University of Alberta

Emergence of wordlikeness in the mental lexicon: Language, population, and task effects in visual word recognition

by

Koji Miwa

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Linguistics

Fall 2013

Edmonton, Alberta

ABSTRACT

Various aspects of our higher-level cognition affect the buttom-up information uptake in perception of objects, faces, and scenes. Such interplay between new information and existing information in our memory can be seen also in rapid visual word recognition. Lexical processing architectures proposed to date, however, have been based mostly on studies with specific characteristics: those investigating monolingual English speakers reading English words, with a lexical decision task demand, and with response times as the primary dependent variable (Libben & Jarema, 2002). Phenomena consistently observed across different linguistic characteristics, individuals, and tasks must surely reflect the core of human language processes (i.e. functional overlap). In this dissertation, I investigated consequences of testing different language, population, and task on visual word recognition processes in three studies: primed Japanese kanji lexical decision with Japanese monolinguals (Chapter 2), eye-tracking Japanese kanji lexical decision with Japanese monolinguals (Chapter 3), and eye-tracking English lexical decision with Japanese-English bilinguals, who possess knowledge of orthographically different languages (Chapter 4). The three studies collectively show that language-specific properties, individual differences, and variable task demands, by themselves, do not result in completely different pictures with respect to how wordlikeness emerges in visual word recognition.

Keywords: visual word recognition, morphological processing, lexical decision, eye-tracking, bilingual processing, mixed-effects modeling

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Current issues in visual word processing

Various aspects of our higher-level cognition (i.e., knowledge, previous experience, expectation, a goal in a given task) affect the buttom-up information uptake (e.g., Nisbett & Masuda, 2003, on cultural effects on attention; Tong & Nakayama, 1999, on familiarity effects in face recognition; Yabus, 1967, on goal-driven eye-movement patterns in scene recognition). Such interplay between new information and existing information in our memory can be seen also in rapid visual word recognition.

Irrespective of whether it is a newspaper, a novel, or a comic book, readers must identify words in the texts rapidly in order to accomplish efficient reading. Previous studies suggest that word recognition is achieved, within several hundred milliseconds, in the mind by automatically activating and evaluating a large number of lexical properties of a perceived word (as reflected in word frequency, age of acquisition, orthographic complexity, and imageability). Perception of a word also triggers an orchestration of facilitatory and inhibitory effects among a set of competing word candidates in memory (Baayen et al., 2006; Balota et al., 2004). In the mind of bilingual readers, the dynamics of lexical activation extends to words in the other language, adding a further layer of complexity to the already complex nature of word recognition processes. Previous studies indicate that the perceived word automatically activates cross-linguistically similar words in the other language with regard to orthographic form, pronunciation, and meaning (e.g., Dijkstra et al., 1999, 2010, with Dutch-English bilinguals), supporting a bilingual word recognition architecture that automatically activates words of two languages integrated in a single lexical database in memory (e.g., Dijkstra & van Heuven, 2002).

Architecture of visual word recognition system

A model serves a crucial role to interpret known lexical effects and to further predict new phenomena. A 'classic' model for visual word recognition is the Interactive Activation (IA) model for monolingual readers (McClelland & Rumelhart, 1981). This is a localist model with nodes for word units explicitely represented in the model architecture, as opposed to a distributed model without such nodes. It assumes that, in an alphabetic writing system, letter and word representations are activated via the identification of visuoperceptual features. For the Latin alphabet, a simple feature set consisting of 14 horizontal, vertical, and diagonal line features can code 26 letters. Features activate the letters in which they occur, which in turn activate words of which they are part. Activated word candidates are assumed to mutually inhibit each other (called 'lateral inhibition') and feed activation back to the letter representation level (called 'top-down feedback'). In the model, the resting levels of activation for word nodes (i.e., when there is no input) are determined by the words' frequency of occurrence. High frequency words have a higher resting level of activation and therefore have a head start in the recognition process and they have more power than low frequency words to inhibit lexical alternatives. Although it is a localist model with explicit representations for words, the activation of words is not all or none due to the presence of feature and letter nodes: its activation is gradient and accompanies activation of words sharing some of the same features and letters. In other words, words, which are initially meaningless visual features, become more and more word-like as time goes by. For example, when the input word MIND is fed to the IA model, its activation level increases as time goes by while accompanying activation of orthographic neighbours FIND, KIND, and WIND all decrease (Figure 1-1, simulated in jIAM implemented by van Heuven, 2008).

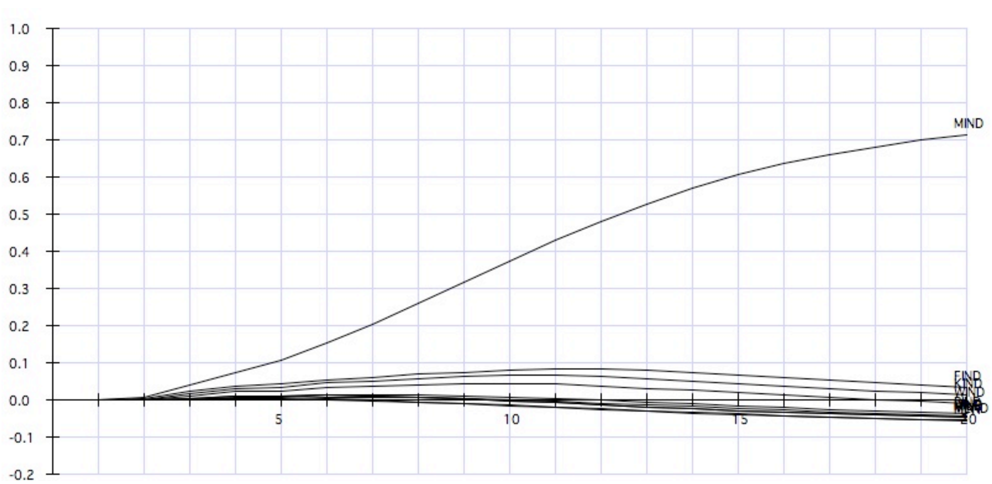


Figure 1-1. The emergence of wordlikeness for the word MIND accompanying activation of orthographic neighbours in the IA model (jIAM, van Heuven, 2008)

While the classical IA model (McClelland & Rumelhart, 1981) is concerned exclusively with orthographic processing, a multiple read-out model by Jacobs, Rey, Ziegler, and Grainger (1998; MROM-P) extended the theoretical framework of localist connectionist modeling to phonological processes as well. Such extension is necessary because phonological information contributes to the recognition of visually presented words (Carreiras, Ferrand, Grainger, & Perea, 2005; Ferrand & Grainger, 1992, 1994; Perfetti, Zhang, & Berent, 1992).

The Bilingual Interactive Activation (BIA) model (Dijkstra & van Heuven, 1998, 2002; Dijkstra, van Heuven, & Grainger, 1998) is a localist connectionist model that extends the monolingual IA model and allows us to conceptualize monolingual and bilingual lexical processes within one theoretical framework. Because the model directly incorporates the IA model in its architecture, it can account for within-language lexical effects in monolinguals. In addition, because the model incorporates both an L2 (second language) and an L1 (first language) lexicon as an integrated whole, it can also account for cross-language lexical effects. While the original BIA model was limited to orthographic lexical aspects (Dijkstra & van Heuven, 1998; Dijkstra, van Heuven, & Grainger, 1998), it has been extended to account for experimental evidence on cross-language phonological and semantic activation. In addition, the extended and slightly revised BIA+ model (Dijkstra & van Heuven, 2002) attempts to explain cross-task variations by incorporating a task/decision system explicitly in the model architecture. Currently, this extra-linguistic decision system is not expected to affect activation in the word identification system, based on the experimental evidence that bilinguals automatically activate two languages or word representations even when this language non-selective activation is not necessary in the task or can even be detrimental (Dijkstra, de Bruijn, Schriefers, & ten Brinke, 2000; Dijkstra & van Hell, 2003; Dijkstra, van Jaarsveld, & ten Brinke, 1998; van Assche et al., in press; van Hell & Dijkstra, 2002).

The three key factors in mental lexicon research

A decade ago, Libben and Jarema (2002) pointed out three key factors in research on the mental lexicon (or the lexical database in memory). These are languages, populations, and tasks depicted in the following organizational framework (Figure 1-2).

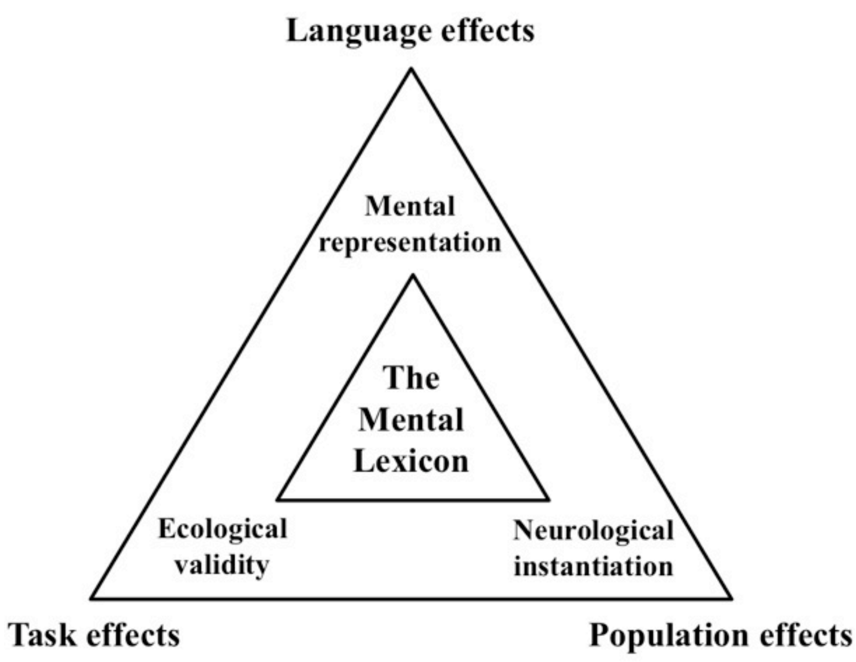


Figure 1-2. A Framework for the mental lexicon research adapted from Libben and Jarema (2002)

Libben and Jarema (2002) studied 58 articles in a special issue of *Brain and Language* and revealed a certain trend: A typical study on the mental lexicon investigated monolingual English speakers reading English words in a context of lexical decision, with response times as the primary dependent variable. In each trial in a lexical decision experiment, readers make a binary decision, as quickly and accurately as possible, regarding whether the presented string of letters is a word or a nonword. Because different words are recognized at different speeds systematically affected by various lexical properties, lexical decision is indeed a useful experimental paradigm to examine how wordlikeness emerges in the course of visual word recognition. However, the past line of research has centred around a certain language (i.e., English) and population (i.e., monolinguals) and needs to be extended to more varieties. In the studies reported below, I investigated consequences of testing a different language, population, and task on visual word recognition processes: primed Japanese *kanji* lexical decision with Japanese monolinguals (Chapter 2), eye-tracking Japanese *kanji* lexical decision with Japanese monolinguals (Chapter 3), and eye-tracking English lexical decision with Japanese-English bilinguals, who possess knowledge of orthographically different languages (Chapter 4). Consideration of diversity in these three factors is crucial for understanding the nature of the human language processing system, which is assumed to be language-, population-, and task-invariant at least at the most abstract level.

Language effects: English alphabet and Japanese *Kanji*/*Katakana*

Seventeeen languages were studied in the 58 articles reviewed by Libben and Jarema (2002), and among these two-thirds were Indo-European languages. Because languages may differ in any domains, such as phonetics/phonology, morphology, syntax, semantics, and orthography, it is important to investigate whether lexical phenomena observed in Indo-European languages can be generalized to other non-Indo-European languages.

Chapters 2 and 3 are devoted to visual word recognition in the Japanese logographic script *kanji*. Unlike an English alphabetic symbol, which is considered to represent a phoneme, each Japanese logographic *kanji* symbol represents a word (e.g., 木 *ki* 'tree'). Some logographic symbols are configurationally simplex (e.g., e.g., 木 *ki* 'tree', 水 *mizu* 'water') while others are configurationally complex (e.g., 海 *umi* 'sea', 枝 *eda* 'branch'). When logographic symbols consist of more than two parts, it is often possible to identify a *semantic radical*, a component denoting an approximate meaning, and a *phonetic radical*, a component denoting a pronunciation. While all logographic *kanji* symbols carry a meaning, not all the symbols occur in isolation. For example, 電 always occurs together with another symbol, as in 電気 *denki* 'electricity' and 停電 *teiden* 'blackout'. In the sense that logographic symbols do not necessarily denote words but do always denote morphemes, Japanese logography is more accurately described as *morphography*.

It is expected that visually dense morphographic symbols may require more processing in early visuo-perceptual processes. However, this multilayered morphographic script may also affect the way morphologically complex words are processed in higher-level cognition and provides an opportunity to reconsider a language-invariant definition of a *morpheme*.

How complex words are processed in the mind is not a trivial issue because it relates to the long-lasting debate on cognitive efficiency with respect to how the mind balances representation and computational loads (McClelland & Patterson, 2002b; Pinker & Ullman, 2002a, 2002b). Intuitively-speaking, if complex words and their morphemes are all represented in the mind, complex words do not necessarily have to be broken down into parts when they are read (i.e., more representation, less computation). If, on the other hand, only morphemes are represented, complex words have to be broken down into morphemic units everytime they are recognized (i.e., less representation, more computation).

Studies in alphabetic languages collectively suggest that morphemes are automatically activated in reading morphologically complex words but that all lexical units are still represented, at least when the issue is conceptualized in a localist manner (Frost, Grainger, & Carreiras, 2008; Giraudo & Grainger, 2001; Taft, 1994), resulting in a hypothesis that the human language processing mechanism is characterized by representational and computational redundancy (Libben, 2006). These findings are, however, colored by characteristics of the Indo-European languages and may not generalize to processing of other languages.

If a *morpheme* is defined as a *minimal meaningful processing unit*, the question then arises: are Japanese semantic radicals automatically processed, similarly to morphemes in alphabetic languages? Chapter 2 tests this prediction. Chapter 3 further investigates the time course of processing of the radicals, the characters, and the word in reading Japanese two-character words to re-evaluate strictly bottom-up processing models proposed for English and Chinese (Taft, 1994; Taft, Zhu, & Peng, 1999).

Population effects: Monolinguals and bilinguals

At the time of Libben and Jarema's (2002) review, two-thirds of the studies tested adult native speakers in one language. However, the same word may not necessarily be processed in an identical manner by monolinguals and bilinguals. Given that the large majority of the current world's population is bilingual, it is intriguing to ask to what extent the past line of evidence for exhaustive lexical activation by monolingual readers generalizes to readers with two lexical databases in the mind. There have been a growing number of studies investigating bilinguals, and these studies collectively suggest that exhaustive lexical activation extends to the non-target language (see Dijkstra & van Heuven, 2002). Yet, the majority of these studies tested bilinguals of languages with the same script (e.g., Dutch-English and French-English bilinguals), and bilinguals of languages with different scripts, such as Japanese-English bilinguals, may perform differently. Chapter 4 tests this prediction of the Bilingual Interactive Activation model (the BIA+, Dijkstra and van Heuven, 2002). While almost all previous studies on bilingual processing tested bilingual readers only (but see Dijkstra, Grainger, & van Heuven, 1999, as an exception), Chapter 4 also tests monolingual readers and bilingual readers using the same stimuli in the same experimental design. In order to evaluate the assumption of the BIA model that the bilingual processing system is a straightforward extension of the monolingual processing system, it is crucial to assess whether observed effects are genuine bilingual effects generated from the bilingual processing system or by-products of target word processing per se.

Task effects: Beyond response times in lexical decision

Libben and Jarema (2002) reported that close to half of the studies they reviewed used lexical decision experiments with response times as the primary dependent variable. However, as in any experimental paradigm, end-point responses in lexical decision experiments may be colored by particular response strategies because lexical decision is a discrimination task as well as a recognition task. One way to cope with potential task effects is to compare lexical effects across different tasks and identify what lexical effects are shared across different tasks and which thus may be more central components of human language architecture (*functional overlap*, Grainger & Jacobs, 1996). Another approach, taken by the following studies in Chapters 3 and 4, is to tap into finer processes in lexical decision beyond the end-point response times and distinguish early supposedly automatic lexical processes and late processes that may be colored more by conscious response strategies. In both Chapters 3 and 4, lexical decision experiments were conducted with eye-tracking. The superimposed eye-tracking provided time-locked early and late eye fixation information at multiple sites within a word. Under the assumption that task-specific strategic processes follow but do not affect early automatic lexical processes (see the BIA+ model in Dijkstra and van Heuven, 2002 and Schmid, 1982 for the assumption that motor response planning and response execution take place late), the early fixation durations should be relatively uncontaminated by conscious starategies for lexicality judgment.

In addition to tracking how task effects contribute in the time-couse of lexical activation, it is also important to track how task-specific response strategies modulate lexical processes throughout an experiment. It is expected that, at the beginning of an experiment after being exposed to some practice trials, readers set certain response strategies and use them globally throughout the experiment. The patterns of responses may also be modulated locally by preceding responses either by auto-correlated time series of responses due to consistency in behaviour (Baayen & Millin, 2010) or by adjusting response criteria (Perea & Carreiras, 2003). For this reason, studies in Chapters 3 and 4 also studied task-effects by tracking how response times and eye fixation durations changed throughout the experiments and how they were affected by response times in the preceding trial.

Statistical considerations

Problems in pre-experimental control designs

Consideration of the above three key factors should go hand in hand with consideration of statistical techniques. For the last several decades, the field of psycholinguistics has been driven primarily by a pre-experimental control approach or the so-called standard experimental approach (i.e., controlling for as many potentially influential variables as possible except one or a few variables of interest to the researcher). Such experimental design has almost necessarily involved dichotomizatin of numerical variables (e.g., categorization of words' frequency of occurrences into high and low groups). Although it is often said that correlation is not causation and that causation can be identified only by such standard experimental approaches with rigid pre-experimental control, this is an idealization with assumptions that all the influential factors have been exhaustively identified for a given phenomenon and that pre-experimental matching for these variables does not induce any statistical nor interpretative biases. In reality, the past line of research has not yet identified all influential factors co-determining visual word recognition process nor clarified what factors are more important than others (except the robust whole word frequency effect always emerging as the significant co-determinant of response speed in visual word recognition experiments). Given the larger number of potentially influential lexical variables (e.g., over 40 variables in the English lexicon project, Balota et al., 2007; over 20 variables in Chapter 3) and the ever-growing number of variables to be considered (e.g., movie subtitle frequency, Brysbaert

& New, 2009a, 2009b; orthographic Levenshtein distance, Yarkoni, Balota, & Yap, 2008), pre-experimental matching, which was already considered problematic three decades ago (Cutler, 1981), is even more problematic today: Extensive pre-experimental control leads to a quantitatively smaller sample size and qualitatively biased samples. From a statistical standpoint, Baayen (2010) warns that pre-experimental matching on numerical covariates may further lead to substantial loss of statistical power (see also Cohen, 1983; MacCallum, Zhang, Preacher, & Rucker, 2002 for the negative consequence of dichotomization of numerical variables).

Proposed solutions

In this thesis, I opted for state-of-the-art mixed-effects statistical modeling, which can assess lexical effects, individual reader effects, and task effects simultaneously in a single statistical model without pre-experimental control (Baayen, Davidson, & Bates, 2008). Although the magnitude of individual reader effects and task effects is generally large relative to language effects, mixed-effects modeling can fortunately capture small yet significant language effects. This technique also allows us to test whether there are idiosyncratic individual differences by means of participant-specific random slopes for lexical effects, in addition to random slopes for different participants (see Andrews & Lo, 2011; Baayen & Milin, 2010; Kuperman & Van Dyke, 2011; Yap, Balota, Sibley, & Ratcliff, 2011, for individual differences on lexical effects). Importantly, because this approach does not require pre-experimental matching but relies on post-experimental statistical control with influential variables explicitly assessed in a regression model, it was possible to start with a relatively large number of lexical stimuli sampled randomly from large lexical databases (708 items in Chapter 3 and 250 items in Chapter 4). Since experimenters, regardless of their level of expertise, can induce experimenter bias when constructing stimuli (Forster, 2000), it is important to minimize experimenters' involvement in preparing experimental materials.

Whether to control confounding effects before or after an experiment, it is crucial to understand the correlational structure among lexical variables. For example, longer words tend to be used less frequently than shorter words (i.e., word length and word frequency negatively correlate) and words used frequently in one language tend to be used equally frequently in another language (i.e., word frequeny of one language positively correalate with frequency of the translation equivanent in another language). It is not uncommon to find strong correlations among multiple lexical variables. In such cases, it is not straightforward to identify an independent contribution of a particular variable (i.e., the problem of *multicollinearity*). When pre-experimental matching is done without considering such correlational structures, the selected materials may not accompany naturally occurring characteristics (e.g., highly frequent long words), which then generate a concern whether such biased samples allow us to come to a generalized conclusion regarding how language is represented and processed in the mind.

Even when we opt for a post-experimental statistical control using regression modeling, the problem of multicollinearity persists. A linear regression model provides an accurate estimate if all the independent variables are indeed independent from each other. When there are multiple independent variables that correlate with each other, one possibility is to limit the number of variables to be considered. However, in order not to ignore any potential influential variables, I opted for a residualization technique in Chapters 3 and 4 and a principal component analysis and a random forest analysis in Chapter 2.

In the next three sections, I present consequences of testing different language, population, and task on visual word recognition processes: primed Japanese kanji lexical decision with Japanese monolinguals (Chapter 2), eye-tracking Japanese kanji lexical decision with Japanese monolinguals (Chapter 3), and eye-tracking English lexical decision with Japanese-English bilinguals, who possess knowledge of orthographically different languages (Chapter 4).

CHAPTER 2

Semantic radicals in Japanese two-character word recognition [†]

Research on morphological processing suggests that complex words are not recognized simply by full-form-to-meaning matching nor are they recognized only through feed-forward combinatorial computations (Bertram, Baayen, & Schreuder, 2000; Frost, Grainger, & Carreiras, 2008; Frost, Grainger, & Rastle, 2005; Libben, 1998; Rastle, Davis, & New, 2004; Taft & Kougious, 2004). Instead, both computational efficiency and storage efficiency seem to be optimized simultaneously (e.g., Kuperman, Schreuder, Bertram, & Baayen, 2009; Libben, 2006).

This optimization of storage and computation should also apply to reading of Japanese and Chinese, languages with a morphographic writing system. In these languages, a large majority of words are represented orthographically by means of two complex characters. There is clear evidence for character activation in the recognition of two-character words (see Joyce 2002; Kawakami, 2002; Tamaoka & Hatsuzuka, 1995, 1998 for Japanese, and Huang, Lee, Tsai, Hung, & Tzeng, 2006; Ji & Gagné, 2007; Zhou, Marslen-Wilson, Taft, & Shu, 1999 for Chinese). Tamaoka and Hatsuzuka (1995) reported that, independently of whole word frequency, the right character's frequency speeds up responses in

[†] A version of this chapter has been published. Miwa, Libben, & Baayen, (2011). Semantic radicals in Japanese two-character word recognition. *Language and Cognitive Processes, 27*, 142-158. This research was supported by a Major Collaborative Research Initiative Grant (#412-2001-1009) from the Social Sciences and Humanities Research Council of Canada to Gary Libben (Director), Gonia Jarema, Eva Kehayia, Bruce Derwing, and Lori Buchanan (Co-

a two-character word lexical decision. Effects of characters' meanings (Tamaoka & Hatsuzuka, 1998) and an effect of the conceptual relation governing the interpretation of two character compounds (Ji & Gagné, 2007) suggest that, as expected, the characters in a two-character compound mediate lexical processing. However, no study on two-character word recognition has assessed the contribution of sub-morphemic components, the orthographic morphemes unique to morphographic orthography known as semantic radicals.

A majority of characters are composed of two radicals: a *semantic radical*, a semantic constituent encoding a basic category meaning, and a *phonetic radical*, a phonological constituent. Phonetic radical activation was witnessed in studies addressing the processing of single-characters (Hsu, Tsai, Lee, & Tzeng, 2009; Lee, Tsai, Huang, Hung, & Tzeng, 2006). The present study focuses on the role of semantic radicals. Semantic radicals function as entries in character dictionaries. They also provide useful classification cues in learning the 1,006 characters taught in primary education (Ministry of Education, Culture, Sports, Science and Technology, 2009).

Experimental evidence for semantic radical activation comes from single-character decision (Feldman & Siok, 1997 on Chinese), single-character decision with priming (Feldman & Siok, 1999 on Chinese), speeded single-character semantic-categorization (Flores d'Arcais & Saito, 1993 on Japanese), and single-character word naming (Flores d'Arcais, Saito, & Kawakami, 1995 on Japanese). Previous research has shown that characters with a semantic radical

investigators). The authors are indebted to Laurie Feldman and Hisashi Masuda for constructive discussion.

occurring in many other characters are read faster. Taft and Zhu (1997) and Saito (1997) assume that semantic radicals play a similar role in the reading of two-character words. As we shall see this assumption is only partially validated by our study, which addresses the role played by semantic radicals in the ecologically important context of two-character words. More specifically, the present study seeks to clarify whether the effects of semantic radicals depend on their position in the left (modifier) versus the right (head) character. We also aim to clarify the role of the semantic transparency of semantic radicals, and to establish the extent to which radical type frequency effects are independent of age of acquisition (AoA, see, e.g., Juhasz, 2005). Finally, we investigate how the commonly-used partial priming manipulation affects complex word recognition.

We implemented an analysis of covariance design, combining several numerical predictors with a factorial treatment, overt priming, as most studies addressing lexical processing in Japanese and Chinese using behavioral measures made use of priming manipulations (see Feldman & Siok, 1999; Joyce, 2002). For the statistical analysis, we made use of a mixed-effects model complemented by random forests (a conditional inference tree-based ensemble method, see, e.g., Strobl, Boulesteix, Kneib, Augustin, & Zeileis, 2008; Strobl, Malley, & Tutz, 2009). This combined approach allows us to evaluate both significance and magnitude of semantic radicals' contribution to response speed relative to the contribution of character and whole word properties.

Method

Participants

Thirty native speakers of Japanese (23 females, mean age = 28.5, *SD* = 7.9) were recruited as paid participants at the University of Alberta and neighboring cities in Alberta, Canada.

Apparatus

The experiment was run with PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993) using a Macintosh iMac computer and an iBook computer operating under OS 9.2. The *S* and *L* keys on a Macintosh keyboard were used for lexical decision responses.

Materials

Forty-six prime-target pairs of two-*kanji* words were constructed. Prime and target pairs shared the semantic radical of their right character, as previous research has shown that the right character, the head, co-determines lexical decision latencies to a greater extent than the left character (Libben, Gibson, Yoon, & Sandra, 2003; Tamaoka & Hatsuzuka, 1995). Forty-six two-character pseudo-homophonous nonwords were prepared by replacing the first character of existing two-character words by another existing homophonic character. We divided the 46 critical word pairs into two sets (A and B) of 23 word pairs each. One group of participants was presented with the words of set A paired with primes sharing the semantic radical, with the words from set B paired with control primes that did not share the radical. A second group of participants was presented with the words from set B paired with primes sharing the right semantic radical and the words from set A paired with non-matching control primes. During the experiment, therefore, participants encountered 92 stimuli in total: 23 primed pairs, 23 control pairs, and 46 nonwords (See Table 2-1. Appendix 2-1 lists the entire list of primed and control pairs). A given word was presented only once to a given participant. Mean semantic similarity, gauged by Latent Semantic Analysis scores (Landauer, Foltz, & Laham, 1998) for the translated prime-target pairs, was 0.07 for both primed and unprimed conditions [$t(45) = 0.12$, $p = 0.91$].

Table 2-1. Types of prime-target pairs used in the present study

Condition	Prime			Target			
	Word	Translation	Radical	Word	Translation	Radical	Shared
Primed	時計	*toke* ‘clock’	言 *gonben*	書記	*shoki* ‘scribe	言 *gonben*	Yes
Control	救急	*kinkyu* ‘emergency’	心 *kokoro*	書記	*shoki* ‘scribe	言 *gonben*	No
Nonword	印刷	*insatsu* ‘fireplace’	刂 *ritto*	渋症	*jusho* ‘N/A’	疒 *yamaidare*	No

Procedure

A trial consisted of a fixation point (*) presented at the centre of the display for 1,000 ms, followed by a 230 ms prime, followed by a backward mask (##) of 200 ms to avoid visual-overlap effects, after which the target word appeared and remained on the screen until participants responded by pressing one of two keys on the keyboard. Participants were instructed to decide, as quickly and accurately as possible, whether the second word (i.e. the target) is an existing word in Japanese by pressing the *L* key for words and *S* key for nonwords. The prime word was always an existing two-character word. The fixation point and masks were presented in Times New Roman 48, and the words in Mincho 48.

Results

Statistical analyses in this study were carried out by using `R` version 2.9.2 (R Development Core Team, 2009). One participant with 43% error rate was excluded from the data analysis. Stimuli that elicited RTs shorter than 300 ms or longer than 3,000 ms were also excluded (8 data points). Furthermore, four target words that elicited more than 30% error rate (113 target data points, 8.5% of the data) were excluded from the analysis. The mean error rate for the remaining 1,213 target responses was 8% (1,121 correct, 92 incorrect). For these 1,213 data points, the quantiles of the target error rates were 0% (minimum), 0% (1st quartile), 5% (median), 14% (3rd quartile), and 28% (maximum). The corresponding quantiles of the subject error rates were 0% (minimum), 5% (1st quartile), 5% (median), 12% (3rd quartile), and 26% (maximum). A reciprocal transformation was applied to RTs (-1000/RT) to remove the skew characterizing the distribution of the raw RTs. Only correct responses, 1,121 data points, were considered for the response time analyses. All predictors with a noticeably skewed distribution were logarithmically transformed.

Table 2-2. Eighteen original predictors of target words and their loadings on the seven target PCs

Type	Predictors	PC1*	PC2	PC3*	PC4*	PC5*	PC6	PC7*
S	LeftKanjiStrokes	0.05	-0.23	-0.13	**0.42**	0.04	**-0.39**	**-0.33**
S	LeftKanjiRadicalStrokes	-0.03	-0.24	**-0.36**	-0.01	-0.19	**-0.44**	-0.16
S	RightKanjiStrokes	0.28	0.09	0.04	**-0.33**	-0.25	-0.12	**-0.36**
S	RightKanjiRadicalStrokes	**0.30**	0.09	0.24	-0.17	-0.29	-0.24	-0.07
R	LogLeftKanjiRadicalCombinability	0.06	0.22	0.27	**0.47**	0.21	-0.04	-0.27
R	LogLeftKanjiRadicalTokenFreq	-0.04	**0.41**	0.16	**0.33**	0.24	-0.13	-0.10
R	LogRightKanjiRadicalCombinability	**-0.38**	-0.01	-0.11	0.17	-0.22	0.23	-0.13
R	LogRightKanjiRadicalTokenFreq	**-0.38**	-0.10	-0.03	0.08	-0.20	0.31	0.00
R	LogRightKanjiOtherRadicalFreq	0.11	0.13	**0.30**	0.14	-0.19	-0.20	**0.64**
R	RightKanjiRadicalTransparency	0.03	0.17	-0.06	**0.31**	**-0.49**	0.25	-0.07
R	RightKanjiRadicalUsefulness	0.23	0.21	0.08	0.14	**-0.50**	0.12	-0.12
C	LogLeftKanjiNeighbour	-0.19	**0.41**	-0.22	-0.15	-0.08	-0.20	0.13
C	LogLeftKanjiTokenFreq	-0.22	**0.43**	-0.09	-0.17	0.10	-0.14	-0.14
C	LeftKanjiAoA	**0.30**	**-0.34**	0.23	0.10	0.10	0.21	-0.01
C	LogRightKanjiNeighbour	-0.29	-0.21	0.24	0.01	-0.23	-0.17	-0.12
C	LogRightKanjiTokenFreq	-0.29	-0.14	**0.44**	-0.11	-0.07	-0.13	-0.08
C	RightKanjiAoA	**0.33**	0.10	**-0.30**	-0.03	0.11	**0.31**	-0.13
W	LogWholeWordTokenFreq	-0.09	0.10	**0.37**	**-0.33**	0.11	0.20	**-0.37**
Variance accounted for by each PC		0.23	0.16	0.13	0.11	0.09	0.07	0.05
Cumulative variance accounted for		0.23	0.39	0.52	0.63	0.72	0.79	0.84

Note. The * mark represents significantly influential predictors in the regression model summarized in Table 2-3. The bolded values represent relatively large loadings (exceeding 0.30 or smaller than -0.30). S = Stroke, R = Radical, C = Character, W = Word

Table 2-2 lists the predictors in our model, ordered by linguistic levels. At the level of radicals, the level of our primary interest, previous studies have shown that radicals used across many characters are processed faster (Feldman & Siok, 1997, 1999; Taft & Zhu, 1997). Following the terminology used by Feldman and Siok (1997, 1999), radical combinability counts (*LogLeftKanjiRadicalCombinability, LogRightKanjiRadicalCombinability*) represent the type count of basic *kanji* characters sharing a given semantic radical with any of the other 1,945 basic *kanji* characters. Radical token frequency measures (*LogLeftKanjiRadicalTokenFreq, LogRightKanjiRadicalTokenFreq*) represent the cumulative token frequency of the *kanji* characters sharing a given semantic radical (Tamaoka, Kirsner, Yanase, Miyaoka, & Kawakami, 2002; Yokoyama, Sasahara, Nozaki, and Long, 1998).

Three further radical measures were considered: the position of the semantic radical in the right character (*RightKanjiRadicalPosition*, *left* vs. *other*) and type frequency of the non-semantic radical in the right character (*LogRightKanjiOtherRadicalFreq*, Saito, Kawakami, & Masuda, 1995, 1997). Finally, we considered a factor distinguishing between prime-target pairs sharing the semantic radical in the same position and those in which the semantic radical appeared at different positions (*PrimeTargetRadicalPositionConsistency*). The positional measures did not reach or even approach significance in our analysis (see also Figure 2-2), and hence will not be reported below.

We obtained two semantic measures for the right *kanji* radicals. *RightKanjiRadicalTransparency* is a measure based on twenty-one native

Japanese readers' evaluation on a seven-point scale of the congruity between the meaning of the character and the meaning of the component radical. *RightKanjiRadicalUsefulness* gauges how useful a given semantic radical is in predicting the character meaning. Thirty native Japanese readers were given a sheet of paper with single-characters with their semantic radical portion visible and all other portions ink-blobbed. This task measured the independent meaningfulness of semantic radicals. The rating scores in *RightKanjiRadicalTransparency* and *RightKanjiRadicalUsefulness* were significantly correlated ($r = 0.53$, $p < 0.01$). Since the reliability of the semantic information provided by a given semantic radical varies across the characters of the language, we expect that semantic radicals with greater independent meaningfulness play a more substantial role in reading.

As predicted at the level of features, we encoded constituent complexity using stroke counts (*PrimeRightKanjiStrokes, LeftKanjiStrokes, LeftKanjiRadicalStrokes, RightKanjiStrokes, RightKanjiRadicalStrokes*). At the character level, we considered written token frequency (Amano & Kondo, 2003) for the left and right characters of target words (*LogRightKanjiTokenFreq, LogLeftKanjiTokenFreq*) and for the right character of the prime words (*LogPrimeRightKanjiTokenFreq*) whose semantic radical is the target of the present the priming manipulation. The character neighbour measures (*LogPrimeRightKanjiNeighbour, LogLeftKanjiNeighbour, LogRightKanjiNeighbour*) represent a given character's position-specific morphographic family size (Joyce & Ohta, 2002). Our age of acquisition (AoA) measures (*PrimeRightKanjiAoA, LeftKanjiAoA, RightKanjiAoA*) represents the school grades at which the characters in our prime and target words are first taught. They are an objective measure for AoA. Lastly, we considered the token frequency of the whole word (*LogWholeWordTokenFreq,* Amano & Kondo, 2003).

Given substantial multicollinearity, we made use of two statistical techniques: the parametric technique of mixed effects modeling with orthogonalization of predictors through a principal component analysis (PCA, Belsley, Kuh, & Welch, 2004) and a non-parametric technique, random forests.

Principal component regression analysis

Principal component orthogonalization was applied separately to the eighteen target word properties and to the five prime word properties. We then selected those PCs that accounted for at least 5% of the variance. Among target PCs, the top seven PCs cumulatively accounted for 84.1% of the variance (See Table 2-2). For the prime PCs, the three accounted for 88.4% of the variance. These PCs were entered as predictors in a mixed-effects analysis of covariance (Baayen, Davidson, & Bates, 2008; Bates, Maechler, & Dai, 2007). A backward stepwise variable selection procedure identified five target PCs as significant: PC1, PC3, PC4, PC5, and PC7.

Table 2-3 lists the estimate, standard error, upper and lower limits of Highest Posterior Density (HPD) confidence interval, as well as t-values and p-values based on Markov chain Monte Carlo (MCMC) sampling for these five

PCs. Each PC's contribution is visualized in Figure 2-1. In order to clarify the nature of the PCs, we inspected the loadings of the original lexical predictors on these PCs. Table 2-2 presents the predictors with the largest loadings in bold (the predictors with high loadings exceeding 0.30 or smaller than -0.30 are shown in bold). In our discussion, we highlight only those predictors that have the greatest loadings and carry the main trends. We will complement this discussion with a random forest analysis that evaluates the contribution of individual predictors in a non-parametric way.

Table 2-3. Influential principal components (PCs) with their estimate, standard error, upper and lower limits of HPD confidence interval, t-value and p-value based on 10,000 Markov chain Monte Carlo (MCMC) samples from the posterior distributions of the parameters.

	Estimate	Std.Error	HPD95lower	HPD95upper	t-value	pMCMC
(Intercept)	-1.3898	0.0389	-1.4550	-1.3247	-35.77	0.0001
TargetPC1	0.0428	0.0104	0.0242	0.0613	4.10	0.0001
TargetPC3	-0.0355	0.0138	-0.0110	-0.0603	-2.57	0.0064
TargetPC4	0.0789	0.0149	0.1052	0.0518	5.30	0.0001
TargetPC5	0.0227	0.0176	0.0539	0.0092	1.29	0.1482
Condition (primed)	-0.0322	0.0153	-0.0627	-0.0020	-2.10	0.0356
TargetPC7	0.0643	0.0217	0.0265	0.1033	2.97	0.0014
TargetPC5 x Condition (primed)	-0.0376	0.0130	-0.0134	-0.0642	-2.90	0.0034

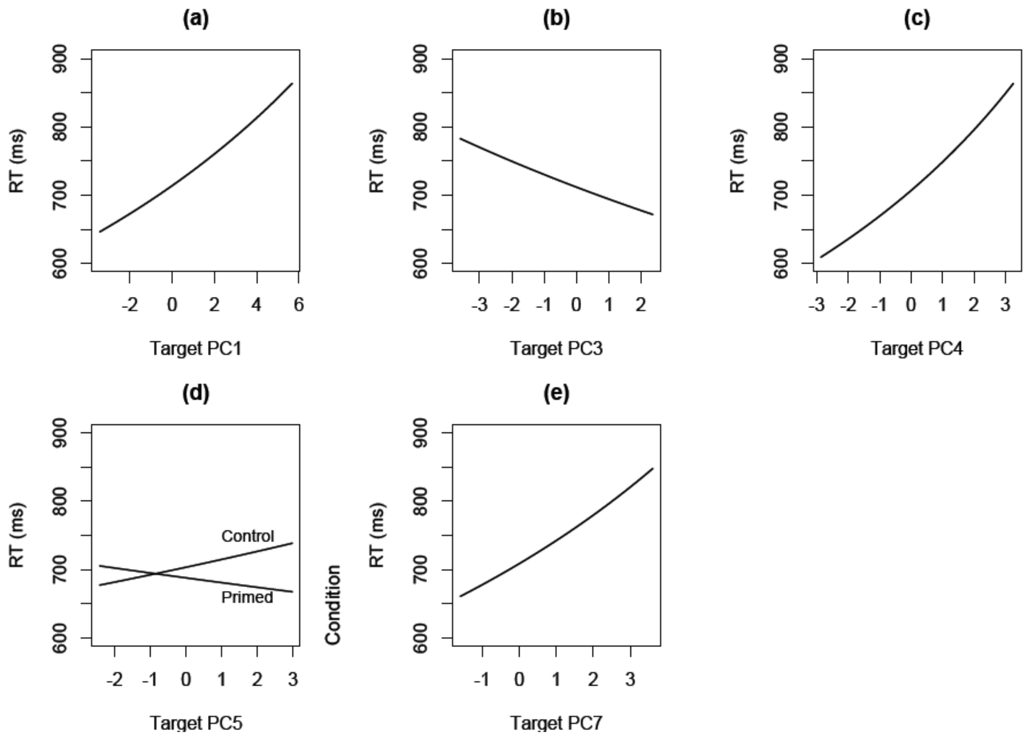


Figure 2-1. Partial effects of four influential principal components (PCs) on response latencies in the primed lexical decision. PC1: Larger right character semantic radical combinability/ token frequency (negative loadings on PC1) indicate shorter RTs (panel a); PC3: Larger word/ right character frequency (positive loadings on PC3) indicate shorter RTs (panel b); PC4: Larger left character semantic radical combinability/ token frequency (positive loadings on PC4) slow responses (panel c); PC5: Larger right character radical transparency/ usefulness (negative loadings on PC5) indicate shorter RTs in the control condition but longer RTs in the primed condition (panel d). PC7: Larger right character non-semantic radical type frequency (positive loadings on PC7) indicates longer RTs.

The inhibitory predictor PC1 (Figure 2-1 panel a), effect size (range) 218 ms, quantifies primarily the combinability and token frequency of the right character's semantic radical (*LogRightKanjiRadicalCombinability* and *LogRightKanjiRadicalTokenFreq*). Since these predictors correlate negatively with an inhibitory PC, their effect on RTs is facilitatory. Right characters that contain a semantic radical with high combinability or high token frequency are read faster. In contrast to the above facilitatory predictors, *RightKanjiAoA* loaded positively on PC1 and hence their effects on RTs are inhibitory. Words with a character that has been taught at a later age are read less quickly, as expected.

PC3 is a facilitatory predictor (panel b, effect size 111 ms). Both *LogWholeWordTokenFreq* and *LogRightKanjiTokenFreq* loaded positively on PC3 (see Table 2-2) and hence indicate facilitation. *LeftKanjiRadicalStroke* has a large negative loading on PC3, indicating increasing RTs with increased orthographic complexity of the semantic radical of the left character.

PC4 is an inhibitory predictor (panel c, effect size 254 ms), and it is characterized by *LeftKanjiStrokes* and *LogLeftKanjiRadicalCombinability*, which have the largest positive loadings, indicating inhibition. The inhibitory effect of *LogLeftKanjiRadicalCombinability* on the RTs contrasts with the facilitatory contribution of the right character's semantic radicals as witnessed by the effects of *LogRightKanjiRadicalCombinability* and *LogRightKanjiRadicalTokenFreq* loading on PC1.

As shown in panel d, PC5 speeded up responses in the primed condition (effect size 38 ms) but slowed down responses in the unprimed control condition (effect size 61 ms). PC5 is characterized by large negative loadings of the semantic radical properties, *RightKanjiRadicalTransparency* and *RightKanjiRadicalUsefulness*. In the unprimed control condition, characters with a semantically more transparent and more useful radical elicited shorter response latencies. This pattern reverses when a related prime is presented. In other words, target words with radicals with low transparency and usefulness benefit more from the priming manipulation.

We note here that a main effects model with just the priming manipulation as predictor yielded a facilitatory effect of priming (mean RT control 718 ms, mean RT primed 703 ms, effect size 15 ms) that just failed to reach significance at $\alpha = 0.05$ ($p = 0.0570$ using Markov chain Monte Carlo sampling, $p = 0.0501$ using the upper bound for the degrees of freedom for the t-test). Importantly, our analysis of covariance allowed us to clarify that the priming effect is indeed a semantic effect, as expected for semantic radicals. Furthermore, the analysis of covariance also clarified that the priming effect increased for decreasing transparency of semantic radicals. Finally, the analysis of covariance also allowed us to bring the priming effect into perspective with respect to other distributional predictors: compared to the other predictors in our model, the effect size of the priming manipulation is modest.

The effect of PC7 was inhibitory (effect size 187 ms). Since the loadings on PC7 are dominated by *LogRightKanjiOtherRadicaTypeFreq* (0.64), PC7 represents an inhibitory effect of type frequency of the non-semantic radical component. This inhibitory effect of the non-semantic radical contrasts with the facilitation observed for the frequency of the semantic radical represented by PC1 (*LogRightKanjiRadicalCombinability* and *LogRightKanjiRadicalTokenFreq* have negative loadings on PC1). This result supports theories that distinguish between semantic and non-semantic radicals (e.g., Feldman & Siok, 1997, 1999).

The position of the semantic radical in the right character (*RightKanjiRadicalPosition*) was not predictive. Similarly, the factor specifying whether the radical occupied the same position across prime and target (*PrimeTargetRadicalPositionConsistency*) failed to reach significance as well. This allows us to conclude that the priming effect is semantic in nature and was not driven by positional overlap.

Prime PCs, representing the properties of the prime words, as well as participants' characteristics (i.e., age, sex, months of stay in Canada), did not emerge as significant predictors. We also investigated whether words with different ON(Chinese-origin)-KUN(Japanese-origin) pronunciations affected the results. Removal of the relevant words and re-analyses did not change the results.

Random forest analysis

Our regression model does not inform us about the specifics of the relative importance of the individual lexical variables that loaded onto various PC. For example, *RightKanjiRadicalTransparency* and *RightKanjiRadicalUsefulness* have very similar loadings on PC5, and their individual contributions cannot be teased apart. Furthermore, *LogWholeWordTokenFrequency* has high loadings on three PCs (PC3, PC4, and PC7) and hence its contribution is larger than one would expect on the basis of inspection of individual PCs. In order to obtain better insight into the relative importance of the individual variables, we made use of a random forests recursive partitioning method, a technique particularly useful for assessing a large number of predictors with small samples (*small n large p* problem, see Strobl, Malley, & Tutz, 2009). Conditional inference trees are grown for subsets of observations and subsets of predictors. Predictions are obtained by an ensemble method in which the votes of individual trees are collected. Variable importance is gauged by evaluating reduction in prediction accuracy when a given predictor is not considered (Breiman, 2001; Strobl, Boulesteix, Zeileis, & Hothorn, 2007).

Variable importance rankings (using the conditional permutation scheme proposed by Strobl, Boulesteix, Kneib, Augustin, & Zeileis, 2008, implemented in the `cforest` function in the `party` package of Hothorn, Buehlmann, Dudoit, Molinaro, & Van Der Laan, 2006; Strobl, et al., 2007, 2008) for the random forests fitted separately to the primed and the control conditions are shown in Figure 2-2. Basically, the variable importance plotted on the horizontal axes is a measure of the drop in prediction accuracy when the predictor is withheld from the model specification. For important predictors, failure to include them results in a large loss of prediction accuracy. For irrelevant factors, on the other hand, it does not matter when they are not included in the model specification.

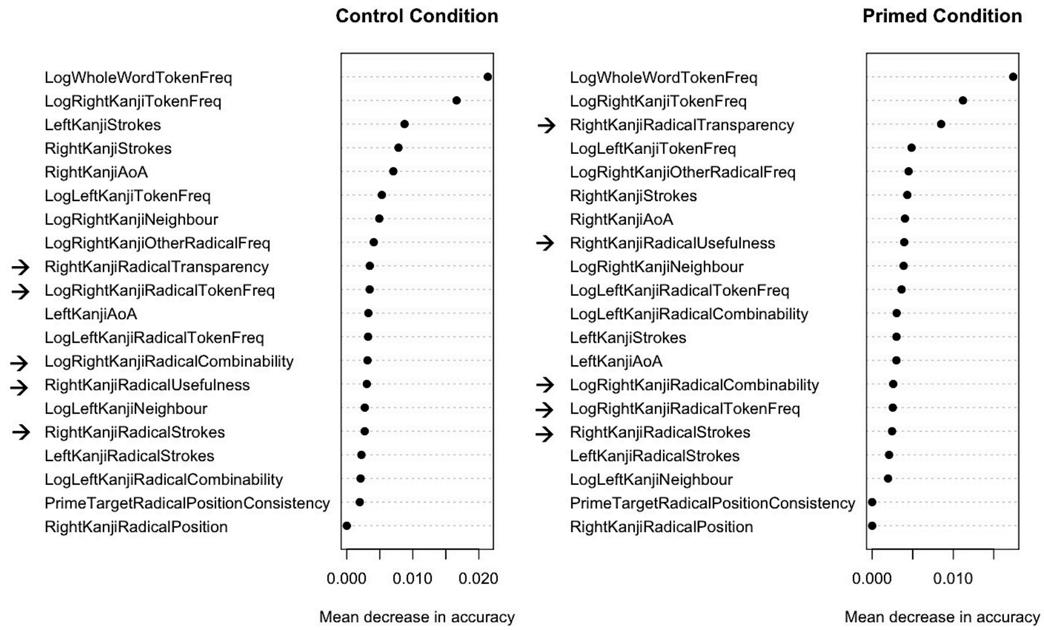


Figure 2-2. A random forest's variable importance ranking for the primed and the control conditions. Variable importance is assessed in terms of mean decrease in accuracy.

Figure 2-2 reveals a pattern consistent with the PCA regression analysis. *LogWholeWordTokenFreq* is identified, as the most important variable, which dovetails well with the observation that *LogWholeWordTokenFreq* has high loadings on several different PCs. The next most important variable is the token frequency of the right character (*LogRightKanjiTokenFreq*) in both primed and control conditions. The right character's importance in lexical decision is consistent with the result of Tamaoka and Hatsuzuka (1995). While the PCA did not distinguish between contributions of semantic radical combinability and cumulative token frequency, the random forests analysis suggests that token frequency (*LogRightKanjiRadicalTokenFreq*) is more important than combinability (*LogRightKanjiRadicalCombinability*) in the unprimed condition. Note that *RightKanjiRadicalTransparency* and *RightKanjiRadicalUsefulness* are ranked higher in the primed condition, further confirming that, in the primed condition, the properties of the semantic radical afforded a processing advantage. Note that, as in the regression analysis, the positional predictors *RightKanjiRadicalPosition* and *PrimeTargetRadicalPositionConsistency* were irrelevant. Finally, *LogRightKanjiOtherRadicaTypeFreq* ranks among the top half of the importance ranked predictors.

General discussion

We identified several graded effects of semantic radical properties in two-character *kanji* word recognition in addition to the effects of character frequency and whole word frequency. Interestingly, the effect of a semantic radical's type and token frequencies depends on its position in the two-character word: inhibitory in modifier position and facilitatory in head position. This suggests that the facilitation observed in single-character studies does not generalize straightforwardly to the modifier position in two-character words. This asymmetrical effect of radical frequency was not modulated by the priming manipulation, which emerged only in interaction with *RightKanjiRadicalTransparency* and *RightKanjiRadicalUsefulness*. The facilitation characterizing heads and the inhibition observed for modifiers may be due to the semantic radical functioning as a kind of classifier. The semantic class indicated by the semantic radical of the head is congruent with that of the compound as a whole. For the modifier, by contrast, the semantic class indicated by its semantic radical is at odds with that of the compound as a whole.

Turning to this interaction, which expressed itself on PC5 in our PCA regression model, we observed facilitation for words with a right semantic radical with lower values on either semantic measure. This suggests that semantic radicals are not mere orthographic units but orthographic morphemes combining form and meaning properties. Random forests clarified that *RightKanjiRadicalTransparency* is more predictive than *RightKanjiRadicalUsefulness*, suggesting that the compositional part-whole relation between a semantic radical and its character is more influential in a visual lexical decision task than the intrinsic semantic richness of this radical.

In the PCA regression, AoA loaded on the same principal component as semantic radical type and token frequencies. Random forest analysis indicated that *LogRightKanjiTokenFreq* and *LogLeftKanjiTokenFreq* outperformed *RightKanjiAoA* and *LeftKanjiAoA* respectively in both priming conditions. Nevertheless, random forest analysis indicates that AoA is a robust predictor. Note that the random forest, using a measure of AoA which is not based on human ratings or performance but on the age at which characters are learned in school, replicates the finding that the frequency effect cannot be reduced to AoA (cf. Brysbaert, Lange, & van Wijnendaele, 2000; Morrison & Ellis, 2000).

The present results indicate that describing the reading of two-character words as analogous to alphabetic compound word processing still underestimates the orthography-specific morphographic complexity of reading in Japanese. *Kanji* characters are themselves morphologically complex. The semantic radical can be viewed as a purely orthographic morpheme, combining a visual form with rich semantics, without support from a phonological/acoustic form. The experimental fingerprint of the semantic radical emerging from our study resembles in many ways the experimental fingerprint of standard morphemes as observed for many European languages, with position-dependent effects, with graded effects of semantic transparency, with effects of AoA, and with greater combinability affording faster processing (see, e.g., Moscoso del Prado Martín et al., 2004). This suggests that in Japanese, the orthography provides an additional layer of morphological complexity to the already complex classificatory system provided by the spoken language.

APPENDIX 2-1

A list of word pairs used in the present study. Items marked with * were excluded from the final analyses based on error rate.

RadicalShared	CriticalPrime	CriticalTarget	RadicalShared	ControlPrime	ControlTarget
Yes	豆粒	妖精	No	薄味	妖精
Yes	救急	得意	No	家財	得意
Yes	犠牲	放牧	No	学校	放牧
Yes	反転	五輪	No	犠牲	五輪
Yes	時計	書記	No	救急	書記
Yes	旅館	*綿飴	No	教授	*綿飴
Yes	道路	跳躍	No	空港	跳躍
Yes	睡眠	明瞭	No	苦悩	明瞭
Yes	筆箱	縦笛	No	潔癖	縦笛
Yes	積荷	*国花	No	自爆	*国花
Yes	満点	高熱	No	車掌	高熱
Yes	自爆	*土煙	No	心霊	*土煙
Yes	釣針	手鏡	No	睡眠	手鏡
Yes	結婚	年始	No	積荷	年始
Yes	空港	血液	No	釣針	血液
Yes	家財	盗賊	No	道路	盗賊
Yes	薄味	合唱	No	時計	合唱
Yes	教授	薬指	No	反転	薬指
Yes	学校	球根	No	筆箱	球根
Yes	車掌	攻撃	No	豆粒	攻撃
Yes	心霊	落雷	No	満点	落雷
Yes	潔癖	治療	No	旅館	治療
Yes	苦悩	習慣	No	結婚	習慣
Yes	物価	同僚	No	海図	同僚
Yes	皆勤	援助	No	角度	援助
Yes	決闘	玄関	No	皆勤	玄関
Yes	樹脂	頭脳	No	壁紙	頭脳
Yes	新鮮	捕鯨	No	景観	捕鯨
Yes	海図	楽園	No	下駄	楽園
Yes	壁紙	電線	No	骨盤	電線

Yes	監獄	密猟	No	決闘	密猟
Yes	悲鳴	*折鶴	No	補聴	*折鶴
Yes	工場	食塩	No	監獄	食塩
Yes	破裂	変装	No	工場	変装
Yes	砂嵐	断崖	No	樹脂	断崖
Yes	乗客	合宿	No	新鮮	合宿
Yes	進展	質屋	No	乗客	質屋
Yes	角度	売店	No	進展	売店
Yes	下駄	実験	No	砂嵐	実験
Yes	骨盤	連盟	No	洗剤	連盟
Yes	補聴	就職	No	卓越	就職
Yes	木造	交通	No	破裂	交通
Yes	洗剤	短剣	No	宅配	短剣
Yes	宅配	発酵	No	悲鳴	発酵
Yes	景観	両親	No	物価	両親
Yes	卓越	隆起	No	木造	隆起

CHAPTER 3

The time-course of lexical activation in Japanese morphographic word recognition: Evidence for a character-driven processing model [†]

Studies on the recognition of complex entities, irrespective of whether these are scenes, objects, or human faces, need to consider how the whole and its parts contribute to our recognition of the input as a coherent meaningful unit (Beck, 1966; Biederman, Mezzanotte, & Rabinowitz, 1982; Greene & Oliva, 2009; Joseph & Tanaka, 2003; Kahneman, Treisman, & Gibbs, 1992; Navon, 1977; Tanaka, Kiefer, & Bukach, 2004; Treisman & Gelade, 1980; Wachsmuth, Oram, & Perrett, 1994). Word recognition is no exception in this respect. Some researchers have argued that morphologically complex words are represented and processed as wholes (Aitchison, 1987; Butterworth, 1983; Caramazza, Laudanna, & Romani, 1988; Janssen, Bi, & Caramazza, 2008). In the word-based supralexical model of Giraudo and Grainger (2001), the activation of the whole word precedes the activation of the constituent parts.

[†] A version of this chapter has been accepted for publication. Miwa, K., Libben, G., Dijkstra, T., & Baayen, R. H. (2012). The time-course of lexical activation in Japanese morphographic word recognition: Evidence for a character-driven processing model. *Quarterly Journal of Experimental Psychology*. This research was supported by a Major Collaborative Research Initiative Grant (#412-2001-1009) from the Social Sciences and Humanities Research Council of Canada to Gary Libben (Director), Gonia Jarema, Eva Kehayia, Bruce Derwing, and Lori Buchanan (Co-investigators) and the Izaak Walton Killam pre-doctoral grant from the Killam Trusts to the first author. We are indebted to Shigeaki Amano, Raymond Bertram, and Rob Schreuder for discussion, to Sally Andrews, Jon Andoni Duñabeitia, Barbara Juhasz, and Sachiko Kinoshita for their constructive feedback on an earlier version of the manuscript.

Many others believe that there is a rapid and automatic morphological decomposition process in recognition and production. In this view, word recognition is not a simple process matching whole word forms to whole word meanings: Sublexical units are posited to exist and also play a role in recognition. There remains, however, an on-going debate over how and at what point in time sublexical units contribute to lexical access (see Frost, Grainger, & Carreiras, 2008; Frost, Grainger, & Rastle, 2005, for overviews). Strict morpheme-based theories of lexical access in reading claim that complex words are decomposed into their constituents and subsequently recombined into a whole word representation (Taft, 2004; Taft & Forster, 1975; Taft & Nguyen-Hoan, 2010). Although interactive activation models allow top-down feedback, bottom-up combinatorial processing is a dominant characteristic of these models as well (McClelland & Rumelhart, 1981; Taft, 1994).

Yet other models proceed on the assumption that the whole and its parts are accessed in parallel (Baayen, Dijkstra, & Schreuder, 1997; Diependaele, Duñabeitia, Morris, & Keuleers, 2011; Frauenfelder & Schreuder, 1992; Kuperman, Schreuder, Bertram, & Baayen, 2009; Pollatsek, Hyönä, & Bertram, 2000). Although efficiency in lexical processing has often been discussed in terms of the dichotomy of computational efficiency and storage efficiency (McClelland & Patterson, 2002a, 2002b; Pinker & Ullman, 2002a, 2002b), it has also been argued that it is efficient to redundantly represent and activate all constituent morphemes, as well as whole word units, thus maximizing opportunities for word identification (Libben, 2006). Previous eye-tracking studies provided partial support for such parallel-route architecture. Pollatsek et al. (2000) tracked eye-movements when Finnish compounds were read in sentences. Although a complete decompositional model predicts a whole compound frequency effect to appear later than an effect of the second constituent frequency, the study found that whole compound frequency effect appears at least as early as the second constituent frequency effect, indicating a race between a decompositonal route to activate the constituents and a direct route to activate the whole compound. Kuperman et al. (2009) more recently combined lexical decision with eye-tracking and observed simultaneous contributions of whole word frequency and morphological constituent frequency already at the first fixation, before the entire word had been scanned. These results challenge strict hierarchical processing models but are compatible with both non-hierarchical multiple route models and with hierarchical models that allow lower level units to connect with higher level units while skipping intermediate levels.

Morphographic word recognition

The writing systems of Chinese and Japanese add various layers of complexities to the current theories developed for English and other related languages. Morphographic orthographies make use of very large numbers of symbols. The minimal basic set of characters taught in Japanese compulsory education comprises 1,945 distinct characters (Japanese Ministry of Education, Culture, Sports, Science and Technology, 2009). The Japanese industrial standard (JIS) list of characters for computers includes 6,353 characters, and ordinary Japanese and Chinese morphographic character dictionaries contain well over 10,000 characters (Coulmas, 2003; Kess & Miyamoto, 1999). Unlike alphabetic letter symbols, Japanese morphographic characters directly encode meaning (e.g., 木 /ki/ 'wood').

Although *kanji* characters have often been compared to morphemes in alphabetic languages, the majority of characters are themselves decomposable into smaller units. The character 海 /kai/ 'sea', for example, consists of a semantic radical 氵 and a phonetic radical 毎. Among 2,965 Japanese Industrial Standard *kanji* characters, 83% of the characters consist of either left and right radicals or top and bottom radicals (Saito, Kawakami, & Masuda, 1995, 1997). Semantic radicals encode a general basic category meaning. The radical 氵 'water', for example, is shared by characters whose meaning is associated with 'water' (e.g., 海 'sea', 液 'liquid', and 酒 'liquor'), although the contribution of the semantic radical to the whole character is not always transparent (e.g., 法 'law' is not related to 'water'). Phonetic radicals, on the other hand, encode approximate information about the pronunciation of the character (e.g., 海 and 悔 are both pronounced /kai/). The different functions of semantic and phonetic radicals are explicitly taught in primary school.

When they encounter unfamiliar words, readers of Japanese can rely on the radicals. For example, an unfamiliar two-character word such as 寒鰤 'winter yellowtail', which appeared only once in 14 years of newspaper texts (Amano &

Kondo, 2003), is relatively well-interpretable thanks to the right character's semantic radical 魚 'fish' and the left character 寒 'cold', even though the reader may not know what the phonological form of the Japanese word is (/kanburi/; the /n/ denotes a moraic nasal). A large majority of Japanese words are written with two *kanji* characters (70% as estimated by Yokosawa & Umeda, 1988). A question addressed in this study is how readers process radical and character information in comprehending relatively familiar two-character words.[1]

Several experimental studies suggest that the characters in two-character words are accessed in reading. Hirose (1992), observing a stronger priming effect of the left character over that of the right character in primed lexical decision, proposed that two-character words are represented in clusters centered around the shared left character, and that they are processed from left to right, with the left character functioning as the retrieval cue. While this perspective appears to be in line with importance of the initial constituent reported by Taft and Forster (1976) for English, Tamaoka and Hatsuzuka (1995) and Zhang and Peng (1992), in contrast, reported that the frequency of the right character facilitates two-character lexical decision responses more than the left character in Chinese and Japanese respectively. Kawakami (2002) reported facilitation from the type frequency of characters in two-character word lexical decision.[2] In addition to character frequency effects, Tamaoka (2005) observed that larger

[1] With respect to two-character compounds, Japanese morphology has been argued to be predominantly right-headed (Kageyama, 2010), although exocentric compounds such as voyage ('ship' + 'sea') seem to occur more often than in English or Dutch.

[2] Although Kawakami interpreted this type count as a measure of orthographic neighbourhood density (cf., Coltheart, Davelaar, Janasson, & Besner, 1977; Forster & Taft, 1994), it can also be

numbers of homophones associated with the left character lead to longer response times in lexical decision and naming. Tamaoka and Hatsuzuka (1998, lexical decision and naming) further reported that semantic/conceptual properties of characters co-determine word recognition responses (cf. Ji & Gagné, 2007, sense-nonsense judgment with English compounds). Considered jointly, these studies suggest that morphographic characters in Japanese two-character words have processing characteristics that closely resemble morpho-semantic effects observed for morphemes in English and related languages.

A separate series of studies has addressed the role of radicals in single-character words. Taft and Zhu (1997) reported that higher type frequency of the right radical speeds up character decision. Feldman and Siok (1997) similarly reported facilitatory effects of radical type frequency, but they considered the function of radicals (i.e., semantic vs. phonetic), rather than their positions. They observed that a greater type frequency of the semantic radical facilitated character decision when the radical is located in the left position of the character. Feldman and Siok (1999) further argued, from primed character decision data, that the meaning of the semantic radical is co-activated. A contribution of radicals also has been reported in speeded semantic categorization (Flores d'Arcais & Saito, 1993) and in word naming (Flores d'Arcais, Saito, & Kawakami, 1995).

---

viewed as a measure of morphological family size (Bertram, Baayen, & Schreuder, 2000; Joyce & Ohta, 2002; Moscoso del Prado Martín et al., 2004; Schreuder & Baayen, 1997).

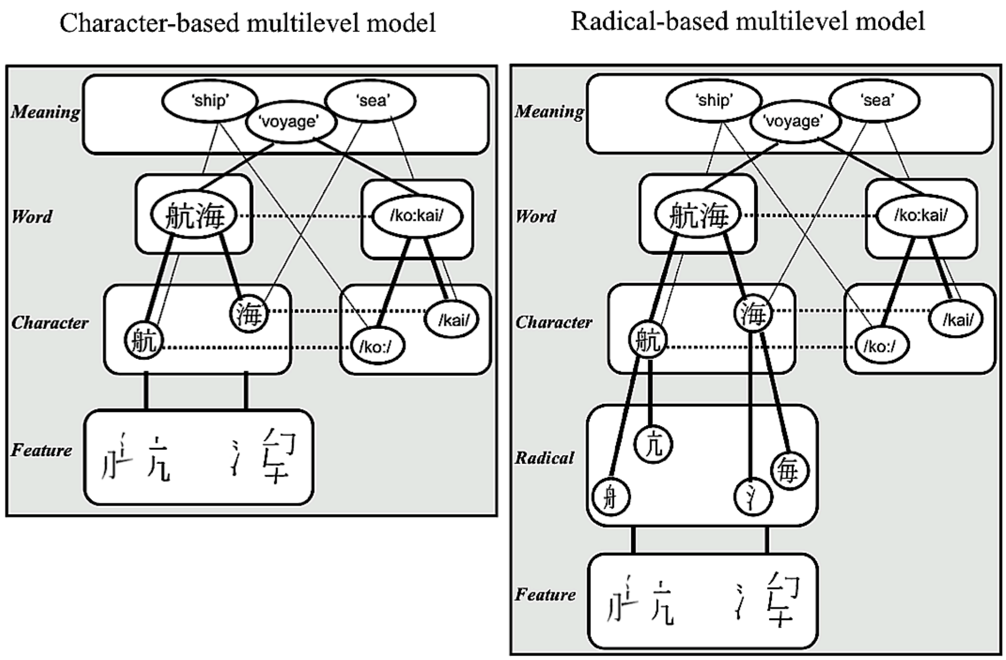


Figure 3-1. Radical-based and character-based multilevel models of morphographic word recognition, summarizing representations and links proposed by Taft, Zhu, and Peng (1999), Saito (1997), and Tamaoka and Hatsuzuka (1998). Activation of neighboring words and characters are not depicted in the figures. Lemma representations in Tamaoka and Hatsuzuka's (1998) model are not shown in the character-based model depicted here (left).

In the present study, we primarily test the predictions of the two hierarchical models of morphographic two-character word recognition shown in Figure 3-1. The character-based model (left, Tamaoka & Hatsuzuka, 1998) claims that characters are the basic lexical units, whereas the radical-based model (right, Ding, Peng, & Taft, 2004; Saito, 1997; Saito, Masuda, & Kawakami, 1998; Taft & Zhu, 1997; Taft, Zhu, & Peng, 1999) assumes that radicals mediate between strokes and characters. Both models presuppose left-to-right scanning of the visual input (Taft & Zhu, 1997; Tamaoka & Hatsuzuka, 1995), and both assume that a higher level unit can only be activated once its lower level constituent units are activated.

The two models diverge with respect to the role of radicals. Taft et al. (1999) and Saito (1997) argue that morphographic characters are initially decomposed into radicals. In models that distinguish characters and radicals, an issue at stake is whether semantic radicals are semantically interpreted as soon as they are activated. Taft et al. (1999) assume that characters form the first level in the hierarchy that provides access to meaning. In other words, in this model, radicals function as purely orthographic access codes. However, there is some experimental evidence suggesting that semantic radicals are interpreted semantically as soon as they have been activated. (Feldman & Siok, 1997, 1999; Miwa, Libben, & Baayen, 2012). The evidence for the two models in Figure 3-1 comes from two distinct streams of research. Evidence for characters as processing units was obtained with experiments using two-character words, while evidence for radicals as processing units was obtained using single-character words. Miwa et al. (2012) performed the first study addressing the role of semantic radicals in the processing of two-character words. In their lexical decision study with partial repetition priming of the semantic radical in the right character, a significant interaction was observed between the priming manipulation and the semantic properties of the semantic radicals, suggesting that even in two-character words, an effect of semantic radicals can be detected.

Goals of this study

The studies reviewed in the previous section involved 15 lexical decision experiments, all based on only 30 to 90 target words (M = 51, SD = 17.8) matched on a limited number of experimental variables. As Cutler (1981) pointed out three decades ago, it is a "confounded nuisance" to pre-experimentally control for the growing number of all potentially important variables. While Tamaoka (2005, 2007) carefully controlled for a relatively large number of 11 and 18 potentially important variables, all the other studies controlled for a much smaller number of variables. Pre-experimental matching on numerical covariates may lead to substantial loss of statistical power (Baayen, 2010; Cohen, 1983; MacCallum, Zhang, Preacher, & Rucker, 2002), and may negatively affect the representativeness of the sampled items. We therefore opted for a regression design analyzed with mixed-effects models (Baayen, 2008; Baayen, Davidson, & Bates, 2008; Baayen & Milin, 2010), assessing subject, item, and task effects jointly to obtain a more comprehensive picture of Japanese visual word recognition with 24 lexical variables, using 708 target words.

All previously mentioned studies relied on chronometric measures. In order to obtain more insight into the microstructure of information processing in lexical decision, we conducted an eye-tracking experiment combined with lexical decision. Previous studies (Hyönä & Pollatsek, 1998; Kuperman et al., 2008, 2009; Pollatsek et al., 2000) suggest that morphological processes can be investigated through eye-movements (but see Andrews, Miller, & Rayner, 2004, for lack of such strong link). Using a regression design with over 500 two-character words, we tested several questions in parallel. First, what is the time course of activation of strokes, radicals, characters, and words? Hierarchical models predict higher level units to become active only once their lower level constituent units have been activated. Hence, these models predict stroke effects to precede radical effects in the eye-movement record, radical effects to precede character effects, and character effects to precede whole word effects. The magnitude of the effects is also expected to vary with time. For instance, radical frequency is expected to have a large effect on initial fixation durations but little or no effect on later fixations. Of special interest here, given the early compound frequency effect observed in Kuperman et al. (2009), is the moment in time at which the effect of compound frequency first emerges.

Second, what is the relative importance of the left and the right characters in two-character word recognition? Does the left character have a privileged status compared to the right character, as argued by Hirose (1992)? If so, does an initial fixation on the right character have a catastrophic effect on comprehension? If, however, the right character is important, as suggested by Tamaoka and Hatsuzuka (1995) and Zhang and Peng (1992), it is worth considering whether the right character's privilege is due to a left-to-right scan process (Tamaoka & Hatsuzuka, 1995) or due to the fact that the right character is the main morpheme that should be processed first, at least in reading modifier-head compounds (Zhang & Peng, 1992). If a left-to-right scanning is preferred for Japanese, as for alphabetic languages (Hyönä & Pollatsek, 1998; Pollatsek et al., 2000), early and late time frames, as determined by eye fixations, should reflect the left and the right characters' contributions respectively.

Third, are semantic radicals interpreted semantically or do they function just as orthographic access codes? In the former case, we expect that the degree to which the semantic radicals contribute to the meaning of the character, as gauged by semantic transparency ratings (Feldman & Siok, 1999; Miwa et al., 2012), should co-determine fixation durations and/or lexical decision speed. If a semantic radical is interpreted semantically, then a next question would be whether a semantic transparency effect appears early, indicating an early morpho-semantic processing. Diependaele et al. (2005, 2011) observed semantic transparency effects in masked priming, such that there was more facilitation for transparent derivations (e.g., *worker*) than an opaque words (e.g., *corner*) (see also Feldman, O'Connor, & Moscoso del Prado Martín, 2009). If Japanese intra-character morphology is functionally comparable to that of multi-morphemic words in alphabetic languages, then a semantic transparency may show facilitation in the earliest time frame.

Fourth, to what extent is the uptake of visual information co-determined by non-linguistic factors? We manipulated the readers' attention by varying the fixation point, which was positioned on the left character, on the right character, or in between the two characters. Kajii, Nazir, and Osaka (2001) report that fixations tend to fall onto the left character in sentential reading. However, the position of fixations seems to be more flexible (left or centre) in Chinese (Yan, Kliegl, Richter, Nuthmann, & Shu, 2010). Furthermore, if the right character is the main morpheme (Zhang & Peng, 1992), then an initial fixation on the right character may be more beneficial. Most previous isolated word reading studies directed the readers' attention to the word centre, which limits generalizability of the results. However, by shifting attention to other positions in the word, the consequences of dis-preferred initial fixation positions can be evaluated.

Predictors

In our study, we made use of a regression design with subjects and items as crossed random-effect factors. This section introduces the fixed-effect factors and covariates that we considered. Unless noted otherwise, we used lexical distributional data as available in the web-accessible database for Japanese characters constructed by Tamaoka et al. (2002) and Tamaoka and Makioka (2004). Table 3-1 summarizes the lexical distributional properties considered in the present study, grouped by different levels of linguistic structure posited by the hierarchical models as developed by Taft et al. (1999) and Saito (1997).

Table 3-1. Lexical predictors, individual differences, and task effects considered in this study

Type	Predictors	
Feature	· LeftKanjiStrokesResid	· RightKanjiStrokesResid
(、、✓)	· LeftKanjiComplex	· RightKanjiComplex
Radical	· LeftKanjiRadicalCombinability	· RightKanjiRadicalCombinability
(氵)	· LeftKanjiRadicalTokenFreqResid	· RightKanjiRadicalTokenFreqResid
Character	· LeftKanjiNeighbourResid	· RightKanjiNeighbourResid
(港)	· LeftKanjiTokenFreq	· RightKanjiTokenFreq
	· LeftKanjiAoAResid	· RightKanjiAoAResid
Word	· WholeWordFreq	· GoogleDocFreqResid
(空港)		
Phonology	· LeftKanjiHomophones	· RightKanjiHomophones
Semantics	· LeftKanjiRadicalTransparency	· RightKanjiRadicalTransparency
	· LeftKanjiTransparency	· RightKanjiTransparency
Individual	· LengthOfStayCanada	
Task	· PreviousRT · PreviousTrialCorrect	· Trial · Fixation
	· PreviousSubgazeDuration	· EyePosition

Feature-level predictors

At the feature level, *LeftKanjiStrokes* and *RightKanjiStrokes* quantify the number of strokes in a character. The stroke count measure is designed to capture what word length captures for alphabetic languages: the complexity of the visual input. Word length generally has an inhibitory effect in chronometric and eye-tracking studies (Balota et al., 2004; Vitu, O'Regan, & Mittau, 1990), although there is some evidence for non-linearity for shorter word lengths (Baayen, 2005; New, Ferrand, Pallier, & Brysbaert, 2006). Similarly, previous studies on Japanese and Chinese suggest that characters with many strokes are processed slower than those with few strokes (Leong, Cheng, & Mulcahy, 1987; Liu, Shu, & Li, 2007). Note, however, that feature level complexity in Japanese manifests itself in the form of the density of visual information within a highly restricted fixed word region. As a consequence, the visual acuity limitation relevant for scanning extended strings of letters in alphabetic languages will not contribute to the visual complexity effects in Japanese.

In addition to stroke counts, *LeftKanjiComplex* and *RightKanjiComplex* represent a broader radical-based complexity measure based on human ratings (Tamaoka et al., 2002). The rated value is 1 if characters perceptually consist of single component (e.g., 口 and 馬) and 2 if they consist of two components (e.g., 欧 = 区 + 欠). In most cases, this rating tends to reflect the number of radicals in the character. We coded the rated character complexity as a factor with the two levels of *Simplex* (complexity rating = 1) and *Complex* (complexity rating > 1).

Radical-level predictors

At the level of radicals, *LeftKanjiRadicalCombinability* and *RightKanjiRadicalCombinability* are the log-transformed type frequency of the semantic radicals, representing how many basic Japanese characters share a given semantic radical. *LeftKanjiRadicalTokenFreq* and *RightKanjiRadicalTokenFreq* are the log-transformed cumulative token frequency of all characters (in the 1,945 basic *kanji* list) sharing a given semantic radical, calculated from Amano and Kondo (2000). Previous studies (Feldman &

Siok, 1997, 1999; Miwa et al., 2012; Taft & Zhu, 1997) suggest that we may expect facilitatory contributions from these type and token frequency measures. The present study considers only semantic radicals because all characters, regardless of their complexity, contain a semantic radical without exception whereas characters need not contain a phonetic radical.

Character-level predictors

At the level of characters, we considered log-transformed character token frequency (*LeftKanjiTokenFreq*, *RightKanjiTokenFreq*) and log-transformed position-dependent character neighbourhood size in two-character words (*LeftKanjiNeighbour* and *RightKanjiNeighbour*). Independent effects of constituent frequency and neighbourhood size in two-character word recognition have been reported by Tamaoka and Hatsuzuka (1995) and Kawakami (2002) respectively.

LeftKanjiAoA and *RightKanjiAoA* refer to the school grade in which a given character was learned, based on the guidelines set by the Japanese Ministry of Education, Culture, Sports, Science and Technology. Although many past studies have reported that words learned earlier in life are processed faster when matched for frequency, the validity of the age of acquisition (AoA) measures used and the role of AoA among many correlated measures are on-going issues (cf. Baayen, Feldman, & Schreuder, 2006; Brysbaert, Lange, & van Wijnendaele, 2000; Havelka & Tomita, 2006; Juhasz, 2005). The AoA measures we use in this study are not based on human reports and provide an objective measure.

Word-level predictors

At the whole word level, we considered log-transformed written frequency based on newspapers published in the 14-year period from 1985 to 1998 (*WholeWordFreq*, Amano & Kondo, 2003). We complemented this frequency measure with the log-transformed Google document frequency as of November 29, 2008. This dispersion measure provides an estimate of the range of different documents (genres, registers) in which a word is used. Contextual diversity of words has been reported as a powerful measure in some recent studies (e.g., Adelman, Brown, & Quesada, 2006; Brysbaert & New, 2009), and we expected this Google dispersion frequency to have an additive effect on top of the standard word frequency effect (see Ji & Gagné, 2007 and Myers, Huang, & Wang, 2006 for previous studies using Google document frequency).

Phonological predictors

In order to assess phonological ambiguity and its effect on reading (Ferrand & Grainger, 2003; Pexman, Lupker, & Jared, 2001; Tamaoka, 2005), we made use of the log-transformed number of homophonous characters (*LeftKanjiHomophones* and *RightKanjiHomophones*). Tamaoka (2005) reported that words with a left character with many homophonic characters, relative to few, elicited longer response times in lexical decision and naming.[3]

Semantic predictors

Given the possibility of a processing advantage for semantically transparent compounds (Libben, 1998; Libben, Gibson, Yoon, & Sandra, 2003), we also included two measures for the semantic transparency of the characters in the compound. Although character activation in compound reading has been argued to be orthographic (Kawakami, 2002; Saito, 1997), other studies suggest that meanings of characters are co-activated (Tamaoka & Hatsuzuka, 1998; Ji & Gagné, 2007). *LeftKanjiTransparency* and *RightKanjiTransparency* gauge the semantic congruity between the meaning of the character and the meaning of the whole word. Both measures are based on mean ratings elicited from six native Japanese readers, using a seven-point scale (Cronbach's alpha > 0.99, $M = 6.0$, $SD = 1.1$ for *LeftKanjiTransparency*; Cronbach's alpha > 0.99, $M = 6.0$, $SD = 1.0$ for *RightKanjiTransparency*, using the *psy* package for R by Falissard, 2007).

Furthermore, in order to test whether semantic radicals are mere orthographic access units or meaningful "orthographic morphemes", we included two measures of semantic radical transparency (*LeftKanjiRadicalTransparency*

---

[3] In Japanese, characters may have two kinds of pronunciations: *ku* (*On*-Reading, Chinese origin) and *sora* (*Kun*-Reading, Japanese origin). In the context of *<u>ku</u>ko* 空港 'airport', the On-Reading is applied, while in the context of *<u>sora</u>iro* 空色 'sky blue', the Kun-Reading is applied. Given that visual lexical decisions are based to a larger extent on orthographic and semantic properties of words, as well as that On-Kun status is finalized only after the whole word is activated, the effect of On-Kun distinction is expected to be small or null in the present study.

and *RightKanjiRadicalTransparency*). These measures represent the degree of semantic congruity between the meaning of the character and the meaning of the radical. Eight native Japanese readers rated similarity in meaning between characters and their semantic radical on a seven-point scale (M = 3.9, SD = 1.7, Cronbach's alpha > 0.99). In the analyses below, we used the mean ratings.

Multicollinearity among lexical predictors

The present set of lexical distributional predictors is characterized by serious multicollinearity. We removed most of this collinearity by residualization of correlated predictors, following Kuperman et al. (2009). For example, because *RightKanjiTokenFreq* and *RightKanjiNeighbor* and *RightKanjiStrokes* all highly correlate with *RightKanjiAoA* (r = -0.56, and r = -0.68 and r = 0.46 respectively, $p < 0.01$ for all), we regressed *RightKanjiAoA* on the former three predictors and used the resulting residuals, *RightKanjiAoAResid*, as a new predictor gauging the right character's AoA uncontaminated by character frequency, character neighbourhood size, and character stroke complexity. As *RightKanjiAoAResid* correlated well with the original measure (r = 0.65, $p < 0.01$), we can interpret it as the independent contribution of AoA devoid of frequency and feature complexity. We followed the same procedure for other pairs of predictors that are highly correlated: *GoogleDocFreqResid* was orthogonalized with respect to *WholeWordFreq* (r = 0.81 for the correlation between *GoogleDocFreqResid* and *GoogleDocFreq*),

---

This was indeed the case in the present study. Hence, this predictor is not mentioned in this paper.

RightKanjiNeighbourResid was orthogonalized with respect to *RightKanjiTokenFreq* (r = 0.88 for the correlation between *RightKanjiNeighbourResid* and *RightKanjiNeighbour*), *RightKanjiRadicalTokenFreqResid* was residualized on *RightKanjiRadicalCombinability* (r = 0.48), and *RightKanjiStrokesResid* was residualized on *RightKanjiNeighbour* (r = 0.92). Because the pattern of multicollinearity among lexical predictors was identical for characters at the left position, the same procedure was followed for computing residualized predictors. As a result, all pairwise correlations among the given lexical properties became less than 0.30, except that between *LeftKanjiTransparency* and *RightKanjiTransparency* (r = 0.59). As for these two predictors, we tested one predictor at a time in a given analysis. As we shall see below, one predictor always outperformed the other, so this correlation was not a problem (see Appendix 3-1 for a correlation matrix for all the numerical predictors considered in this study).

Individual differences and task-related predictors

Although the readers we tested in the present study were all native Japanese readers, they differed in the extent to which they are using Japanese in Canada. As a measure of language proficiency, we included their log-transformed *LengthOfStayCanada* in months as a predictor. This measure correlated positively with age (r = 0.47, p = 0.03) and negatively with log-transformed self-ratings of daily exposure to Japanese (r = -0.52, p = 0.01) and the 100-Rakan Japanese *kanji* reading ability scores (Kondo & Amano, 2001, $r = -0.54$, $p = 0.01$). *LengthOfStayCanada* did not correlate significantly with vocabulary size in Japanese (Amano, Kondo, & Kataoka, 2005) for the readers we tested. Vocabulary size in Japanese, however, correlated positively with 100-Rakan reading ability scores ($r = 0.46$, $p = 0.04$, cf., $r = 0.70$, $N = 1000$; Amano, 2007), which also correlated with *LengthOfStayCanada*. Given this multicollinearity, we opted for *LengthOfStayCanada* as the predictor reflecting various types of individual differences and language proficiency for our statistical analyses, leaving the specific advantages and disadvantages of the other related measures to future research.

Consistency in human behaviour often leads to auto-correlated time series of response times and fixation durations (Baayen & Milin, 2010; de Vaan, Schreuder, & Baayen, 2007; Kuperman et al, 2009; Perea & Carreiras, 2003). We removed the auto-correlation from the errors by including three control predictors: *PreviousRT*, the response time at the previous trial, *PreviousTrialCorrect*, a factor encoding the correctness of the response at the previous trial (levels *Correct* and *Incorrect*), and *Trial*, the rank of the item in the experimental list.

A further predictor was *Fixation*, a factor specifying whether the initial fixation was directed to the *Left* character, the *Central* position between the two characters, or the *Right* character.

In the eye-movement analyses, we considered *PreviousSubgazeDuration*, the subgaze duration at the previously fixated region, and *EyePosition*, a factor encoding the current eye position (levels *Left* and *Right* character regions).

Experiment 1: Lexical decision with eye-tracking

Method

Participants. Twenty-one native Japanese speakers (18 female, 3 males; mean age = 21.2 years old, *SD* = 2.9) were recruited at the University of Alberta. All participants had normal or corrected-to-normal vision, and their mean score on the 100-Rakan *kanji* word reading test was 48.7 out of 100 (*SD* = 19.9), which is comparable to the larger population mean (*M* = 49.6, *SD* = 19.6, *N* = 1000; Amano, 2007). The participants had been in Canada for 25.9 months on average (*SD* = 26.9, range 0 to 76 months).

Apparatus. An SR Research EyeLink II head-mounted eye-tracker was used to track participants' eye-movements. The pupil-only mode was used to track eye movement with a sampling rate of 250 Hz. Words were presented on a 20-inch display controlled by SR Research Experiment Builder.

Materials. Target words in this lexical decision experiment were randomly sampled from a subset of the NTT lexical database (Amano & Kondo, 2003). This subset was compiled from the database by imposing the following restrictions. First, the words should occur at least 100 times in the newspaper corpus. Second, only common nouns were selected. Third, the words with homophonous neighbors were excluded. Fourth, the words should not contain a duplicated character (e.g. *oriori* 折々 ‘occasional’ where 々 indicates that the left character is repeated) nor a *kanji* numeral (e.g. *hachinin* 八人 ‘eight people’). Fifth, the words should not be restricted in their use to fixed or idiomatic phrases (e.g., *katabo* 片棒 ‘a bar’ normally occurs in an idiom *katabo wo katsugu* ‘take part in’). Sixth, relatively unfamiliar two-character words that are not listed in *Kojien Japanese Dictionary* (Nimura, 2002) were excluded as well (e.g., *konkaku* 混獲 ‘mass capturing’). From the resulting subset, we randomly sampled 708 two-character words.

We also prepared 708 nonwords falling into four different types: (1) 60 nonwords were created by switching the order of two characters, (2) 60 nonwords were created by replacing the first constituent with another homophonous character, (3) 60 nonwords were created by replacing the second constituent with another homophonous character, (4) the remaining 528 nonwords were created by randomly combining characters. The first three sets of nonwords were included as part of a separate study not reported here.

Procedure. The experiment consisted of three sessions conducted on different days. Each session lasted for approximately 90 minutes, except for the first session that lasted for 120 minutes. At the beginning of the first session, participants completed the 100-Rakan test and the vocabulary size estimation test.

In the lexical decision experiment, participants were asked to indicate whether the presented word is a legitimate Japanese two-character word or not by pressing buttons on a Microsoft SideWinder game pad with their left (= No) and right (= Yes) index fingers. Their eye-movements were tracked by an EyeLink II head-mounted eye-tracker. For each trial, a fixation point (an asterisk * in 60 point Verdana bold font), which was also used for drift correction, was presented for at least 500 ms, followed by a target two-character word in white Mincho font, size 130, on a black background. With a viewing distance of 70 cm from the screen, the visual angle was 5.3° for each character. The word remained on the screen until the participant responded. A drift correction was performed at every trial; a target word did not appear until participants had fixated on the fixation point. The location of the fixation point was varied across different sessions such that participants were presented with a fixation either at the central position of the screen, at a position slightly towards the left (i.e., where a left character was presented), or at a position slightly towards the right (i.e., where a right character was presented). The order of sessions with different fixation points was counter-balanced within subjects.

The lexical decision experiment started with 12 practice trials in each session, followed by 472 experimental trials ((708 + 708)/3) containing two breaks. After the practice trials and at each break point, participants were given feedback as to how fast (ms) and accurately (correct %) they had been responding so far. Throughout the entire experiment, the left eye was tracked for the half of the participants and the right eye was tracked for the rest of the participants. The words were presented in a different randomized order to each subject.

Results

Statistical analyses were carried out using R version 2.13.2 (R Development Core Team, 2011). Data from two participants were excluded from the subsequent RT and eye-movement analyses due to high error rates (exceeding 35%). All predictors with a skewed distribution (i.e., frequency-based predictors and the readers' length of stay in Canada) were logarithmically transformed.

As dependent variables, we considered response times (RTs), as well as first and second subgaze durations. Total fixation durations were virtually identical to response times and are not analyzed separately. Subgaze duration was defined as the cumulative first-pass fixation duration that fell into pre-specified character regions before the eye departed to another character region. We opted for the subgaze duration based on character regions, as visual inspection of the on-line eye-movements and density plots for fixations suggested that the eye-movements were character-based and not radical-based. In trials with two and three fixations, 70% of the eye-movements moved to the other character region (71.3%, 65.3%, and 73.3% for the left, central, and right fixation conditions respectively).

Response time analysis. For the response time analysis, data points with response time shorter than 300 ms or longer than 3,000 ms were excluded from the dataset. In addition, all data points of those words that elicited over 40% incorrect responses were removed. Furthermore, remaining individual data points with an incorrect response were excluded as well. The analysis was restricted to those two-character words for which the lexical distributional properties were available for both the left and right characters. This resulted in a dataset with 9,228 data points for 555 different words. Because the distribution of RTs was highly skewed with a long right tail, a reciprocal transformation (-1000/RT) was applied to the RTs. Using a linear mixed-effects model with subject and word as crossed random-effect factors (Baayen, 2008; Baayen et al., 2008; Bates, Maechler, & Dai, 2007), we first fitted a simple main effects model with lexical properties at all levels of the hierarchy listed in Table 3-1.[4] We then considered interactions with respect to *Fixation*, *PreviousTrialCorrect*, and *LengthOfStayCanada*. After removing non-significant predictors to obtain the most parsimonious yet adequate model, we removed as potentially harmful outliers data points with standardized residuals exceeding 2.5 standard deviation units, and then refitted the model. The random effect structure of the final model comprised random intercepts for item (*SD* = 0.12) and subject (*SD* = 0.22), by-subject random slopes for centralized *Trial* (*SD* = 0.01), by-subject random slopes for centralized *PreviousRT* (*SD* = 0.07), and by-subject random contrasts for *LeftKanjiComplex* (*SD* = 0.03). The random contrasts for *LeftKanjiComplex* showed greater variance for words with a complex left character. Other random slopes were tested, and none were significant. The standard deviation of the residual error was 0.26. Table 3-2 summarizes the coefficients of this model and Figure 3-2 visualizes the interactions. Predictors that did not reach significance at the 5% level are not listed in Table 3-2.

Table 3-2. Estimate, standard error, t-value, p-value, and effect size of influential predictors for the lexical decision response times.

	Type	Estimate	SE	t-value	p-value	Effect Size (ms)
(Intercept)		-1.044	0.115	-9.06	< 0.0001	
PreviousRT	T	0.142	0.018	8.04	< 0.0001	179
Trial	T	-0.010	0.005	-1.92	0.0545	-155
Fixation (Left)	T	0.084	0.022	3.86	0.0001	12
Fixation (Right)	T	0.096	0.020	4.74	< 0.0001	14
PreviousTrialCorrect (Incorrect)	T	0.192	0.055	3.49	0.0005	6
LengthOfStayCanada	I	0.040	0.033	1.20	0.2292	85
LeftKanjiStrokesResid	F	0.009	0.002	5.58	< 0.0001	94
LeftKanjiNeighbourResid	C	-0.013	0.006	-2.40	0.0165	-38
RightKanjiTokenFreq	C	-0.009	0.004	-2.14	0.0324	-35
RightKanjiAoAResid	C	0.010	0.005	2.16	0.0308	40
WholeWordFreq	W	-0.057	0.004	-13.39	< 0.0001	-178
GoogleDocFreqResid	W	-0.051	0.006	-8.43	< 0.0001	-173
RightKanjiHomophones	P	0.028	0.007	3.90	0.0001	56
RightKanjiTransparency	S	-0.008	0.006	-1.42	0.1545	-24
RightKanjiTokenFreq x PreviousTrialCorrect (Incorrect)	C x T	-0.018	0.005	-3.68	0.0002	Figure 3-2 (a)
RightKanjiTransparency x Trial	S x T	-0.002	0.001	-2.82	0.0048	Figure 3-2 (b)
LengthOfStayCanada x Fixation (Left)	I x T	-0.027	0.008	-3.59	0.0003	Figure 3-2 (c)
LengthOfStayCanada x Fixation (Right)	I x T	-0.031	0.007	-4.55	< 0.0001	Figure 3-2 (c)

Note: T = Task, I = Individual, F = Feature, R = Radical, C = Character, W = Word, P = Phonology, S = Semantics

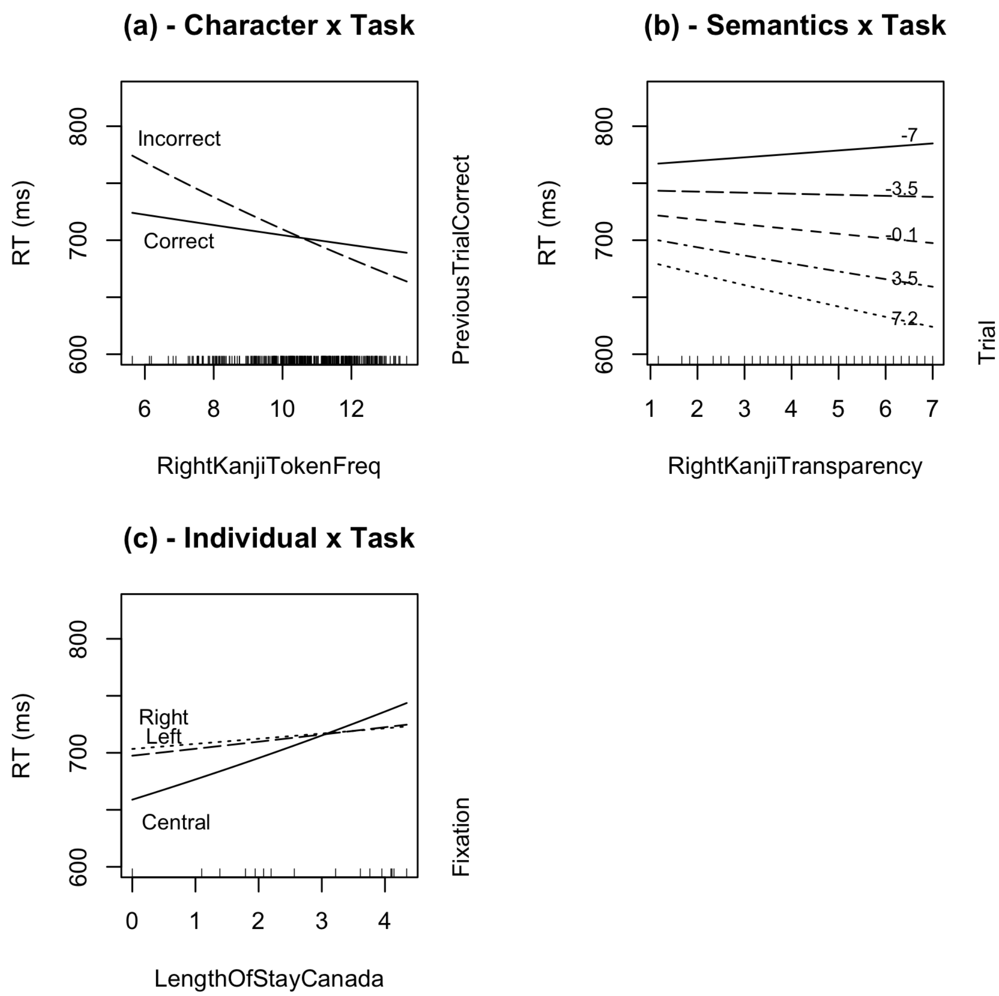


Figure 3-2. Interactions co-determining the lexical decision response times and the number of subgazes

Feature-level effects. Lexical distributional properties at all levels of the hierarchy emerged as significant predictors of the response times. Words with greater left character feature complexity (*LeftKanjiStrokesResid*) elicited longer response times (effect size = 94 ms). The absence of significant effects of

RightKanjiStrokesResid and *RightKanjiComplex* is consistent with theories that assume processing to proceed from left to right (Hirose, 1992; Taft & Zhu, 1997; Tamaoka & Hatsuzuka, 1995).

Character-level effects. The effect of *RightKanjiTokenFreq* was facilitatory, particularly when the response at the previous trial was incorrect (Figure 3-2, Panel a). We suspect that after readers make an error, they pay special attention to the head character, as this will help them to make a correct lexicality decision: In order to reject a stimulus such as *cloudchair*, the readers have to assess whether *cloudchair* is an existing kind of chair. If this interpretation is correct, the effect of *RightKanjiTokenFreq* is a late, conceptual, effect.

RightKanjiAoAResid, the residualized age of acquisition of the right character, revealed the expected inhibitory effect (effect size = 40 ms). If the right character was learned later, the response to the target word was slowed down, indicating that childhood experience has a lasting effect on reading in adulthood, independently of token frequency. The left character's AoA was not a significant predictor.

Word-level effects. *WholeWordFreq* and *GoogleDocFreqResid* both facilitated responses (effect sizes = -178 ms and -173 ms). The presence of the additive effect of *GoogleDocFreqResid* suggests a need to consider contextual diversity of words as an important factor in understanding how words are entrenched in memory (Adelman et al., 2006; Brysbaert & New, 2009). Adelman et al. (2006) reported for English that when frequency is residualized on contextual diversity, it is no longer a significant predictor. For the present data, this did not hold: Both residualized frequency and *GoogleDocFreq* contribute independently to the model, both $p < 0.0001$).

Phonological effects. The number of homophones of the right character slowed down responses as well (effect size = 56 ms), as expected. This finding contrasts with Tamaoka's (2005) observation of an inhibitory morphemic homophony effect for the left character only. This difference might be due to the way nonwords were constructed. In Tamaoka's (2005) study, nonwords were pseudo-homophones with homophonic left characters only. In the present study, the pseudo-homophones appeared in both positions, while in addition many nonwords were random combinations of characters. As a consequence, the role of the right constituent as the head is more important in the present study. This morphemic homophony effect may reflect a rebounding effect of phonology to orthography (Pexman et al., 2001; Tamaoka, 2005, 2007). Alternatively, it may reflect competition between different meanings associated with homophonic alternatives. We will return to the homophone effect below when discussing the second subgaze durations.

Semantic effects. The semantic transparency of the right character slowed down responses at the beginning of the experiment and became facilitatory towards the end (Figure 3-2, Panel b), suggesting that the criteria for discriminating between words and nonwords were adjusted in the course of the experiment. In this task, it is not trivial to discriminate real transparent compounds such as *handbag* from nonwords such as *toebag*. In the course of the experiment, the reader becomes more proficient at discriminating the words from the nonwords, apparently by relying more on the presence of a transparent semantic relation between the head and the modifier in memory, which is not available for nonwords. As a consequence, the expected facilitation from the head transparency emerges later in the experiment. These effects of the character transparency emerged only the reaction time analysis and were absent in the analyses of subgaze durations. This suggests that the effect occurs late, after the eye has completed extracting information from the individual characters.

Individual differences. Finally, individual differences were present (Figure 3-2, Panel c), notably for trials with the fixation mark placed at the central position. As can be seen in Panel c, the central fixation position elicited faster response times, suggesting that this central position is the optimal viewing position for isolated compound reading. For readers who have stayed longer in Canada, however, the advantage of this optimal viewing position became increasingly smaller. Recall that *LengthOfStayCanda* is correlated with other predictors (e.g., the amount of exposure to Japanese, age, and reading ability), hence a precise interpretation of this effect requires further research (cf. Goral et al., 2008, for dissociation of age and linguistic effects in lexical attrition). Table 3-2 also lists the contribution of *LeftKanjiNeighbourResid*: Response times decreased (effect size = -38 ms) with increasing *LeftKanjiNeighbourResid*. We discuss the interpretation of this effect below in the analyses of the subgaze durations.

First subgaze duration analysis. As in the analysis of the response times, items and subjects with a large number of incorrect responses were removed from the dataset. For the remaining data points, the number of fixations elicited varied from 1 to 15 per trial, with the mode at 3 fixations (3,203 trials), followed by 2 fixations (2,772 trials) and 4 fixations (1,348 trials). A small minority of 428 trials elicited only one fixation. In the subsequent analyses, we focused on subgaze durations. Subgaze counts varied from one to eight with the mode at two subgazes (72% of the subgazes). In the subsequent subgaze duration analyses, we focus on the trials with exactly two subgazes, which represent the large majority of data points.[4]

For the analysis of the first subgaze durations (3,711 data points), initial fixations shorter than 100 ms were removed. In a quantile-quantile plot of the first subgaze durations, these short fixations patterned differently from the remaining durations. Trials that elicited incorrect responses for the lexical decision and trials with a blink were also excluded. The remaining durations were subsequently log-transformed to adjust for non-normality.

[4] The analysis of the subgaze counts indicates that this subset is biased slightly towards words preceded by trials with a short response latency, words responded to by readers who had only recently left Japan, words with fewer strokes, and words with higher frequencies.

Table 3-3. Estimate, standard error, t-value, p-value, and effect size of influential predictors for the first subgaze durations for trials with two subgazes.

	Type	Estimate	SE	t-value	p-value	Effect Size (ms)
(Intercept)		6.215	0.077	80.25	< 0.0001	
Trial	T	0.000	0.000	-2.05	0.0400	-65
PreviousRT	T	0.046	0.013	3.64	0.0003	42
EyePosition (Right)	T	-0.827	0.142	-5.83	< 0.0001	-94
LeftKanjiStrokesResid	F	0.014	0.002	7.21	< 0.0001	96
RightKanjiStrokesResid	F	-0.009	0.002	-5.68	< 0.0001	-57
LeftKanjiComplex (Simple)	F	0.013	0.021	0.62	0.5373	0
LeftKanjiRadical Combinability	R	0.014	0.005	2.94	0.0033	18
RightKanjiRadical Combinability	R	-0.020	0.005	-4.08	< 0.0001	-26
RightKanjiRadical TokenFreqResid	R	0.000	0.009	0.05	0.9589	1
LeftKanjiTokenFreq	C	-0.033	0.004	-7.77	< 0.0001	-99
LeftKanjiNeighbourResid	C	-0.120	0.053	-2.24	0.0252	-58
LeftKanjiAoAResid	C	0.018	0.004	4.27	< 0.0001	46
RightKanjiTokenFreq	C	0.005	0.004	1.21	0.2251	16
WholeWordFreq	W	-0.008	0.004	-2.07	0.0381	-18
GoogleDocFreqResid	W	-0.010	0.005	-2.24	0.0252	-24
RightKanjiRadical Transparency	S	0.008	0.003	2.83	0.0047	16
LeftKanjiStrokesResid x LeftKanjiComplex (Simple)	F x F	0.018	0.005	3.78	0.0002	Fig 3-3 (a)
LeftKanjiStrokesResid x EyePosition (Right)	F x T	-0.042	0.004	-10.21	< 0.0001	Fig 3-3 b)
RightKanjiStrokesResid x EyePosition (Right)	F x T	0.027	0.004	5.99	< 0.0001	Fig 3-3 (c)
RightKanjiRadical TokenFreqResid x EyePosition (Right)	R x T	0.067	0.028	2.42	0.0156	Fig 3-3 (d)
LeftKanjiNeighbourResid x LeftKanjiTokenFreq	C x C	0.000	0.004	-0.02	0.9826	Fig 3-3 (e, f)
LeftKanjiNeighbourResid x EyePosition (Right)	C x T	-0.235	0.110	-2.15	0.0318	Fig 3-3 (e, f)
LeftKanjiNeighbourResid x LeftKanjiTokenFreq x EyePosition (Right)	C x C x T	0.029	0.010	2.86	0.0042	Fig 3-3 (e, f)
LeftKanjiTokenFreq x EyePosition (Right)	C x T	0.040	0.010	4.03	0.0001	Fig 3-3 (g)
LeftKanjiNeighbourResid x RightKanjiTokenFreq	C x C	0.008	0.003	2.44	0.0149	Fig 3-3 (h)

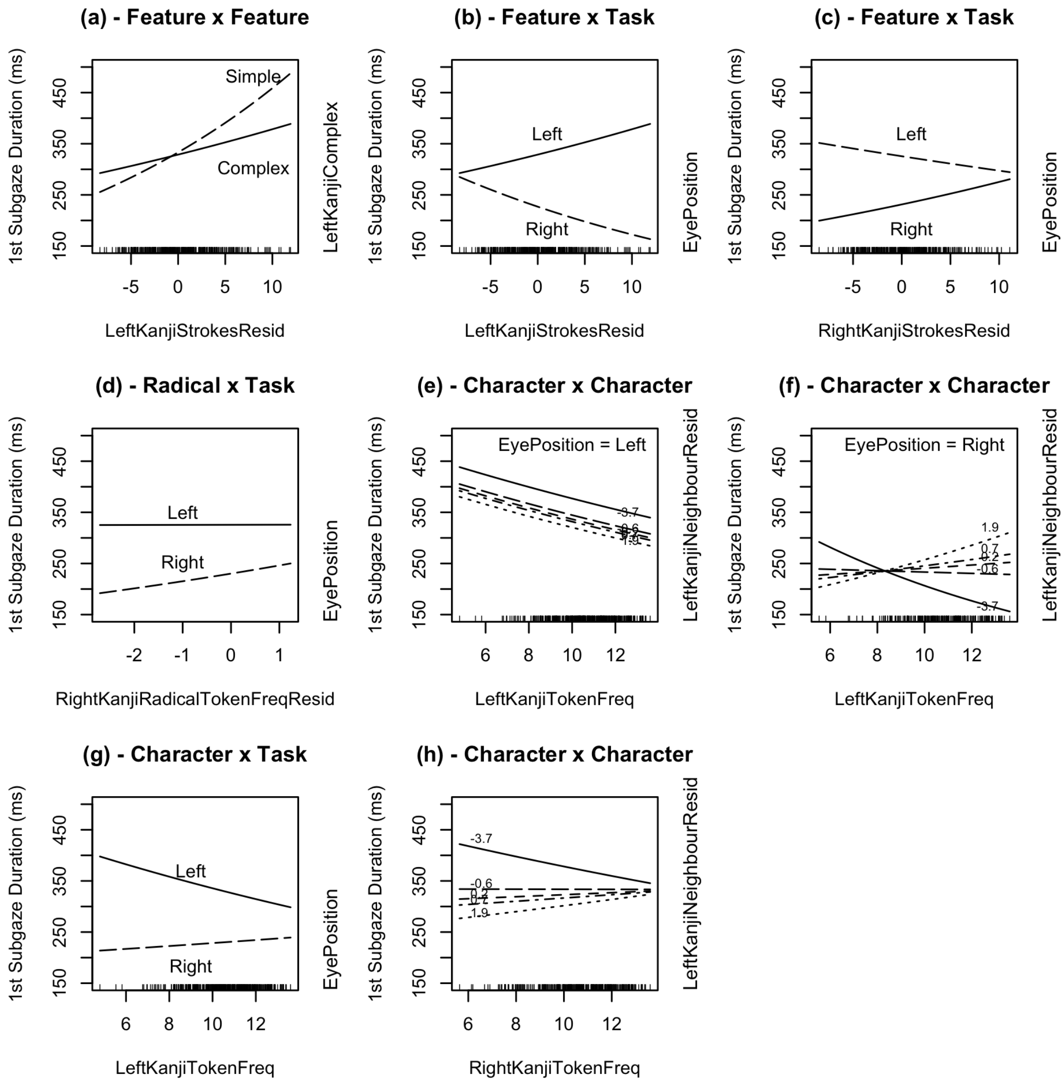


Figure 3-3. Interactions co-determining the first subgaze durations in trials with two subgazes

We fitted a mixed-effects model with subjects and items as crossed random effect factors to the first subgaze durations. We considered all pairwise interactions and removed unsupported coefficients from the model specification.

To safeguard against adverse effects of outliers, data points with absolute standardized residuals exceeding 2.5 were removed and the model was refitted. The coefficients of this model are summarized in Table 3-3, and the interactions are visualized in Figure 3-3. The random effect structure of this model comprised random intercepts for item (*SD* = 0.07) and subject (*SD* = 0.18), by-subject random slopes for *Trial* (*SD* = 0.0003), and by-subject random contrasts for *EyePosition* (*SD* = 0.39). The random contrasts for *EyePosition* capture the heteroscedasticity characterizing the two eye positions, with greater variance when the eye is fixating on the right character. The standard deviation of the residual error was 0.25.

Feature-level effects. As expected, feature-level complexity contributed substantially to the first subgaze durations. The effect of *LeftKanjiStrokesResid* was attenuated for complex characters (Figure 3-3, Panel a). For words with a complex left character (e.g., 根 and 針), the effect of stroke complexity was muted (the solid line in Panel a), compared to simplex characters with the same stroke counts (e.g., 馬 and 骨). In other words, if a character can be broken down into radicals, then processing proceeds more quickly than if the character is an undecomposable whole. What we are seeing here is the advantage of compositionality. An advantage of compositionality has been reported as well for complex words in auditory comprehension of Danish (Balling & Baayen, 2008). This interaction can be understood as arising due to a redundancy gain (see Raab, 1962, on statistical facilitation) in a dual-route architecture for perceptual identification of characters. Complex characters would then have two identification routes: a direct route using full character access representations, as well as a decompositional route using access representations of constituent radicals. The processing advantage of complex characters would then arise due to overlap in the distribution of the completion times of the two routes, with the slower route occasionally winning the race for identification whenever the faster route happens to be slow. We note here that the mathematics of statistical facilitation has been worked out for independent channels only. For inter-dependent channels, we hypothesize that a similar advantage still holds (see Dijkstra, Frauenfelder, & Schreuder, 1993, and Miller, 1982).

Character complexity also interacted with the location of the fixation (*EyePosition*) illustrated for *LeftKanjiStrokesResid* in Panel b and *RightKanjiStrokesResid* in Panel c. More complex characters elicited longer subgaze durations when the character was currently fixated on, but shorter subgaze durations when the character was not fixated on. This pattern resembles parafoveal-on-foveal effects as reported in sentence reading, with complexity and difficulty in the parafoveal region attracting attention and shortening the time the eye remains on the current constituent (Hyönä & Bertram, 2004; Kennedy & Pynte, 2005; Kliegl, Nuthmann, & Engbert, 2006; Pynte, Kennedy, & Ducrot, 2004). The processing of the non-fixated information unit indicates that the strict eye-mind assumption is too restrictive.

Radical-level effects. The type frequency of the characters' radicals, *LeftKanjiRadicalCombinability* and *RightKanjiRadicalCombinability*, was inhibitory for the left character (effect size = 18 ms) and facilitatory for the right character (effect size = -26 ms). The asymmetrical contributions of the left and the right radicals arose possibly because the semantic class marked by the modifier's radical was incompatible with that of the whole word (see also Miwa et al., 2012, for asymmetrical contribution of the left and the right radicals). In addition, *RightKanjiRadicalTokenFreqResid* co-determined the first subgaze durations but in an attention-dependent manner (Panel d): Its inhibitory contribution was evident only when the eye was on the right character. Note that although radical properties co-determined the first subgaze durations, the magnitudes of their effects were small or only *EyePosition*-specific.

Character-level effects. An effect of *LeftKanjiTokenFreq* was present in an interaction with *EyePosition* and *LeftKanjiNeighbourResid*, the type count of the number of two-character words sharing the left character. When the eye was fixating on the left character (Panel e), regardless of the number of the left kanji's neighbours, *LeftKanjiTokenFreq* speeded up recognition. When the eye was fixating on the right character, a cross-over interaction was observed (Panel f). Words with few *LeftKanjiNeighbourResid* showed facilitation from the left character's frequency. As the number of completions increased, this facilitation disappeared and reversed into inhibition. Panels (e) and (f) together illustrate a general preference for processing the left character regardless of the initial eye position. *LeftKanjiTokenFreq* therefore shows an expected facilitatory effect when the character is attended (Panel g).

In addition to the effect of *LeftKanjiTokenFreq*, an effect of *RightKanjiTokenFreq* was present but only in an interaction with

LeftKanjiNeighbourResid (Panel h): When there are few possible completions on the right (low *LeftKanjiNeighbourResid*), facilitation by the right character's frequency was observed. However, in the presence of greater uncertainty about the identity of the right character in a dense neighbourhood, readers cannot utilize *RightKanjiTokenFreq*. This is in line with Hyönä, Bertram, and Pollatsek's (2004) report that the second constituent is processed more deeply when it is more constrained. In their sentential reading study with an eye-movement–contingent display change technique, the change effect associated with the second constituent was stronger for words with a first constituent with low frequency and small family size. The effect of *RightKanjiTokenFreq* for both eye positions is consistent with the previously discussed effect of parafoveal preprocessing of feature properties (Panels b and c). There was also a significant effect of *LeftKanjiAoAResid*: Characters that were learned later required longer subgazes (effect size = 46 ms), adding to the early special importance of the left constituent.

Word-level effects. More frequent compounds elicited shorter first subgaze durations, as reflected by the negative coefficients of *WholeWordFreq* (-18 ms) and *GoogleDocFreqResid* (-24 ms). Such an early contribution of whole word frequency was also reported by Kuperman et al. (2009) for Dutch. As we shall see below, the effect of compound frequency became stronger at the second subgaze.

Phonological effects. Character phonology, *LeftKanjiHomophones* and *RightKanjiHomophones*, did not co-determine the first subgaze duration. This is consistent with the hypothesis that homophonic effects in visual word recognition are due to rebounding activation from phonology to orthography (Tamaoka, 2005; Pexman, Lupker, & Jared, 2001). If this line of reasoning is correct, we should be able to observe phonological effects at the second subgaze duration (see below). *LengthOfStayCanada* was not significant either.

Semantic effects. Furthermore, there was an inhibitory effect of *RightKanjiRadicalTransparency* (16 ms). If the radical is more transparent, it is more effective in activating its own typically more general meaning (e.g., 月 ‘body part’ in 脳 ‘brain’), which will compete with the meaning denoted by its character. Unlike in the analysis of response times, *LeftKanjiTransparency* and *RightKanjiTransparency*, both of which evaluate the semantic contribution of the character to the meaning of the two-character compound, did not reach significance for the first subgaze duration. Apparently, at the first subgaze, it is only a local semantic relation, transparency of the radical and its character, that is available for processing.

Second subgaze duration analysis. 3,731 data points for the second subgaze durations in the trials with two subgazes were analyzed in a mixed-effects model with subjects and items as crossed random effect factors. A square root transformation was used to adjust non-normality in the distribution of the subgaze durations. The random effect structure of the final model comprised random intercepts for item (*SD* = 0.97) and subject (*SD* = 1.13), by-subject random slopes for centralized *Trial* (*SD* = 0.003), centralized *PreviousRT* (*SD* = 0.57), centralized *PreviousSubgazeDuration* (*SD* = 0.15), and by-subject random contrasts for *EyePosition* (*SD* = 1.11). The standard deviation of the residual error was 2.86. Table 3-4 lists the coefficients of the model and Figure 3-4 visualizes the interactions.

Table 3-4. Estimate, standard error, t-value, p-value, and effect size of influential predictors for the second subgaze durations for trials with two subgazes.

	Type	Estimate	SE	t-value	p-value	Effect size (ms)
(Intercept)		20.773	1.118	18.58	< 0.0001	
PreviousSubgazeDuration	T	-0.718	0.040	-17.88	< 0.0001	-743
PreviousRT	T	1.095	0.203	5.40	< 0.0001	107
Trial	T	0.000	0.001	-0.16	0.8748	-5
EyePosition (Right)	T	-1.945	0.743	-2.62	0.0089	-18
LengthOfStayCanada	I	-0.149	0.257	-0.58	0.5616	-23
LeftKanjiStrokesResid	F	0.109	0.041	2.67	0.0076	78
RightKanjiStrokesResid	F	-0.121	0.044	-2.74	0.0062	-82
RightKanjiTokenFreq	C	-0.200	0.046	-4.35	< 0.0001	-57
RightKanjiAoAResid	C	0.146	0.052	2.82	0.0048	41
RightKanjiNeighbourResid	C	-0.249	0.067	-3.72	0.0002	-49
WholeWordFreq	W	-0.134	0.104	-1.29	0.1960	-79
GoogleDocFreqResid	W	-0.335	0.058	-5.83	< 0.0001	-82
LeftKanjiHomophones	P	-0.199	0.077	-2.58	< 0.0001	-28
RightKanjiHomophones	P	0.425	0.079	5.37	< 0.0001	61
LeftKanjiRadicalTransparency	S	0.353	0.170	2.08	0.0374	10
LeftKanjiStrokesResid x EyePosition (Right)	F x T	-0.135	0.042	-3.22	0.0013	Fig 4 (a)
RightKanjiStrokesResid x EyePosition (Right)	F x T	0.178	0.045	3.94	0.0001	Fig 4 (b)
LeftKanjiRadicalTransparency x WholeWordFreq	S x W	-0.060	0.025	-2.42	0.0156	Fig 4 (c)
LengthOfStayCanada x EyePosition (Right)	I x T	0.605	0.262	2.31	0.0210	Fig 4 (d)

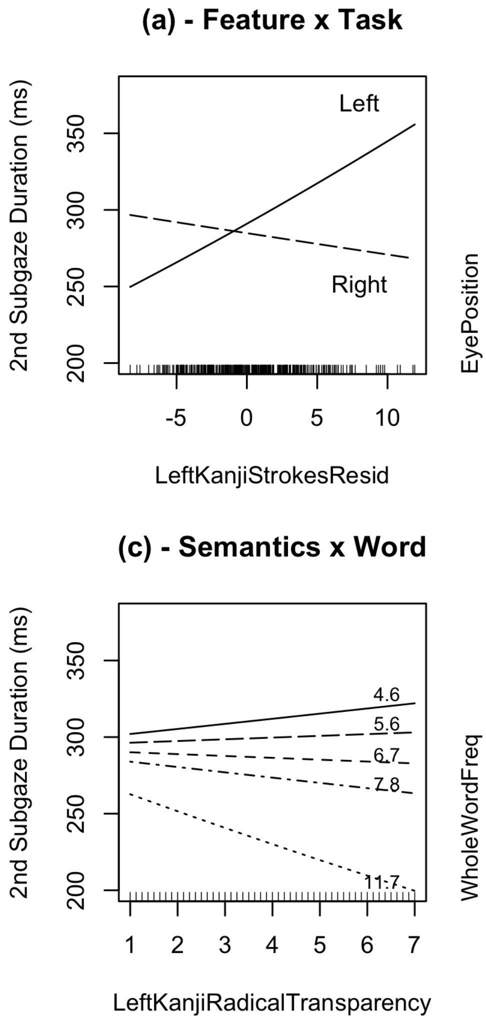


Figure 3-4. Interactions co-determining the second subgaze durations in trials with two subgazes

Feature-level effects. As can be seen in Figure 3-4, Panels a and b, the effects of character stroke complexity, *LeftKanjiStrokesResid* and *RightKanjiStrokesResid*, depended on the location of the eye fixation. The general patterns of these interactions are comparable to those observed for the first subgaze duration (Figure 3-3, Panels b and c). However, at this second subgaze, if the eye fixated on the left character, *LeftKanjiStrokesResid* greatly slowed down the second subgaze (the solid line, Panel a), while if the eye fixated on the right character, the effect of *LeftKanjiStrokesResid* was muted. The effects of *RightKanjiStrokesResid* showed a reversed pattern (Panel b). Interestingly, the effects of the two character stroke complexities are small when the eye rests on the right character, but large when the eye rests on the left character. This difference may be due to the preferential processing path from left to right. If the reader starts at the left, the second subgaze duration concerns the right character. At this point, a substantial amount of information is already available from the first character, smoothing the way for reading the second character. However, if the reader starts from the right character, then the second subgaze duration concerns the left character, the normal starting position for reading, and therefore inviting more in-depth processing of the left character. *LeftKanjiComplex* and *RightKanjiComplex*, the measures gauging the number of radicals in a character, did not contribute to the second subgaze duration.

Character-level effects. The contributions of *RightKanjiTokenFreq* (-57 ms), *RightKanjiNeighbourResid* (-49 ms), and *RightKanjiAoAResid* (41 ms) are comparable to the corresponding effects of the left character at the first subgaze. Whereas *LeftKanjiTokenFreq*, *LeftKanjiNeighbourResid*, and *LeftKanjiAoAResid* contributed at the first subgaze, they did not reach significance at the second subgaze. This suggests that the weight of importance shifts from the left character to the right character in this later time frame.

Word-level effects. As expected, the effects of frequency and contextual diversity of the whole word, *WholeWordFreq* and *GoogleDocFreqResid*, became larger at the second subgaze (-79 ms and -82 ms respectively; compare with the corresponding effect sizes of -18 ms and -24 ms at the first subgaze). As will be discussed below, *WholeWordFreq* interacted with *LeftKanjiRadicalTransparency.*

Phonological effects. Significant contributions of the numbers of homophonic characters were present for both the left and the right characters (*LeftKanjiHomophones* and *RightKanjiHomophones*, -28 ms and 61 ms respectively). Consistent with the analysis of response times (Table 3-2, 56 ms), *RightKanjiHomophones* was inhibitory. Furthermore, there was a smaller facilitatory effect of *LeftKanjiHomophones*, which contrasted with the inhibitory effect of *LeftKanjiHomophones* reported in Tamaoka's (2005) lexical decision study. This difference may be due to the different kinds of nonwords that we used, which included two-character words with illegal left characters. The late emergence of these homophone effects is consistent with the hypothesis that homophonic characters are activated only after the target character's phonological representation has been activated (rebounding activation; Tamaoka, 2005).

Semantic effects. The semantic congruity between the characters and their semantic radical, *LeftKanjiRadicalTransparency* co-determined the second subgaze durations (Figure 3-4, Panel c). The processing advantage for words with semantically transparent constituents is consistent with the results of Libben et al. (2003). However, facilitation was restricted to higher frequency words and disappeared for low frequency words. *LeftKanjiRadicalTransparency* facilitates the recognition only when *WholeWordFreq* is high. Conversely, the effect of *WholeWordFreq* was strongest for words with high *LeftKanjiRadicalTransparency*. This interaction suggests that whole word frequency effect is at least in part a semantic effect.

Individual differences. Finally, Panel d shows individual differences such that an expected effect of *LengthOfStayCanada* was observed when the eye was on the right character (i.e., the end point of a preferential left-to-right scan): A processing advantage was seen in readers with short *LengthOfStayCanada*.

The kinds of the effects observed at the second subgaze are qualitatively similar to those observed for the lexical decision response times. Interestingly, however, not only the second but also the first subgaze durations correlated with the RTs ($r = 0.34$, $p < 0.0001$, for the first subgaze duration; $r = 0.51$, $p < 0.0001$ for the second subgaze duration) with comparable β in the regression analysis ($\beta = 0.14$ and $\beta = 0.16$ respectively).

Discussion

Overall, the analysis of the first gaze durations identified contributions of lexical distributional properties at all levels of the morphographic structural hierarchy shown in Table 3-1. Although whole word frequency, constituent frequency, and radical frequency all co-determined first subgaze durations, the magnitude of their contributions differed. Properties of characters contributed robustly to a larger extent than properties of radicals and properties of whole word units, as diagnosed by their feature complexity, frequency, or transparency. The large contributions of characters suggest that the characters, rather than radicals, are the dominant recognition units for two-character words. Importantly, the above effects were observed across all subjects because we carefully checked for random-effect slopes for subject for our predictors. The present findings are more consistent with the character-based models of two-character word recognition (Tamaoka & Hatsuzuka, 1998; Joyce, 2004). However, the presence of both whole word frequency and radical effects at the first subgaze indicates that models positing that lexical access would proceed by first accessing the character and only then accessing the radical and the whole word representation are too restrictive.

With regard to the relative importance of the left and the right constituents, the properties of the left character contributed more than those of the right character at the first subgaze. This suggests that it is more effective to parse two-character words from left to right, although when read from right to left, the properties of the right character come into play as well, albeit to a lesser extent.

Thus far, we have interpreted the second subgaze in the same way as the first subgaze duration. However, in trials with more than one subgaze, the last subgaze was interrupted by the button press, which terminated the trials. This raises the question of to what extent the second subgaze is interpretable as a measure of information extraction and lexical access. The response time and the second subgaze duration incorporate the time required for motor response planning and response execution, estimated to be on the order of magnitude of 200 ms by Schmidt (1982). Given that the mean lexical decision response time in trials with two subgazes was 653 ms, it is estimated that the lexical decision was finalized around 653 – 200 = 453 ms post stimulus onset, i.e., after the first subgaze (M = 323 ms) but well before the end of the second subgaze. Assuming that the response execution time is constant, apart from random execution noise, and independent from lexical properties, then only the intercept of the regression model for the second subgaze is affected.[5]

The larger contributions of character properties compared to radical properties, particularly during the early processing stages, indicate that two-character words are processed in a character-driven manner, rather than by strictly combinatorial processes. However, joint contributions of morphographic units at all levels of the linguistic structure suggest that the character-based model is not sufficient to fully capture the complexity of morphographic word recognition at its current state. With respect to relative importance of the left and the right characters, eye-tracking highlighted their contributions at early and late processing stages respectively. Although the right character contributes more prominently to lexical decision responses, this was not because the right character is the primary access unit but because it contributes late when lexical

[5] The assumption that response planning and execution time is constant and does not vary with lexical properties may involve a simplification. For instance, Abrams and Balota (1991) observed that word frequency affects not only the timing but also the force with which the response is executed. As we asked our participants to keep their fingers on the response buttons during the experiment, the consequences of the differences in the force with which lexical decisions may have been executed for the estimates of the lexical decision speed and second subgaze durations are negligible.

decisions are made. Furthermore, semantic transparency effects for radicals indicate that radicals are not mere orthographic components.

Finally, it was also notable that the magnitude and the direction of lexical effects were modulated by readers' locus of attention in a left-to-right preferred processing path such that lexical properties of the fixated and non-fixated characters showed asymmetrical joint contributions.

It might, however, be argued that the character-driven processes we observed were induced by the large inter-character space that goes hand in hand with the relatively large character font size. Similarly, the small whole word frequency effects observed during the early time frame might be merely due to visual acuity limitation. Bertram and Hyönä (2003) investigated an effect of word length on morphological processes in Finnish and suggested that a decompositional route dominates over a direct route when processing long compounds. If a direct route to the compound representation also exists in Japanese, a smaller font size may trigger a substantially larger whole compound frequency effect at the early stage.

In addition to the font size, it might also be argued that the small contributions of radicals during the early stage of lexical processing in Experiment 1 were due to the nature of the nonwords. The nonwords in Experiments 1 were random combinations of characters. Hence, readers would not have to zoom in on radicals to distinguish words from nonwords. We evaluated the font size and nonword type accounts in Experiment 2.

Experiment 2: Evaluation of the font size and nonword type accounts

In Experiment 2, we tested whether the pattern of lexical activation we observed in Experiment 1 generalizes to words presented in the more commonly used 40-sized fonts (visual angle = 1.64º). The 40-size font represents a typical font size used in previous isolated word lexical decision studies (e.g., 1.38º in Feldman & Siok, 1999; 1.6º in Miwa et al., 2012; 1.23º in Myers et al., 2006; 2.05º in Shen & Forster, 1999; 1.6º in Taft & Zhu, 1997; 2.78º in Zhou, Marslen-Wilson, Taft, & Shu, 1999, where that the viewing distance was assumed to be 70 cm unless reported otherwise). In Experiment 2, we also used nonwords containing a non-existing character, with the aim of forcing readers to pay closer attention to intra-character components. Under those circumstances, the effect of radicals may emerge more prominently. However, if reading Japanese two-character compounds is fundamentally character-driven, then this manipulation of the nonwords should not affect the main patterns of results.

Method

Participants. Twenty-one native Japanese readers (17 females, mean age = 23.3 years old, *SD* = 5.9) participated at the University of Alberta, Canada.

Materials. Two hundred words were sampled randomly from the set of words used in Experiment 1, equally across ten frequency-ordered bins. An equal number of nonwords were prepared by replacing either the left or the right character's intra-character component with an existing constituent to make a non-existing character. Half the nonwords contained a non-existing left character, and the other half contained a non-existing right character.

Procedure. The procedure was identical to that in Experiment 1, but words were presented in smaller 40-size font (visual angle for each character = 1.64º).

Results

Response time analysis. The data were trimmed, and the response times were transformed in the same way as in Experiment 1. A mixed-effects model was fitted to inversely transformed response times for 192 words (3,559 data points). In our final model, the random effect structure comprised random intercepts for item ($SD = 0.10$) and subject ($SD = 0.18$), and by-subject random slopes for centralized *Trial* ($SD = 0.05$) and *PreviousRT* ($SD = 0.08$). The standard deviation of the residual error was 0.25.

As fixed effects, we identified *WholeWordFreq* ($p < 0.0001$, effect size = -92 ms) and *GoogleDocFreqResid* ($p < 0.0001$, effect size = -153 ms) as dominant lexical effects. The left and the right characters contributed to a comparable extent: *LeftKanjiTokenFreq* (p < 0.0284, effect size = -37 ms) and *RightKanjiTokenFreq* ($p < 0.0158$, effect size = -40 ms). Importantly, although the task forced the readers to attend to the intra-character structure, only a *Trial*-dependent small effect of *LeftKanjiRadicalTransparency* was observed (effect size changed from -16 ms to 28 ms, as the experiment went by).

LengthOfStayCanada did not have a significant main effect, as in Experiment 1 (see Appendix 3-2 for the full summary of the significant fixed effects).

First fixation duration analysis. For analyses of eye movements, data points excluded for the response time analysis were excluded here as well. Words were scanned with two fixations most of the time (10% for a single fixation, 65% for two fixations, 20% for three fixations, and 3% for four fixations), and fixation counts ranged from 1 to 6 ($M = 2.2$, $SD = 0.7$). Since two fixations constituted the majority of the trials, we analyzed first and second fixation durations in trials with exactly two fixations.[6]

As in Experiment 1, only the trials with a correct response that elicited two fixations were analyzed (192 words, 2,272 data points). Initial fixations shorter than 100 ms and longer than 800 ms were removed (5 data point). In our final model fitted to the log-transformed first fixation durations, the random effect structure comprised random intercepts for item ($SD = 0.06$) and subject ($SD = 0.10$) and by-*Trial* random slopes for subjects ($SD = 0.03$). The standard deviation of the residual error was 0.17. The fixed effect structure comprised a small yet significant effect of *GoogleDocFreqResid* ($p = 0.006$, effect size = -36 ms) and large contributions of the left character properties, such as *LeftKanjiTokenFreq* ($p < 0.0001$, effect size = -120 ms). Importantly for the purpose of Experiment 2, radical properties did not contribute prominently: the observed radical effect of *LeftKanjiRadicalCombinability* was small and inhibitory ($p = 0.0469$, effect size = 15 ms) and is comparable to its effect observed the first subgaze duration analysis in Experiment 1. In Experiment 2, the right characters' properties contributed more prominently than Experiment 1. Interestingly, as in Experiment 1, the left character effects and the right character effects were asymmetrical, and the magnitudes of effects for the former were larger: For example, *LeftKanjiStrokesResid* inhibited (p = 0.0001, effect size = 135 ms) while *RightKanjiStrokesResid* facilitated (p = 0.0001, effect size = -60 ms), and *LeftKanjiTokenFreq* facilitated (p = 0.0001, effect size = -120 ms) while *RightKanjiTokenFreq* inhibited (p = 0.0079, effect size = 30 ms). All of these effects well replicate the findings in Experiment 1. Analyses of subgaze durations replicated the character-driven processing pattern (see Appendix 3-2 for the full summary of the significant fixed effects).

Second fixation duration analysis. We fitted a mixed-effects model to the second fixation durations in the subset of trials analyzed above. In our final model, the random effect structure comprised random intercepts for item (*SD* = 0.63) and subject (*SD* = 1.20), and by-subject random slopes for centralized *Trial* (*SD* = 0.36). The standard deviation of the residual error was 2.26.

As fixed effects, as in Experiment 1, properties of the right character and the whole compound unit dominated: *WholeWordFreq* (p = 0.0001, effect size = -43 ms), *GoogleDocFreqResid* (p = 0.0001, effect size = -88 ms), *RightKanjiTokenFreq* (p = 0.0001, effect size = -48 ms), *RightKanjiAoAResid* (p = 0.0014, effect size = 33 ms). Left character effects did not reach significance. Note that, at this later fixation, whole word effects are large in magnitude, and

---

[6] The analysis of the fixation counts indicates that this subset is biased slightly towards words preceded by trials with a short response latency, words presented later in the experiment, words

RightKanjiTokenFreq shows a standard facilitatory frequency effect. Interestingly, this later time frame was also co-determined by the *Trial*-dependent effect of *LeftKanjiRadicalTransparency*, as seen in the response time analysis (See Appendix 3-2 for the summary of all significant predictors).

Discussion

Experiment 2 largely replicates the main findings of Experiment 1. Even when words are presented in smaller-size font and together with different nonwords, the effects of character properties were more prominent than those of radicals properties during the early processing stages. Experiment 2 also replicates that a whole word frequency effect emerges already in the early time frame with small yet significant effects, and contributes more strongly in the later time frame. The small effect of the frequency of the whole word unit at the first fixation in Experiment 2 suggests that the small effect size associated with the early whole word frequency effect in Experiment 1 was not due to a visual acuity constraint, but is an essential characteristic of morphological processing in Japanese observable across all subjects (i.e., random slopes for subjects were not justified for a whole word frequency effect). Importantly, when the subset of data in Experiment 1 with the left fixation position was analyzed (185 words for each subject), the pattern of results remained unchanged, suggesting that the fixation position and statistical power did not contaminate the comparison between the two experiments. With respect to relationship between non-linguistic task demand and lexical processing, the above results are in line with

---

with fewer strokes, and words with higher whole word and right character frequencies.

Kaakinen and Hyönä (2010)'s eye-tracking sentential reading study with a manipulation of task demands. In their study, depending on whether the task was comprehension or proof-reading, readers adjusted eye movements already at the first fixation according to the given task demand, with regard to the landing position and the fixation duration. However, lexical effects were not modulated by the task demands during this early time frame, while they were in the later time frame proved by the gaze duration analysis. Experiment 2 of the present study similarly demonstrated that, even when the task requires attention to intra-character radical components and the font size motivates fewer eye movements, character-driven processing remains unaffected.

General discussion

In this visual lexical decision with eye-tracking study, we tested several hypotheses in parallel: namely, whether the processing of morphographic two-character words proceeds strictly from the smallest units to large units in a bottom-up combinatorial manner, whether the right character is quantitatively and qualitatively more important than the left character, whether semantic radicals are processed semantically, and how non-linguistic variables affect lexical processes.

First, we studied the temporal order in which a two-character word and its constituent characters and radicals are activated in the course of lexical access. During the earliest time frame, both in Experiment 1 and 2, we observed a larger effect of left character frequency than those of radical combinability and whole word frequency. The early emergence of whole word frequency effect replicates the previous findings for Dutch and Finnish (Kuperman, Bertram, & Baayen, 2008; Kuperman et al., 2009). During the later time frame, the effect of the frequency of the left character disappeared and was replaced by a large effect of the frequency of the right character. The magnitude of the whole word frequency effect increased in this later time frame. The early large effects of character frequency in combination with a small effect of whole word frequency, as well as later predominant effects of right character and whole word frequency effects, were replicated when words were presented in smaller fonts and presented with different types of nonwords. This indicates that the present character-driven processing signature does not depend on font sizes nor nonword-induced task demand in lexical decision.

Second, we studied the different contributions of the left and the right characters to the lexical decision responses. On the basis of the lexical decision response times alone, using frequency as a diagnostic for access to lexical representations, one would have to conclude that the right character is more important than the left (Experiment 1) or both are equally important (Experiment 2). Interestingly, the eye-tracking record revealed clear and strong frequency effects of the left character in the early time frame and those of the right character in the later time frame. This indicates that the right character advantage reflected in the response times arises not because the right character is the main morpheme to be processed first. Instead, response times predominantly reflect later processes (i.e., later information uptake and subsequent decision processes; cf., Tamaoka & Hatsuzuka, 1995).

The time-course of the left-then-right constituent activation observed in the present study is comparable to that in eye-tracking studies on alphabetic compound processing (Hyönä & Pollatsek, 1998; Kuperman et al., 2008; Pollatsek et al., 2000). It should be noted, however, that the two constituents do not simply facilitate processing at different points in time; We observed that one inhibited processing while the other was facilitatory in nature (see also Vergara-Martínez, Duñabeitia, Laka, & Carreiras, 2009, for qualitatively different EEG signatures between left and right constituents in Basque compound word reading).

Third, we were interested in the depth to which semantic radicals are processed. Slight yet significant contributions of semantic radical transparency were observed in both eye movement and response time analyses, providing further support for Feldman and Siok's (1997, 1999) and Miwa et al.'s (2012) claim that semantic radicals contribute to the semantic interpretation of words. An issue that should be considered in parallel is whether initial morphological decomposition is morpho-orthographic (Rastle, Davis, Marslen-Wilson, & Tyler, 2000; Taft & Nguyen-Hoan, 2010) or morpho-semantic in nature (Diependaele et al., 2011; Diependaele, Sandra, & Grainger, 2005; Feldman, O'Connor, & Moscoso del Prado Martín, 2009). The radical transparency effects observed in the earliest time frame (Experiment 1) partially support the latter claim that morpho-semantic processing is involved at very early processing stages. Note, however, that the early radical transparency effect we observed was not facilitatory but inhibitory, suggesting that the processing of semantic radicals was not in harmony with normal comprehension. A further research is required to confirm how robust this effect is.

Fourth, by manipulating the location of the fixation point and tracking eye movements, for the first time as an isolated word reading study, we found effects of a locus of attention on lexical processing. Strong parafoveal-on-foveal effects emerged, with the sign and the magnitude of stroke complexity effects modulated by the fixation location. When the eye attends to one character first, it is attracted to the other character when that character is highly complex, indicating the need for allocating processing resources to the other character (Kliegl et al., 2006; Pynte et al., 2004; Hyönä & Bertram, 2004). As a consequence, the greater the complexity of the unfixated character, the shorter the eye rests on the fixated character. Font size and task demand manipulations left the above pattern unchanged.

In what follows, we assess how well current models of morphological processing explain the temporal order and the magnitude of effects of whole word, character, and radical activation. The supra-lexical model of Giraudo and Grainger (2001) predicts the whole word to be activated before its constituents (i.e., strong effects of whole word frequency, weaker effects of character frequency, and the weakest effects of radicals in the earliest time frame). However, the time-course of activation that emerges from our data is one in which the character is activated first, followed by the activation of the whole word on one hand and the activation of radicals on the other: In the early time frame, strong character frequency effects pair with a small whole word frequency effect and small radical combinability effects, indicating an initial access to character representations and subsequent spreading activations to radical and whole word representations.

The multilevel interactive activation model proposed for Japanese by Tamaoka and Hatsuzuka (1998) correctly predicts, in the earliest time frame, that whole word effects should be smaller than character effects. It also correctly predicts rebounding phonological effects, which appeared late in our data. However, in this interactive activation model, semantic radicals are not represented by separate nodes. Given that combinability and transparency of semantic radicals affect lexical processes, albeit with small effect sizes, nodes for semantic radicals need to be incorporated in the model architecture.

Adding radical nodes to the model of Tamaoka and Hatsuzuka (1998) leads to the interactive activation architecture proposed for Chinese by Taft and Zhu (1997) and Taft et al. (1999) and for Japanese by Saito (1997). These models predict a time-course of activation that is exactly the opposite of the time-course predicted by the supra-lexical model. Now, radicals are supposed to be activated before characters, and characters before whole words. This architecture, however, is challenged by our eye-tracking data in that, in the earliest processing stages, effects of characters dominate over those of radicals.

Within the general interactive activation approach to lexical processing in Japanese, our data suggest a modification of both the model of Tamaoka and

Hatsuzuka (1998) on one hand and that of Saito (1997) on the other. The compromise presented in Figure 3-5 incorporates nodes for radicals, characters, and words as in the model of Saito (1997) but, unlike this model, it includes connections from the feature level that link up directly to the character level, by-passing the radical nodes. Consequently, radicals can be activated, either by receiving rebounding activation from the character level or by receiving activation from the feature level (the dotted line in Figure 3-5). Our current data do not allow us to estimate the relative importance of these two routes for activation of radicals. However, given that radical effects are not modulated by frequency of the characters nor by word frequency, processing of radicals proceeds independently, with character activation taking precedence at least in early processing stages. By including level-skipping links from features to characters, the model accounts for the fact the characters are the most prominent units from the earliest time frame onward: Characters receive more bottom-up support than radicals.

Interestingly, this level-skipping assumption we propose for Japanese is comparable to direct whole word activation routes assumed to function in parallel to sequential decompositional routes in recent morphological processing models for alphabetic languages (Diependaele et al., 2005, 2011; Pollatsek et al., 2000). Diependaele et al. (2005, 2011) observed facilitatory semantic transparency effects in masked priming. In order to account for this arguably early morpho-semantic processing, Diependaele et al. (2011) reason that direct whole word access routes should be assumed. While the primary evidence for our level-skipping assumption came from different magnitudes of lexical effects associated with different morphological levels, the above comparison to models for alphabetic languages allow us to predict that an early semantic transparency effect in masked priming may be observed similarly for Japanese and Chinese if complex characters are functionally comparable to multimorphemic words in alphabetic languages.

The question remains why character representations emerge as primary access units. Our hypothesis is that characters carry the greatest amount of information for a word's intended meaning, compared to radicals and compounds. Radicals often serve two purposes, depending on their position. They may denote general semantic categories, but they may also provide information about a character's pronunciation. As a consequence, they are unreliable cues to a word's meaning. Conversely, many two-character compounds are semantically at least partially transparent and compositional. The greater their compositionality, the more the burden of interpretation rests with the characters. In other words, characters may be the most important cues to meaning, compared to radicals (which are ambiguous and less useful cues) and compared to whole words (due to compositionality). We leave the validation of this hypothesis, for instance within the naive discrimination learning framework proposed by Baayen et al. (2011) to future research.

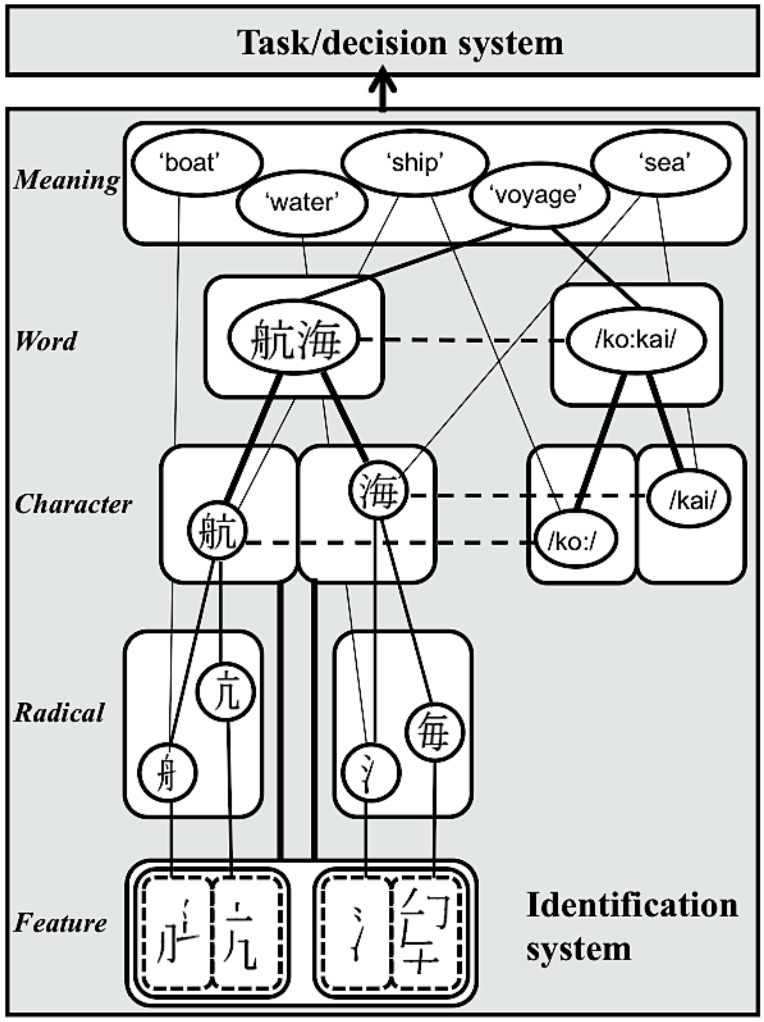


Figure 3-5. A character-driven processing model of Japanese two-character word recognition with semantic radicals as orthographic morphemes. The activations of morphographic neighbours, phonological neighbours, and semantic associates are not specified in the figure.

In addition to the level-skipping assumption, there are two other differences between the architecture proposed in Figure 3-5 and the models proposed in the literature. First, we take semantic radicals to be the smallest meaningful units in the identification system; in other words, we consider semantic radicals to be orthographic morphemes. In Figure 3-5, semantic radicals therefore have out-going connections that link up to the semantic representations. These links are motivated by the significant radical transparency effect observed in our data, consistent with the results of Feldman and Siok (1997, 1999) and Miwa et al. (2012). Although radical morphemes, unlike morphemes in alphabetic languages, do not function as primary recognition units, they nevertheless contribute to a word's meaning percept.

Second, Figure 3-5 makes it explicit that task demands and decision making strategies co-determine responses and potentially affect lexical processing at later processing stages. In Experiment 1, the effect of the accuracy on the previous trial, in interaction with right character frequency, on the RTs indicates changes in local response criteria, while the interaction between trial and right character transparency is indicative of changes in global response criteria. These interactions involving lexical distributional predictors indicate that the two systems are not strictly staged but function in parallel at least at later processing stages. This assumption of late involvement of the non-linguistic system is based on that in the Bilingual Interactive Activation (BIA+) model (Dijkstra & van Heuven, 2002), which makes it explicit that bilinguals cannot suppress activation of two languages even when activation of one language is sufficient for completion of a given task.

With regard to lexical predictors to be considered for the visual recognition of Japanese morphographic words, we are fully aware that the present study considered only 22 lexical variables and that it remains important to extend the range of predictors to include, for instance, imageability (e.g., Balota et al., 2004; McMullen & Bryden, 1987), visuoperceptual features and geometrical complexity (Grainger, Rey, & Dufau, 2008; Huang & Wang, 1992), collocational N-gram frequency (Arnon & Snider, 2010; Tremblay, Derwing, Libben, & Westbury, 2011), and whether a compound is endocentric (and right-headed) or exocentric (see Joyce, 2002 for consideration of compound formation principles).

Future research should also assess potential effects of individual differences on lexical access (see Andrews & Lo, 2011; Kuperman & Van Dyke, 2011; Yap, Balota, Sibley, & Ratcliff, 2011). Because we carefully checked for random-effect slopes for subject for our predictors, the main effects reported in the present study are very unlikely to be due to individual differences. Furthermore, our models can be used to extrapolate to domestic Japanese readers by setting the value of *LengthOfStayCanada* slightly below zero, in order to predict their expected response times. As the effects of *LengthOfStayCanada* are small, no major differences for domestic readers are anticipated. We leave it to future research to disentangle the precise contributions of length of stay, age, daily exposure to the language, and reading ability.

In conclusion, the present study documents processing consequences from all levels of morphographic structure, namely the radicals, the character, and the whole word. Eye-movements revealed that two-character words in Japanese are preferentially processed from the left character to the right character, with whole word frequency exerting an effect already from the earliest time frame. Importantly, the effects of character properties were robust and larger in magnitude than those of radicals and whole word properties at early processing stages. The patterns observed in all data combined led us to propose a character-driven architecture with a level-skipping assumption: Connections from the feature level by-pass the lower radical level and link up directly to the higher character level, allowing character effects to dominate early processing stages irrespective of font sizes and task demands.

Appendix 3-1. A correlation matrix among numerical predictors considered in this study. Predictors with the superscript ^{R} were end-products of a residualization procedure. The significant correlations at the 0.01 level are bolded.

Predictors	1	1^R	2	2^R	3	4	5	5^R
1. LeftKanjiStrokes	**1**	**0.95**	0	-0.01	**0.16**	-0.02	-0.01	**-0.15**
1^R. LeftKanjiStrokesResid		**1**	0	-0.01	**0.14**	-0.03	0.01	**-0.16**
2. RightKanjiStrokes			**1**	**0.92**	0.04	0.09	0.04	-0.03
2^R. RightKanjiStrokesResid				**1**	0.04	0.04	0.05	-0.02
3. LeftKanjiRadical Combinability					**1**	-0.04	**0.85**	0
4. RightKanjiRadical Combinability						**1**	-0.02	0.02
5. LeftKanjiRadical TokenFreq							**1**	**0.47**
5^R. LeftKanjiRadical TokenFreqResid								**1**
6. RightKanjiRadical TokenFreq								
6^R. RightKanjiRadical TokenFreqResid								
7. LeftKanjiNeighbour								
7^R. LeftKanjiNeighbourResid								
8. RightKanjiNeighbour								
8^R. RightKanjiNeighbourResid								
9. LeftKanjiTokenFreq								
10. RightKanjiTokenFreq								
11. LeftKanjiAoA								
11^R. LeftKanjiAoAResid								
12. RightKanjiAoA								
12^R. RightKanjiAoAResid								
13. WholeWordFreq								
14. GoogleDocFreq								
14^R. GoogleDocFreqResid								
15. LeftKanjiHomophones								
16. RightKanjiHomophones								
17. LeftKanjiRadical Transparency								
18. RightKanjiRadical Transparency								
19. LeftKanjiTransparency								
20. RightKanjiTransparency								

Predictors	6	6^R	7	7^R	8	8^R	9	10
1. LeftKanjiStrokes	-0.03	-0.02	**-0.31**	**-0.17**	-0.03	-0.03	**-0.31**	-0.01
1^R. LeftKanjiStrokesResid	-0.04	-0.03	0.00	0.09	-0.03	-0.03	**-0.15**	-0.01
2. RightKanjiStrokes	-0.03	-0.09	0.01	-0.05	**-0.38**	**-0.30**	0.09	**-0.26**
2^R. RightKanjiStrokesResid	-0.05	**-0.14**	0.00	-0.05	0.00	0.04	0.08	-0.08
3. LeftKanjiRadical Combinability	-0.05	-0.04	-0.09	-0.01	0.00	0.01	**-0.15**	0.00
4. RightKanjiRadical Combinability	**0.84**	0.00	-0.03	-0.05	**-0.15**	-0.10	0.01	**-0.13**
5. LeftKanjiRadical TokenFreq	-0.01	0.00	0.05	-0.02	0.00	-0.01	**0.12**	0.04
5^R. LeftKanjiRadical TokenFreqResid	0.06	0.07	-0.01	-0.01	0.02	0.01	0.00	0.03
6. RightKanjiRadical TokenFreq	**1**	**0.48**	-0.03	-0.06	-0.06	**-0.14**	0.03	**0.13**
6^R. RightKanjiRadical TokenFreqResid		**1**	-0.01	-0.02	-0.09	**-0.11**	0.01	0.00
7. LeftKanjiNeighbour			**1**	**0.85**	-0.01	-0.02	**0.53**	0.01
7^R. LeftKanjiNeighbourResid				**1**	0.01	0.04	0.00	-0.04
8. RightKanjiNeighbour					**1**	**0.88**	-0.04	**0.48**
8^R. RightKanjiNeighbourResid						**1**	-0.09	0.00
9. LeftKanjiTokenFreq							**1**	0.09
10. RightKanjiTokenFreq								**1**
11. LeftKanjiAoA								
11^R. LeftKanjiAoAResid								
12. RightKanjiAoA								
12^R. RightKanjiAoAResid								
13. WholeWordFreq								
14. GoogleDocFreq								
14^R. GoogleDocFreqResid								
15. LeftKanjiHomophones								
16. RightKanjiHomophones								
17. LeftKanjiRadical Transparency								
18. RightKanjiRadical Transparency								
19. LeftKanjiTransparency								
20. RightKanjiTransparency								

Predictors	11	11^R	12	12^R	13	14	14^R	15
1. LeftKanjiStrokes	**0.54**	0.00	0.05	0.05	-0.01	-0.04	-0.04	0.00
1^R. LeftKanjiStrokesResid	**0.37**	0.00	0.04	0.03	-0.01	-0.03	-0.04	-0.05
2. RightKanjiStrokes	-0.05	-0.03	**0.46**	0.00	-0.04	-0.03	-0.01	-0.06
2^R. RightKanjiStrokesResid	-0.06	-0.04	**0.22**	0.00	-0.04	-0.03	0.00	-0.05
3. LeftKanjiRadical Combinability	**0.19**	0.10	0.02	0.02	-0.01	-0.07	-0.07	-0.06
4. RightKanjiRadical Combinability	-0.02	-0.04	**0.14**	0.03	-0.04	-0.03	0.00	0.04
5. LeftKanjiRadical TokenFreq	-0.02	0.06	0.00	0.00	0.04	-0.03	-0.07	-0.04
5^R. LeftKanjiRadical TokenFreqResid	-0.09	-0.06	-0.04	-0.03	-0.01	-0.04	-0.04	0.03
6. RightKanjiRadical TokenFreq	-0.02	-0.02	0.02	0.05	0.04	0.02	0.00	0.03
6^R. RightKanjiRadical TokenFreqResid	-0.01	0.00	0.05	0.04	0.02	0.02	0.01	0.00
7. LeftKanjiNeighbour	**-0.60**	0.00	-0.03	-0.05	0.00	0.02	0.03	**-0.15**
7^R. LeftKanjiNeighbourResid	**-0.33**	0.00	-0.04	-0.05	**-0.14**	-0.08	0.00	**-0.18**
8. RightKanjiNeighbour	-0.01	-0.03	**-0.68**	0.00	0.01	0.03	0.02	0.04
8^R. RightKanjiNeighbourResid	-0.03	-0.08	**-0.47**	0.00	**-0.13**	-0.04	0.05	0.07
9. LeftKanjiTokenFreq	**-0.59**	0.00	0.01	0.00	**0.23**	**0.17**	0.04	0.00
10. RightKanjiTokenFreq	0.02	0.08	**-0.56**	0.00	**0.26**	**0.13**	-0.03	-0.03
11. LeftKanjiAoA	**1**	**0.66**	0.07	**0.13**	-0.01	**-0.10**	**-0.12**	-0.01
11^R. LeftKanjiAoAResid		**1**	0.08	**0.14**	0.09	-0.05	**-0.12**	-0.09
12. RightKanjiAoA			**1**	**0.65**	-0.05	-0.08	-0.06	-0.04
12^R. RightKanjiAoAResid				**1**	0.05	-0.03	-0.08	-0.02
13. WholeWordFreq					**1**	**0.59**	0.00	0.03
14. GoogleDocFreq						**1**	**0.81**	0.02
14^R. GoogleDocFreqResid							**1**	0.00
15. LeftKanjiHomophones								**1**
16. RightKanjiHomophones								
17. LeftKanjiRadical Transparency								
18. RightKanjiRadical Transparency								
19. LeftKanjiTransparency								
20. RightKanjiTransparency								

Predictors	16	17	18	19	20
1. LeftKanjiStrokes	-0.07	**-0.19**	0.02	0.08	0.07
1^R. LeftKanjiStrokesResid	-0.06	**-0.12**	-0.01	0.07	0.07
2. RightKanjiStrokes	0.02	-0.08	**-0.20**	0.02	-0.01
2^R. RightKanjiStrokesResid	0.01	-0.07	**-0.12**	0.04	0.01
3. LeftKanjiRadical Combinability	0.01	-0.03	-0.05	0.07	0.08
4. RightKanjiRadical Combinability	0.04	-0.04	0.07	-0.02	0.03
5. LeftKanjiRadical TokenFreq	0.01	-0.07	-0.07	0.08	0.09
5^R. LeftKanjiRadical TokenFreqResid	0.03	-0.07	-0.03	-0.01	0.02
6. RightKanjiRadical TokenFreq	0.07	-0.07	0.01	-0.02	0.03
6^R. RightKanjiRadical TokenFreqResid	0.03	-0.01	**-0.11**	-0.05	-0.03
7. LeftKanjiNeighbour	0.05	**0.27**	-0.09	-0.03	-0.04
7^R. LeftKanjiNeighbourResid	0.09	**0.34**	-0.07	-0.09	-0.05
8. RightKanjiNeighbour	-0.02	0.04	**0.23**	0.05	0.05
8^R. RightKanjiNeighbourResid	-0.06	**0.10**	**0.26**	0.01	0.01
9. LeftKanjiTokenFreq	-0.05	-0.03	-0.06	0.08	0.02
10. RightKanjiTokenFreq	0.07	-0.10	0.01	0.08	0.09
11. LeftKanjiAoA	-0.03	**-0.22**	0.00	-0.02	0.04
11^R. LeftKanjiAoAResid	0.00	**-0.12**	-0.09	-0.05	0.02
12. RightKanjiAoA	-0.05	-0.06	**-0.28**	-0.03	-0.09
12^R. RightKanjiAoAResid	-0.07	-0.07	**-0.19**	0.03	-0.06
13. WholeWordFreq	-0.04	**-0.13**	-0.05	0.03	0.02
14. GoogleDocFreq	-0.09	-0.06	-0.01	0.03	-0.03
14^R. GoogleDocFreqResid	-0.08	0.02	0.03	0.02	-0.06
15. LeftKanjiHomophones	**-0.16**	-0.07	-0.01	0.07	0.04
16. RightKanjiHomophones	**1**	0.05	0.05	-0.09	**-0.14**
17. LeftKanjiRadical Transparency		**1**	0.04	-0.10	-0.09
18. RightKanjiRadical Transparency			**1**	-0.01	0.07
19. LeftKanjiTransparency				**1**	**0.51**
20. RightKanjiTransparency					**1**

Appendix 3-2. Fixed effects in the models for the response times, first fixation durations, and second fixation durations in Experiment 2

Response time	Type	Estimate	SE	t-value	p-value	Effect size
(Intercept)		-1.120	0.098	-11.38	< 0.0001	
PreviousRT	T	0.124	0.022	5.61	< 0.0001	117
Trial	T	-0.082	0.015	-5.43	< 0.0001	-84
PreviousTrialCorrect (Incorrect)	T	0.086	0.016	5.39	< 0.0001	30
LeftKanjiStrokesResid	F	0.007	0.002	3.12	0.0018	41
LeftKanjiTokenFreq	C	-0.012	0.006	-2.19	0.0284	-37
LeftKanjiNeighbourResid	C	-0.023	0.010	-2.44	0.0147	-38
RightKanjiTokenFreq	C	-0.014	0.006	-2.42	0.0158	-40
RightKanjiAoAResid	C	0.021	0.007	3.24	0.0012	46
WholeWordFreq	W	-0.043	0.007	-6.45	< 0.0001	-92
GoogleDocFreqResid	W	-0.059	0.007	-8.71	< 0.0001	-153
LeftKanjiHomophones	P	-0.025	0.010	-2.47	0.0137	-34
LeftKanjiRadicalTransparency	S	0.005	0.005	0.92	0.3587	9
LeftKanjiRadicalTransparency x Trial	S x T	0.006	0.002	2.71	0.0067	-16:28

First fixation duration	Type	Estimate	SE	t-value	p-value	Effect size
(Intercept)		6.124	0.065	93.69	< 0.0001	
Trial	T	0.000	0.000	-0.81	0.4180	-8
LeftKanjiStrokesResid	F	0.020	0.002	13.08	< 0.0001	135
RightKanjiStrokesResid	F	-0.009	0.002	-4.93	< 0.0001	-60
LeftKanjiRadicalCombinability	R	0.011	0.006	1.99	0.0469	15
LeftKanjiTokenFreq	C	-0.037	0.004	-9.33	< 0.0001	-120
LeftKanjiAoAResid	C	0.028	0.004	6.21	< 0.0001	71
LeftKanjiNeighbourResid	C	-0.025	0.006	-3.99	0.0001	-42
RightKanjiTokenFreq	C	0.010	0.004	2.66	0.0079	30
RightKanjiAoAResid	C	-0.013	0.005	-2.85	0.0044	-30
RightKanjiNeighbourResid	C	0.016	0.006	2.78	0.0055	32
GoogleDocFreqResid	W	-0.013	0.005	-2.75	0.0060	-36

Second fixation duration	Type	Estimate	SE	t-value	p-value	Effect size
(Intercept)		74.774	1.535	48.70	< 0.0001	
PreviousFixationDuration	T	-9.340	0.233	-40.14	< 0.0001	-865
PreviousRT	T	0.788	0.152	5.19	< 0.0001	64
Trial	T	-0.006	0.001	-4.38	< 0.0001	-49
PreviousTrialCorrect (Incorrect)	T	0.689	0.185	3.73	0.0002	21
RightKanjiTokenFreq	C	-0.188	0.047	-4.03	0.0001	-48
RightKanjiAoAResid	C	0.172	0.054	3.19	0.0014	33
WholeWordFreq	W	-0.217	0.052	-4.17	< 0.0001	-43
GoogleDocFreqResid	W	-0.384	0.056	-6.83	< 0.0001	-88
LeftKanjiHomophones	P	-0.230	0.082	-2.80	0.0051	-28

LeftKanjiRadicalTransparency	S	0.040	0.040	0.99	0.3210	8
LeftKanjiRadicalTransparency x Trial	S x T	0.001	0.000	2.15	0.0315	-12:23

CHAPTER 4

Reading English with Japanese in mind: Effects of frequency, phonology, and meaning in different-script bilinguals [†]

While a business person enjoys a cup of *coffee* to start a day in New York, a university professor in Amsterdam lectures on how *koffie* stimulates neural activity, and a student with drowsy eyes in Tokyo sips on another cup of コーヒー to concentrate on his homework. The words *coffee* /kʰɔfi/, *koffie* /kɔfi/, and コーヒー /koohii/ (with the double vowels representing a moraic long vowel) are examples of cognates. These are word pairs with a significant degree of semantic, orthographic and/or phonological form overlaps across languages, which often reflects a cross-linguistic historical link or lexical borrowing. Although it is easy to see that cross-language semantic similarity motivates orthographic and phonological resemblance (e.g., *coffee* in English, *koffie* in Dutch, and *caffè* in Italian), the orthographic similarity of the word pairs do not always correlate even when the phonology does (e.g., *coffee* in English alphabet, コーヒー /koohii/ in Japanese *katakana*, and 咖啡 /kafei/ in Chinese *hanzi*, with phonological resemblance maintained across these languages).

[†] A version of this chapter has been accepted for publication. Miwa, K., Dijkstra, T., Bolger, P., & Baayen, R. H. (2013). Reading English with Japanese in mind: Effects of frequency, phonology, and meaning in different-script bilinguals. *Bilingualism: Language and Cognition*. The authors are indebted to Marc Brysbaert, Wouter Duyck, Victor Ferreira, Sachiko Kinoshita, Judith Kroll, and Sarah White for their constructive feedback on an earlier version of this manuscript. The authors would also like to thank David Allen, Chad Marsolek, and Mariko Nakayama for discussion.

When comparing typologically different languages, lexical borrowings are the main source of cross-language phonological similarity. In Japanese, lexical borrowing is ubiquitous. Words borrowed from other languages are mostly written in the angular *katakana* script (e.g., /inntabjuu/ インタビュー 'interview'), instead of in the complex *kanji* (e.g., /kaikenn/ 会見) or in the curvy *hiragana* scripts (e.g., /inntabjuu/ いんたびゅー). Lexical borrowing is an on-going phenomenon in Japanese, and the presence of *katakana* scripts makes it possible for any foreign word to be absorbed into the Japanese lexicon irrespective of whether it is a frequent word across cultures (e.g., ウォーター for water) or a proper noun (e.g., サリー for Sally).[7]

The large number of borrowings in Japanese provides a unique opportunity for investigating how the specific characteristics of English and Japanese writing systems affect bilingual visual word recognition. Resolving this issue is important for the characterization of the human language-processing architecture, which at its most abstract level may arguably be language-independent. We addressed this issue by investigating to what extent *katakana* word knowledge is activated when Japanese-English bilinguals perform a visual lexical decision task on English words.

---

7 Nakayama, Kiryu, and Yamaguchi (2007) studied a corpus covering 10 years of a nation-wide newspaper from 1994 to 2003 and counted 8,226 katakana words that appeared at least 100 times. Typical Japanese dictionaries of katakana words list more than three times as many katakana words: the Personal Katakana Words Dictionary (Kindaichi, 1999) lists 28,000 words,

Visual word processing in bilinguals of languages with different scripts

Readers of Dutch learn English easily and immediately recognize cognates such as *coffee* (*koffie* in Dutch), *milk* (*melk*), and *water* (*water*), thanks to the L2 English words' orthographic and phonological similarities to their translation equivalent in L1 Dutch. This suggests that bilinguals in languages with the same script could profit during the recognition of an L2 word by making use of the already existing L1 word representation in lexical memory. This intuition is indeed in line with the available literature. Although some initial studies proposed that reading a word in one language might lead to a restricted activation of words only in that language (the so-called 'language-selective access view', Gerald & Scarborough, 1989; Rodriguez-Fornells, Rotte, Heinze, Nösselt, & Münte, 2002; Scarborough, Gerard, & Cortese, 1984), the majority of experimental studies indicates that a presented visual word input leads to activation of word candidates in both languages (the so-called 'language non-selective access view', Dijkstra & van Heuven, 1998, 2002; van Heuven, Dijkstra, & Grainger, 1998).

A direct consequence of such a non-selective access process is that cognates, due to representations linked in memory, are processed more quickly. Such cognate facilitation effects have been reported for a variety of experimental tasks. Dijkstra, Grainger, and van Heuven (1999) tested Dutch-English bilinguals in English lexical decision and progressive demasking tasks. They observed facilitatory effects of cross-linguistically shared semantic and lexical

---

and the Concise Dictionary of Katakana Words (Sanseido Henshujo, 1994) lists as many as 56,300 words.

orthographic representations and inhibitory effects of shared phonology. Cognate effects have further been reported in word association (e.g., van Hell & de Groot, 1998), in word naming (de Groot, Borgwaldt, Bos, & van den Eijnden, 2002), in picture naming (Hoshino & Kroll, 2008; Kohnert, 2004), in sentence reading with eye-tracking (Duyck, van Assche, Drieghe, & Hartsuiker, 2007; van Assche, Duyck, Hartsuiker, & Diependaele, 2009; van Assche, Drieghe, Duyck, Welvaert, & Hartsuiker, 2011), in vocabulary learning (Otwinowska-Kasztelanic, 2009), and in cross-language priming studies (e.g., de Groot & Nas, 1991, with Dutch-English bilinguals; Voga & Grainger, 2007, with Greek-French bilinguals). Collectively, these studies indicate that the lexical processing architecture of bilingual readers utilizes lexical distributional properties of two languages in lexical memory even when processing words in one language.

Perhaps surprisingly, there is a growing amount of data supporting exhaustive cross-language lexical activation even in bilinguals of languages with different scripts. Gollan, Forster, and Frost (1997), Nakayama, Hino, Sears, and Lupker (in press), and Kim and Davis (2003) tested Hebrew-English bilinguals, Japanese-English bilinguals, and Korean-English bilinguals respectively with a masked cross-script priming paradigm and reported that a word in one language still activated its phonologically and semantically related cognate in another orthographically distinct language (see also Morford et al., 2011 for ASL-English bilinguals; Voga & Grainger, 2007 for Greek-French bilinguals; Zhang, van Heuven, & Conklin, 2011 for Chinese-English bilinguals).

Note that the observation of language non-selective lexical activation in bilinguals of languages with different scripts does not imply that the underlying lexical processing architecture is the same as in bilinguals of languages with identical scripts. In fact, given differences in orthography, there must be some differences in the organization of the first stages of visual word processing for the two languages involved. At more abstract levels, however, the underlying processing mechanisms are likely to be similar. For instance, the role and interaction of lexical-phonological and semantic information sources might be analogous in Japanese-English and Dutch-English bilinguals. The present study aims to clarify what mechanisms remain the same and what must be different to account for lexical processing across languages with different scripts. In this study, we will consider this issue from the theoretical perspective of a localist connectionist framework.

Extending the BIA+ model to languages with different scripts

The Bilingual Interactive Activation (BIA) model (Dijkstra & van Heuven, 1998, 2002; Dijkstra, van Heuven, & Grainger, 1998) is a localist connectionist model that extends the monolingual IA model (McClelland & Rumelhart, 1981) and allows us to conceptualize monolingual and bilingual lexical processes within one theoretical framework. Because the model directly incorporates the IA model in its architecture, it can account for within-language lexical effects in monolinguals. In addition, because the model incorporates both an L2 and an L1 lexicon as an integrated whole, it can also account for cross-language lexical effects. While the original BIA model was limited to orthographic lexical aspects (Dijkstra & van Heuven, 1998; Dijkstra, van Heuven, & Grainger, 1998), it has been extended to account for experimental evidence on cross-language phonological and semantic activation (the BIA+ model, Dijkstra & van Heuven, 2002). In addition, the BIA+ model attempts to account for cross-task variations by incorporating a task/decision system explicitly in the model architecture. Currently, this extra-linguistic system is not expected to immediately affect lexical activation in the word identification system, based on the experimental evidence that bilinguals automatically activate two languages or word representations even when this language non-selective activation is not necessary in the task and, in fact, can even be detrimental (Dijkstra, de Bruijn, Schriefers, & ten Brinke, 2000; Dijkstra & van Hell, 2003; Dijkstra, van Jaarsveld, & ten Brinke, 1998; van Assche et al., 2011; van Hell & Dijkstra, 2002).

For languages sharing the same script (e.g., Dutch-English and French-English bilinguals with Latin alphabets), the BIA and BIA+ models predict that orthographic features of the input word immediately activate orthographic lexical representations in the two languages simultaneously. In contrast, for languages with different scripts, identification of script-specific orthographic features is not expected to facilitate activation of words in both languages. When the BIA+ model is generalized to languages with different scripts, the model may provide different predictions. In the example shown in Figure 4-1, the input word *interview* activates the corresponding letter nodes I, N, T, E, R, V, W, and these letter features then activate the word node *interview*. For Japanese-English bilinguals, however, it is expected that the feature set coding visuoperceptual features of Latin alphabets does not directly encode Japanese *katakana* script (the dotted line (a) in Figure 4-1).[2] Consequently, the orthography-driven non-selective lexical access across languages is not expected for bilinguals of languages with different scripts at the earliest processing stages. The following sections summarize our hypotheses for Japanese-English bilinguals with respect to the BIA+ architecture in Figure 4-1.

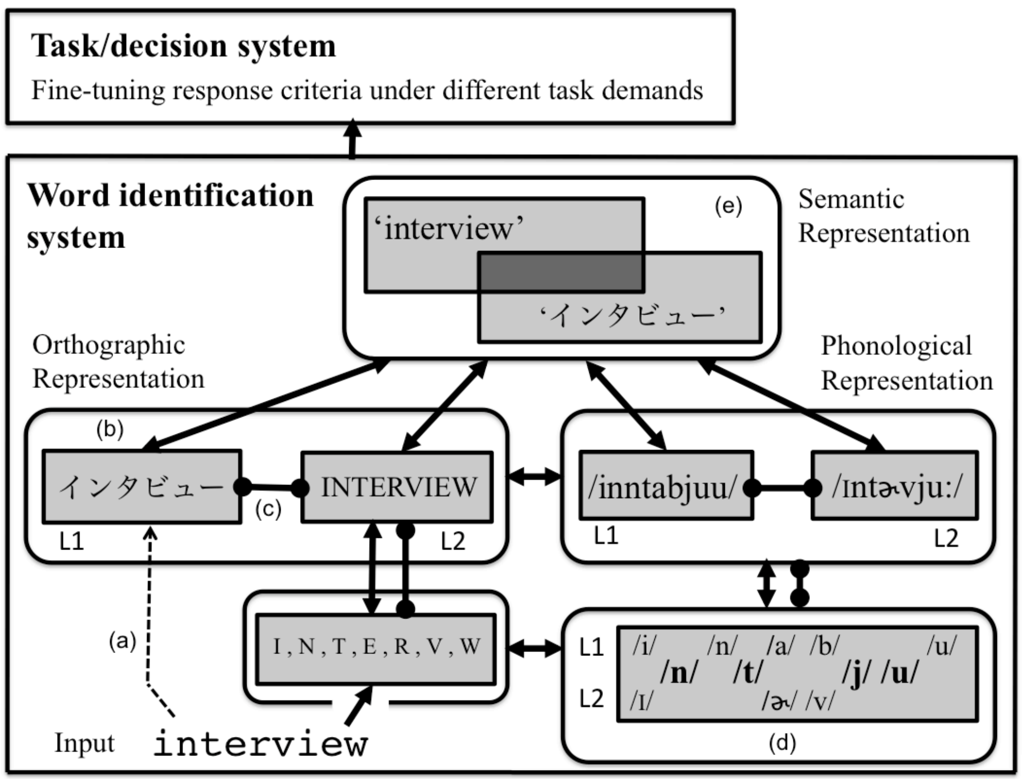


Figure 4-1. A bilingual interactive activation (BIA+) architecture applied for Japanese-English bilinguals' processing of an L2 English word. Arrows represent facilitatory links and circular connectors represent inhibitory links.

Cross-language phonological similarity

Phonological effects in visual word recognition have been studied predominantly in monolingual word recognition research (Carreiras, Ferrand, Grainger, & Perea, 2005; Ferrand & Grainger, 1992, 1994; Perfetti et al., 1992), but there is growing evidence for cross-language phonological activation in bilingual visual word recognition as well (Brysbaert, 2003; Brysbaert, van Dyck, & van de Poel, 1999; Dijkstra et al., 1999; Duyck, 2005; Duyck, Diependaele, Drieghe, & Brysbaert, 2004; Schwartz, Kroll, & Diaz, 2007).

For Japanese-English bilinguals, if activation of orthographic lexical representations of L1 words is mediated only by the conceptual route, then a cross-language phonological similarity effect should appear late in time. However, phonology-driven sublexical language non-selective access is theoretically still possible (the route to the box (d) in Figure 4-1). In the latter case, an effect of cross-language phonological similarity may appear early. For Japanese-English bilinguals, it is expected that the activation of English phonemes leads to activation of the corresponding Japanese phonemes and syllabic and/or moraic phonological nodes (e.g., the activation of phonemes /i/, /n/, /t/, /a/, /b/, /j/, and /u/ facilitates the activation of syllabic representations or moraic representations of /i/, /nn/, /ta/, /bju/, and /u/). The phonemic similarity between English and Japanese may lead to a larger global activation in the lexicon, just like a greater degree of orthographic similarity between L1 and L2 words matters for languages with the same script. We used rated phonological similarity as a diagnostic measure of phonological similarity.

Relative word frequencies in two languages

Like the monolingual IA model, the BIA+ model maintains inhibitory connections among words in the target language and also assumes inhibitory connections between orthographic lexical representations in two languages (Figure 4-1, line c). One consequence of such inhibitory connections is that the magnitude of the expected facilitatory effect of target English word frequency will be smaller when the Japanese translation equivalent has a high frequency of occurrence, because an English lexical orthographic representation and the Japanese lexical orthographic representation become co-active at some point in the course of word identification. It has been reported that word frequencies in two languages interact for interlingual homographs, words that share orthography across languages but not meaning (Dijkstra et al., 2000; Dijkstra, Moscoso del Prado Martin, Schulpen, Baayen, & Schreuder, 2005; Dijkstra, van Jaarsveld, & ten Brinke, 1998; Kerkhofs, Dijkstra, Chwilla, & de Bruijn, 2006).

The BIA+ model predicts that, for Japanese-English bilinguals, any orthographic cross-language inhibition can only occur later in the recognition process, because the Japanese orthographic representation becomes activated only by virtue of cross-language phonological or conceptual mediation. We use English word frequency and Japanese word frequency as diagnostic measures of strength of the activation of lexical orthographic representations to test this prediction.

Semantic similarity

Translation equivalents in two languages occasionally have different shades of meaning. For example, unlike the English word *interview* which is used unrestrictedly in various contexts, the use of the Japanese translation equivalent インタビュー is restricted to 'mass media interviews' and not typically used for 'job interviews.' A question relevant to bilingual word processing is whether such cross-language semantic similarity contributes to recognition of L2 words. In an unprimed English progressive demasking task with Dutch-English bilinguals, Dijkstra et al. (2010) reported a processing advantage for English words with higher semantic similarity. As shown in Figure 4-1 box (e), we predicted that semantic similarity contributes relatively late in the course of word recognition. We used rated semantic similarity as the diagnostic measure to test this prediction.

Cognate status

While the BIA and BIA+ models account for cognate facilitation effects by means of cross-language orthographic, phonological, and semantic overlaps alone, it has been suggested that a special representation for cognate may underlie the effects (Davis et al., 2010; Lalor & Kirsner, 2000; Sánchez-Casas & García-Albea, 2005). The present study considers cognate effects as an orchestration of gradient lexical effects but also considers a dichotomized factor coding cognate status to test the special representation view. If cognates are qualitatively special, then a factor encoding cognate status should emerge significant on top of relevant numerical predictors. A regression model provides an opportunity to straightforwardly test this theoretical prediction.

Extra-linguistic variables

Although the BIA and BIA+ models have been frequently discussed together in relation to language non-selective lexical access, the latter is distinguished from the former with respect to its explicit consideration of a non-linguistic system co-determining responses in a given task. In order to fully test the BIA+ model, consideration of extra-linguistic variables is therefore crucial. The BIA+ model currently assumes that the non-linguistic system does not modulate lexical processes at the earliest processing stages. This assumption will be falsified if response-based data and data sampled from the earliest time frame both reveal the same interactions between lexical and task-related variables. To the best of our knowledge, however, no study has identified such interactions.

Methodological concerns, statistical considerations, and goals of this study

Previous priming studies for bilinguals with different scripts have provided evidence supporting automatic language non-selective lexical activation within an integrated lexicon, as implemented in the BIA model. It should be noted, however, that a Japanese lexical orthographic representation is not yet activated at the earliest moment Japanese-English readers encounter an English word due to the orthographic dissimilarity (Figure 4-1 line a). In the context of cross-script priming, however, a lexical orthographic representation of one language (Figure

4-1 box b) is artificially pre-activated when a word of the other language is perceived, raising a concern as to what extent language non-selective activation holds in a task without priming (as exceptions, see Thierry & Wu, 2004, 2007 and Wu & Thierry, 2010 for implicit priming to study bilinguals of languages with different scripts). While masked priming is one of the most popular techniques to test subconscious and automatic lexical processes, researchers have not reached a consensus on an interpretation of a masked priming effect (e.g., see Forster, 1998 for a lexical pre-activation account, Kinoshita & Norris, 2010 for a non-lexical account, and Marsolek, 2008 for an antipriming account). At the moment, studies without priming for bilinguals with different scripts are scarce, and we investigate this issue without using priming.

In the present study, combining the lexical decision task with eye-tracking on simplex English words, we tested the above predictions of the extended BIA+ model on how lexical distributional properties of target English words and Japanese translation equivalents are utilized by Japanese-English bilinguals (Experiment 1) and by native monolingual readers (Experiment 2). Lexical decision was chosen, as this has been the most widely used experimental task (Libben & Jarema, 2002). Eye-tracking was used because previous studies employing lexical decision with eye-tracking (Kuperman, Schreuder, Bertram, & Baayen, 2009; Miwa, Libben, Dijkstra, & Baayen, 2013) reported that early and late lexical processes systematically co-determine the initial and late eye-fixation measures respectively.

Although the vast majority of psycholinguistic studies on bilingual processing have analyzed bilingual-specific effects solely in bilingual readers, we also compared the behavioral performance and eye-movements of late bilinguals and native monolinguals to (1) make sure that effects of interest are genuine bilingual effects, rather than artifacts arising from statistical correlation across languages and general processing mechanisms in reading English words, (2) to confirm that there is a good amount of functional overlap between bilinguals and monolinguals because part of the BIA/BIA+ model architecture, in theory, accounts for monolingual readers' lexical processing, and (3) to explore the 'expert' ability in reading, as native proficiency also provides a benchmark of an expert reader, and the acquisition of such 'expert' ability is often viewed as a goal of late bilinguals.

In order to test the above predictions of the BIA+ model for Japanese-English bilinguals, we opted for a mixed-effects analysis (Baayen, Davidson, & Bates, 2008). While the vast majority of previous studies on bilingual language processing dichotomized numerical predictors and opted for pre-experimental control (note de Groot et al., 2002; Dijkstra et al., 2010; Lemhöfer et al., 2008, as exceptions), regression analyses with post-experimental control provide more power and precision to observe lexical effects of interest (Baayen, 2010; Baayen & Milin, 2010; Cohen, 1983; MacCallum, Zhang, Preacher, & Rucker, 2002) and is resistant to potential effects of sampling bias (Cutler, 1981, Forster, 2000). In this study, we randomly sampled our materials from a large lexical database in order not to introduce any selection bias and opted for post-experimental statistical control. Mixed-effects modeling allows us to test lexical distributional properties, participants' characteristics, and task-related variables in a single statistical model. Only a few previous studies on bilingual processing have used regression to assess multiple lexical effects simultaneously (see Dijkstra et al, 2011, and Lemhöfer et al., 2008, for Dutch-English bilinguals), and this is the first study that makes use of this technique for bilinguals of languages with different scripts.

Experiment 1: English lexical decision with Japanese-English bilingual readers

Predictors considered in this study

Table 4-1 summarizes the lexical predictors, individual differences, and task effects considered in the present study. The lexical predictors are classified into those specific to Japanese-English bilinguals and those of English target words. All predictors were centered for the regression analyses.

Japanese-English-bilingual-specific lexical predictors. Rated Phonological similarity (*PhonologicalSimilarityJPN)* is a rated cross-language phonological similarity measure obtained by 10 Japanese speakers. In order to safeguard *PhonologicalSimilarityJPN* from potential confound arising from other lexical knowledge, we also considered objective Levenshtein distance coding phonological similarity (*PhonologicalDistance*). *PhonologicalDistance* gauges the number of operations required to transform Japanese words into the corresponding English translation equivalents in their phonologically transcribed form (Levenshtein, 1966; Yarkoni, Balota, & Yap, 2008; Dijkstra et al., 2010; Gooskens & Heeringa, 2004; Schepens, Dijkstra, & Grootjen, 2011) based on the `sdists` function available in the `R` package `cba` (Buchta & Hahsler, 2009). In order to compare words of different lengths, we normalized the phonological Levenshtein distance based on the length of target English words (M = 4.3, SD = 1.4).

SemanticSimilarity was based on four Japanese-English bilingual readers' ratings on cross-language conceptual similarity. English and Japanese words were presented to the raters side-by-side in two columns in a spreadsheet. Using a seven-point scale (1 = very different, 7 = identical), the raters assessed the extent to which *katakana* loanwords in Japanese were different in meaning to the corresponding English target words and whether any Japanese katakana words were completely unfamiliar to them.

Table 4-1. Lexical predictors, individual differences, and task effects considered in this study. The range and mean are presented for their original values before residualization and centralization procedures. The superscripts represent a transformation method used for the given predictor. The values for *Individual* and *Task* variables are those in Experiment 1 (Japanese-English bilinguals).

Type	Predictor	Range (Mean, SD) / Levels
Japanese-English	PhonologicalSimilarityJPN	2.3: 6.0 (*M* = 4.5, *SD* = 0.7)
	PhonologicalDistance	0.2 : 1.2 (*M* = 0.6, *SD* = 0.2)
	FreqJPN_resid (log)	0 : 10.4 (*M* = 4.2, *SD* = 2.8)
	GoogleFreqJPN_resid (log)	5.7 : 19.2 (*M* = 14.8, *SD* = 2.4)
	SemanticSimilarity_resid	3.0 : 7.0 (*M*= 5.9, *SD* = 0.7)
	Cognates	Levels: Cognate, NotCognate
English	Length	6 : 9 (*M* = 6.8, *SD* = 0.9)
	OLD20_resid	1.3 : 4.5 (*M* = 2.4, *SD* = 0.5)
	FreqHAL (log)	7.8 : 13.0 (*M* = 9.6, *SD* = 1.1)
	SUBTLCD_resid (log)	0 : 4.4 (*M* = 1.9, *SD* = 0.9)
	Imageability_resid	1.6 : 7.0 (*M* = 4.7 , *SD* = 1.7)
Individual	LengthOfStayCanada (log)	1.1 : 5.3 (*M* = 2.7, *SD* = 1.3)
Task	PreviousResponseCorrect	Levels: Correct, Error
	Trial (/100)	0.11 : 4.60 (*M* = 2.4, *SD* = 1.3)
	PreviousRT_resid (-1000/)	-1.1 : 1.1 (*M* = 0, *SD* = 0.4)
	FirstSubgazeDuration_resid (log)	4.9 : 7.8 (*M* = 6.2, *SD* = 0.4)

FreqJPN reflected how many times Japanese *katakana* words appeared in a Japanese newspaper corpus containing over three million words and covering the 14-year period from 1985 to 1998 (Amano & Kondo, 2003). *FreqJPN* was log-transformed, as its distribution had a long right tail. Note that it is often also possible to translate English words to logographic *kanji* or moraic *kana* words, as well as *katakana* loanwords. However, because the log-transformed frequency of *kanji* or *kana* translations obtained from the same corpus was not a significant predictor, we do not further discuss it.

Although *FreqJPN* comprises two distributions due to zero frequency of occurrence for some words, the corresponding log-transformed Google document frequency measure (*GoogleFreqJPN*) does not indicate such qualitative difference among the set of *katakana* words (see Appendix 4-1). This suggests that the zero frequencies in *FreqJPN* are not due to qualitative differences with respect to words' lexical status, such as transliterations and translations, but due to the fact that the written word corpus is conservative for the purpose of the present study (i.e., it provides frequency counts for *katakana* words only up to the year 1998 and only in the context of newspaper texts).

Finally, a factor *Cognate* (levels: *Cognate* and *NotCognate*) was considered in addition to the above mentioned numerical predictors. Because considerable semantic and phonological overlap and decent word frequencies are expected for such special cognate representations to emerge in the first place, words with larger-than-the-average values in all *FreqJPN*, *PhonologicalSimilarity*, and the conceptual similarity were categorized into *Cognate* ($N = 58$).

Lexical distributional predictors of the target English words. As orthographic predictors, we considered word length (*Length*) and orthographic Levenshtein distance (*OLD20*, Yarkoni et al., 2008). A low *OLD20* score indicates that a given word is located within a dense orthographic space. To measure English word frequency, we used log-transformed *FreqHAL* (HAL:

Hyperspace Analogue to Language, Lund & Burgess, 1996; Burgess, & Livesay, 1998, as available in Balota et al., 2007). *SUBTLCD* is a log-transformed context diversity measure based on a number of films in which a given word had been used (Adelman, Brown, & Quesada, 2006; Brysbaert & New, 2009a, 2009b). We also considered ratings of word *Imageability*. Because *SemanticSimilarity* is expected to vary with imageability, with a larger cross-linguistic variance for abstract concepts relative to concrete concepts, *Imageability* safeguards our rated *SemanticSimilarity* measure. We obtained *Imageability* scores rated by ten native English readers, using a seven-point scale (1 = not imageable, 7 = very imageable).

Task-related predictors and individual differences. In the response time analyses, we considered the following variables: *PreviousRT*, inversely transformed RT in the previous trial; *Trial*, the number of preceding trials; and *PreviousResponseCorrect*, whether the responses in the preceding two trials were correct (see Baayen & Millin, 2010 for autocorrelation in the time-series of response times). In the eye-movement analyses, we also considered *PreviousFixationDuration* for second fixation duration analyses to account for potential spillover effects from the previous fixation. *PreviousRT Trial,* and *PreviousResponseCorrect* were also considered in the eye-movement analyses.

Consideration of readers' L2 proficiency is also important because such individual differences potentially lead to distinct processing mechanisms (Kroll & Stewart, 1994; Potter, So, von Eckhardt, & Feldman, 1984). In our sample, we considered log-transformed participants' months of stay away from Japan for each participant (*LengthOfStayCanada*) in our regression analyses as a measure of L2 English proficiency. *LengthOfStayCanada* naturally brings in several other components of language proficiency; it highly correlated with age (r = 0.68, p = 0.001), with their vocabulary size in English measured by X_Lex The Swansea Levels Test (Meara, 2005, r = 0.48, p = 0.03). We leave the specific advantages and disadvantages of other related measures to future research.

Multicollinearity and residualization. In studying independent contributions of lexical distributional properties, it is crucial to resolve multicollinearities among them. While there are a number of methods proposed to handle this problem (Belsley, Kuh, & Welch, 1980; Strobl, Malley, & Tutz, 2009), we opted for a residulization procedure, as in Kuperman et al. (2009). For example, *OLD20* highly correlates with *Length*, *SUBTLCD*, and the number of meanings in WordNet (Miller, 1990). *OLD20* was therefore regressed on these three variables, and the residuals were used as a new predictor *OLD20_resid*. The new predictor correlated significantly and strongly with the original predictor (r = 0.67, p < 0.01 with *OLD20*). To safeguard our measures, the same residualization procedure was applied to the following variables with highly significant inter-correlations: *SUBTLCD* (regressed on *FreqHAL* and a number of meanings); *Imageability* (regressed on *FreqHAL*, *Length*, number of meanings); *FreqJPN* (regressed on *FreqHAL*, *SUBTLCD*, and number of meanings); *SemanticSimilarity* (regressed on *FreqJPN* and *GoogleFreqJPN*). After the residualization procedure, all the new predictors correlated significantly with the respective original predictors: r = 0.67 for *OLD20_resid* and *OLD20,* $r = 0.75$ for *SUBTLCD_resid* and *SUBTLCD*, and $r = 0.95$ for *FreqJPN_resid* and *FreqJPN*. *GoogleFreqJPN* and *PhonologicalDistance* were not included in statistical models together with *FreqJPN* and *PhonologicalSimilarityJPN* but considered separately to assess whether the pattern of results remains unchanged when one predictor is replaced with another. Task-related variables *PreviousRT* and *FirstSubgazeDuration* were similarly regressed on correlated predictors (*Trial* for the former and *Trial*, *PreviousRT*, *FreqHAL*, *Length*, and *SUBTLCD* for the latter), resulting in *PreviousRT_resid* and *FirstSubgazeDuration_resid*. [8]

Method

Participants. Nineteen Japanese-English late-bilingual readers (three males, mean age = 25.1, *SD* = 5.8) were recruited at the University of Alberta. All participants had normal or corrected-to-normal vision. The participants had stayed in Canada for 33 months on average (*SD* = 45.6) and acquired English as their second language.

Materials. We sampled 250 words from English Lexicon Project database (Balota et al., 2007) based on the following criteria: (1) the word length was between 6 and 9 letters; (2) the word frequency was greater than 2,000 in the *FreqHAL* frequency distribution; (3) the morphological status was simplex; (4)

---

[8] As in any statistical techniques, however, residualization has limitation in teasing apart individual effects of correlated predictors. For example, it is up to the researcher to determine what to be regressed on what, and a residualized predictor is, strictly speaking, no longer an original predictor. We believe, however, that this is only a minor concern in this study because no two variables were highly correlated, as seen in the strong correlations between the residualized and original predictors.

the part of speech was noun; and (5) the mean accuracy rates in lexical decision and naming were at least 0.9 (see Appendix 4-2 for the list of words used in the present study). We also sampled 200 nonwords from the ARC nonword database (Rastle, Harrington, & Coltheart, 2002) to make a total of 450 letter strings. These nonwords were similar to the words: 6 to 9 letters long, with existing onsets, existing bodies, and legal bigrams.

Apparatus. The experiment was designed and controlled by SR Research Experiment Builder software. Words were presented on a 20-inch display. Eye-movements were tracked by an EyeLink II head-mounted eye-tracker (SR Research, Canada) in the pupil-only mode with a sampling rate of 250 Hz. Calibrations were conducted with horizontal three points.

Procedure. In this lexical decision with eye-tracking experiment, participants were asked to decide as quickly and accurately as possible whether the letter strings presented on the computer display were legitimate existing English words or non-existent words (nonwords) by pressing the right and left buttons, respectively, of a Microsoft SideWinder game-pad. The words were presented following a fixation circle, on which participants were asked to fixate their eyes. This fixation circle served as a drift correct point allowing the researcher to correct for head drifts between trials. Given our understanding of the optimal viewing position in sentence reading and in isolated word reading (Brysbaert & Nazir, 2005; Farid & Grainger, 1996; O'Regan & Jacobs, 1992; Vitu, O'Regan, & Mittau, 1990), the location of the fixation mark was slightly shifted horizontally so that the first fixations were positioned slightly (1.5 letters) to the left of the word centre. The target words were presented in Courier New 44-point white font on a black background. At a viewing distance of 70 cm from the screen, the visual angle was estimated to be 1.1° for each letter (i.e., for six-letter words and nine-letter words, such as *camera* and *interview*, the visual angles were approximately 6.7º and 10.0º respectively). A testing session contained two breaks, one after every 150 words. Participants saw a summary of their performance regarding accuracy and response speed after the practice trials and at each break point. The experiment took roughly 90 to 120 minutes. The right eye was tracked throughout the experiment.

Results

Response latency. R version 2.13.2 (R Development Core Team, 2011) was used for statistical analyses. We opted for a mixed-effects regression analysis with subjects and items as crossed random effects (Baayen et al., 2008; Bates, Maechler, & Dai, 2007). In the sections that follow, we report analyses of response latencies, first subgaze durations, and last fixation durations. Response accuracy across subjects ranged from 0.86 to 0.98 (M = 0.92, SD = 0.04). Therefore, no participant was excluded from the analyses. Of all 4,750 trials, those with response latency either shorter than 300 ms or longer than 3,000 ms were excluded (14 trials). Twenty-one words with more than 30% erroneous responses and one word (*heroin*) with a coding error were excluded from the analysis (414 trials, 8.7 % of the remaining data). The following analyses are based on the remaining 228 words, amounting to 4,099 trials after excluding the 5.1% of trials that contained incorrect responses.

A reciprocal transformation (-1000/RT) was applied to the RTs to attenuate skewness in their distribution, based on the appropriate exponent suggested by the Box-Cox power transformation technique (Box & Cox, 1964; Venables & Ripley, 2002). The Box-Cox transformation technique was applied to all dependent measures in the rest of this paper).

We first fitted a model with the factor *Cognate* only, with subject and item as crossed random effects. In this model, response times for words with a *Cognate* translation equivalent were shorter (effect size = -23 ms, $p = 0.03$) in line with cognate facilitation effects reported in previous studies. However, *Cognate* was no longer significant once task-related predictors and target word properties were added to the model ($p = 0.55$), indicating that the variance explained by *Cognate* was absorbed by the numerical predictors. We then fitted a simple main effects model with all predictors and also considered all pairwise interactions. We finally tested interactions between task-related predictors and lexical predictors. Potentially influential outliers were removed with standardized residuals exceeding 2.5 standard deviation units (1.8% of the data points). Table 4-2 summarizes the fixed-effects in our final model, and Figure 4-2 visualizes significant interactions. The reference level for the factor *PreviousResponseCorrect* was *Correct* throughout this study. The random-effect structure of this model consists of random intercepts for item ($SD = 0.09$) and subject ($SD = 0.16$), by-subject random slopes for *Trial* ($SD = 0.02$), and by-subject random contrasts for *PreviousResponseCorrect* (*SD* = 0.06). The standard deviation of the residual error was 0.24. Other random slopes for subjects did not reach significance.

Table 4-2. Estimate, standard error, t-value, p-value, and effect size of the fixed effects in the model of Japanese-English bilingual readers' lexical decision response times.

Response time	Type	Estimate	SE	t-value	p-value	Effect size
(Intercept)		-1.477	0.037	-40.43	< 0.0001	
PreviousResponseCorrect (Error)	Task	0.064	0.017	3.69	0.0002	28
Trial	Task	-0.085	0.006	-13.52	< 0.0001	-168
PreviousRT_resid	Task	0.135	0.012	10.81	< 0.0001	124
Length	Engl	0.046	0.008	5.79	< 0.0001	65
OLD20_resid	Engl	0.033	0.021	1.55	0.1209	26
FreqHAL	Engl	-0.064	0.007	-9.61	< 0.0001	-129
SUBTLCD_resid	Engl	-0.061	0.010	-5.89	< 0.0001	-111
FreqJPN_resid	Jpn-Engl	-0.003	0.003	-1.04	0.2966	-14
PhonologicalSimilarityJPN	Jpn-Engl	-0.027	0.011	-2.38	0.0172	-44
SemanticSimilarity_resid	Jpn-Engl	-0.037	0.011	-3.27	0.0011	-70
PreviousRT_resid * Trial	Jpn-Engl * Task	-0.028	0.009	-2.98	0.0029	No plotted
PhonologicalSimilarityJPN * Trial	Jpn-Engl * Task	-0.011	0.005	-2.28	0.0227	Fig 4-2 (a)
FreqHAL * FreqJPN_resid	Engl * Jpn-Engl	0.006	0.003	2.53	0.0116	Fig 4-2 (b)
SemanticSimilarity_resid * PreviousRT_resid	Jpn-Engl * Task	-0.046	0.016	-2.92	0.0035	Fig 4-2 (c)
FreqHAL * OLD20_resid	Engl * Engl	-0.050	0.018	-2.74	0.0061	Fig 4-2 (d)

Note. Task = task-related predictors, *Engl* = English target word properties, *Jpn-Engl* = Japanese-English-bilingual-specific predictors. The effect sizes refer to the magnitude of an effect calculated as the difference between the model's prediction for the minimum and the maximum back-transformed values of a given predictor

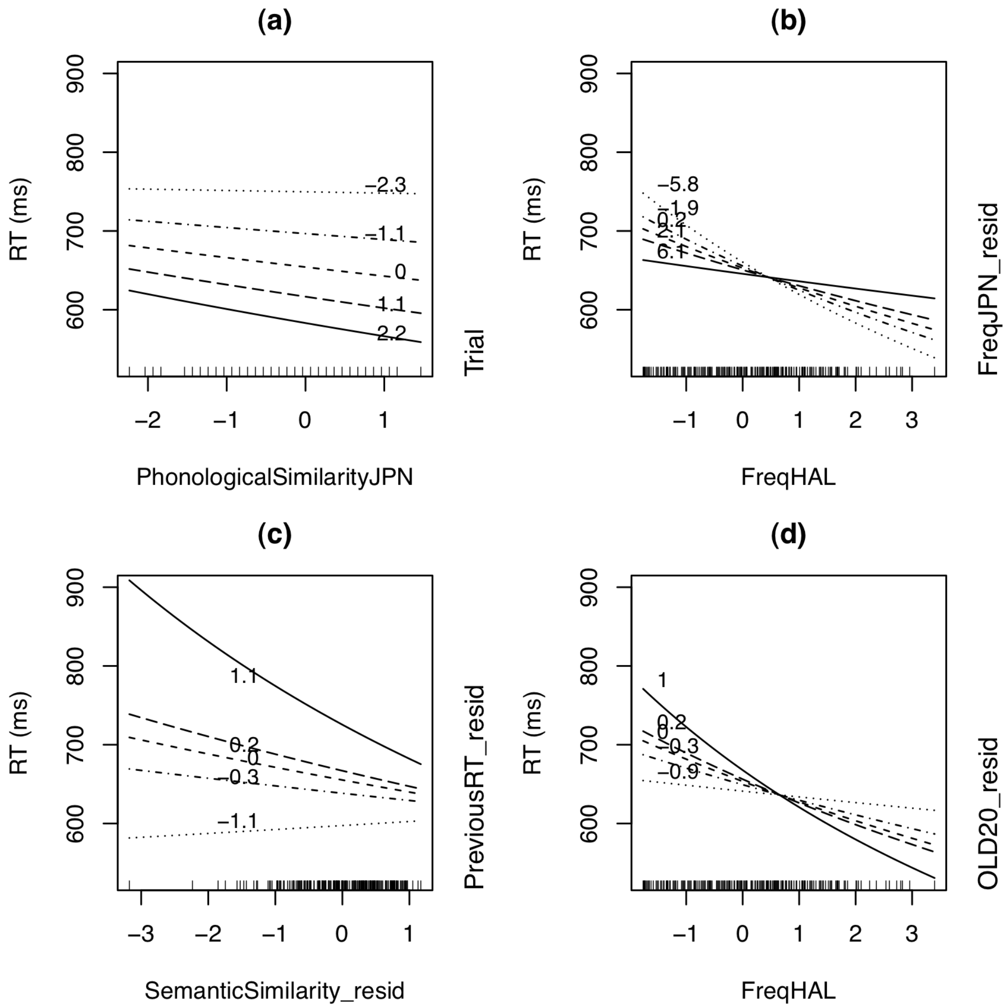


Figure 4-2. Lexical interactions in the mixed-effects model for Japanese-English bilinguals' lexical decision response times. Different lines represent quantiles, and the rug in the x-axes represents the pattern of distribution.

Interestingly, all bilingual-specific predictors co-determined the lexical decision response latencies (Figure 4-2, Panels a, b, and c). Although the effect of *PhonologicalSimilarityJPN* reached significance in a simple main-effects model, a model allowing interactions into the model specification further clarified that *PhonologicalSimilarityJPN* facilitated responses later in the experiment (Panel a). Replacing *PhonologicalSimilarityJPN* with *PhonologicalDistance* successfully replicated the pattern of this interaction.

The effect of *FreqJPN_resid* emerged in an interaction with target word frequency *FreqHAL* (Panel b). For words with large *FreqJPN_resid* (the solid line in Panel b), the magnitude of the *FreqHAL* effect was attenuated. When *FreqJPN_resid* was replaced with *GoogleFreqJPN_resid*, a virtually identical interaction was obtained.

SemanticSimilarity_resid had a small yet significant facilitatory effect in a simple main-effects model. Upon close inspection, its interaction with the task-related variable, *PreviousRT* (Panel c) indicated that when the response latency in the previous trial was long (the solid line in Panel c), cross-language *SemanticSimilarity_resid* facilitated the response more strongly. This is in line with the finding that the cross-language semantic similarity effect is more likely to be observed when the task induces longer response latencies (Dijkstra et al., 2010, using a progressive demasking word identification task).

Several other lexical distributional properties of the English target words co-determined response latencies. The magnitude of the *FreqHAL* facilitatory effect was the greatest for words with large *OLD20_resid*, which are words situated in a sparse orthographic space (the solid line in Figure 4-1 Panel d). As summarized in Table 4-2, the context diversity of words gauged by *SUBTLCD_resid* contributed beyond *FreqHAL*, and *Length* inhibited responses.

The effect of *SemanticSimilarity_resid* was not modulated by *LengthOfStayCanada*, suggesting that, for Japanese-English bilinguals, the recruitment of cross-language semantic activation did not vary across readers with different L2 proficiency.

It is likely that the above effects reflected in response latencies also guide eye-movements. We investigated the time-course of the above effects by studying whether they load onto the early or late fixations, or a combination of both. Japanese-English bilinguals read words with a single fixation only 1% of the time (36% showed two fixations, 39% three fixations, and 17% four fixations, with a mode at three fixations). In order to include as many data points as possible, for all trials with at least two fixations, we analyzed (1) first fixation durations, (2) first subgazes (in this study, sum of all non-final fixations, which were ended by a saccade to the next location) and (3) last fixations as measures of very early processing, relatively early processing, and late processing respectively. The first fixation and subgaze durations were measured from the onset of target word presentation. It is assumed that the first subgaze duration is less contaminated by conscious lexical decision response strategies than the last fixation, which was ended with a button press.

First fixation duration. For the analysis of the first fixation durations, data points with a first fixation shorter than 100 ms and longer than 850 ms, those before a blink, and those with an incorrect response were excluded from the analysis. A log-transformation was applied to attenuate skewness in the distribution of the fixation durations (M = 5.7, SD = 0.3, raw median = 280 ms).

As in the response time analysis, we tested simple main effects and pairwise interactions, as well as interactions between lexical and task-related predictors. Potentially influential outliers were removed with standardized residuals exceeding 2.5 standard deviation units (2.0% of the data points). The final model is summarized in Table 4-3. The random-effect structure of this model comprised random intercepts for item (*SD* = 0.05) and subject (*SD* = 0.11), and by-subject random slopes for *Trial* (*SD* = 0.01). The standard deviation of the residual error was 0.21.

As expected, the first fixation duration was co-determined by the signature of early buttom-up orthographic processing. The inhibitory effect of *OLD20_resid* indicates that words in sparse orthographic neighbourhood receive a longer first fixation. A word frequency effect (*FreqHAL*) was observed already as well, replicating the early word frequency effects in previous studies (Kuperman et al., 2009; Miwa et al., 2013).

Interestingly, the bilingual-specific predictor *PhonologicalSimilarityJPN* already contributed at this earliest time frame, indicating that a sublexical cross-language phonological decoding route is used (the route to the box d in Figure 4-1). Unlike its effect in the RT analysis, however, its effect was inhibitory. It should also be noted that the objective *PhonologicalDistance* measure, when replaced with *PhonologicalSimilarityJPN*, did not reach significance.

Table 4-3. Estimate, standard error, t-value, p-value, and effect size (ms) of the fixed effects in the model of Japanese-English bilingual readers' first fixation durations, first subgaze durations, and last fixation durations.

First fixation duration	Type	Estimate	SE	t-value	p-value	Effect size
(Intercept)		5.673	0.025	228.98	< 0.0001	
Trial	Task	0.011	0.004	2.93	0.0034	14
PreviousRT_resid	Task	-0.025	0.011	-2.23	0.0256	-16
OLD20_resid	Engl	0.040	0.013	2.99	0.0028	23
FreqHAL	Engl	-0.009	0.004	-1.99	0.0467	-13
PhonologicalSimilarityJPN	Jpn-Engl	0.017	0.007	2.29	0.0219	18
First subgaze duration	**Type**	**Estimate**	**SE**	**t-value**	**p-value**	**Effect size**
(Intercept)		6.215	0.037	166.13	< 0.0001	
Trial	Task	-0.091	0.009	-9.77	< 0.0001	-190
PreviousRT_resid	Task	0.127	0.016	7.76	< 0.0001	134
Length	Engl	0.083	0.008	10.25	< 0.0001	131
OLD_resid	Engl	0.044	0.021	2.05	0.0404	39
FreqHAL	Engl	-0.066	0.007	-9.64	< 0.0001	-147
SUBTLCD_resid	Engl	-0.070	0.010	-6.78	< 0.0001	-138
PhonologicalSimilarityJPN	Jpn-Engl	-0.027	0.012	-2.32	0.0206	-47
FreqJPN_resid	Jpn-Engl	-0.004	0.003	-1.27	0.2029	-20
FreqHAL * FreqJPN_resid	Engl * Jpn-Engl	0.005	0.003	1.99	0.0468	Fig 4-3 (a)
Last fixation duration	**Type**	**Estimate**	**SE**	**t-value**	**p-value**	**Effect size**
(Intercept)		13.282	0.261	50.89	< 0.0001	
FirstSubgazeDuration _resid	Task	-8.006	0.145	-55.23	< 0.0001	-507
PreviousResponseCorrect (Error)	Task	0.667	0.130	5.14	< 0.0001	18
Trial	Task	0.257	0.038	6.83	< 0.0001	31
PreviousRT_resid	Task	-1.643	0.158	-10.41	< 0.0001	-104
FreqHAL	Engl	0.142	0.071	1.98	0.0474	20
SUBTLCD_resid	Engl	0.240	0.112	2.14	0.0323	27
FreqJPN_resid	Jpn-Engl	0.016	0.031	0.52	0.6042	0
PhonologicalSimilarityJPN	Jpn-Engl	0.006	0.124	0.05	0.9609	5
SemanticSimilarity_resid	Jpn-Engl	-0.497	0.121	-4.1	< 0.0001	-49
PhonologicalSimilarityJPN * FreqJPN_resid	Jpn-Engl * Jpn-Engl	-0.104	0.044	-2.38	0.0175	Fig 4-3 (b)
SemanticSimilarity_resid * Phon.SimilarityJPN	Jpn-Engl * Jpn-Engl	0.444	0.199	2.23	0.0259	Fig 4-3 (c)
SemanticSimilarity_resid * PreviousRT_resid	Jpn-Engl * Task	-0.616	0.199	-3.09	0.0020	Fig 4-3 (d)

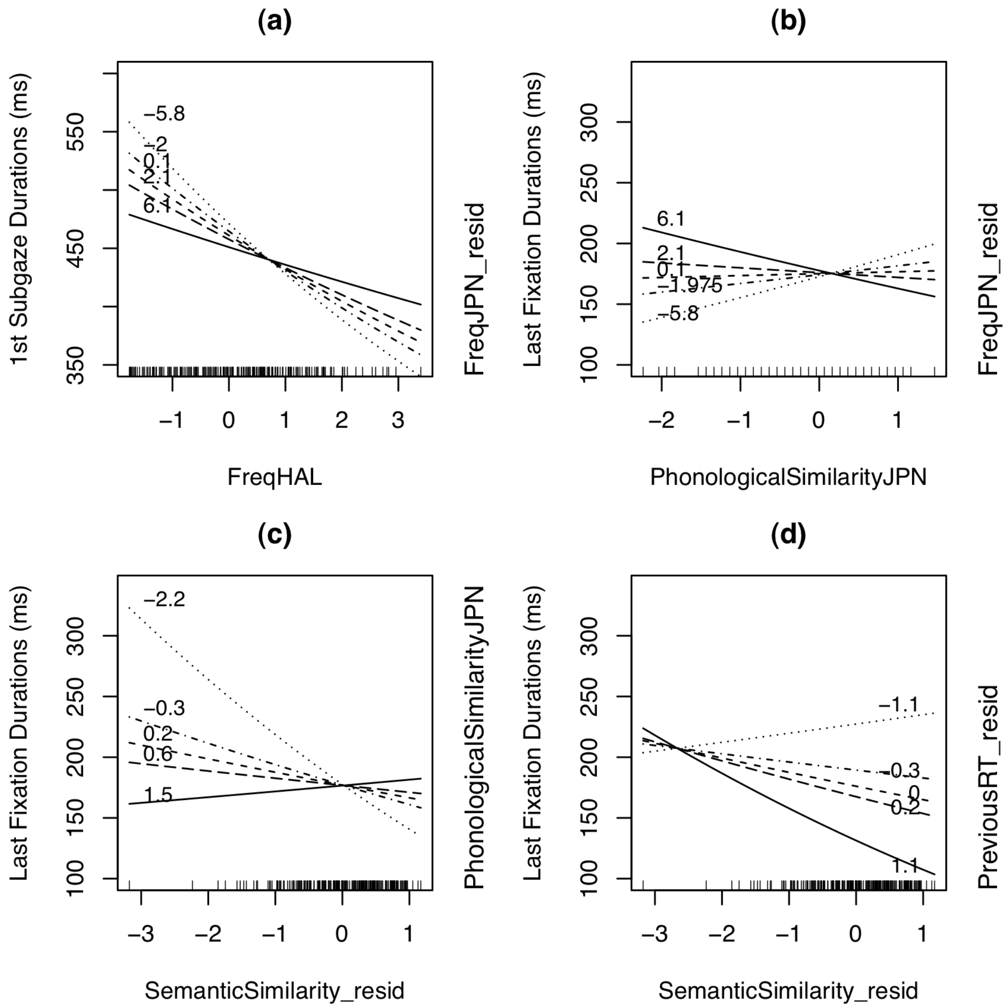


Figure 4-3. Lexical interactions in the mixed-effects model for Japanese-English bilinguals' first subgaze and last fixation durations in trials with two fixations. Different lines represent quantiles, and the rug in the x-axes represents the pattern of distribution.

First subgaze duration. For the analysis of the first subgaze durations, data points with a first fixation shorter than 100 ms, those before a blink, and those with an incorrect response were excluded from the analysis. A log-transformation was applied to attenuate skewness in the distribution of the subgaze durations (*M* = 6.2, *SD* = 0.4, raw median = 500 ms).

After testing simple main effects, pairwise interactions, and interactions between lexical and task-related predictors, potentially influential outliers were removed with standardized residuals exceeding 2.5 standard deviation units (1.2% of the data points). The final model is summarized in Table 4-3. The random-effect structure of this model comprised random intercepts for item (*SD* = 0.08) and subject (*SD* = 0.16), and by-subject random slopes for *Trial* (*SD* = 0.04). The standard deviation of the residual error was 0.32.

At the first subgaze, some bilingual-specific effects were observed. Whereas at the first fixation, *PhonologicalSimilarityJPN* was inhibitory, at the second fixation, *PhonologicalSimilarityJPN* became facilitatory. It should be noted, however, that *PhonologicalDistance* did not reach significance when replaced with *PhonologicalSimilarityJPN*. Interestingly, the interaction between *FreqHAL* and *FreqJPN_resid* observed in the response time analysis was also observed at the first subgaze in a virtually identical form (Figure 4-3, Panel a). This interaction was successfully replicated when *FreqJPN_resid* was replaced with *GoogleFreqJPN_resid.*

Cognate and *SemanticSimilarity_resid* were not significant predictors. *LengthOfStayCanada* was not a significant predictor and did not modulate any lexical effects either.

Last fixation duration. For the analysis of the last fixation duration, the words excluded in the analysis of response latencies were excluded here as well.

Trials with second fixation durations longer than 900 ms, those before or after a blink, and those with an incorrect response were also excluded from the analysis. A square-root transformation was applied to attenuate the skewness in the distribution of the fixation durations (*M* = 13.5, *SD* = 4.3, raw median = 188 ms). Unlike the first fixation and subgaze durations, which were terminated by the eye moving to another location in the word, the last fixation duration was terminated by the readers' button-press. Consequently, we expected that the second fixation durations would reflect, in addition to lexical predictors, variables associated with response planning and execution.

We tested simple main effects, pairwise interactions, and tested interactions among task-related predictors and lexical predictors. As in the response time analysis, with all relevant numerical predictors considered in a model, *Cognate* did not contribute significantly to the model fit. The final model for Japanese-English bilinguals' second fixation durations is summarized in Table 4-4. Figure 4-4 presents the significant lexical interactions in the model. Potentially influential outliers were removed with standardized residuals exceeding 2.5 standard deviation units (1.6% of the data points). The random-effect structure of the final model comprised random intercepts for item (*SD* = 0.93) and subject (*SD* = 1.08). The standard deviation of the residual error was 2.97.

Several bilingual-specific effects co-determined the last fixation durations. *FreqJPN_resid* interacted with *PhonologicalSimilarity*. Larger *PhonologialSimilarityJPN* shortened the fixation duration for words with high

FreqJPN_resid (the solid line in Figure 4-3, Panel b), suggesting that words with higher cross-language *PhonologicalSimilarity* and *FreqJPN* were perceived as more word-like. In a model in which *PhonologicalDistance* was used instead of *PhonologicalSimilarityJPN*, a virtually identical interaction was obtained.

SemanticSimilarity_resid facilitated processing. The magnitude of facilitation was greater for words with low cross-language phonological similarity (Panel c). This interaction may indicate a L1-based response strategy to rely on either phonological similarity or semantic similarity to make a lexical decision response. *SemanticSimilarity_resid* also interacted with *PreviousRT_resid* (Figure 4-3, Panel d). Recall that the facilitatory effect of *SemanticSimilarity* for words preceded by long *PrevioiusRT_resid* was also observed in the RT analysis. Since *SemanticSimilarity_resid* was absent both at the first fixation and subgaze, we can conclude that this semantic effect emerges late.

The last fixation seems to be qualitatively different from the first fixation and subgaze, as indicated by the atypical inhibitory effects of *FreqHAL* (20 ms) and *SUBTLCD_resid* (27 ms). The inhibitions from L2 word properties may be due to a response strategy to rely on L1 word properties. Involvement of a conscious response strategy is evident from the significant effect of *PreviousResponseCorrect*. When participants make an error, they usually become aware of it immediately after a button press and try to be cautious in the following trials. *LengthOfStayCanada* did not reach significance and did not modulate lexical effects either.

Discussion

Lexical decision measures emerged, from our analyses, as a composite measure amalgamating processing costs that arise at different stages of information uptake. First fixations reflected early bottom-up processing as witnessed by orthographic neighbourhood density and target word frequency. First subgazes reflected lexical effects in the word identification system not affected by conscious response strategies, followed by last fixations, which were more dedicated to response planning and execution in the task/decision system.

Importantly, using a regression technique, we observed all the expected bilingual-specific effects in the reaction times (i.e., phonological similarity, interaction between L1 and L2 frequencies, and semantic similarity) simultaneously in a task without a prior presentation of a L2 word. Interestingly, these effects also co-determined eye-movements but at different points in time. The early contribution of cross-language phonological similarity is in line with previous masked priming studies. It appears that, even for languages with different scripts and in a task without priming, sublexical phonological decoding immediately takes place. It should be noted, however, that its effect was inhibitory in the earliest time frame and facilitatory in later time frames. Although this result indicates that there is mechanism to generate phonological inhibition (see Grainger et al., 1999 for an inhibitory phonological similarity effect), it also raises a concern that the facilitatory phonological similarity effect reported in a masked priming paradigm may not be driven purely by early pre-activation mechanisms in lexical access. We leave this issue to future research (see Kinoshita & Norris for an evidence accumulation account).

It was also notable that the interaction between L1 and L2 frequencies was found relatively early at the first subgaze. Under the assumption that response planning and execution takes approximately 200 ms (Schmidt, 1982), it is likely that the cross-language competition was not due to readers' conscious response strategy but rather part of central lexical processing mechanism, as assumed in the BIA+ model.

The time-course of lexical activation characterized by the relatively early contribution of cross-language phonological similarity, followed by the competition between L1 and L2 words, and then by a cross-language semantic similarity effect is in line with the predictions of the BIA+ model.

In Experiment 2, we tested monolingual readers of English to ensure that the bilingual-specific effects observed in Experiment 1 arose from a bilingual-specific processing mechanism, as well as to explore whether the within-language English lexical distributional properties are utilized similarly by the two groups of readers.

Experiment 2: English lexical decision with English monolingual readers

Method

Participants. Nineteen monolingual English readers (7 males, mean age = 21.6, *SD* = 7.1) were recruited at the University of Alberta. There was no significant difference between the late bilinguals in Experiment 1 and the monolingual readers with respect to age. All participants had normal or corrected-to-normal vision. Monolingual readers were defined here as native readers of English with more than 80% daily exposure to English relative to the amount of exposure to their second languages at the time of the experiment, as reported by the participants. None of the participants had Japanese as their second or third language.

Materials, Apparatus, and Procedure. The same as in Experiment 1.

Results

Response latency. In the analyses of Experiment 2, we excluded the words which elicited higher error rates in Experiment 1 to ensure that the comparisons of Experiment 1 and 2 are based on the same set of stimuli. Response accuracy rate ranged from 0.96 to 1.00 (mean = 0.99, *SD* = 0.01) for English monolingual readers, therefore no subject was excluded from the analyses. Of all trials (4,750 data points), data points with response latency shorter than 300 ms or longer than 2,000 ms were excluded (11 data points). Twenty-two words that were excluded in the analyses for Japanese-English bilinguals in Experiment 1 were also excluded from the analyses for monolingual readers of English in order to compare the two groups of readers based on the identical set of words (8.7 % of the remaining data points). The following analyses are based on the same 228 words analyzed in Experiment 1 with correct responses (4,275 data points, after excluding 1.1% of trials with incorrect responses). A reciprocal transformation (-1000/RT) was applied to RTs to attenuate the skew in its distribution.

Table 4-4. Estimate, standard error, t-value, p-value, and effect size (ms) of the fixed effects in the model of English monolingual readers' lexical decision response times.

Response time	Type	Estimate	SE	t-value	p-value	Effect size
(Intercept)		-1.660	0.065	-25.66	< 0.0001	
PreviousResponseCorrect (Error)	Task	0.085	0.018	4.82	< 0.0001	23
Trial	Task	-0.020	0.029	-0.68	0.4976	-46
PreviousRT_resid	Task	0.181	0.020	9.27	< 0.0001	124
Length	Engl	0.029	0.007	4.33	< 0.0001	23
OLD20_resid	Engl	0.260	0.153	1.70	0.0889	3
Freq_HAL	Engl	-0.030	0.006	-5.22	< 0.0001	-37
SUBTLCD_resid	Engl	-0.058	0.009	-6.64	< 0.0001	-63
Imageability_resid	Engl	-0.011	0.004	-2.81	0.0049	-19
OLD20_resid * Trial	Engl * Task	0.172	0.076	2.28	0.0228	Fig 4-4 (a, b)
Freq_HAL * Trial	Engl * Task	-0.002	0.003	-0.72	0.4749	Fig 4-4 (a, b)
Freq_HAL * OLD20_resid	Engl * Engl	-0.027	0.016	-1.70	0.0886	Fig 4-4 (a, b)
Freq_HAL * OLD20_resid * Trial	Engl * Engl * Task	-0.016	0.008	-2.08	0.0380	Fig 4-4 (a, b)

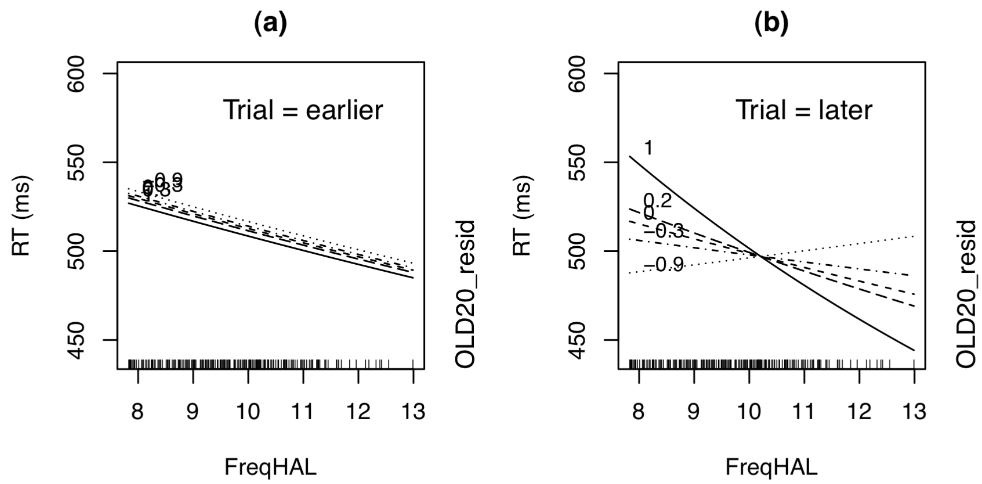


Figure 4-4. Lexical interactions in the mixed-effects model for monolingual English readers' response latencies. Different lines represent quantiles, and the rug in the x-axes represents the pattern of distribution.

We first fitted the model obtained for the Japanese-English bilinguals' response times in Experiment 1 to those of the monolingual readers of English. Importantly, the bilingual-specific lexical effects (*Freq_HAL* * *FreqJPN_resid*, *PhonologicalSimilarityJPN* * *Trial*, and *SemanticSimilarity_resid* * *PreviousRT_resid*) did not reach significance for the English monolinguals.[9] We then tested simple effects, all pairwise interactions, and finally interactions between task effects and lexical effects. Potentially influential outliers were removed with standardized residuals exceeding 2.5 standard deviation units (1.9% of the data points). Table 4-4 summarizes the coefficients of the final

[9] SemanticSimilarity_resid had a significant main effect, but its effect was no longer significant once Imageability_resid was included in the model.

model. The random-effect structure of this model comprised random intercepts for item (*SD* = 0.07) and subject (*SD* = 0.14), and by-subject random slopes for *Trial* (*SD* = 0.03) and *PreviousRT_resid* (*SD* = 0.07). The standard deviation of the residual error was 0.25.

There were several notable differences and commonalities regarding what lexical predictors co-determine response latencies of bilinguals and monolinguals. In Experiment 2, a three-way interaction of *FreqHAL* by *OLD20_resid* by *Trial* was observed. As illustrated in Panels a and b of Figure 4-4, *FreqHAL* provided facilitation without interacting with *OLD20_resid* early in the experiment (*Trial* was dichotomized for the purpose of visualization only in Figure 4-4). However, later in the experiment, the interaction between *FreqHAL* and *OLD20_resid* emerged (Panel b), as observed for Japanese-English bilinguals in Experiment 1. Unlike Experiment 1, *Imageability_resid* co-determined response times as well, such that the responses were delayed for words with low imageability (Balota et al., 2004; see also Prado & Ullman, 2009, for discussion of imageability as a diagnostic of storage next to frequency).

As reported by previous bilingual processing studies (e.g., Duyck et al., 2008; Gollan et al., 2008), the magnitude of the English word frequency effect was larger for Japanese-English late bilinguals than for English monolingual readers. This larger frequency effect for non-native readers can be explained by the negative decelerating functional form of the word frequency effect, indicating that each additional log unit of frequency provides smaller and smaller processing benefits (Baayen, Feldman, & Schreuder, 2006; Duyck et al., 2008). Baayen and Milin (2010) reported the same pattern between subject within monolinguals. Namely, a frequency effect was found only for slow responders. The slower responses of bilinguals allowed more time for the frequency effect to emerge. Overall, monolingual English readers' response latencies (M = 546 ms, SD = 138) in Experiment 2 were faster than those of Japanese-English bilingual readers (M = 733 ms, SD = 248) in Experiment 1 ($p < 0.0001$, mixed-effects model not shown).

First fixation duration. Native English speakers read words with a single fixation only 8% of the time (66% showed two fixations, 23% three fixations, and 2% four fixations). The eye-movements were therefore analyzed in the same manner as in Experiment 1.

For the analysis of the first fixation durations, data points with a first fixation shorter than 100 ms and longer than 850 ms, those before a blink, and those with an incorrect response were excluded from the analysis. A log-transformation was applied to attenuate skewness in the distribution of the fixation durations (M = 5.6, SD = 0.2, raw median = 264 ms). We first fitted the model for Japanese-English bilinguals first and confirmed that *PhonologicalSimilarityJPN* was not significant, as expected.

The final model is summarized in Table 4-5. Potentially influential outliers were removed with as criterion standardized residuals exceeding 2.5 standard deviation units (2.7% of the data points). The random-effect structure of this model comprised random intercepts for item (SD = 0.03) and subject (SD =

0.11), and by-subject random slopes for *Trial* (*SD* = 0.01) and *Length* (*SD* = 0.03). The standard deviation of the residual error was 0.16.

For the expert readers, Length was the only significant lexical predictor (recall that *Length* was not significant for Japanese-English bilinguals). *FreqHAL* and *OLD_resid*, which co-determined the bilingual readers' first fixation durations in Experiment 1, were not significant. An item-wise correlation between the first fixation durations of Japanese-English bilingual readers (Experiment 1) and those of English monolingual readers was significant but weak ($r = 0.19$, $p < 0.01$). Irrespective of such qualitative difference between native monolinguals and late bilinguals, the first fixations of the former group of readers were not significantly faster than those of the latter. We conclude that the processing advantage for native readers was not apparent at the first fixation.

Table 4-5. Estimate, standard error, t-value, p-value, and effect size (ms) of the fixed effects in the model of English monolingual readers' first fixation durations, first subgaze durations, and last fixation durations.

First fixation duration	Type	Estimate	Std.Error	t-value	p-value	Effect size
(Intercept)		5.620	0.024	230.71	< 0.0001	
Trial	Task	0.007	0.004	1.69	0.0911	8
PreviousRT_resid	Task	-0.023	0.008	-2.98	0.0029	-17
Length	Engl	0.031	0.007	4.51	< 0.0001	26
First subgaze duration	**Type**	**Estimate**	**Std.Error**	**t-value**	**p-value**	**Effect size**
(Intercept)		-3.157	0.094	-33.5	< 0.0001	
Trial	Task	-0.047	0.018	-2.61	0.0092	-11
PreviousRT_resid	Task	0.136	0.036	3.82	0.0001	32
Length	Engl	0.223	0.017	13.48	< 0.0001	76
FreqHAL	Engl	-0.043	0.014	-3.12	0.0018	-20
SUBTLCD_resid	Engl	-0.089	0.021	-4.18	< 0.0001	-34
Length * Trial	Engl *Task	-0.024	0.010	-2.44	0.0147	Not plotted
Last fixation duration	**Type**	**Estimate**	**Std.Error**	**t-value**	**p-value**	**Effect size**
(Intercept)		14.164	0.184	76.91	0.0000	
FirstSubgazeDuration_resid	Task	-3.050	0.063	-48.11	0.0000	-1080
PreviousResponseCorrect (Error)	Task	0.530	0.167	3.17	0.0015	15
Imageability_resid	Engl	-0.100	0.042	-2.36	0.0181	-19

First subgaze duration. Data points before a blink and trials with incorrect responses were excluded from the analysis. Trials with a first subgaze duration shorter than 100 ms were excluded, and an inverse transformation was applied to the remaining fixation durations to attenuate skewness in the distribution (M = -3.2, SD = 0.9, raw median = 308).

We first fitted the first subgaze model in Experiment 1 and confirmed that the effects of *PhonologicalSimilarityJPN* and the interaction between *FreqHAL* and FreqJPN_resid did not reach significance. This ensures that the two effects are due to the bilingual processing system.

Table 4-5 summarizes the final model for English monolinguals, after testing all main effects and pairwise interactions. Potentially influential outliers were removed with standardized residuals exceeding 2.5 standard deviation units (1.3% of the data points). The random-effect structure of this model comprised random intercepts for item (*SD* = 0.14) and subject (*SD* = 0.4), and by-subject random slopes for *Trial* (*SD* = 0.07). The standard deviation of the residual error was 0.73. The random-effect structure of this model comprised *Length*, *OLD20_resid*, *FreqHAL*, and *SUBTLCD_reisd*, all of which co-determined Japanese-English bilinguals' first subgaze in Experiment 1. A significant item-wise correlation between the first subgaze durations of Japanese-English bilingual readers (Experiment 1) and those of English monolingual readers (r = 0.46, $p < 0.01$) also indicate strong commonality in information uptake during the relatively early stage of word recognition. The first subgaze durations of the monolingual readers were significantly faster than those of the bilingual readers ($p < 0.0001$, effect size = 163 ms, mixed effects model not shown), indicating a processing advantage for native readers during this time frame.

Last fixation duration. In the same subset of words analyzed above, we excluded last fixation durations longer than 900 ms and applied a square root transformation to attenuate a skew in the distribution (*M* = 14.2, *SD* = 3.7, raw median = 216). Fixations before or after a blink and trials with incorrect responses were excluded from the analysis.

We first fitted the same model fitted to the last fixation durations of the late bilinguals. The bilingual-specific effects we observed in Experiment 1 (*PhonologicalSimilarity* * *FreqJPN_resid*, *SemanticSimilarity_resid* * *PhonologicalSimilarityJPN, SemanticSimilarity_resid* * *PreviousRT_resid*) were not significant, as expected, and were therefore removed from the model specification. The final model is summarized in Table 4-5. Potentially influential outliers were removed with standardized residuals exceeding 2.5 standard deviation units (1.9% of the data points). The random-effect structure of this model comprised random intercepts for item (*SD* = 0.81) and subject (*SD* = 0.75), and by-subject random slopes for *FirstSubgazeDuration_resid* (*SD* = 0.61). The standard deviation of the residual error was 2.35.

The *Imageability_resid* was the only lexical predictor co-determining monolingual English readers' last fixation durations. Recall that *Imageability_resid* also co-determined the same readers' response times to a comparable magnitude. Interestingly, as in Expeirment 1, *PreviousResponseCorrect* was significant here as well. When participants make an error response (which they are usually aware of), their immediately following responses become slower to be more cautious. Such strategic effects do not seem to co-determine early measures and only inflated last fixation durations in the study.

An item-wise correlation between the last fixation durations of Japanese-English bilingual readers (Experiment 1) and those of English monolingual readers was significant but weak ($r = 0.29$, $p < 0.01$). Furthermore, the last fixations of the monolingual English readers were not significantly faster than those of the late bilinguals. Apparently, the two groups of readers used different yet equally optimal response strategies for their benefit.

Discussion

Experiment 2 confirmed that there was commonality in processing between monolingual readers and bilingual readers: the lexical processes proceed from visuo-perceptual and sublexical orthographic effects (*Length* and *OLD20_resid*) to orthographic lexical effects (*FreqHAL* and *SUBTLCD_resid*) and then to semantic processes (*Imageability_resid* and *SemanticSimilarity_resid*). Importantly, however, the bilingual-specific effects obtained in the mixed-effects models in Experiment 1 did not reach significance in Experiment 2. This indicates that the bilingual-specific effects of our interest were genuine bilingual effects arising from the theoretical bilingual-specific architecture rather than artifacts arising from processes of target words per se. The functional overlap across different groups of readers and the significance of bilingual-specific effects only for bilingual readers are in line with the BIA and BIA+ models, which assume that the bilingual lexical processing system is straightforward extension of the monolingual lexical processing system.

Experiment 2 also identified differences between expert monolingual readers and non-expert late bilinguals. Expertise in reading, as witnessed in native monolingual readers' faster responses, was not apparent at the first fixation nor at the last fixation. This indicates that expertise in reading does not specifically reflect quicker decoding of visuo-perceptual features based on perceptual learning nor post-lexical strategic factors. Instead, expertise in reading is attributed to faster processes in the word identification system, as seen in the shorter first subgaze durations for monolingual readers.

Finally, the significant effects of *PreviousResponseCorrect* on last fixation durations and response times, but not first fixations and subgaze durations, of both monolingual readers and bilingual readers suggest that this is a language-general phenomenon and that readers' conscious strategy to respond more cautiously affects only the late processes. This, in turn, indicates that the first fixations and subgazes are relatively strategy-free measures for automatic lexical processing. The late involvement of a response strategy and the word identification system's insensitivity to a strategic factor is line with the unidirectional processing flow from the word identification system to the task/decision system in the BIA+ model.

General discussion

This study addressed the question of the time-course of lexical access in Japanese-English bilinguals. To this end, we combined lexical decision with eye-tracking because eye-tracking, unlike behavioral responses, affords insight into the time-course of lexical activation. Importantly, because the task did not involve cross-script priming, the lexical orthographic representation (Figure 4-1, box b) was not artificially activated when the L2 target word was presented, allowing more natural interpretation of the results. We observed clear effects of all the bilingual-specific lexical predictors (cross-language phonological similarity, L1 word frequency interacting with L2 word frequency, and cross-language semantic similarity), except the factor coding cognate status, as predicted by the BIA+ model.

First, cross-language phonological similarity facilitated the lexical decision responses of Japanese-English bilinguals. The eye-tracking record clarified that the phonological similarity effect emerged already at the first fixation. Within the framework of BIA+ model, we interpret this effect as sketched in Figure 4-1 box (d). Once the alphabetic letter representations I, N, T, E, R, V, and W are activated, based on the written input *interview*, the corresponding phonemes (e.g, /ɪ/, /n/, /t/, /ɚ/, /v/, /j/, and /u/) are activated to derive the appropriate lexical phonological representation /ɪntɚvju:/. It is conceivable that, at the same time, the activation of English phonemes can lead to (at least partial) activation of the corresponding Japanese phonemes (e.g., /i/, /n/, /t/, /a/, /b/, /j/, and /u/), eventually leading to the activation of /inntabjuu/.

For bilinguals of languages with the same orthographic script, a cross-language phonological effect has been reported to arise early at a short prime duration of 42 ms (Brysbaert et al., 1999; Duyck et al., 2004), and such a masked cross-language phonological priming effect was also reported for Japanese-English bilinguals (Nakayama et al., 2011). In this lexical decision study without priming, too, a clear effect of phonological similarity appeared in the early time frames. However, its effect was inhibitory at the first fixation, facilitatory at the first subgaze, and facilitatory also at the moment of the response. Previous studies reported mixed results: Dijkstra, Grainger, and van Heuven (1999) reported an inhibitory effect of phonological similarity, and Lemhöfer and Dijkstra (2004) and Haigh and Jared (2007) reported facilitatory effects. Our results indicate that this is not an either-or problem but that inhibition and facilitation manifest themselves at different points in time.

Interestingly, the magnitude of the facilitatory effect of cross-language phonological similarity on lexical decision responses increased throughout the experiment in Experiment 1. Because the co-activation of an L1 Japanese word via its phonological overlap with the L2 English word is a reasonable criterion for a 'yes' response in lexical decision, participants may have fine-tuned their global response criteria so that phonology received progressively more weight to optimize responses.

Second, L1 Japanese word frequency co-determined lexical decision responses in an interaction with L2 English word frequency. The English word frequency effect was progressively attenuated as Japanese frequency increased. Importantly, this interaction was also observed at the first subgaze, but not at the first and last fixations. This finding can be understood within the BIA+ model as follows. The BIA+ model posits, as does any interactive activation model, inhibitory links between non-identical orthographic/phonological lexical representations (see Figure 4-1, line c, and previous studies by Dijkstra et al., 1998; Kerkhofs et al., 2006). Because there are no links projecting from the English letter units to the Japanese *katakana* word representation (i.e., the dashed line a in Figure 4-1 is not active), the input word *interview* cannot activate the Japanese lexical *katakana* representation インタビュー at an early stage. However, the English input word *interview* activates the alphabetic letter representations I, N, T, E, R, V, and W, and consequently the English orthographic lexical representation INTERVIEW relatively early. The activation of the L1 Japanese lexical orthographic representation of a cognate can and does occur but only via indirect activation mediated by sublexical and lexical phonological representations or via top-down activation from the conceptual representation (see, for instance, the significant cross-language masked priming effect for non-cognate translation equivalents reported by Grainger & Frenck-Mestre, 1998; see also Duñabeitia, Perea, & Carreiras, 2010; Pecher, 2001; Perea, Duñabeitia, & Carreiras, 2008). The absence of the interaction between English and Japanese frequencies at the last fixation indicates that the effect is not due to a response strategy but part of bilingual lexical identification system.

Third, a small significant facilitatory contribution of semantic similarity was observed. Its effect arose late, at the last fixation and in the lexical decision responses. In the present study, the semantic similarity effect did not depend on readers' L2 proficiency, gauged by the duration of their stay in Canada. This finding is not in line with models postulating asymmetrical form-to-meaning mapping for L1 and L2, such as the Revised Hierarchical Model (RHM; Kroll & Stewart, 1994), but is more consistent with models assuming a strong form-to-meaning mapping for both L1 and L2, allowing rapid semantic activation in L2 word processing (Duyck & Brysbaert, 2004).

Fourth, regression modeling allowed us to straightforwardly test the view that cognate facilitation arises from special morphological representations for cognates (Davis et al., 2010; Lalor & Kirsner, 2000; Sánchez-Casas & García-Albea, 2005). If a given factor encodes a theoretically informative representation, then it should emerge as a significant predictor with other relevant predictors included in the regression model (e.g., a significant effect of orthographic identity status with the presence of a significant effect of orthographic similarity in Dijkstra et al., 2010). Although models with the factor *Cognate* as a sole lexical predictor replicated standard cognate facilitation effects, it was no longer a significant predictor once the above-mentioned numerical predictors were included in the regression equations. This is in harmony with Voga and Grainger's (2007) conclusion; a cognate facilitation effect is not due to special morphological representations but due to shared form overlap.

Finally, we observed similarities and differences in how monolingual readers and bilingual readers make lexicality judgments. On one hand, the lexical decision processes of the two groups were comparable, as indicated by the lexical distributional properties that similarly co-determined the lexical decision responses of the two groups of readers. Such a functional overlap between groups is consistent with the general architecture of the BIA+ model, which is a generalization of the monolingual interactive activation model. On the other hand, the lexical decision process of the bilingual readers diverged from that of the monolingual readers, beyond the bilingual-specific lexical effects mentioned above. This happened via the fine-tuning of response criteria throughout the experiment, as Japanese-English bilinguals apparently adjusted the response threshold for lexical decision with respect to Japanese word properties.[10] The first fixations of the two groups of readers were different as well: *Length* was significant only for monolinguals, and *OLD_resid* was significant only for bilinguals). This may reflect the fact that the perceptual span of proficient readers is wider than that of less proficient readers (Rayner, Slattery & Belanger, 2010). We also observed that monolingual English readers' expertise in reading manifested itself at the subgaze but not at the first fixation and at the last fixation. This suggests that what is termed *language proficiency* comprises highly skilled processing in the lexical identification system more than at the perception and decision making stages.

Future research should further investigate potential consequences of quality of lexical predictors, experimental manipulations, and individual differences. As for quality of lexical predictors, phonological and semantic similarity measures rated by Japanese readers were used in this study, as in many previous studies. However, the objective edit distance measure replicated the late effect, but not early effects, of the rated phonological similarity successfully.[11] It

---

[10] Alternatively, it may also be possible to interpret the trial-dependent lexical effects in terms of resting-level shift in the word identification system. One concern about this assumption of resting-level shift in the word identification system is that it is less straightforward to explain why only L1 Japanese properties are affected, rather than L2 target word properties. Particularly given that a word frequency effect was larger for Japanese-English bilinguals, we could have observed FreqHAL * Trial interaction in this case.

[11] This may be due to Japanese-English late bilinguals' incomplete mental phonological representations of English words, which was not assumed for PhonologicalDistance. When a rated phonological similarity based on an assessment by 10 native English speakers was considered (PhonologicalSimilarityENGL), this variable behaved much like the objective PhonologicalDistance albeit the fact that PhonologicalSimilarityENGL correlated with

is important to further investigate what constitute rated measures and how to code them objectively. As for experimental manipulations, the font size chosen in this study was relatively larger than that in normal reading. Given that font size is not known to affect general patterns of lexical effects (Chauncey, Holcomb, & Grainger, 2008 with ERP and masked priming; Miwa et al., 2013 with lexical decision and eye-tracking), we expect that the current results generalize to smaller font sizes. While readers made multiple fixations most of the time in this study, an analysis of eight native English speakers' reading with small font (visual angle per letter = 0.4˚) revealed that multiple eye-movements were used 73% of the time (median = 2, range = 1:5), indicating that isolated word reading itself triggers a task-specific pattern of eye-movements. As for individual differences, we only consider a potential effect of readers' length of stay in Canada, and we leave it to future research to disentangle contributions of various related measures (see Andrews & Lo, 2011; Kuperman & Van Dyke, 2011; Yap, Balota, Sibley, & Ratcliff, 2011). It is likely that the by-subject random intercepts, which capture between subject variability that cannot be traced back to the predictors included in the present study, comprise variation that can be explained by more refined measures characterizing the individual subjects, including measures for their skills in their second language. We leave this issue to further research.

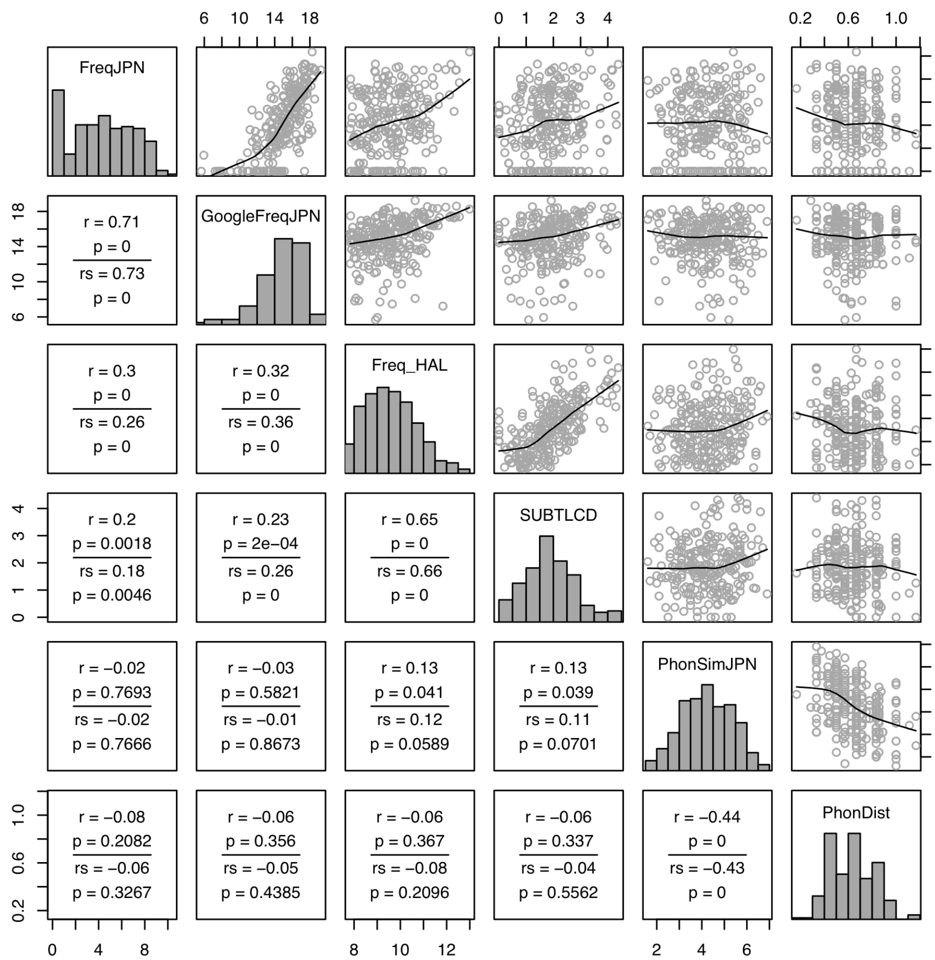
PhonologicalSimilarityJPN ($r = 0.64$, $p < 0.01$) more strongly than PhonologicalDistance ($r = 0.4$, $p < 0.01$). In the PhonologicalSimilarityENGL rating task, the English speakers saw an English word on the computer display, while they heard a corresponding Japanese word recorded by a native female Japanese reader. Using a seven-point scale (1 = identical pronunciation, 7 = very different), they rated how similar or different the auditory Japanese translation equivalent was from the pronunciation of the visually presented English word. The obtained ratings were

In conclusion, without using a priming technique, the present study tapped into the time-course of lexical activation by observing eye movements to test various predictions of the BIA+ model. The bilingual-specific lexical processes characterized by early cross-language phonological similarity to an interaction between L1 and L2 frequency, and then to late cross-language semantic similarity is in line with the BIA+ model. The absence of a significant contribution of a factor coding cognate status indicates that a cognate facilitation effect arises from the orchestration of effects of numerical predictors coding form and meaning in two languages. The localist connectionist framework of the BIA+ model, which thus far has been guided by research on bilinguals with the same script, can be modified to account for lexical processes of Japanese-English bilinguals, under the straightforward assumption that English letter units do not project onto Japanese word units.

---

subtracted from 8 to reverse the scale to code the degree of phonological similarity (*M* = 4.3, *SD* = 1.1).

Appendix 4-1. A matrix of scatterplots for Japanese and English frequencies and cross-language phonological similarity measures.



Appendix 4-2. English words used in the present study and their Japanese *katakana* translation and phonology. The flap /ɽ/ was used to encode English approximants /r/ and /l/. /ɸ/ represents a voiceless bilabial fricative. Vowels and consonants were repeated in the Japanese phonological transcriptions to encode the Japanese-specific moraic long vowels, moraic nasals, and moraic obstruents. Words marked with * were excluded from the analyses.

English word	Japanese katakana	Japanese phonology	English word	Japanese katakana	Japanese phonology
accent	アクセント	akusennto	lesson	レッスン	ɽessunn
account	アカウント	akaunnto	letter	レター	ɽetaa
advance	アドバンス	adobannsu	library	ライブラリー	ɽaibuɽaɽii
advantage	アドバンテージ	adobannteedʒi	license	ライセンス	ɽaisennsu
advice	アドバイス	adobaisu	magazine	マガジン	magadʒinn
agenda	アジェンダ	adʒennda	manifest	マニフェスト	maniɸesuto
amateur*	アマチュア	amatʃua	manner	マナー	manaa
anchor	アンカー	annkaa	marble	マーブル	maabuɽu
answer	アンサー	annsaa	margin	マージン	maadʒinn
appeal	アピール	apiiɽu	massage	マッサージ	massaadʒi
arcade*	アーケード	aakeedo	matrix	マトリックス	matoɽikkusu
architect	アーキテクト	aakitekuto	measure	メジャー	medʒaa
aspect	アスペクト	asupekuto	medicine	メディスン	medisunn
attempt	アテンプト	atemmputo	merchant	マーチャント	maatʃannto
auction	オークション	ookuʃonn	message	メッセージ	messeedʒi
autumn	オータム	ootamu	method	メソッド	mesoddo
avenue	アベニュー	abenjuu	minister	ミニスター	minisutaa
balloon	バルーン	baɽuunn	miracle	ミラクル	miɽakuɽu
ballot*	バロット	baɽotto	mirror	ミラー	miɽaa
basket	バスケット	basuketto	mission	ミッション	miʃʃonn
blanket	ブランケット	buɽannketto	moment	モーメント	moomennto
bottom	ボトム	botomu	monster	モンスター	monnsutaa
bracket	ブラケット	buɽaketto	morning	モーニング	mooniŋŋgu
breast	ブレスト	buɽesuto	motion	モーション	mooʃonn
breath	ブレス	buɽesu	mountain	マウンテン	maunntenn
bronze	ブロンズ	buɽonnzu	muscle	マッスル	massuɽu
buffalo	バッファロー	baɸɸaɽoo	museum	ミュージアム	mjuudʒiamu
buffer	バッファー	baɸɸaa	nature	ネーチャー	neetʃaa
bullet	ブレット	buɽetto	needle	ニードル	niidoɽu
bulletin	ブリテン	buɽitenn	notice	ノーティス	nootisu
bundle	バンドル	banndoɽu	notion	ノーション	nooʃonn
burden	バードン	baadonn	number	ナンバー	nammbaa

business	ビジネス	bidʒinesu	occasion	オケーション	okeeʃonn
butter	バター	bataa	office	オフィス	oɸisu
cabinet	キャビネット	kjabinetto	opinion	オピニオン	opinionn
camera	カメラ	kameɽa	opponent	オポーネント	opoonennto
candle	キャンドル	kjanndoɽu	option	オプション	opuʃonn
cannon	キャノン	kjanonn	palace	パレス	paɽesu
career	キャリア	kjaɽia	parade	パレード	paɽeedo
cartoon	カートゥーン	kaatuunn	paradise	パラダイス	paɽadaisu
castle	キャッスル	kjassuɽu	paradox	パラドックス	paɽadokkusu
catalog	カタログ	kataɽogu	pencil	ペンシル	pennʃiɽu
cathedral*	キャシードラル	kjaʃiidoɽaɽu	peninsula*	ペニンスラ	peninnsuɽa
cattle	キャトル	kjatoɽu	period	ピリオド	piɽiodo
ceiling*	シーリング	ʃiiɽiŋŋgu	personnel	パーソネル	paasoneɽu
century	センチュリー	senntʃuɽii	phantom	ファントム	ɸanntomu
challenge	チャレンジ	tʃaɽenndʒi	planet	プラネット	puɽanetto
champion	チャンピオン	tʃammpionn	plastic	プラスチック	puɽasutʃikku
chance	チャンス	tʃannsu	pocket	ポケット	poketto
channel	チャンネル	tʃanneɽu	poison	ポイズン	poizunn
chapter	チャプター	tʃaputaa	police	ポリス	poɽisu
character	キャラクター	kjaɽakutaa	politics	ポリティクス	poɽitikkusu
charter	チャーター	tʃaataa	poverty	パーバティー	paabatii
cherry	チェリー	tʃeɽii	priest	プリースト	puɽiisuto
chocolate	チョコレート	tʃokoɽeeto	prince	プリンス	puɽinnsu
church	チャーチ	tʃaatʃi	principle	プリンシプル	puɽinnʃipuɽu
circuit*	サーキット	saakitto	prison	プリズン	puɽizunn
circus	サーカス	saakasu	privilege	プリビレッジ	puɽibiɽeddʒi
cluster	クラスター	kuɽasutaa	profile	プロフィール	puɽoɸiiɽu
college	カレッジ	kaɽeddʒi	program	プログラム	puɽoguɽamu
comment	コメント	komennto	promise	プロミス	puɽomisu
complaint	コンプレイント	kommpuɽeinnto	protest	プロテスト	puɽotesuto
component	コンポーネント	kommpoonennto	rabbit	ラビット	ɽabitto
condition	コンディション	konndiʃonn	receipt	レシート	ɽeʃiito
conflict	コンフリクト	konnɸuɽikuto	recipe	レシピ	ɽeʃipi
content	コンテント	konntennto	rescue	レスキュー	ɽesukjuu
corner	コーナー	koonaa	result	リザルト	ɽizaɽuto
couple	カップル	kappuɽu	rocket	ロケット	ɽoketto
course	コース	koosu	salary	サラリー	saɽaɽii
courtesy*	カーテシー	kaateʃii	sample	サンプル	sammpuɽu
credit	クレジット	kuɽedʒitto	satellite*	サテライト	sateɽaito
crystal	クリスタル	kuɽisutaɽu	scheme*	スキーム	sukiimu
culture	カルチャー	kaɽutʃaa	school	スクール	sukuuɽu
damage	ダメージ	dameedʒi	search	サーチ	saatʃi
danger	デンジャー	denndʒaa	secretary	セクレタリー	sekuɽetaɽii

debate	ディベート	dibeeto	sentence	センテンス	senntennsu
defense	ディフェンス	diɸennsu	session	セッション	seʃʃonn
degree	ディグリー	diguɽii	shadow	シャドー	ʃadoo
design	デザイン	dezainn	shield*	シールド	ʃiiɽudo
diagram	ダイアグラム	daiaguɽamu	soccer	サッカー	sakkaa
diamond	ダイアモンド	daiamonndo	socket	ソケット	soketto
dilemma*	ジレンマ	dʒiɽennma	soldier	ソルジャー	soɽudʒaa
disaster	ディザスター	dizasutaa	source	ソース	soosu
disease	ディジーズ	didʒiizu	speech	スピーチ	supiitʃi
district	ディストリクト	disutoɽikuto	sponsor	スポンサー	suponnsaa
doctrine*	ドクトリン	dokutoɽinn	square	スクエア	sukuea
domain	ドメイン	domeinn	stance	スタンス	sutannsu
donkey	ドンキー	doŋŋkii	statue*	スタチュー	sutatʃuu
dragon	ドラゴン	doɽagonn	status	ステータス	suteetasu
dungeon*	ダンジョン	danndʒonn	street	ストリート	sutoɽiito
effort	エフォート	eɸooto	strength	ストレングス	sutoɽenngusu
elephant	エレファント	eɽeɸannto	string	ストリング	sutoɽiŋŋgu
embassy	エンバシー	emmbaʃii	studio	スタジオ	sutadʒio
emergency	イマージェンシー	imadʒennʃii	summer	サマー	samaa
emperor*	エンペラー	emmpeɽaa	surface	サーフェス	saaɸesu
episode	エピソード	episoodo	syndrome	シンドローム	ʃinndoɽoomu
example	エグザンプル	eguzammpuɽu	system	システム	ʃisutemu
expert	エキスパート	ekisupaato	talent	タレント	taɽennto
fashion	ファッション	ɸaʃʃonn	target	ターゲット	taagetto
fatigue*	ファティーグ	ɸatiigu	technique	テクニック	tekunikku
fellow	フェロー	ɸeɽoo	template	テンプレート	temmpuɽeeto
finance	ファイナンス	ɸainansu	temple	テンプル	temmpuɽu
flavor	フレーバー	ɸuɽeebaa	territory	テリトリー	teɽitoɽii
flight	フライト	ɸuɽaito	texture	テクスチャ	tekusutʃa
friend	フレンド	ɸuɽenndo	theatre	シアター	ʃiataa
garbage	ガービッジ	gaabiddʒi	thread	スレッド	suɽeddo
garlic*	ガーリック	gaaɽikku	threshold*	スレッシュホールド	suɽeʃʃuhooɽudo
gender	ジェンダー	dʒenndaa	toilet	トイレット	toiɽetto
grease*	グリース	guɽiisu	traffic	トラフィック	toɽaɸikku
guitar	ギター	gitaa	tragedy	トラジェディー	toɽadʒedii
hazard	ハザード	hazaado	treaty	トリーティー	toɽiitii
helmet	ヘルメット	heɽumetto	tunnel	トンネル	tonnneɽu
heroin*	ヘロイン	heɽoinn	twilight	トワイライト	towaiɽaito
horizon	ホライズン	hoɽaizunn	vanilla	バニラ	baniɽa
husband	ハズバンド	hazubanndo	vehicle	ビークル	biikuɽu
impact	インパクト	immpakuto	venture	ベンチャー	benntʃaa
incentive	インセンティブ	innsenntibu	version	バージョン	baadʒonn
industry	インダストリー	inndasutoɽii	veteran*	ベテラン	beteɽann

insect	インセクト	innsekuto	village	ビレッジ	biɽeddʒi
instinct	インスティンクト	innsutinnkuto	violin	バイオリン	baioɽinn
interest	インタレスト	inntaɽesuto	vitamin	ビタミン	bitaminn
interval	インターバル	inntaabaɽu	volume	ボリューム	boɽjuumu
interview	インタビュー	inntabjuu	weather	ウェザー	wezaa
jacket	ジャケット	dʒaketto	whistle	ホイッスル	hoissuɽu
leisure	レジャー	ɽedʒaa	witness	ウィットネス	wittonesu

CHAPTER 5

General discussion

Libben and Jarema (2002) proposed that it is important not to derive a conclusion on human language processing solely from typical psycholinguistic studies with monolingual English speakers reading English words with their speed of lexicality judgment as the primary dependent variable. One decade later, the present thesis tested different languages and different populations with different measures, including readers' eye movements recorded during lexical decision experiments (i.e., primed Japanese *kanji* lexical decision with Japanese monolinguals in Chapter 2, Japanese *kanji* lexical decision with Japanese monolinguals in Chapter 3, and eye-tracking English lexical decision with Japanese-English bilinguals in Chapter 4). A primary goal of this thesis was to assess the impact of the three key factors in the mental lexicon research (i.e., language, population, and task) in order to sort out some of the aspects that are independent of language, population, and task variations from ones that are dependent on such factors. In the sections that follow, specific findings will be summarized with respect to the three key factors.

Language effects: Latin alphabet vs. morphographic scripts

Chapters 2 and 3 investigated whether and to what extent the characteristics of the orthographic scripts (i.e., alphabetic scripts encoding syllabic units vs. Japanese morphographic script) affect response times and eye movements in visual complex word recognition experiments. There are several commonalities and differences between the results obtained in the present Japanese word recognition studies and those obtained in the previous alphabetic word recognition studies.

Commonalities across languages

It is notable that morphemes (or morphographic characters) were activated in reading two-character compounds, in much the same way as they are activated in alphabetic compound recognition, with the left and right morphemes co-determining early and late eye fixation durations respectively (Hyönä & Pollatsek, 1998; Kuperman et al., 2008, 2009; Pollatsek et al., 2000).

The preferential left to right processing pattern may be due to the general left-to-right reading pattern but may also be due to the necessity to activate phonological information. Indeed, both left and right morphemes' phonological information, gauged by the number of homophonous characters, co-determined eye movements and response speed. The involvement of phonological information even in reading of a morphographic kanji, which encodes a meaning by itself, is in line with the theory that phonological information is automatically and always activated regardless of orthographic scripts (the *universal phonological principle*, Perfetti, Zhang, & Berent, 1992). With respect to the left and right morphemes, their asymmetrical contributions also deserve attention. Studies on alphabetic language processing found that foreseen difficulty (e.g., high complexity and low frequency of the stimulus) in a parafoveal region shorten the on-going processing (Hyönä & Bertram, 2004; Kennedy & Pynte, 2005; Kliegl, Nuthmann, & Engbert, 2006; Pynte, Kennedy, & Ducrot, 2004) while difficulty at the foveal region slows down the processing. The same pattern was observed in Chapter 3 (*LeftKanjiTokenFreq* and *RightKanjiTokenFreq*, as well as *LeftKanjiStrokesResid* and *RightKanjiStrokesResid*, showed asymmetrical effects in the earliest time frame). These results indicate that the parafoveal processing difficulty that motivates the current processing is script-independent.

The early effect of frequency of the whole compound is also notable. At the first subgaze/fixation in the lexical decision study in Chapter 3, a small yet significant whole word frequency effect was observed, as in alphabetic Dutch compound recognition (Kuperman et al., 2009). This indicates that, regardless of the orthographic script, the processing of compound words does not proceed strictly in a bottom-up combinatorial manner (i.e., in a bottom-up model, a standard facilitatory frequency effect of the whole compound word should appear only after facilitatory frequency effects of its morphemes) nor strictly in a supra-lexical manner, although a precise mechanism by which such early compound frequency effects take place is yet to be investigated. For example, it may be due to Gestalt perceptual pattern recognition and/or an increase in higher-level cognitive processes driven by rapid bottom-up information flow, depending on the locus of a word frequency effect.

Differences across languages

Many past studies of alphabetic word recognition suggest that morphemes are automatically activated at the earliest processing stages (e.g., Rastle, Davis, & New, 2004; Kuperman et al., 2009: Taft & Nguyen-Hoan, 2010). In many cases, however, a morpheme has been defined only vaguely as a minimal meaningful unit. Under this definition of a *morpheme*, morphological processing of Japanese kanji words differed from that of alphabetic words. While meanings of semantic radicals were found to characterize wordlikeness in Chapter 2, the effect of radical frequency was, when it was significant at all, small and inhibitory (i.e., not a standard frequency effect) in the earliest time frame in Chapter 3. In contrast, character frequency effects were robust and large relative to radical frequency effects at early processing stages, regardless of font sizes and any other lexical properties. These results indicate that semantic radicals (or purely *orthographic morphemes*) are psycholinguistically not comparable to conventional *morphemes* which are initial access units with orthographic, phonological, and semantic representations in memory.

Although visual complexity of Japanese kanji words vary with respect to stroke counts (or visuo-perceptual density within a fixed square region) rather than word length, higher visual complexity elongated processing time, as in processing of alphabetic words (Balota et al., 2004; Vitu, O'Regan, & Mittau, 1990). While this result is intuitively reasonable, the cross-linguistic difference should be taken into account to interpret the result in the context of the classical IA architecture (McClelland & Rumelhart, 1981). The IA model assumes a featural template (or a simple feature set) to which input letters are initially matched. If feature-level processing is solely based on template matching in an all-or-none fashion, there should not be any effects of character strokes. In addition, if such a template is assumed for morphographic character symbols, the template must be complex enough to incorporate all characters with various degrees of complexity, such as 龍, 島, and 四. All in all, the result indicates that identification of orthographic symbols is achieved by hierarchically organized layers of features (see Grainger, Rey, & Dufau, 2008 for a review).

Population effects: monolinguals and bilinguals of different languages

One decade ago, a majority of studies tested monolingual readers (of English), although the large majority of world's population is bilingual. While the number of bilingual studies has certainly increased since then, the past line of bilingual research has been dominated by studies on bilinguals with the same orthographic scripts, such as Dutch-English bilinguals and French-English bilinguals. This dissertation studied commonalities and differences between same-script bilinguals and different-script bilinguals, as well as monolinguals and bilinguals in general.

Same-script vs. different-script bilinguals

Chapter 3 tested Japanese-English bilinguals to test the predictions of the BIA+ model (Dijkstra & van Heuven, 2002): bilinguals have their native L1 and non-native L2 language knowledge integrated into the same lexical database in memory (i.e., a bilingual mental lexicon can be understood as a straightforward extension of a monolingual one). Although past studies with bilinguals with the same script have pointed to this conclusion with respect to lexical representations and the corresponding non-selective activation in lexical processing, it is yet to be clarified whether language non-selective activation is population-independent. The lexical decision experiment in Chapter 4 revealed that language non-selective activation also occurs for Japanese-English bilinguals in much the way as it occurs for bilinguals with the same script, albeit the fact that orthography-based non-selective lexical access cannot take place for this group of bilinguals. This suggests that a phonology-based non-selective lexical access route does exist and is sufficient to trigger cross-language non-selective lexical activation ultimately even at the levels of orthography and semantics.

Monolinguals vs. bilinguals

Chapter 4 tested monolingual English readers as an expert control group to test whether the effects observed in the preceding paragraph are indeed genuine bilingual effects. This approach also sheds further light on commonalities and differences between monolinguals and bilinguals in general.

For example, whole word frequency co-determined lexical processing more strongly in a later time frame in lexical processing of both monolinguals and bilinguals (Chapter 4). This suggests, although a locus of a word frequency effect has not been identified either in a particular domain nor in a particular time frame, that a word frequency effect is predominantly a product of later processing. Interestingly, the magnitude of a word frequency effect was modulated by the target words' orthographic neighbourhood density in both monolinguals and bilinguals' lexical decision responses, such that a word frequency effect is larger for words with smaller orthographic neighbourhood density. These commonalities between monolinguals and bilinguals support the BIA/BIA+ models' assumption that the bilingual lexical processing architecture is an extension of the monolingual processing architecture and that a certain part of language processing architecture is shared between monolinguals and bilinguals.

On the other hand, some differences were observed between these two groups of readers. First, although longer words were responded to more slowly than shorter words by both monolinguals and bilinguals, word length co-determined the duration of only native monolingual readers' first fixation but not that of bilingual readers. This suggests that expert monolingual readers and non-expert bilinguals utilize qualitatively different information, perhaps at the level of lower-level visuo-perceptual processing.

Monolingual and bilingual readers also utilized different types of information to make their responses more efficient throughout the experiment. Bilingual readers, for example, made use of cross-language phonological similarity more and more throughout the experiment (viz., the effect of *PhonologicalSimilarityJPN* became larger as *Trial* went by in Chapter 4).

Finally, the two groups of readers' eye movements were guided by different factors. While orthographic neighbourhood density co-determined lexical decision responses for both monolinguals and bilinguals, it co-determined eye movements of the bilingual readers only. Given this, although we cannot conclude that orthographic neighbourhood density information was not used by monolingual readers at all (because it certainly co-determined their lexical decision responses), it is possible to conclude that attentional patterns can be guided by different factors.

All in all, the above differences suggest that the bilingual is not two monolinguals in one person (Grosjean, 1989). Although the differences between monolinguals and bilinguals appear to be in conflict with the commonalities observed above, they are not. It is expected that differences with respect to word length, response strategy, and eye movement patterns are the products of the early visuo-perceptual processing, late response planning and execution, and attentional guidance respectively. Unfortunately, these are not explicitely considered by the BIA+ model in its current state. Given the degree of commonalities, it is clear that the bilingual lexical processing architecture shares some functional overlaps with the monolingual processing architecture. As the BIA+ model extended the original BIA model with respect to reader-specific response planning and execution, the BIA+ model should be extended similarly to account for reader-specific variations in other domains.

Task effects: Microstructure of lexical decision

Throughout this dissertation, a lexical decision task was used because this has been a most commonly used experimental paradigm to study degree of wordlikeness and consequently how words are represented in the mind. However, it should be kept in mind that a lexical decision paradigm may color experimental results. In searching for a functional overlap, it is one way to perform a cross-task comparison (Grainger & Jacobs, 1996). In this study, however, I opted for a close observation of a single task instead. In Chapters 3 and 4, regression models explicitly included the number of preceding trials up to the target word to study how readers changed patterns of their responses throughout the experiments.

In Chapter 3, the facilitatory effect of semantic transparency of the right *kanji* character was magnified as the experiment progressed. In Chapter 4, the facilitatory effect of cross-language phonological similarity was again magnified for Japanese-English bilinguals as the experiment went by. Although English monolinguals read the same stimulus words, the degree of the interaction between orthographic neighbourhood density and word frequency was magnified instead as the experiment went by. Importantly, these interactions between the number of preceding trials and lexical variables were not present in an early time frame. This suggests that task demands do affect how readers take on lexicality judgment tasks, but such response strategies do not appear to modulate the automatic word identification system. For this reason, the character-driven processing model in Chapter 3 and the BIA+ model in Chapter 4 specify that the identification system affects the task/decision system but not vice versa.

The introduction of the number of preceding trials as a predictor in a regression model is a relatively recent trend, and consequently future research should investigate how this variable interacts with lexical predictors to study whether lexical decision responses by comparable participants with nonwords also trigger a consistent response pattern.

Lexical decision tasks in this dissertation were also accompanied by a super-imposed eye-tracking component, following Kuperman et al.'s lexical decision eye-tracking study with Dutch compounds (2009). This procedure revealed that, regardless of whether words are logographic, small, or morphologically simple, isolated word lexical decision always involves attentional shifts, as gauged by eye movements. Attention-wise, therefore, a lexical decision task is not a mere miniature version of so-called "natural" sentential reading, in which morphologically simple words usually receive only one fixation.

Finally, Chapter 2 raised a concern about the common interpretation of a priming effect: recognition of a prime word partially completes certain processing required for a succeeding target word (i.e., certain nodes are pre-activated in the IA model framework). In Chapter 2, inclusion of the priming manipulation as a factor in the regression model made it clear that the supposedly neutral control condition was also affected by the priming procedure. As Marsolek (2008) suggested, a so-called control condition established by legimate word primes does not serve as an unbiased baseline of comparison to evaluate an effect of piming. Because a great number of studies used priming in the past without discussing much about how the priming occurs, future research should further clarify the mechanism of priming itself. For example, is priming really beneficial? Computational simulations for recognition of the word WORD by the IA model (Figure 5-1) reveal that, whether it is a partial repetition priming by an orthographic neighbour WORK, which is expected to produce a facilitatory priming effect (Grainger & Ferrand, 1996), or antipriming by orthographically dissimilar yet legitimate word MILK, the activation level for WORD reaches the recognition threshold of 0.7 faster when there is no priming (note that WORD was presented from the fifth cycle in all cases). Understanding more precise mechanism of priming should give us a better picture of the mental lexicon (see also Milin, Filipović Đurđević, & Baayen, 2008 for negative consequence of priming)

Remaining issues and future implications

Figure 5-2 summarizes the mental lexicon and its components within a larger context of cognitive science. The components of the mental lexicon and human cognition discussed in this dissertation are highlighted in grey. Because the present dissertation dealt with recognition of words in isolation only, it is apparent that there are many other issues yet to be investigated and not highlighted in the figure.

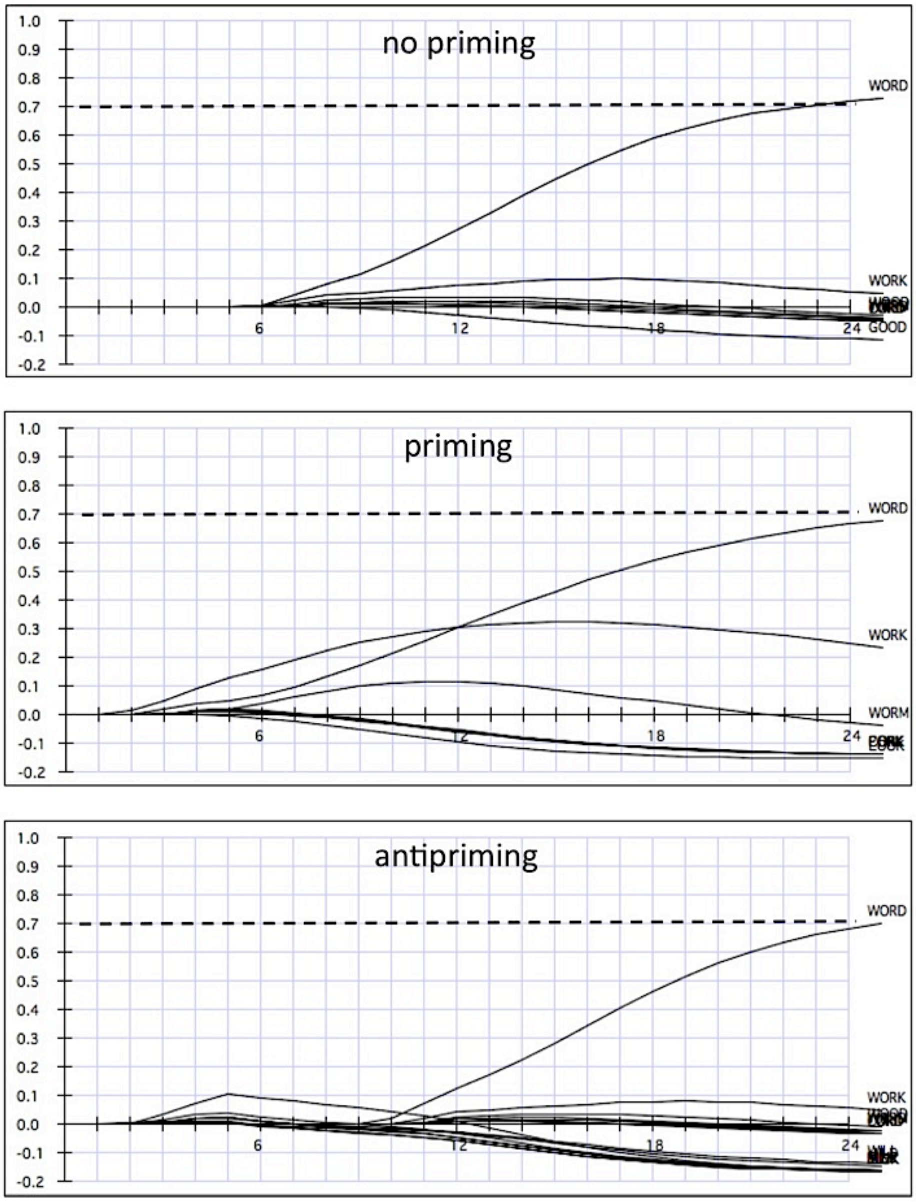


Figure 5-1. No priming, priming, and antipriming procedures simulated in the IA model (jIAM, van Heuven, 2008). Top (No Priming): WORD was presented from the fifth cycle. Middle (Priming): orthographically similar WORK was presented at the first cycle followed by WORD at the fifth cycle. Bottom (Antipriming): MILK was presented at the first cycle followed by WORD at the fifth cycle.

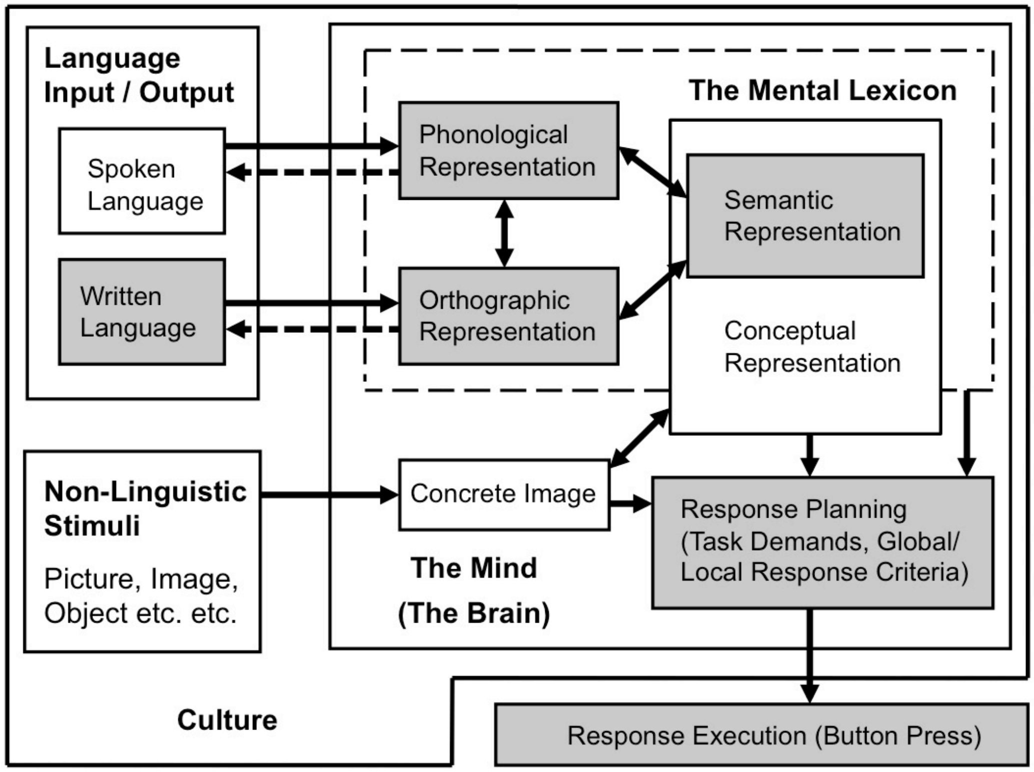


Figure 5-2. A graphical summary of the mental lexicon within a larger context of cognitive science.

First, a degree of a functional overlap between language comprehension and production processes (the dotted arrows in Figure 5-2) should be identified. This relates to a theoretical issue of whether information flows simply in the revese directions within the same processing architecture for comprehension and production, as speculated in Figure 5-2, or qualitatively distinct models are necessary. General information flow from the conceptual/semantic level to the phonological level, as assumed in some models of word production (see Levelt, 1999), indicate that the former perspective is reasonable. However, the finding that the locus of the word frequency effect in word production is at the lexeme (phonological) level (Jescheniak & Levelt, 1994), rather than at the more central lemma level, is not in line with the former perspective because, if this is the case, a word frequency effect should have been observed more strongly in the earliest time frame in this study.

Second, comprehension of visual words should be compared with comprehension of spoken words. Just as early studies of visual word recognition did not pay much attention to processing of phonological information, as reflected in the classical IA model (McClelland & Rumelhart, 1981), orthographic processing did not received much attention in studies of spoken word processing in the past, as reflected in the corresponding interactive activation model (e.g., the TRACE model: McClelland, & Elman, 1986). Studies in Chapters 3 and 4 of this dissertation both show that phonology intrudes on what could have been pure orthogrophic processing, indicating that there is a link from orthography to phonology. However, whether orthographic effects intrude on what might be expected to be pure phonological processing is a question that received attention only recently, and more work is needed for this area (Rastle, McCormick, Bayliss, & Davis, 2011; see also Escudero, Hayes-Harb, & Mitterer, 2008 for evidence from second language acquisition).

Third, comprehension of visual words can be compared with comprehension of non-linguistic stimuli, such as pictures, image, and objects. In other words, while the present dissertation accessed the mental lexicon from linguistic inputs, it is possible to investigate its function through recognition of non-linguistic inputs with an explicit linguistic task demand (e.g., picture naming) or with a non-linguistic task demand. In the latter case, a study tests whether language affects apparently non-linguistic cognitive processes and relates to the issue of linguistic relativity (Boroditsky, 2003; Whorf, 1956). A study without a linguistic task demand can reveal what components of the mental lexicon is more closely tied to non-linguistic cognition. For example, while studies on linguistic relativity with a large number of lexical distributional properties are currently scarce, a picture comparison study of Miwa, Libben, Rice, & Baayen (2008) assessed contributions of a large number of lexical predictors in a mixed-effects regression analysis. They found that two pictures are perceived to be less similar when frequency counts of the left and right pictures' word translation equivalents differ more greatly. This indicates that word frequency is semantic/conceptual in nature (Baayen, Feldman, & Schreuder, 2006).

Fourth, it should be born in mind that culture affects cognitive processes and consequently that it may also affect lexical processing in the mind. While there is a growing number of studies on bilingual processing and cross-language comparison, the two language samples are often confounded with two distinct cultures. For example, given the accumulated evidence that North Americans tend to process information holistically and East Asians tend to process information analytically (e.g., Nisbett & Masuda, 2003), the early word length effect for English monolinguals but not for Japanese-English bilinguals observed in Chapter 4 may be explained by culture-driven perceptual patterns inherent in these groups.

Finally, although the results were interpreted in a context of interactive activation framework throughout this dissertation, this line of chronometric studies should be evaluated at the intersection with the corresponding neurological research. It has been commonly believed that localist models, unlike parallel distributed processing models, do not resemble how the brain actually works, despite their explanatory values in psychology. Bowers (2009, 2010), however, suggest that localist models provide a better account of single-cell recording studies and that grandmother cells in neuroscience are comparable to localist representations in psychology (see also van Heuven & Dijkstra, 2010 for how localist connectionist models based on chronometric research can be understood in relation to the brain imaging)

Phenomena consistently observed across different orthographic scripts, individuals, and tasks must surely reflect the core of human language processes (i.e. functional overlap). The three studies in this dissertation collectively show that language-specific properties, individual differences, and variable task demands, by themselves, do not result in completely different pictures with respect to how wordlikeness emerges in visual word recognition (e.g., the time-couse of word frequency effect, left-to-right preferential morphological processing, and language non-selective activation). In addition, these studies also show that localist models still serve as a useful organizational framework. While the localist models described in this dissertation are currently still underspecified with respect to language, population, and task variations, it is left for future research to clarify what makes humans *human* by tracking functional overlaps not only in psycholinguistics but also in neuroscience, as well as across linguistic and non-linguistic processes.

REFERENCES

Abrams, R. A., & Balota, D. A. (1991). Mental chronometry: Beyond reaction time. *Psychological Science, 2,* 153-157.

Adelman, J. S., Brown, G. D. A., & Quesada, J. F. (2006). Contextual diversity, not word frequency, determines word-naming and lexical decision times. *Psychological Science, 17,* 814-823.

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

Amano, S. (2007). *Ketaishiyoto gengonoryoku*. In T. Kobayashi, S. Amano, & N. Masataka (Eds.), Mobairushakaino genjoto yukue: Riyojittaini motoduku hikarito kage (pp. 124-148). Tokyo: NTT Publishing.

Amano, S., & Kondo, K. (2000). *Nihongo-no goi tokusei* [Lexical properties of Japanese]. Tokyo: Sanseido.

Amano, S., & Kondo, K. (2003). *NTT database series: Lexical properties of Japanese* (version 2) [CD-ROM]. Tokyo: Sanseido.

Amano, S., Kondo, T., & Kataoka, R. (2005). *Tangoshinmitsudowo riyoshita goisusokute: Intanettoniyoru daikibochosa* [Vocabulary size estimation based on word familiarity: A large-scale web-based study]. In Proceedings of The 22nd Annual Meeting of the Japanese Cognitive Science Society. Kyoto, Japan, 58-59.

Andrews, S., & Lo, S. (2011). Not all skilled readers have cracked the code: Individual differences in masked form priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. doi: 10.1037/a0024953

Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology, 16*, 285-311.

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language, 62*, 67-82.

Baayen, R. H. (2005). *Data mining at the intersection of psychology and linguistics*. In Cutler, A. (Ed). Twenty-First Century Psycholinguistics: Four Cornerstones. (pp. 69-83). Hillsdale, NJ: Erlbaum.

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

Baayen, R. H. (2010). A real experiment is a factorial experiment? *The Mental Lexicon, 5,* 149-157.

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.

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

Baayen, R. H., Feldman, L. B., & Schreuder, R. S. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language, 55.* 290-313.

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. International *Journal of Psychological Research, 3*, 12-28.

Baayen, R. H., Milin, P., Filipović Đurđević, D., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review, 118*, 438-481.

Balling, L., & Baayen, R.H. (2008). Morphological effects in auditory word recognition: Evidence from Danish. *Language and Cognitive Processes, 23,* 1159-1190.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S., Spieler, D., & Yap, M. J. (2004). Visual word recognition for single-syllable words. *Journal of Experimental Psychology: General, 133,* 283–316.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods, 39*, 445-459.

Bates, D., Maechler, M., & Dai, B. (2007). *lme4: Linear mixed-effects models using S4 classes*. R package version 0.999375-27. http://lme4.r-forge.r-project.org/

Beck, J. (1966). Effects of orientation and of shape similarity on perceptual grouping. *Perception and Psychophysics, 1*, 311-312.

Belsley, A., Kuh, E., & Welsch, R. E. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity*. New York: Wiley.

Belsley, D. A., Kuh, E., & Welch, R. E. (2004). *Regression diagnostics: Identifying influential data and sources of collinearity*. Toronto: Wiley.

Bertram, R., Baayen, R. H., & Schreuder, R. (2000). Effects of family size for complex words. *Journal of Memory and Language, 42*, 390-405.

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.

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology, 14,* 143-177.

Boroditsky, L. (2003). Linguistic Relativity. In Nadel, L. (Ed.) *Encyclopedia of Cognitive Science.* MacMillan Press: London, UK, pages 917-921.

Bowers, J. S. (2009). On the biological plausibility of grandmother cells: implications for neural network theories in psychology and neuroscience. *Psychological Review, 116*, 220-251.

Bowers, J. S. (2010). More on grandmother cells and the biological implausibility of PDP models of cognition: a reply to Plaut and McClelland (2010) and Quian Quiroga and Kreiman (2010). *Psychological review, 117*, 300-306.

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

Breiman, L. (2001). *Random forests*. Machine Learning, 45, 5–32.

Brysbaert, M. (2003). *Bilingual visual word recognition: Evidence from masked phonological priming*. In S. Kinoshita & S. J. Lupker (Eds.). Masked priming: State-of-the-art (pp. 323-344). Hove, UK: Psychology Press.

Brysbaert, M. & Nazir, T. (2005). Visual constraints on written word recognition: Evidence from the optimal viewing position effect. *Journal of Research in Reading, 28*, 216-228.

Brysbaert, M., Lange, M., & Van Wijnendaele, I. (2000). The effects of age-of-acquisition and frequency-of-occurrence in visual word recognition: further evidence from the Dutch language. *European Journal of Cognitive Psychology, 12,* 65-85.

Brysbaert, M., & New, B. (2009a). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41,* 977-990.

Brysbaert, M., & New, B. (2009b). *SUBTLEXusExcel2007.xlsx*. Retrieved January 30, 2010 from Behavior Research Methods Supplementary Material: http://brm.psychonomic-journals.org/content/supplemental.

Brysbaert, M., van Dyck, G., & van de Poel, M. (1999). Visual word recognition in bilinguals: Evidence from masked phonological priming. *Journal of*

Experimental Psychology: Human Perception and Performance, 25, 137-148.

Buchta, C., & Hahsler, M. (2009). *cba: Clustering for Business Analytics*. R package version 0.2-6.

Burgess, C., & Livesay, K. (1998). The effect of corpus size in predicting reaction time in a basic word recognition task: Moving on from Kučera and Francis. *Behavior Research Methods, Instruments, & Computers, 30*, 272-277.

Butterworth, B. (1983). Lexical representation. In B. Butterworth (Ed.), Language production (Vol. 1, pp. 257–294). San Diego, CA: Academic Press.

Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition, 28,* 297-332.

Carreiras, M., Ferrand, L., Grainger, J., & Perea, M. (2005). Sequential effects of phonological priming in visual word recognition. *Psychological Science, 16*, 585-589.

Chauncey, K, Holcomb, P. J., & Grainger, J. (2008). Effects of stimulus font and size on masked repetition priming: An event-related potentials (ERP) investigation. *Language and Cognitive Processes, 23*, 183-200.

Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement, 7,* 249-254.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). Psyscope: A interactive graphic system for designing and controlling experiments in the psychology laboratory using macintosh computers. *Behavioral Methods, Research, Instruments, and Computers, 25*, 257-271.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). *Access to the internal lexicon*. In S. Dornic (Ed.), Attention and performance VI (pp. 535-555). Hillsdale, NJ: Erlbaum.

Coulmas, F. (2003). *Writing systems: An introduction to their linguistic analysis*. Cambridge: Cambridge University Press.

Cutler, A. (1981). Making up materials is a confounded nuisance, or: will we be able to run any psycholinguistic experiments at all in 1990? *Cognition, 10*, 65-70.

Davis, C., Sánchez-Casas, R., García-Albea, J., Guasch, M., Molero, M. & Ferré, P. (2010). Masked translation priming: Varying language experience and word type with Spanish–English bilinguals. *Bilingualism: Language and Cognition, 13*, 137-155.

de Groot, A. M. B., Borgwaldt, S., Bos, M., & van den Eijnden, E. (2002). Lexical decision and word naming in bilinguals: Language effects and task effects. *Journal of Memory and Language, 47*, 91-124.

de Groot, A. M. B., & Nas, G. L. J. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language, 30*, 90-123.

de Vaan, L., Schreuder, R., & Baayen, R. H. (2007). Regular morphologically complex neologisms leave detectable traces in the mental lexicon. *The Mental Lexicon, 2,* 1-23.

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.

Diependaele, K., Sandra, D., & Grainger, J. (2005). Masked cross-modal morphological priming: Unravelling morpho-orthographic and morpho-semantic influences in early word recognition. *Language and Cognitive Processes, 20*, 75-114.

Dijkstra, T., de Bruijn, E., Schriefers, H., & ten Brinke, S. (2000). More on interlingual homograph recognition: Language intermixing versus explicitness of instruction. *Bilingualism: Language and Cognition, 3,* 69-78.

Dijkstra, T., Frauenfelder, U. H., & Schreuder, R. (1993). Grapheme-phoneme interaction in a bimodal detection task. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 1-20.

Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and language, 41,* 496-518.

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappeli, M., & Baayen, R. H. (2010). How cross-language similarity affects cognate recognition. *Journal of Memory and Language, 62*, 284-301.

Dijkstra, T., Moscoso del Prado Martin, F., Schulpen, B., Baayen, R. H., & Schreuder, R. (2005). A roommate in cream? Morphological family size effects on interlingual homograph recognition. *Language and Cognitive Processes. Special issue on morphology, 20*, 7-41.

Dijkstra, T., & van Hell, J. G. (2003). Testing the language mode hypothesis using trilinguals. *International Journal of Bilingual Education and Bilingualism, 6,* 2-16.

Dijkstra, T., & van Heuven, W. J. B. (1998). *The BIA model and bilingual word recognition*. In J. Grainger & A. Jacobs (Eds.), Localist connectionist approaches to human cognition (pp. 189–225). Hillsdale NJ: Erlbaum.

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5*, 175–197.

Dijkstra, T., van Heuven, W. J. B., & Grainger, J. (1998). Simulating cross-language competition with the bilingual interactive activation model. *Psychologica Belgica, 38*, 177-196.

Dijkstra, T., van Jaarsveld, H., & ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition, 1,* 51-66.

Ding, G., Peng, D., & Taft, M. (2004). The nature of the mental representation of radicals in Chinese: A priming study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 530-539.

Duñabeitia, J. A., Perea, M., & Carreiras, M. (2010). Masked translation priming effects with highly proficient simultaneous bilinguals. *Experimental Psychology, 57,* 98-107.

Duyck, W. (2005). Translation and associative priming with cross-lingual pseudohomophones: Evidence for nonselective phonological activation in bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 1340-1359.

Duyck,W., & Brysbaert, M. (2004). Forward and backward number translation requires conceptual mediation in both balanced and unbalanced bilinguals. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 889–906.

Duyck, W., Diependaele, K., Drieghe, D., & Brysbaert, M. (2004). The size of the cross-lingual masked phonological priming effect does not depend on second language proficiency. *Experimental Psychology, 51*, 116-124.

Duyck, W., van Assche, E., Drieghe, D., & Hartsuiker, R. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 663-679.

Duyck, W., Vanderelst, D., Desmet, T., & Hartsuiker, R. J. (2008). The frequency effect in second-language visual word recognition. *Psychonomic Bulletin and Review, 15*, 850-855.

Escudero, P., R. Hayes-Harb, & H. Mitterer. 2008. Novel second-language words and asymmetric lexical access. *Journal of Phonetics, 36*, 2, 345-360.

Falissard, B. (2007). *psy: Various procedures used in psychometry*. R package version 0.7.

Farid, M., & Grainger, J. (1996). How initial fixation position influences visual word recognition: A comparison of French and Arabic. *Brain and Language, 53*, 351-368.

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

Feldman, L. B., & Siok, W. W. T. (1999). Semantic radicals contribute to the visual identification of Chinese characters. *Journal of Memory and Language, 40*. 559-576.

Feldman, L. B., O'Connor, P. A., & Moscoso del Prado Martín, F. (2009). Early morphological processing is morpho-semantic and not simply morpho-orthographic: A violation of form-then-meaning accounts of word recognition. *Psychonomic Bulletin & Review, 16*, 684-691.

Ferrand, L., & Grainger, J. (1992). Phonology and orthography in visual word recognition: Evidence from masked nonword priming. *Quarterly Journal of Experimental Psychology, 45A*, 353–372.

Ferrand, L., & Grainger, J. (1994). Effects of orthography are independent of phonology in masked form priming. *Quarterly Journal of Experimental Psychology, 47A,* 365–382.

Ferrand, L., & Grainger, J. (2003). Homophonic interference effects in visual word recognition. *Quarterly Journal of Experimental Psychology, 56A*, 403-419.

Flores d'Arcais, G. B., & Saito, H. (1993). Lexical decomposition of complex kanji characters in Japanese readers. *Psychological Research, 55*, 52-63.

Flores d'Arcais, G. B., Saito, H., & Kawakami, M. (1995). Phonological and semantic activation in reading kanji characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 34-42.

Forster, K. I. (1998). The pros and cons of masked priming. *Journal of Psycholinguistic Research, 27,* 203-233.

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

Forster, K. I., & Taft, M. (1994). Bodies, antibodies, and neighborhood-density effects in masked form priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 844-863.

Frauenfelder, U. H., & Schreuder, R. (1992). *Constraining psycholinguistic models morphological processing and representation: The role of productivity*, In G. E. Booij van Marle (Ed.), Yearbook of morphology (pp. 165-185). Dordrecht: Foris.

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

Frost, R., Grainger, J., & Rastle, K. (2005). Current issues in morphological processing: An introduction. *Language and Cognitive Processes, 20*, 1-5.

Gerard, L. D., & Scarborough, D. L. (1989). Language-specific access of homographs by bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 305-315.

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

Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 1122-1139.

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

Gooskens, C., & Heeringa, W. (2004). Perceptive evaluation of Levenshtein dialect distance measurements using Norwegian dialect data. *Language variation and change, 16,* 189-208.

Goral, M., Libben, G., Obler, L. K., Jarema, J., & Ohayon, K. (2008). Lexical attrition in younger and older bilingual adults. *Clinical Linguistics & Phonetics, 22,* 509-522.

Grainger, J., & Ferrand, L. (1996). Masked Orthographic and Phonological Priming in Visual Word Recognition and Naming: Cross-Task Comparisons. *Journal of Memory and Language, 35*, 623-647.

Grainger, J., & Frenck-Mestre, C. (1998). Masked priming by translation equivalents in proficient bilinguals. *Language and Cognitive Processes, 13,* 601-623.

Grainger, J., Rey, A., & Dufau, S. (2008). Letter perception: From pixels to pandemonium. *Trends in Cognitive Sciences, 12*, 381-387.

Greene, M. R., & Oliva, A. (2009). Recognition of natural scenes from global properties: Seeing the forest without representing the trees. *Cognitive Psychology, 58*, 137-176.

Haigh, C. A., & Jared, D. (2007). The activation of phonological representations by bilinguals while reading silently: Evidence from interlingual homophones. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 623-644.

Havelka, J., & Tomita, I. (2006). Age of acquisition in naming Japanese words. *Visual Cognition, 13*, 981-991.

Hirose, H. (1992). Using the priming paradigm to investigate word recognition for kanji compound words. *The Japanese Journal of Psychology, 63*, 303-309.

Hoshino, N., & Kroll, J. F. (2008). Cognate effects in picture naming: Does crosslanguage activation survive a change of script? *Cognition, 106*, 501-511.

Hothorn, T., Buehlmann, P., Dudoit, S., Molinaro, A., & Van Der Laan, M. (2006). Survival ensembles. *Biostatistics, 7*, 355-373.

Hsu, C. H., Tsai, J. L., Lee, C., Y., & Tzeng, O. J. L. (2009). Orthographic combinability and phonological consistency effects in reading Chinese phonograms: An event-related potential study. *Brain and Language, 108,* 56-66.

Huang, H. W., Lee, C. Y., Tsai, J. L., Hung, D. L., & Tzeng, O. J. L. (2006). Orthographic neighborhood effects in reading Chinese two-character words. *NeuroReport, 17,* 1061-1065.

Huang, J. T., & Wang, M. Y. (1992). *From unit to Gestalt: Perceptual dynamics in recognizing Chinese characters*. In H. C. Chen & O. J. L Tzeng (Eds.), Language processing in Chinese (pp. 3-35). Amsterdam: North-Holland.

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.

Hyönä, J., Bertram, R., & Pollatsek, A. (2004). Are long compound words identified serially via their constituents? Evidence from an eye-movement-contingent display change study. *Memory and Cognition, 32*, 523-532.

Hyönä, J., & Pollatsek, A. (1998). Reading Finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception & Performance, 24,* 1612-1627.

Janssen, N., Bi, Y., & Caramazza, A. (2008). A tale of two frequencies: Determining the speed of lexical access for Mandarin Chinese and English compounds. *Language and Cognitive Processes, 23*, 1191-1223.

Japanese Ministry of Education, Culture, Sports, Science and Technology. (2009). *Joyo kanji hyo* [a list of joyo kanji characters]. Retrieved February 18, 2009 from http://www.mext.go.jp/b_menu/hakusho/nc/k19811001001/k19811001001.html

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word Frequency Effects in Speech Production: Retrieval of Syntactic Information and of Phonological Form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 824-843.

Ji, H., & Gagné, C. L. (2007). Lexical and relational influences on the processing of Chinese modifier-noun compounds. *The Mental Lexicon, 2:3,* 387-417.

Joseph, R. M., & Tanaka, J. (2003). Holistic and part-based face recognition in children with autism. *Journal of Child Psychology and Psychiatry, 44*, 529-542.

Joyce, T. (2002). Constituent-morpheme priming: Implications from the morphology of two-kanji compound words. *Japanese Psychological Research, 44,* 79-90.

Joyce, T. (2004). *Modeling the Japanese mental lexicon: Morphological, orthographic and phonological consideration*. In S. P. Shohov (Ed.). Advances in Psychological Research: Volume 31. (pp. 27-61). Hauppauge, NY: Nova Science.

Joyce, T., & Ohta, N. (2002). Constituent morpheme frequency data for the morphology of two-kanji compound words. *Tsukuba Psychological Research, 24,* 111-141.

Juhasz, B. J. (2005). Age-of-Acquisition effects in word and picture identification. *Psychological Bulletin, 131,* 684-712.

Kaakinen, J. K., & Hyönä, J. (2010). Task effects on eye movements during reading. *Journal of Experimental Psychology: Learning, Memory & Cognition, 36,* 1561-1566.

Kageyama, T. (2010). Variation between endocentric and exocentric word structures. *Lingua, 120,* 2405-2423.

Kahneman, D., Treisman, A., & Gibbs, B. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24,* 175-219.

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.

Kawakami, M. (2002). Effects of neighborhood size and kanji character frequency on lexical decision of Japanese kanji compound words. *The Japanese Journal of Psychology, 73,* 346-351.

Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research, 45,* 153-168.

Kerkhofs, R., Dijkstra, T., Chwilla, D. J., & de Bruijn, E. R. A. (2006). Testing a model for bilingual semantic priming with interlingual homographs: RT and N400 effects. *Brain Research, 1068,* 170-183

Kess, J. F., & Miyamoto, T. (1999). *The Japanese mental lexicon: Psycholinguistic studies of kana and kanji processing*. Philadelphia: John Benjamins.

Kim, J., & Davis, C. (2003). Task effects in masked cross-script translation and phonological priming. *Journal of Memory and Language, 49,* 484-499.

Kindaichi, H. (1999). *Personal katakana words dictionary*. Tokyo : Gakushu Kenkyusha

Kinoshita, S., & Norris, D. (2010). Masked priming effect reflects evidence accumulated by the prime. *The Quarterly Journal of Experimental Psychology, 63,* 194-204.

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General, 135,* 12-35.

Kohnert, K. (2004). Cognitive and cognate-based treatments for bilingual aphasia: A case study. *Brain and Language, 91,* 294–302.

Kondo, T., & Amano, S. (2001). *Score distribution of the reading ability test for kanji words "100-Rakan": What can be measured*. Proceedings of the 65th Annual Meeting of the Japanese Psychological Association (pp. 489). Tsukuba, Japan: Japanese Psychological Association.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33,* 149-174.

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

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.

Kuperman, V., & Van Dyke, J. A. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language, 65,* 42-73.

Lalor, E., & Kirsner, K. (2000). Cross-lingual transfer effects between English and Italian cognates and noncognates*. International Journal of Bilingualism, 4,* 385-398.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to Latent Semantic Analysis. *Discourse Processes, 25,* 259-284

Lee, C. Y., Tsai, J. L., Huang, H. W., Hung, D. L., & Tzeng, O. J. L. (2006). The temporal signatures of semantic and phonological activations for Chinese sublexical processing: An event-related potential study. *Brain Research, 1121,* 150-159.

Lemhöfer, K. M. L., & Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: Effects of code similarity in generalized lexical decision. *Memory and Cognition, 32,* 533-550.

Lemhöfer, K. M. L., Dijkstra, T., Schriefers, H. J., Baayen, R. H., Grainger, J. & Zwitserlood, C. M. E. (2008). Native language influences on word recognition in a second language: A mega-study. *Journal of Experimental Psychology A-Learning, Memory and Cognition, 34,* 12-31.

Leong, C. K., Cheng, P. W., & Mulcahy, R. (1987). Automatic processing of morphemic orthography by mature readers. *Language and Speech, 30,* 181-196.

Levelt, W. J. M. (1999). Models of word production. *Trends in Cognitive Science, 3*, 223-232.

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics Doklady, 10,* 707-710.

Libben, G. (1998). Semantic transparency in the processing of compounds: Consequences for representation, processing, and impairment. *Brain and Language, 61,* 30-44.

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., Gibson, M., Yoon, Y. B., & Sandra, D. (2003). Compound fracture: The role of semantic transparency and morphological headedness. *Brain and Language, 84,* 50–64.

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

Liu, Y., Shu, H., & Li, P. (2007). Word naming and psycholinguistic norms: Chinese. *Behavior Research Methods, 39,* 192-198.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers, 28,* 203-208.

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.

Marsolek, C. J. (2008). What antipriming reveals about priming. *Trends in Cognitive Sciences, 12,* 176-181.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*, 1-86.

McClelland, J. L., & Patterson, K., (2002a). 'Words or Rules' cannot exploit the regularity in exceptions. *Trends in Cognitive Sciences, 6,* 464-465.

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.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception. Part 1. An account of basic findings. *Psychological Review, 88,* 375–405.

McMullen, P. A., & Bryden, M. P. (1987). The effects of word imageability and frequency on hemispheric asymmetry in lexical decisions. *Brain and Language, 31,* 11-25.

Meara, P. (2005). *X_Lex: the Swansea Vocabulary Levels Test*. v2.05. Swansea: Lognostics.

Milin, P., Filipović Đurđević, D., & Baayen, R. H. (2008, October). *Prime and punishment*. Poster session presented at the 6th International Conference on Mental Lexicon, Banff, Canada.

Miller, G. A. (1990). WordNet: An on-line lexical database. International *Journal of Lexiography, 3,* 235-312.

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology, 14,* 247-279.

Ministry of Education, Culture, Sports, Science and Technology. (2009). *Joyo kanji hyo* [a list of kanji characters for everyday use]. Retrieved February 18, 2009 from http://www.mext.go.jp/b_menu/hakusho/nc/k19811001001/k19811001001.html

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

Miwa, K., Libben, G., Dijkstra, T., & Baayen, R. H. (2013). The time-course of lexical activation in Japanese morphographic word recognition: Evidence for a character-driven processing model. *Quarterly Journal of Experimental Psychology.* DOI: 10.1080/17470218.2013.790910

Miwa, K., Libben, G., Rice, S., & Baayen, R. H. (2008, July). *Lexical activations in picture comparison: A cross-linguistic approach to the relation between language and thought in the mental lexicon*. In Proceedings of the 6th International Conference of Cognitive Science, 230-232.

Morford, J. P., Wilkinson, E., Villwock, A., Piñar, P. & Kroll, J. F. (2011). When deaf signers read English: Do written words activate their sign translations? *Cognition, 118,* 286-292.

Morrison, C. M., & Ellis, A. W. (2000). Real age of acquisition effects in word naming and lexical decision. *British Journal of Psychology, 91,* 167-180.

Moscoso del Prado Martín, F., Bertram, R., Häikiö, T., Schreuder, R., & Baayen, R. H. (2004). Morphological family size in a morphologically rich language: The case of Finnish compared with Dutch and Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 1271–1278.

Myers, J., Huang, Y. C., & Wang, W. (2006). Frequency effects in the processing of Chinese inflection. *Journal of Memory and Language, 54,* 300-323.

Nakayama, E., Kiryu, R., & Yamaguchi, M. (2007). *Shinbun-ni miru kikangairaigo* [Basic loan words in newspapers]. The National Institute for

Japanese Language, 126. Retrieved February 10, 2011 from http://www.ninjal.ac.jp/gairaigo/Report126

Nakayama, M., Sears, C. R., Hino, Y., & Lupker, S. J. (in press). Cross-script Phonological Priming for Japanese-English Bilinguals: Evidence for Integrated Phonological Representations. *Language and Cognitive Processes.*

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9,* 353-383.

New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2006). Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin & Review, 13,* 45-52.

Nimura, I. (2002). *Kojien* (5th ed.). Tokyo: Iwanami Shoten.

Nisbett, R. E., & Masuda, T. (2003). Culture and point of view. *Proceedings of the National Academy of Sciences of the United States of America, 100*, 11163-11175.

O'Regan, J. K., & Jacobs, A. (1992). Optimal viewing position effects in word recognition: A challenge to current theory. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 185-197.

Otwinowska-Kasztelanic, A. (2009). Raising awareness of cognate vocabulary as a strategy in teaching English to Polish adults. *Innovation in Language Learning and Teaching, 3,* 131-147.

Pecher, D. (2001). Perception is a two-way junction: Feedback semantics in word recognition. *Psychonomic Bulletin & Review, 8,* 545-551.

Perea, M., & Carreiras, M. (2003). Sequential effects in the lexical decision task: The role of the item frequency of the previous trial. *The Quarterly Journal of Experimental Psychology, 56A,* 385-401.

Perea, M., Duñabeitia, J. A., & Carreiras, M. (2008). Masked associative/semantic priming effects across languages with highly proficient bilinguals. *Journal of Memory and Language, 58,* 916-930.

Perfetti, C. A., Zhang, S., & Berent, I. (1992). *Reading in English and Chinese: Evidence for a "universal" phonological principle*. In R. Frost & L. Katz (Eds.), Advances in Psychology: Orthography, phonology, morphology, and meaning (Vol. 94, pp. 227-248). Amsterdam: North-Holland.

Pexman, P. M., Lupker, S. J., and Jared, D. (2001). Homophone effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 139-56.

Pinker, S., & Ullman, M. T. (2002a). The past and future of the past tense. *Trends in Cognitive Sciences, 6,* 456-463.

Pinker, S., & Ullman, M. T. (2002b). Combination and structure, not gradedness, is the issue. *Trends in Cognitive Sciences, 6,* 472-474.

Plaut, D., & McClelland, J. (2010). Locating object knowledge in the brain: Comment on Bowers's (2009) attempt to revive the grandmother cell hypothesis. *Psychological Review, 117,* 284-288.

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.

Potter, M. C., So, K. F., von Eckhardt, B., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and more proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior, 23,* 23-38.

Prado, E. L., & Ullman, M. T. (2009). Can imageability help us draw the line between storage and composition? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 849–866.

Pynte, J., Kennedy, A., & Ducrot, S. (2004). The influence of parafoveal typographical errors on eye movements in reading. *European Journal of Cognitive Psychology, 16,* 178–202.

R Development Core Team (2009). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria, URL http://www.R-project.org.

R Development Core Team. (2011). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, URL http://www.R-project.org.

Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences. 24,* 574-590.

Rastle, K., Davis, M. H., Marslen-Wilson, W. D., & Tyler, L. K. (2000). Morphological and semantic effects in visual word recognition: A time-course study. *Language and Cognitive Processes, 15,* 507–537.

Rastle, K., Davis, M. H., & New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin and Review, 11,* 1090-1098.

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC Nonword Database. *Quarterly Journal of Experimental Psychology, 55A,* 1339-1362.

Rastle, K., McCormick, S. F., Bayliss, L., & Davis, C. J. (2011). Orthography influences the perception and production of speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 1588-1594.

Rayner, K. Slattery, T. J., & Belanger, N. N. (2010). Eye movements, the perceptual span, and reading speed. *Psychonomic Bulletin & Review, 17,* 834-839.

Rodriguez-Fornells, A., Rotte, M., Heinze, H. J., Nösselt, T., & Münte, T. F. (2002). Brain potential and functional MRI evidence for how to handle two languages with one brain. *Nature, 415,* 1026-1029.

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

Saito, H. (1997). *Shintekijisho* [Mental lexicon]. In Y. Ohtsu & T. Gunji (Eds.), Linguistic Sciences: Vol. 3 (pp. 93–153). Tokyo: Iwanami Shoten.

Saito, H., Kawakami, M., & Masuda, H. (1995). Frequency of semantic and phonetic components of radical types in complex left-right Kanji. *Johobunka Kenkyu* [Studies in Informatics and Sciences], *1*, 113-134.

Saito, H., Kawakami, M., & Masuda, H. (1995). Variety of phonetic components of radical types in complex left-right kanji. *Johobunka kenkyu* [Studies in Informatics and Sciences], *2,* 89-115.

Saito, H., Kawakami, M., & Masuda, H. (1997). Frequency of components of radical types in complex (top-bottom) kanji. *Johobunka kenkyu* [Studies in Informatics and Sciences], *6*, 115-130.

Saito, H., Masuda, H., & Kawakami, M. (1998). Form and sound similarity effects in kanji recognition. *Reading and Writing: An Interdisciplinary Journal, 10*, 323-357.

Sánchez-Casas, R., & García-Albea, J. E. (2005). *The representation of cognate and noncognate words in bilingual memory: Can cognate status be characterized as special kind of morphological relation?* In J. F. Kroll & A. M. B. de Groot (Eds.), Handbook of Bilingualism: Psycholinguistic approaches (pp. 226-250). Oxford: Oxford University Press.

Sanseido Henshujo. (1994). *Concise dictionary of katakana words*. Tokyo: Sanseido.

Scarborough, D. L., Gerard, L., & Cortese, C. (1984). Independence of lexical access in bilingual word recognition. *Journal of Verbal Learning and Verbal Behavior, 23,* 84-99.

Schepens, J., Dijkstra, T., & Grootjen, F. (2011). Distribution of cognates in Europe as based on Levenshtein distance. *Bilingualism: Language and Cognition*. DOI: 10.1017/S1366728910000623.

Schmidt, R. A. (1982). *More on motor programs*. In J. A. S. Kelso (Ed.), Human motor behavior: An introduction (pp. 189-217). Hillsdale, NJ: Erlbaum.

Schreuder, R. & Baayen, R. H. (1997). How complex simplex words can be. *Journal of Memory and Language, 37,* 118-139.

Schwartz, A. I., Kroll, J. F., & Diaz, M. (2007). Reading words in Spanish and English: Mapping orthography to phonology in two languages. *Language and Cognitive Processes, 22,* 106-129.

Shen, D., & Forster, K. I. (1999). Masked phonological priming in reading Chinese words depends on the task. *Language and Cognitive Processes, 14,* 429-459.

Strobl, C., Boulesteix, A. L., Kneib, T., Augustin, T., & Zeileis, A. (2008). Conditional variable importance for random forests. *BMC Bioinformatics, 9.* URL http://www.biomedcentral.com/1471-2105/9/307.

Strobl, C., Boulesteix, A. L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics, 8.* URL http://www.biomedcentral.com/1471-2105/8/25.

Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: Rationale, application and characteristics of classification and regression trees, bagging and random forests. *Psychological Methods, 14,* 323-348.

Taft, M. (1994). Interactive activation as a framework for understanding morphological processing. *Language and Cognitive Processes, 9,* 271-294.

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

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior, 14,* 638-647.

Taft, M., & Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of verbal learning and verbal behavior, 15,* 607-620.

Taft, M., & Kougious, P. (2004). The processing of morpheme-like units in monomorphemic words. *Brain and Language, 90,* 9-16.

Taft, M., & Nguyen-Hoan, M. (2010). A sticky stick? The locus of morphological representation in the lexicon. *Language and Cognitive Processes, 25,* 277-296.

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

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

Tamaoka, K. (2005). The effect of morphemic homophony on the processing of Japanese two-kanji compound words. *Reading and Writing, 18,* 281-302.

Tamaoka, K. (2007). Rebounding activation caused by lexical homophony in the processing of Japanese two-kanji compound words. *Reading and Writing, 20,* 413-439.

Tamaoka, K. & Hatsuzuka, M. (1995). The effects of kanji printed-frequency on processing Japanese two-morpheme compound words. *The Science of Reading, 39,* 121-137.

Tamaoka, K. & Hatsuzuka, M. (1998). The effects of morphological semantics on the processing of Japanese two-kanji compound words. *Reading and Writing: An Interdisciplinary Journal, 10,* 293-322.

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.

Tanaka, J. W., Kiefer, M., & Bukach, C. M. (2003). A holistic account of the own-race effect in face recognition: evidence from a cross-cultural study. *Cognition, 93,* B1-B9.

Tamaoka, K., Kirsner, K., Yanase, Y., Miyaoka, Y., & Kawakami, M. (2002). A web-accessible database of characteristics of the 1,945 basic Japanese kanji. *Behavior Research Methods, Instruments, & Computers, 34,* 260-275.

Thierry, G., & Wu, Y. J. (2004). Electrophysiological evidence for language interference in late bilinguals. *NeuroReport, 15,* 1555-1558.

Thierry, G., & Wu, Y. J. (2007). *Brain potential reveal unconscious translation during foreign-language comprehension.* Proceedings of the National Academy of Sciences of the United States of America, 104, 12530-12535.

Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12,* 97-136.

Tremblay, A., Derwing, B., Libben, G., & Westbury, C. (2011). Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Language Learning, 61,* 569-613.

van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., & Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. *Journal of Memory and Language, 64,* 88-107.

van Assche, E., Duyck, W., Hartsuiker, R. I., & Diependaele, K. (2009). Does bilingualism change native-language reading? *Psychological Science, 20,* 923-927.

van Hell, J. G., & De Groot, A. M. B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition, 1,* 193-211.

van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review, 9,* 780-789.

van Heuven, W. (2008). The Interactive Activation Model (jIAM). Retrieved September 28, 2012, from http://www.psychology.nottingham.ac.uk/staff/wvh/jiam/.

van Heuven, W. J. B., & Dijkstra, T. (2010). Language comprehension in the bilingual brain: fMRI and ERP support for psycholinguistic models. *Brain Research Reviews, 64,* 104-22.

van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language, 39,* 458-483.

Venables, W. N. & Ripley, B. D. (2002). *Modern Applied Statistics with S*. Fourth Edition. Springer, New York.

Vergara-Martínez, M., Duñabeitia, J. A., Laka, I., & Carreiras, M. (2009). ERP correlates of inhibitory and facilitative effects of constituent frequency in compound word reading. *Brain Research, 1257*, 53-64.

Vitu, F., O'Regan, J. K., & Mittau, M. (1990). Optimal landing position in reading isolated words and continuous text. *Perception & Psychophysics, 47,* 583-600.

Voga, M., & Grainger, J. (2007). Cognate status and cross-script translation priming. *Memory and Cognition, 35,* 938-952.

Wachsmuth, E., Oram, M. W., & Perrett, D. I. (1994). Recognition of objects and their component parts: Responses of single units in the temporal cortex of the macaque. *Cerebral Cortex, 4,* 509-522.

Whorf, B. L. (1956). Science and linguistics. In J. B. Carroll (Ed.). *Language, thought and reality: Selected writings of Benjamin Lee Whorf*. Cambirdge, MA: MIT Press.

Wu, Y. J., & Thierry, G. (2010). Chinese-English bilinguals reading English hear Chinese. *The Journal of Neuroscience, 30,* 7646-7651.

Yan, M., Kliegl, R., Richter, E. M., Nuthmann, A., & Shu, H. (2010). Flexible saccade-target selection in Chinese reading. *The Quarterly Journal of Experimental Psychology, 63,* 705-725.

Yap, M. J., Balota, D. A., Sibley, D. E., & Ratcliff, R. (2011). Individual differences in visual word recognition: Insights from the English lexicon project. *Journal of Experimental Psychology: Human Perception and Performance.* doi: 10.1037/a0024177.

Yarkoni, T., Balota, D. A., & Yap, M. J. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity, *Psychonomic Bulletin & Review, 15,* 971-979.

Yokosawa, K., & Umeda, M. (1988). *Processes in human kanji-word recognition*. In J. Xinsong (Ed.), Proceedings of the 1988 IEEE international conference on systems, man, and cybernetics (pp. 377-380), New York: International Academic Publishers.

Yokoyama, S., Sasahara, H., Nozaki, H., & Long, E. (1998). *Shinbun denshi media no kanji: Asahi shinbun CD-ROM niyoru kanji hindo hyoo* [Japanese kanji characters in the newspaper media: Kanji printed-frequency index of Asahi newspaper on CD-ROM]. Tokyo: Sanseido.

Zhang, B., & Peng, D. (1992). *Decomposed storage in the Chinese lexicon*. In H. C. Chen & O. J. L Tzeng (Eds.), Language processing in Chinese (pp. 131-149). Amsterdam: North-Holland.

Zhang, T., Van Heuven, W. J. B., & Conklin, K. (2011). Fast automatic translation and morphological decomposition in Chinese-English bilinguals. *Psychological Science, 22,* 1237-1242.

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.