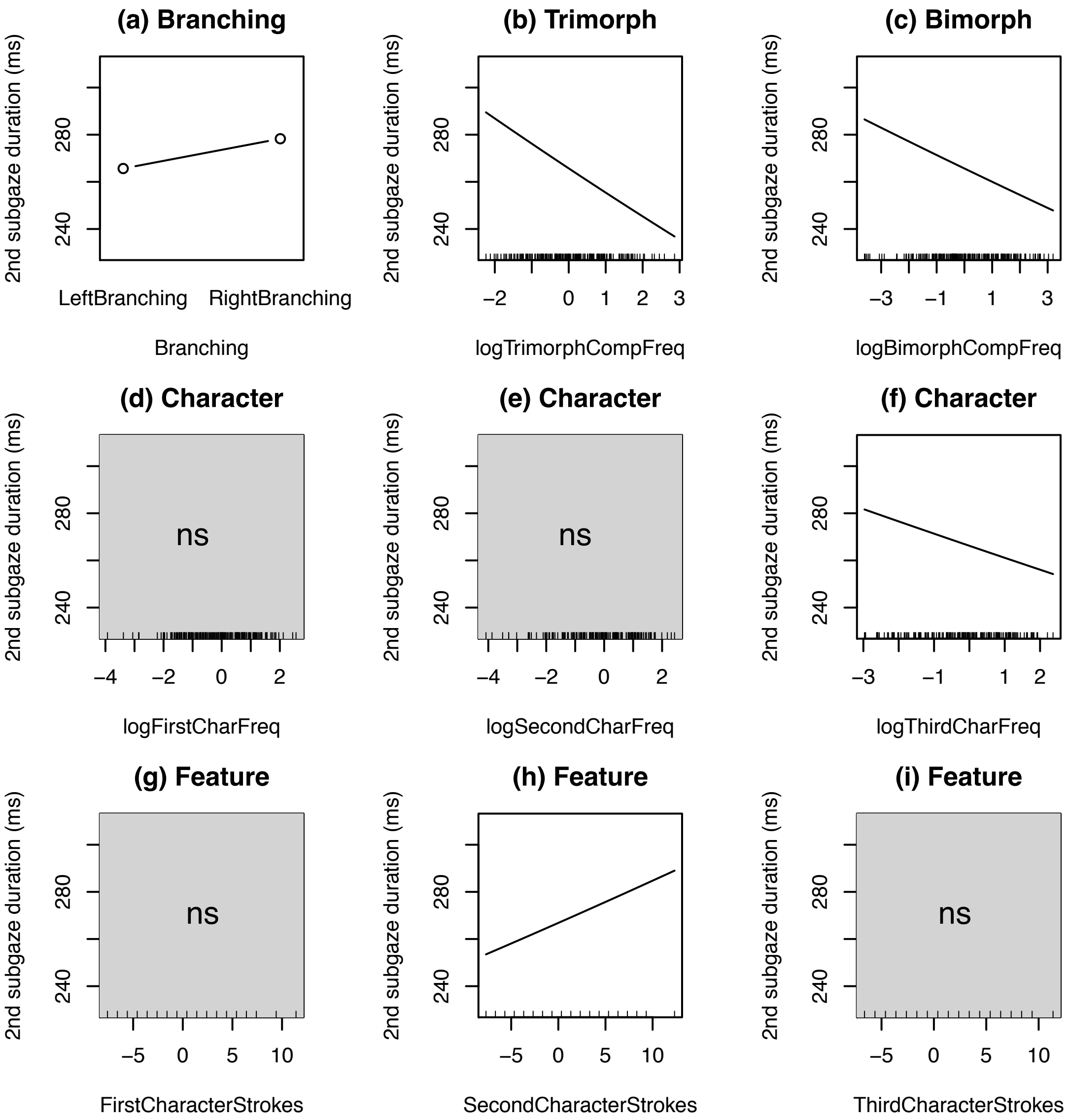
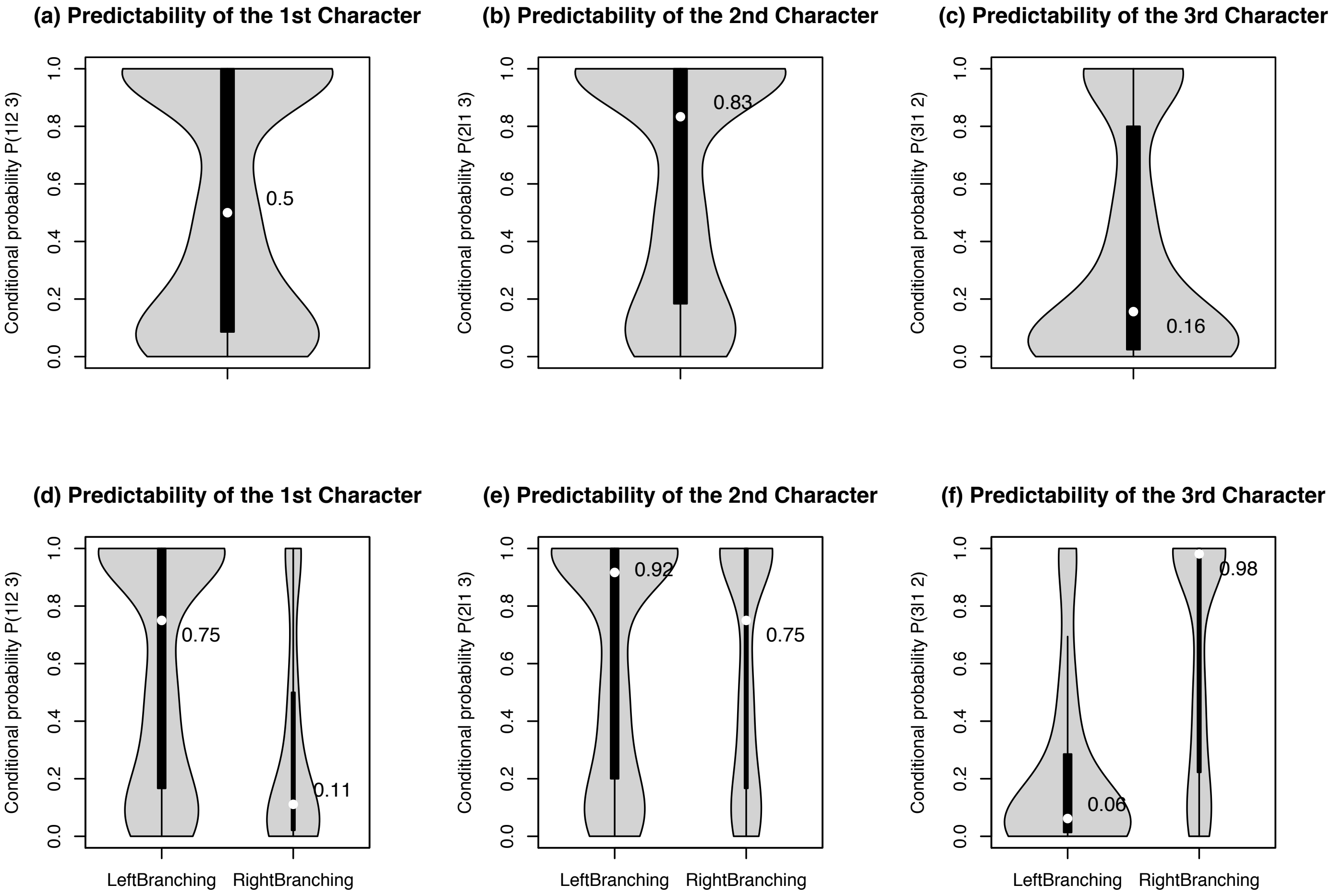

(a) Branching
2nd subgaze duration (ms)
240
280
LeftBranching
RightBranching
Branching
(b) Trimorph
2nd subgaze duration (ms)
240
280
−2
0
1
2
3
logTrimorphCompFreq
(c) Bimorph
2nd subgaze duration (ms)
240
280
−3
−1
1
3
logBimorphCompFreq
(d) Character
2nd subgaze duration (ms)
240
280
ns
−4
−2
0
2
logFirstCharFreq
(e) Character
2nd subgaze duration (ms)
240
280
ns
−4
−2
0
2
logSecondCharFreq
(f) Character
2nd subgaze duration (ms)
240
280
−3
−1
1
2
logThirdCharFreq
(g) Feature
2nd subgaze duration (ms)
240
280
ns
−5
0
5
10
FirstCharacterStrokes
(h) Feature
2nd subgaze duration (ms)
240
280
−5
0
5
10
SecondCharacterStrokes
(i) Feature
2nd subgaze duration (ms)
240
280
ns
−5
0
5
10
ThirdCharacterStrokes

Figure 5


(a) Predictability of the 1st Character
Conditional probability P(1|2 3)
0.0
0.2
0.4
0.6
0.8
1.0
0.5
(b) Predictability of the 2nd Character
Conditional probability P(2|1 3)
0.0
0.2
0.4
0.6
0.8
1.0
0.83
(c) Predictability of the 3rd Character
Conditional probability P(3|1 2)
0.0
0.2
0.4
0.6
0.8
1.0
0.16
(d) Predictability of the 1st Character
Conditional probability P(1|2 3)
0.0
0.2
0.4
0.6
0.8
1.0
0.75
0.11
LeftBranching
RightBranching
(e) Predictability of the 2nd Character
Conditional probability P(2|1 3)
0.0
0.2
0.4
0.6
0.8
1.0
0.92
0.75
LeftBranching
RightBranching
(f) Predictability of the 3rd Character
Conditional probability P(3|1 2)
0.0
0.2
0.4
0.6
0.8
1.0
0.06
0.98
LeftBranching
RightBranching