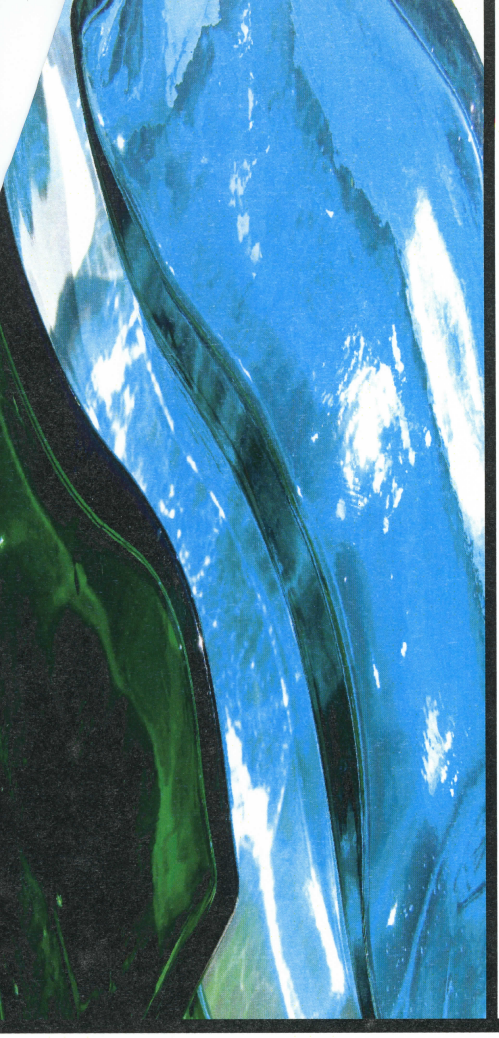


Nova Science Publishers, Inc.
400 Oser Avenue, Suite 1600, Hauppauge, N. Y. 11788-3619
Phone (631) 231-7269 * Fax (631) 231-8175
E-mail: Novascience@earthlink.net * Novasci1@aol.com
http://www.novapublishers.com

Terry Joyce

Modeling the Japanese Mental Lexicon: Morphological, Orthographic and Phonological Consideration

In: "Advances in Psychology Research"
Volume 31
Serge P. Shohov (Editor)
ISBN: 1-59033-985-1
2004

In: *Advances in Psychology Research, Volume 31*
Editor: Serge P. Shohov, pp. 27-61

ISBN 1-59033-985-1

Chapter 2

MODELING THE JAPANESE MENTAL LEXICON: MORPHOLOGICAL, ORTHOGRAPHIC AND PHONOLOGICAL CONSIDERATIONS

Terry Joyce
Research Institute for Languages and Cultures of Asia and Africa,
Tokyo University of Foreign Studies, Japan

ABSTRACT

Being heavily influenced by aspects of the Japanese writing system—often described as the most complicated system of writing ever created—the Japanese mental lexicon presents special insights and interesting challenges for our understanding of the organization of lexical representations and the mechanisms of lexical retrieval. This paper discusses important morphological, orthographic and phonological issues related to the Japanese language and the complexity of its writing system in the context of the Japanese lemma unit model advocated by Joyce (1999, 2002a). A special feature of this model is the incorporation of lemma unit representations as connections or way-stations mediating the links between both orthographic and phonological access representations and semantic representations.

Part 1 of the paper outlines the Japanese lemma unit model and how it accounts for the diversity in the morphology of two-kanji compound words. Focusing on orthographic and phonological considerations, Part 2 discusses the model's potential to handle, on the one hand, orthographic issues arising from the multi-script nature of the Japanese writing system and, on the other hand, phonological issues due to its dual-reading system. Part 2 presents a series of two cross-script and two cross-modal versions of constituent-morpheme priming experiments conducted to investigate these aspects. Finally, the paper reflects briefly on the role of the lemma units as interfacing or mediating links in connectionist modeling to differentiate components of cognitive processing.

INTRODUCTION

Descriptions of the Japanese writing system by scholars of scripts are remarkably consistent in their use of superlatives and in their emphasis on the complexity of the system. For instance, Coulmas (1989) has observed that the Japanese writing system "is often said to be the most intricate and complicated writing system ever used by a sizeable population" (p. 122), while DeFrancis (1989), rather more bluntly, remarks that the Japanese "ended up with one of the worst overall systems of writing ever created" (p. 138). More recently, Fischer (2001) comments that the mixture of scripts which are "written together following arbitrary rules perhaps embody the most complicated form of writing ever devised" (p. 167) and Sproat (2000) notes that "Japanese is surely the most complex modern writing system, and the hardest to force into any taxonomic mold" (p. 132).[1]

Setting aside questions about whether these descriptions are totally justified,[2] the Japanese writing system is undeniably complex. The first reason for this complexity is the multi-script nature of the system, employing in largely complementary ways morphographic *kanji*,[3] two sets of syllabic *kana* symbols (*hiragana*, used mainly for inflectional grammatical elements, and *katakana*, employed primarily in writing foreign words and names), Arabic numerals and increasingly the Latin alphabet. The second reason is the dual system of *on-readings* (borrowed Sino-Japanese pronunciations) and *kun-readings* (Native Japanese pronunciations) for kanji. However, as Yamada (1967) astutely points out few writing systems "are as fitting subjects for the development of a theory of characters" (p. 705). As echoed in Sproat's (2000) reference to the problems and debate concerning its classification, the Japanese writing system undoubtedly presents special insights and challenges for our understanding of writing and of how writing systems function. Shifting the perspective from the external form to the internal representation, as the nature of a writing system will be reflected to some degree in the structure of the literate mental lexicon, one may claim that the Japanese mental lexicon presents important insights and challenges for our understanding of the organization of lexical representations and the mechanisms of lexical retrieval in the human mental lexicon.

This paper is about modeling the Japanese mental lexicon, and is about important morphological, orthographic and phonological issues related to the language and the complexity of the Japanese writing system. While reference is made to a number of hypotheses and models, these issues are primarily examined in the context of the Japanese lemma unit model advocated by Joyce (1999, 2002a), which is a modified adaptation of a Chinese model, proposed by Taft, Liu and Zhu (1999) within the multi-level interactive-activation framework (Taft, 1991, 1994). A special feature of this model is the incorporation

[1] It should be noted, for a more balanced perspective, that not all regard the complexity negatively. Backhouse (1984), for example, remarks that the mixture of scripts "makes for a potential flexibility of orthography on a scale that is inconceivable in the case of more familiar writing systems" (p. 220).

[2] The fact that the Japanese writing system is capable of meeting the functional needs of a modern, literate society (Smith, 1996) seems to suggest that the difficulties might be somewhat overstated, although the question of just what functional literacy might mean for Japanese clearly warrants more research than it has received so far.

[3] It is extremely unfortunate that so many scholars of writing systems continue to use the term logographic, even while acknowledging the term morphographic is more precise (Coulmas, 1996; Daniels, 1996, 2001). As Joyce (2002b) argues, this revision is not merely a matter of emphasis, because the shift in focus that this terminological revision requires is of fundamental significance for our understanding of kanji in the Japanese writing system, writing systems in general, the mental lexicon, and even language itself.

of lemma unit representations as connections or way-stations mediating the links between both orthographic and phonological access representations and semantic representations.

A fundamental issue for models of the Japanese mental lexicon is to account for the diversity in the morphology of two-kanji compound words.[4] After briefly illustrating this morphological diversity and noting its significance for modeling the Japanese mental lexicon, Part 1 of this paper outlines the Japanese lemma unit model. Part 1 also describes two constituent-morpheme priming experiments (Joyce, 1999, 2002a) that compared, in terms of handling this morphological diversity, the lemma unit model with Hirose's (1992, 1994, 1996) hypotheses about the representation and retrieval of two-kanji compounds. Part 2 on orthographic and phonological considerations discusses the possibilities provided by the incorporation of lemma units to handle, on the one hand, orthographic issues arising from the multi-script nature of the Japanese system and, on the other hand, phonological issues due to its dual-reading system. Specifically, Part 2 presents a series of two cross-script and two cross-modal versions of the constituent-morpheme priming experiments conducted to investigate these aspects of the Japanese lemma unit model. In the final part, the paper reflects briefly on the nature of lemma unit representations and their potential to function as interfacing or mediating links within connectionist models seeking to further differentiate the components of cognitive processing.

1. MORPHOLOGICAL CONSIDERATIONS

1.1 The Morphology of Two-Kanji Compound Words

The combination of extensive lexical borrowing from Chinese and productive native word-formation processes has resulted in a rich diversity in the morphological structure of compound words. One interesting consequence of this, for instance, is the co-existence of Sino-Japanese and Native Japanese compound words of almost identical meaning, such as the Sino-Japanese word 登山 /tozan/ and the Native Japanese word 山登り /yamanobori/, both combinations of the morphemes 登 'climb' and 山 'mountain,' meaning 'mountain climbing.'[5]

While classifications inevitably differ to some degree both in their scope and level of detail, about nine main principles or relationships are generally accepted as underlying the formation of two-kanji compound words (see Nomura, 1988; also, Kageyama, 1982; Tamamura, 1985; Ozaki, Todome, Nishioka, Yamada, & Yamada, 1992; Shibatani, 1990). This paper, however, concentrates on only five of these principles; the three syntactic

[4] As an agglutinative language, Japanese also possesses a rich inflectional morphology. However, the morphology of two-kanji compound words is singled out for special attention because it is the most common word structure in the Japanese language (Nomura, 1988). According to one estimate, two-kanji compounds account for 70 percent of Japanese words (Yokosawa & Umeda, 1988). Given this, and the fact that many of the same structures are also seen in other polymorphemic words, clearly the matter is one of special concern for models of the Japanese mental lexicon.

[5] While not without problems (Nomura, 1988), it is common to classify the Japanese lexicon according to origin, and in the case of two-kanji compound words to distinguish between Native Japanese and Sino-Japanese words. Largely reflecting this etymological distinction, Sino-Japanese compounds are pronounced according to the on-readings of the constituent kanji—Chinese pronunciations borrowed into Japanese together with the kanji—whereas Native Japanese compounds are generally pronounced according to the kun-readings—native Japanese readings applied to kanji as translation glosses.

principles of modifier + modified, verb + complement, and complement + verb, and the two semantic relationships of associative pairs and synonymous pairs[6] (Joyce & Ohta, 1999), examples of which are shown in Table 1.

Table 1. Examples of Word-Formation Principles for Two-Kanji Compound Words

Compound words	Pronunciation	Element meanings	Compound meaning
Modifier + Modified			
山桜	/yamazakura/	'mountain' + 'cherry'	mountain cherry
高値	/takane/	'high' + 'cost'	high cost
Verb + Complement			
登山	/tozan/	'climb' + 'mountain'	mountain climbing
殺人	/satsujin/	'kill' + 'person'	murder
Complement + Verb			
外食	/gaishoku/	'outside' + 'eat'	eating out
毒殺	/dokusatsu/	'poison' + 'kill'	kill by poison
Associative Pairs			
親子	/oyako/	'parent' + 'child'	parent and child
生死	/seishi/	'live' + 'die'	life and death
Synonymous Pairs			
山岳	/sangaku/	'mountain' and 'mountain'	mountains
変化	/henka/	'change' + 'change'	change

Modifier + modified compounds have a right-hand head structure, conforming to the syntax of modification with modifiers proceeding modified nouns (Kageyama, 1982). The majority of these compounds are noun + noun compounds, as in 山 'mountain' modifying 桜 'cherry' in 山桜 /yamazakura/ 'mountain cherry.' However, lexicalized examples of adjectival and verbal modification with inflectional endings omitted are also extremely common (Tamamura, 1985). Compounds of the verb + complement pattern are Sino-Japanese, being of Chinese origin rather than native to Japanese. There are semantic restrictions with this left-hand-head pattern, for, as Kageyama (1982) observes, there is an absence of indirect objects and transitive subjects as complements, with the majority of these compounds being combinations of verb + direct object, such as 登 'climb' + 山 'mountain' in 登山 /tozan/ 'mountain climbing.' In contrast, Sino-Japanese compound words with the reversed structure of complement + verb involve examples with subject + predicate, indirect object and adverb complements (Ozaki, Todome, Nishioka, Yamada, & Yamada, 1992), such as 外 'outside' + 食 'eat' in 外食 /gaishoku/ 'eating out.' This order of elements is also a native Japanese syntactic structure, and, accordingly, there are many Native Japanese compound words with this kind of relationship. However, most of these native compounds words are not two-kanji compound words, in the strictest sense, being normally written with two kanji plus a hiragana character for the nominal inflection ending of the verb, for example, as with 山登り /yamanobori/ 'mountain climbing.' A characteristic of semantic compounds is that the elements are of the same word class, and these can be divided into either associative pairs or synonymous pairs depending on the relationship between the elements. In the case of

[6] The selection of these five principles is influenced by a number of factors. While the other principles of affixation, repetition, abbreviation, and phonetic borrowing would undoubtedly also provide data of interest, they are, for varying reasons, less suitable for the constituent-morpheme priming paradigm employed in the experiments discussed here. Stemming directly from interest in the organization of compound words in the mental lexicon, another factor was the desire to include contrasting patterns, such as the reversed syntactic patterns of verb + complement and complement + verb. In this respect, these principles break somewhat with Nomura's (1988) classification, where these syntactic patterns are relegated under a basic pattern of complements.

associative pair compounds, the conjunction of the elements may be either inclusive, such as 親 'parent' + 子 'child' in 親子 /oyako/ 'parent and child,' or exclusive, with many examples where the elements are of opposite meanings. In contrast, synonymous pairs involve elements that have similar meanings, such as 山 'mountain' and 岳 'mountain' combined in 山岳 /sangaku/ 'mountains.'

Given that models of the Japanese mental lexicon must capture in some way the morphological relations between related words, for example, such as 登山 and 山登り as well as numerous other compound words sharing the morphemes 登 and 山, the diversity in compound-word structure undoubtedly poses important questions about how the representations of compound words might be stored and linked together, and, reflecting this arrangement, about the nature of lexical retrieval.

1.2 Modeling Morphological Knowledge

As an important part of our linguistic knowledge, the representation of morphological information relating to the structure of polymorphemic words is a fundamental issue for all models of the mental lexicon (Feldman, 1995; Jarema, Kehayia, & Libben, 1999; Sandra & Taft, 1994; Taft, 1991). This is clearly true not only because of the vast numbers of polymorphemic words that exist in all languages and the relative ease with which language users produce and understand both existing and novel forms (Sandra, 1994), but also because the issue has important implications for the organization of lexical representations and for lexical processing. The form and organization of lexical representations and the mechanisms of lexical retrieval are mutually defining aspects of the mental lexicon, so notions about lexical representations—ranging from full listings (e.g., Rubin, Becker, & Freeman, 1979; Butterworth, 1983) to decomposed storage (e.g., Taft & Forster, 1975, 1976)—will directly determine the mechanisms assumed to underlie lexical retrieval—whether search (e.g., Forster, 1976; 1989) or activation (e.g., McClelland, & Rumelhart, 1981), as well as the extent of morphological involvement—as separate parsing routes (e.g., Caramazza, Laudanna, & Romani, 1988) or as intermediate-level units (e.g., Taft, 1994).

Mental lexicon research into the nature of lexical representations and retrieval generally takes the form of singling out one of these aspects for investigation in some experiments and in comparing alternative accounts in terms of their plausibility and their compatibility with the results obtained. In this part of the paper on morphological considerations in modeling the Japanese mental lexicon, the main focus is on the contrast between search and activation as mechanisms of lexical retrieval. The alternative proposals, which will be outlined next, are Hirose's (1992, 1994, 1996) hypotheses, which evoke search mechanisms, and the Japanese lemma unit model advocated by Joyce (1999, 2002a), which assumes activation as the mechanism underlying lexical retrieval. And, the experimental results on which these two proposals are compared comes from two constituent-morpheme priming experiments that controlled for the word-formation principle of the two-kanji compound word targets, which are discussed shortly below.

1.3 Two Proposals as Models of the Japanese Mental Lexicon

Hirose's (1992, 1994, 1996) studies deserve our attention because they are among the first to look beyond the single kanji level and to consider the organization of lexical representations for two-kanji compound words. Hirose does not actually propose a detailed model, but he does present clear hypotheses about the 'serial search' of 'clustered' arrangements of representations for two-kanji compound words, based on the results of a series of constituent-morpheme priming experiments using the lexical decision task.

Although Hirose found significant priming in both constituent prime conditions compared to an unrelated prime condition, because reaction times were significantly faster in the first-constituent condition than in the second-constituent condition, he interpreted this result as evidence of a serial (from left to right) processing of the compound word targets. In another experiment comparing priming from the first-constituent when it was the first kanji of many compound words to when it was the first kanji of few compound words, facilitation was greater in the few-compounds condition. Based on these results, Hirose suggested that the mental lexicon for two-kanji compound words is structured so that words sharing the same first kanji are linked in clusters, with the first kanji serving as a retrieval cue, but that words sharing the same kanji as a second element are not.

In contrast to the serial search mechanism hypothesized by Hirose (1992, 1994, 1996), a number of models assuming activation have also been proposed for the Japanese mental lexicon, such as the companion-activation model (CAM) advocated by Saito (1997) (see also Saito, Masuda, & Kawakami, 1998) and an interactive-activation model suggested by Tamaoka and Hatsuzuka (1998).[7] However, those models also face the problems associated with representational redundancy, homographs, and varying degrees of semantic transparency that motivated Taft, Liu and Zhu to make a recent modification to a Chinese version of the multilevel interactive-activation framework. Accordingly, the activation model discussed here is the Japanese lemma unit model proposed by Joyce (1999, 2002a), which is an adaptation of Taft et al's revised model for the Japanese mental lexicon.

In adapting the latest version of the Chinese model to the Japanese mental lexicon Joyce (1999, 2002a) has proposed modifications, in order to accommodate unique aspects of the Japanese writing system, which are discussed in Part 2. However, because the basic hierarchical structure is the same, it is useful to briefly trace the development of the Chinese model.

Essentially, the interactive-activation framework (McClelland, & Rumelhart, 1981; Rumelhart & McClelland, 1982) proposes a hierarchical structure of different levels of representational units, or nodes, linked by both facilitatory and inhibitory connections. Taft (1991, 1994) adopted this framework to explain the processing of polymorphemic words in English and, subsequently with colleagues, has expanded on this to account for the visual word recognition of Chinese words (Taft, Liu, & Zhu, 1999; Taft & Zhu, 1995, 1997a, 1997b). In extending the framework to Chinese, Taft and Zhu have argued that above the feature-level representations corresponding to the strokes of Chinese characters there are sub-lexical level representations for radicals (both semantic and phonetic markers) that are

[7] Although Tamaoka and Hatsuzuka (1998) incorporate what they refer to as syntactic feature representations, or lemmas, in their model, these do not actually replace compound-word level representations, but rather serve as a means of representing word-class information within the model.

positionally sensitive (1997a), and that above these there are Chinese character, or morpheme-level, representations (1995, 1997b). These earlier versions of the model also included two-character (polymorphemic) word level representations with corresponding phonological representations. However, in an important recent modification to the model Taft et al. (1999) have proposed the inclusion of lemma units, as abstract modality-free units mediating the links between orthographic, phonological, and semantic units. As Taft et al. (1999) point out, these lemma units, which emerge from the regular co-occurrence of orthographic forms with meaning, are similar to the 'concept nodes' in the model of morphological processing proposed by Schreuder and Baayen (1995).

This revision was motivated by two unsatisfactory aspects of the previous models. The first was the redundancy in having both orthographic and phonological representations at both the constituent and the compound word levels. This redundancy is eliminated with the lemma units, which mediate links between orthographic and phonological access representations. The second aspect was the problem of homographs, orthographic units associated with more than one meaning. Again, incorporating lemma units into the model can overcome this problem, for a homograph orthographic unit can be linked to separate meanings via different lemmas. The lemma units also provide an effective way to capture varying degrees of semantic transparency between polymorphemic words. As in Schreuder and Baayen's (1995) model of morphological processing, the lemma units function as a means of differentiating and addressing concepts. Accordingly, semantic transparency for a compound is defined as the degree of overlap between the sets of semantic primitives activated by the constituent elements and the set activated by the compound word itself. In this way, morphological relations are expressed in terms of the shared semantic and syntactic properties that underlie morphological families, rather than in spatial notions such as 'clusters'. Taft et al. (1999) draw on this notion in their account of differences in priming for semantically opaque and transparent compound words, suggesting that activation in the lemma units can be reset. When an opaque compound word is presented, because there is little or no overlap between the meaning of the compound word and the meanings of the elements, activation for irrelevant lemma units is reset to baseline levels. However, in the case of transparent compound words, activation is maintained which can serve to facilitate related words.

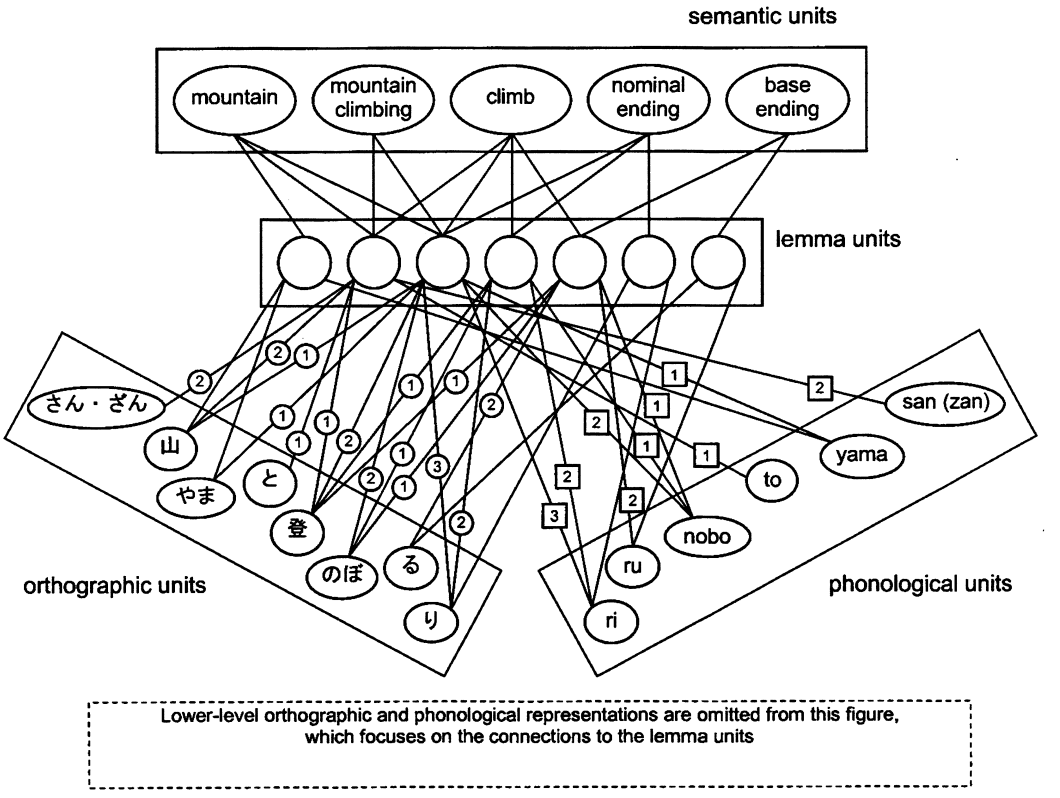


Figure 1. The Japanese lemma unit model (Joyce, 1999, 2002a): The numbers for some of the connections from orthographic and phonological units indicate the ordering of the constituents of polymorphemic words.

Given that the problems of representational redundancy, homographs, and varying degrees of semantic transparency are also concerns for models of the Japanese mental lexicon within the interactive-activation framework, such as with Saito's (1987) CAM model and Tamaoka and Hatsuzuka's (1998) model, the incorporation of lemma unit representations to handle these problems in the Japanese lemma unit model (Joyce 1999, 2002a), as illustrated in Figure 1, is clearly appealing. In the Japanese lemma unit model, lexical retrieval of a two-kanji compound word is seen as the result of activation passed to the lemma unit representing the compound word from the orthographic representation units for the constituent kanji characters. That is, activation of the lemma unit for a two-kanji compound word will depend on the combined activation from both the first-constituent and second-constituent. This activation does not involve serial processing, because activation is passed on from the constituent kanji in parallel.

Having briefly outlined two proposals for models of the Japanese mental lexicon, the next section reflects on the proposals' plausibility in accounting for the morphological diversity underlying two-kanji compound words and describes two constituent-morpheme priming experiments conducted to examine the differing predictions from the two proposals concerning patterns of priming after controlling for the word-formation principle of the two-kanji compound word target.

1.4 Constituent-Morpheme Priming (Joyce, 1999, 2002a)

As any model of the Japanese mental lexicon must in some way capture the morphological relations that exist between polymorphemic words, one may ask how well these two proposals—Hirose's (1992, 1994, 1996) hypotheses and the Japanese lemma unit model (Joyce, 1999, 2002a)—cope with the diversity inherent in the morphological structure of two-kanji compound words, described above.

From a storage perspective, the morphological diversity would seem to pose problems for Hirose's (1992) hypothesis that compound words are linked in clusters based on the first kanji. For instance, if clusters are only based on a shared first kanji, then semantically-related compounds sharing a common second kanji, such as modifier + modified compounds like 山桜 /yamazakura/ 'mountain cherry,' 葉桜 /hazakura/ 'cherry tree in leaf,' and 夜桜 /yozakua/ 'cherry bloosoms at night' will not be linked. Moreover, words from different semantic areas, such as 青空 /aozora/ 'blue sky,' 青物 /aomono/ 'greens, (green) vegetables,' and 青年 /seinen/ 'adolescence (lit. green + years),' would be linked simply because they share the same first character. Another problematic area for this hypothesis arises from the reversed syntactic structures of verb + complement and complement + verb. While Hirose does not explicitly discuss Native Japanese compound words like 山登り, presumably the clustering based on the first kanji applies to these compound words as well. However, this would lead to a strange situation where 登山 and 山登り would not be linked in the mental lexicon because of the reversed order of elements, even though they are almost identical in meaning. Even if native Japanese compound words are treated separately, there are still many semantically-related Sino-Japanese compound words, such as 殺人 /satsujin/ 'murder' ('kill' + 'person') and 毒殺 /dokusatsu/ 'poison' ('poison' + 'kill'), that would not be linked according to Hirose's hypothesis. On the other hand, the Japanese lemma-unit model has little problem with this morphological diversity. Instead of grouping morphologically related compound words in 'clusters,' the relations underlying morphological families are modeled by the connections between representations.

Hirose's (1992, 1994, 1996) hypotheses also appear to have problems in terms of lexical retrieval. It is slightly puzzling that Hirose's hypotheses are based on results from the constituent-morpheme priming paradigm. This is the same paradigm used by Monsell (1985) in a study involving both semantically transparent (e.g., *tightrope*) and opaque (e.g., *butterfly*) English compound words. In contrast to Hirose's pattern of facilitation, however, Monsell found similar levels of facilitation from the first and the second constituents. Although the Japanese lemma-unit model, where retrieval is based on an activation mechanism, predicts similar levels of priming from both constituent kanji, such a finding would be incompatible with the search mechanisms that Hirose has evoked, given their reliance on the first kanji as a retrieval cue.

As already noted, the storage of lexical information and the mechanism of lexical retrieval are mutually defining aspects of the mental lexicon. Given this and the concerns for the plausibility of Hirose's (1992, 1994, 1996) hypotheses in coping with the diversity in compound-word structure, Joyce (1999, 2002a) asked if the pattern of facilitation found by Hirose (1992) would be observed once compound-word structure is considered. As the two proposals make different predictions concerning constituent-morpheme priming, Joyce (1999, 2002a) examined the pattern of facilitation for two-kanji compound words in two priming

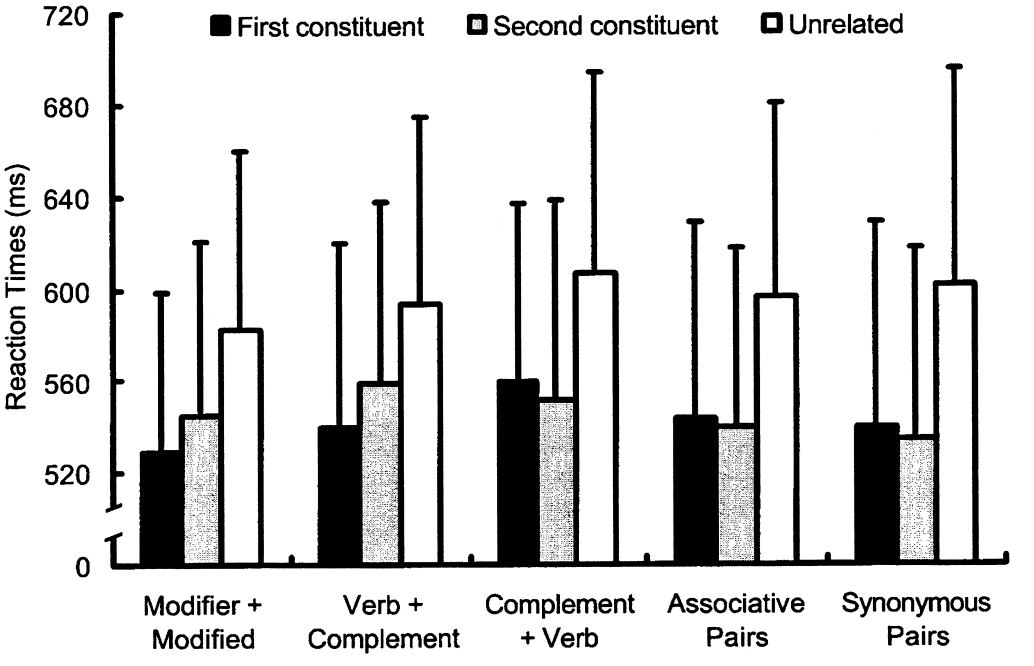
experiments that control for compound-word morphology by including five word-formation principles as experimental conditions.

The experimental design was identical in both experiments: a 3 x 5 two-factor (within-subject) design. The three prime conditions, based on relationship between primes and the two-kanji compound words, were first constituent, second constituent, and unrelated prime. The five word-formation conditions were modifier + modified, verb + complement, complement + verb, associative pair, and synonymous pair, described earlier.[8] Closely following the procedure in Hirose's (1992) study, Experiment 1 has a stimulus onset asynchrony (SOA) of 3000 ms. However, Tamaoka and Hatsuzuka (1998) have criticized Hirose's experiment for its very long SOA, claiming that at such a long delay participants may adopt strategies in their lexical-decision making. Accordingly, Experiment 2 employed a short SOA of 250 ms, which, based on Neely's (1977) seminal study on the distinction between automatic and expectancy-based processing, should be free from strategy-adoption.

The reaction time results for Joyce's (1999; 2002a) Experiment 1 and 2 are shown in Figure 2. In Joyce's Experiment 1, reaction times in both the first-constituent and the second-conditions were faster than the unrelated prime condition across all five word-formation conditions. Moreover, in four of the five word-formation conditions, there were no significant differences between the first-constituent and the second-constituent conditions. Only in the verb + complement condition was there a significant difference with the first-constituent being faster than the second-constituent condition. Indicating that the results of the first experiment were free from participant strategy adoption, the results of Experiment 2 not only closely matched the reaction times obtained in Experiment 1 but also provide a very similar pattern of priming. Reaction times in both the first-constituent and the second-conditions were again faster than the unrelated prime condition across all five word-formation conditions, with the only significant difference between the first-constituent and second-constituent conditions being in the verb + complement word-formation condition.

[8] The two-kanji compound word target stimuli were selected from a corpus of 1,000 two-kanji compounds (Joyce & Ohta, 1999), which was surveyed to obtain native Japanese speaker evaluations for familiarity and for the appropriateness of classifying the compound words according to a particular word-formation principle. Compound words with evaluation scores of 5.5 or over on a 7-point scale for both criteria were selected for the five word-formations principles in both experiments.

Although Joyce and Ohta (1999) did not refer to the notion of semantic transparency, because the classification survey asked Japanese speakers to discern the relationship between the constituents, target compound words with high classification scores may be regarded as being semantically transparent.



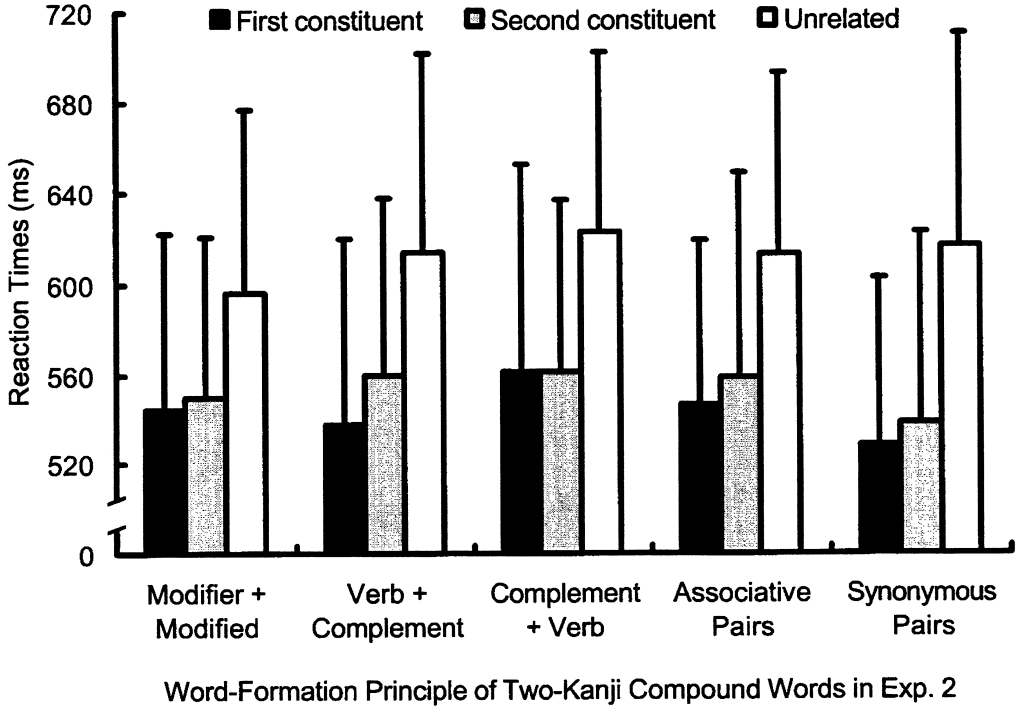


Figure 2. Mean Reaction Times (in Milliseconds) as a Function of Word Formation Principle by Prime Condition in Joyce's (1999, 2002a) Experiment 1 and Experiment 2

The results of Joyce's (1999, 2002a) study indicated that, across all five word-formation conditions, both related prime conditions facilitated responses to the target, and, in the majority of cases, at similar levels. These results are undoubtedly more consistent with the prediction for constituent-morpheme priming from the Japanese lemma unit than the pattern of priming obtained by Hirose (1992), suggesting that a model of the Japanese mental lexicon must be capable of accounting for similar levels of facilitation from both constituents. According to the lemma unit model, the lexical retrieval of a two-kanji compound word is achieved through the activation passing in parallel to the lemma unit representing the compound word from both the orthographic representation units for the constituent kanji characters. One can account for the priming found in these experiments by assuming that lingering activation in the lemma units linked to the orthographic unit for the prime gives them an advantage over other units that have not been activated. When the prime is presented once more as a constituent element of the compound word, this lingering activation leads to faster reactions times compared to the unrelated prime condition. Thus, Joyce's results represent important support for the plausibility of the lemma unit model as a model of the Japanese mental lexicon.

However, one aspect of Joyce's (1999, 2002a) results that requires further comment is the significantly faster reaction times for the first-constituent compared to the second-constituent in the verb + complement word-formation condition, as that was the only word-formation principle where the priming from the constituents was significantly different. Although Taft, Zhu and Peng (1999) argue that character representations are not positionally sensitive within Chinese compound words, the frequency with which a kanji will appear in a given position within a two-kanji compound word is likely to be closely related to its word class, for this will greatly determine the kind of word-formation principles in which the kanji will be involved. Given the central role of frequency as a regulating mechanism within the lemma-unit model, the influences of word class and compound word morphology cannot be ruled out. Lemma units function within the model as a means of differentiating and addressing concepts. Accordingly, morphological relations are expressed in terms of shared semantic and syntactic properties, which would include restrictions on the combinations of morphemes, such as verb and direct object, as well as information about word class and element order. Investigating the issue of positional sensitivity for verbal constituents, Joyce (2003a, 2003b) has recently reported results indicating an interaction between positional frequency of verbal constituents and word-formation principle. Based on constituent-morpheme frequency data (Joyce & Ohta, 2002), Joyce (2003a) calculated positional frequency ratios—representing how frequently a kanji character would appear in a particular position within two-kanji compound words—to contrast verbal constituents with low positional ratios (LPR: less than 20%) to those of high positional ratios (HPR: of over 80%) over the two word-formation conditions of verb + complement and the reversed order of complement + verb. The results showed that there were no significant differences between the first and second constituent condition in the two word-formation conditions with LPR verbs. That is, when the positional ratio was low, both constituents facilitated lexical decisions at similar levels. On the other hand, there were significant differences between the two constituent conditions in the two word-formation conditions with HPR verbs, with reaction times in the verb-constituent conditions being significantly faster than in the respective complement conditions. That is, while the first constituent condition was faster than the second constituent condition in the HPR-verb + complement word-formation condition, the pattern of priming was the opposite

in the reversed word-formation condition of complement + HPR-verb, where the reaction times in the second constituent condition were faster than in the first constituent condition. Seeking to overcome stimuli selection problems with this experimental design by splitting the contrasts over two separate experiments, Joyce (2003b) has replicated the findings for the HPR-verbs with the mirrored pattern of priming in the reversed word-formation conditions of HPR-verb + complement and complement + HPR-verb compound words, adding further support for the suggestion that the mental lexicon representations for verbal constituents with high positional ratios, at least, may be positionally sensitive with respect to two-kanji compound words.

2. Orthographic and Phonological Considerations

As explained in Part 1, the Japanese lemma unit model (Joyce, 1999; 2002a) represents an extremely promising approach to modeling the Japanese mental lexicon. The incorporation of lemma unit representations within the model not only provides an effective means of handling the problems of representation redundancy, homophones, and varying degrees of semantic transparency (Taft, Liu, & Zhu, 1999), but it also provides an appealing way to account for the rich diversity in the morphology of two-kanji compound words. Beyond these important features, however, the inclusion of the lemma units is essential for a model of the Japanese mental lexicon in order to account for the complex relationships between meaning, orthography and phonology that arise from the Japanese writing system's multi-script nature and its dual-reading system. In proposing the lemma unit model for the Japanese mental lexicon, Joyce (1999, 2002a) noted, as a theoretical extension to the functions of the lemma units, the potential of the model, on the one hand, to unify both kanji and kana processing within a single integrated model, and, on the other hand, to provide an elegant method of capturing the nuances in the use of on- and kun-readings. These orthographic and phonological considerations are the focus of this part of the paper, which presents a series of two cross-script and two cross-modal versions of the constituent-morpheme priming experiments conducted to explore these issues and to provide further empirical support for the Japanese lemma unit model.

2.1 Cross-Script Priming: Experiments 1 and 2

As Kess and Miyamoto (1999) observe, one of the most active areas of research within Japanese psycholinguistics has been the study of kanji and kana processing (see also, Chen, 1997; Chen & Zhou, 1999; Flores d'Arcais, 1992; Hatta, & Saito, 1999, 2000; Kaiho, & Nomura, 1983; Leong & Tamaoka, 1998; Paradis, Hagiwara, & Hildebrandt, 1985; Saito, 1997; Sugishita, Otomo, Kabe, & Yunoki, 1992; Tamaoka, 1991, 1994). As it would be beyond the scope of this paper to look at this literature in any detail, we can do little more than note a few studies that illustrate the important lesson that Kess and Miyamoto draw from their review of this rich body of research, which is that simple 'early dichotomies' between semantic routes and phonological routes are inadequate to explain the complex interaction involved in reading Japanese (pp. 196-197).

Some studies have provided evidence of similarities between kanji and kana processing, such as Besner and Hildebrandt (1987) who reported evidence of direct visual processing of kana and Wydell, Patterson, and Humphreys (1993) who have provided evidence that phonological information is also involved in kanji processing. On the other hand, other comparative studies have stressed functional differences between the two scripts, interpreting findings such as those of Shimamura (1987) of faster word comprehension with kanji words and faster word naming with kana as evidence of differences in processing. More recently, Shafiullah and Monsell (1999) suggest that decoding kana and kanji involves different resources based on their finding of switching costs in the results of a series of naming and semantic categorization task experiments where participants had to switch between reading words in kanji and kana at predictable intervals.[9] While the implications of such switching costs for the normal reading of mixed kanji and kana writing are not clear, still, the fact that words that are normally represented in kanji can still be retrieved when written in kana suggests that the two scripts function together as elements of an integrated system rather than as totally separate systems.

Although not addressed by existing multilevel interactive-activation models for Japanese, which focus on kanji processing (Saito, 1997; Saito, Masuda, & Kawakami, 1998; Tamaoka & Hatsuzuka, 1998), one important challenge for a model of the Japanese mental lexicon is to provide some explanation of how both kanji and kana map on to meaning; how, for example, the two orthographic representations of 山 and やま /yama/ are both linked to the meaning of 'mountain.' Joyce (1999, 2002a) suggests that this can be achieved simply within the Japanese lemma unit model by expanding the orthographic access representations to include both kanji and kana (and even alphabet) representations, with alternative representations linking to the same lemma unit.

Bowers and Michita (1998) present evidence that supports the notion that orthographic representations for kanji and hiragana both map onto a common abstract representation. Claiming that words in upper and lower case (e.g., *READ* and *read*) are believed to map onto a common orthographic code, despite their perceptual dissimilarity, Bowers and Michita sought to investigate whether the abstractions are at the letter or the word level by conducting cross-script long-term priming experiments with kanji and hiragana representations that do not share letter-to-letter correspondences. Bower and Michita's experiments had one study-test condition, which compared the form of presentation at the study stage (hiragana, kanji, spoken and non-studied) with the form of presentation at the test stage (Experiment 1, hiragana; Experiment 2, kanji) using the lexical decision task. The results of these experiments showed significant priming in the cross-script conditions (32 msec for kanji-hiragana condition in Exp. 1 and 17 msec for hiragana-kanji condition in Exp. 2) at similar levels to the repetition conditions (32 msec for hiragana-hiragana condition in Exp. 1 and 25 msec for kanji-kanji condition in Exp. 2). However, only minimal (nonsignificant) priming was found in the cross-modal conditions (10 msec for spoken-hiragana condition in Exp. 1 and 3 msec for spoken-kanji condition in Exp. 2). Bowers & Michita interpreted the cross-script priming as evidence that abstract orthographic word codes exist independently of abstract letter codes. They also interpreted the lack of cross-modal priming as evidence that

[9] Shafiullah and Monsell (1999) claim that the cost of switching between kanji and kana was robust, although it was not numerically large, with a mean reaction time difference of 13.1 +/- 1.8 ms between switch and non-switch trials over all of their experiments (p. 595).

semantic or phonological codes have relatively little influence in mediating long-term priming for visually presented words, supporting the suggestion that cross-script priming (and cross-case priming) is mediated by orthographic codes. Thus, Bowers and Michita's findings point to the presence of common abstract orthographic representations, at least at the word level.

The first purpose of the present cross-script priming studies is to provide evidence for the existence of connections from the kana representations of morphemes to the lemma units representing compound words normally represented orthographically as two-kanji compound words. Related to this, a second purpose of these studies was to further explore the implications of the Japanese writing system's dual-reading system for the lemma unit model by using as primes both on-readings, in Experiment 1, and kun-readings, in Experiment 2, for the constituents of two-kanji compound words. Within the Japanese lemma unit model, the different characteristics of on-readings and kun-readings may be explained in terms of the number of lemma units to which the readings are connected and by the strengths of these connections, where the weightings on the connections will be a product of frequency.

The number of connection that may exist for the kanji 犯 'commit [a sin, a crime],' which has an on-reading of /han/ and a kun-reading of /okasu/, may serve to illustrate something of the possible differences in terms of connections for on- and kun-readings. This kanji has a familiarity rating of 6.0 in the NTT database (Amano & Kondō, 1999) and is listed in order of frequency at 849 with 3,535 tokens in a recent survey of newspaper kanji frequencies (Yokoyama, Sasahara, Nozaki, & Long. 1998). According to recently compiled constituent morpheme frequency data for two-kanji compound words (Joyce & Ohta, 2002), this kanji appeared in newspapers as the first constituent of six two-kanji compounds words and as the second constituent of 16 compound words. While it would be unreasonable to assume that all these compound words are represented in the average mental lexicon, in addition to connections for its on-reading and its kun-reading, this kanji could have connections to at least 22 lemma units. Looking at the on-reading /han/, this is shared by at least 33 kanji, consisting of 21 Jōyō kanji (including 9 kyōiku kanji), 9 JIS level 1 kanji, and 3 JIS level 2 kanji.[10] Among these 33 kanji, the kanji 犯 has the sixth highest familiarity rating (Amano & Kondō, 1999). However, counting all the compound words that have the on-reading /han/ as part of their pronunciation in the constituent-morpheme frequency data (Joyce & Ohta, 2002), there is a total of 564, where the kanji sharing this reading are the first constituents of 308 compound words and the second constituents of 256 two-kanji compound words. Again, while it would be even more unreasonable to assume that all these words are represented in the average mental lexicon, still, the number of connections to lemma units for compound words contrasts sharply with the maximum possible 22 connections from the kanji 犯. Looking at the kun-reading /okasu/, there are seven words with this reading (with four of these being inflected forms of base verbs), of which 犯す has the highest familiarity score of 5.94 (Amano & Kondō, 1999).

[10] The numbers of kanji sharing on-readings and kun-readings mentioned here are the numbers of kanji and words provided for selection by an input conversion program for a personal computer (Natural Input 2002, Microsoft, Corp.).
Jōyō kanji are the 1,945 kanji for general usage introduced by the Japanese Cabinet in 1981. Kyōiku kanji refers to a subset of the Jōyō kanji consisting of the 1,006 kanji that are taught during the six years of elementary school. Although newspapers and official documents generally follow the Jōyō kanji guidelines, the total number of kanji in daily use is much higher because many kanji for family and place names are not covered in the list. The Japanese Industrial Standard (JIS) code (JIS X-0208-1990) which defines character codes for computers consists of two sets of kanji totally 6,355 kanji (2,965 level 1 and 3,390 level 2 kanji).

These facts suggest that the levels and patterns of facilitation in the two cross-script priming experiments presenting kana orthography primes associated with the constituent-morphemes of two-kanji compound words might be different from those found by Joyce (1999, 2002) with kanji orthography primes. In Experiment 1 with on-reading primes, while the kana orthography primes can be assumed to activate more lemma units than kanji orthography primes, because a given on-reading is likely to be shared by many kanji,[11] still the activated phonological information will match with part of the pronunciation for the compound word target. In Experiment 2 with kun-reading primes, on the one hand, although the kun-readings may only be shared by a few kanji at the most, there will be no phonological overlap between the kun-reading of the prime and the pronunciation of the two-kanji compound word target. Thus, the finding of priming in Experiment 2 would provide strong evidence of the lemma unit representations mediating the activation of semantically and morphologically related words. Accordingly, the purpose of the cross-script experiments is to investigate whether the on-readings (Exp. 1) and kun-readings (Exp. 2) for the constituent-morpheme kanji of two-kanji compound word presented in kana orthography will facilitate recognition of the two-kanji compound word. Such findings would support the existence of connections between kana representations and lemma units within the lemma unit model. For brevity, Experiments 1 and 2 are described together.

2.1.1 Method

Participants

Ninety native Japanese students (45 in Exp. 1 and 45 in Exp. 2, average age 21.3, SD = 3.9) of the University of Tsukuba participated in the experiments as volunteers.

Design

Both experiments had a 3 x 4 two-factor design, with both factors as within-subject variables. Although essentially the same design as that used in Joyce (1999, 2002a), there are two minor changes to note. The first change is in the prime conditions, where the unrelated kanji prime is replaced with an unprimed condition (a blank screen) as the baseline condition. This change was to eliminate the possibility of priming effects being inflated by inhibition from an unrelated kanji (for discussion of neutral baselines, see De Groot, Thomassen, & Hudson, 1982). The second change is in the number of word-formation principles. Although Joyce (1999, 2002a) used five principle conditions, associative pair compound words were not included in the series of experiments reported in this paper. The main reason for their exclusion is that too few compound words of this kind were of sufficiently high familiarity to match with the other principles.[12]

[11] There is considerable variation in the number of kanji sharing a given on-reading. Looking only at the 1,945 Jōyō kanji, although 61 kanji have on-readings that are not shared by other Jōyō kanji, at the other extreme, the on-reading /shō/ is shared by 65 kanji and the on-reading /kō/ by 64 kanji. The average number of kanji sharing a particular on-reading is 7.2, but the number is much higher if calculated for all the 6,355 JIS kanji.

[12] In Joyce (1999, 2002a), familiarity was controlled based on the results of an familiarity evaluation survey (Joyce and Ohta, 1999), but with the publication of the more reliable familiarity data within the NTT database (Amano & Kondō, 1999), that data was used to control for familiarity effects in this series of experiments. According to this data, few of the associative pair compounds surveyed by Joyce and Ohta (1999) were found to be of sufficiently high familiarity compared to the other principles, and so this principle was excluded.

Experiment 1 Materials

Fujii (1996) observes that there are 296 syllable string types as on-readings for the 1,945 Jōyō kanji. Moreover, 229 of these have a two mora structure, where the second syllable is from a small set of just seven syllables; い /i/, う /u/, き /ki/, く /ku/, ち /chi/, つ /tsu/, and ん /n/ (p. 100). Because of this rather limited variety in on-readings, and because many kanji share the same on-readings, it was not possible to make this experiment so that an on-reading represented in kana was only used once for one particular kanji. To control as far as possible for the effects of repeated presentations of the same on-reading primes to different kanji as constituents of two-kanji compound word targets, two measures were taken. The first measure was to ensure that the average number of kanji associated with a given on-reading was matched closely over the four word-formation principles. Thus, for each of the four word-formation principles, 30 compound words were selected with classification scores (Joyce & Ohta, 1999) and familiarity scores (Amano & Kondō, 1999) of 5.5 or more on a 7-point scale for both criteria, which were also closely matched over the conditions in terms of the average number of kanji per on-reading. The mean scores for both criteria over the four word-formation principles, together with the average number of kanji associated with an on-reading, are shown in Table 2. The second measure to limit the effects of repeated presentations of the same on-reading primes was to restrict the maximum number of kanji sharing an on-reading to two within a particular word-formation principle and to six over all four principles. The stimuli were divided into two sets, with a maximum of three kanji sharing an on-reading in each. For both sets, three presentation lists were prepared to counterbalance the stimulus items over the three prime conditions. Each participant was run on two presentation lists with a break between blocks, and by counterbalancing the combinations of all six presentation lists, which were randomized for each participant, each participant would, at most, only see a prime repeated twice, once in each block. For the lexical decision task, 72 non-word combinations of two kanji were selected from the set of items created by Joyce (1999, 2002a).

Table 2. Mean Classification Scores and Familiarity Scores (NTT database) for Target Compound Words and Mean Number of Kanji per On-reading Primes Used in Experiment 1 and Experiment 3

Principle	Classification score	Familiarity score	Number of kanji per on-reading prime
Modifier + Modified	6.62 (0.23)	5.93 (0.31)	18.23
Verb + Complement	6.72 (0.22)	5.93 (0.20)	18.33
Complement + Verb	6.57 (0.32)	5.93 (0.25)	18.35
Synonymous Pairs	6.38 (0.25)	5.93 (0.24)	18.25

Note: Both surveys used a 7-point scale. The figures in parenthesis are the standard deviations.

Experiment 2 Materials

The problem of repeated presentations of a prime was not an issue in Experiment 2, because it was possible to counterbalance such primes over the different prime conditions across the three presentation lists. However, in contrast to the one or two syllable on-readings, there was greater variation in the lengths of the kun-readings. Because this is related to word class, with kun-readings tending to be longer for adjectives and verbs than for nouns, it was

not possible to perfectly match for this, but the kun-readings for first and second constituents were matched as closely as possible, while maintaining controls for average familiarity and classification scores. For each of the four word-formation principles, 30 compound words were selected with classification scores (Joyce & Ohta, 1999) and familiarity scores (Amano & Kondō, 1999) of 5.5 or over on a 7-point scale for both criteria. The stimuli were also closely matched over the conditions in terms of the average length of the kun-readings. The mean scores for both criteria over the four word-formation principles, together with the average lengths of the kun-readings for the constituents, are shown in Table 3. For the lexical decision task, 72 non-word combinations of two kanji were selected from the set of items created by Joyce (1999, 2002a).

Table 3. Mean Classification Scores and Familiarity Scores (NTT database) for Target Compound Words and Mean Length of Kun-reading Primes as Function of Position Used in Experiment 2 and Experiment 4

Principle	Classification score	Familiarity score	Mean prime length	
			First	Second
Modifier + Modified	6.62 (0.23)	5.93 (0.28)	2.8	2.3
Verb + Complement	6.64 (0.37)	5.93 (0.24)	3.0	2.4
Complement + Verb	6.39 (0.44)	5.93 (0.26)	2.8	3.0
Synonymous Pairs	6.18 (0.31)	5.93 (0.31)	3.0	3.3

Note: Both surveys used a 7-point scale. The figures in parenthesis are the standard deviations. First = first constituent and second = second constituent.

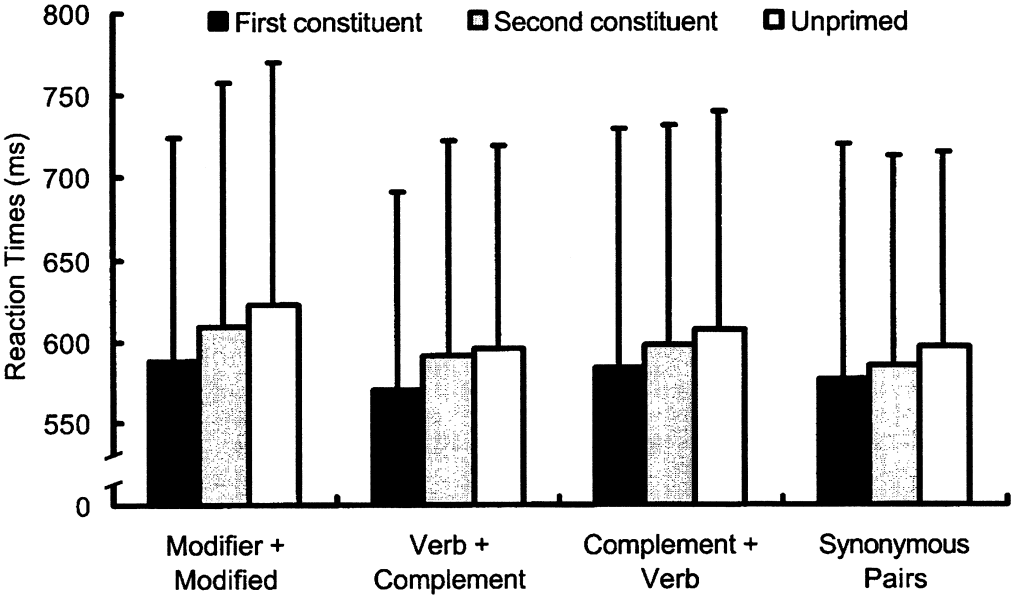
Apparatus

Super Lab Pro (version 2.0, Cedrus Corporation), running on a personal computer (Performance, Gateway), controlled the presentation of stimulus items and recorded lexical decisions collected via a response box (Cedrus Corporation, RB-600). Stimuli at a font size of 36 points were displayed on the computer screen at a viewing distance of approximately 50 cm.

Procedure

The presentation procedure for both experiments was almost identical; the primary difference was that on-reading primes were presented in Experiment 1 and kun-reading primes were presented in Experiment 2.

At the start of a trial, a plus symbol (+) appeared in the center of the screen as a fixation point for 250 ms. This was followed by the kana orthography prime displayed for 250 ms, and then by a string of asterisk-like symbols (※※※ in Exp. 1, ※※※※※ in Exp. 2) as a mask for 50 ms. The target stimulus item was then displayed and remained on the screen until the participant pressed a button on the response box for the lexical decision. There was a 1,500 ms inter-trial interval. Participants were instructed to press a green button for a compound word and a red button for a non-word as quickly and as accurately as possible. The instructions to the participants were printed on a card, which they were asked to read. After briefly repeating the instructions verbally to confirm that they had been understood, there was a practice session of ten trials, and both experiments took between 15 to 20 minutes to complete.



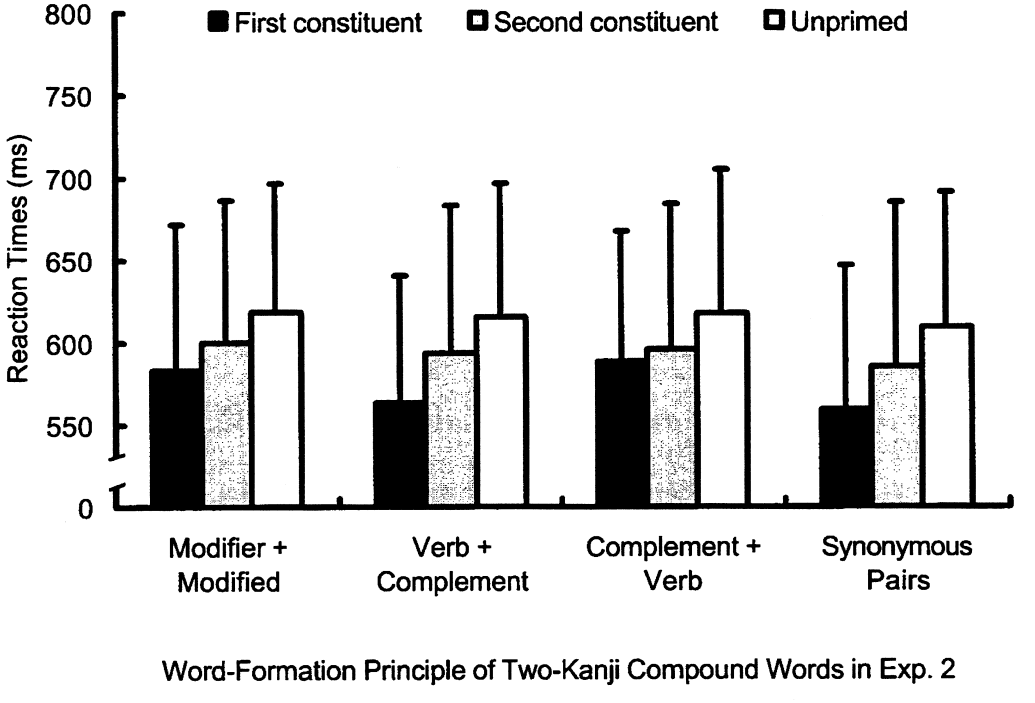


Figure 3. Mean Reaction Times (in Milliseconds) as a Function of Word Formation Principle by Prime Condition in Experiment 1 and Experiment 2

2.1.2 Results

Analyses of variance (ANOVAs) were carried out for the effects of prime and word-formation principle both by subject (both factors as within-subject variables) and item (prime as a within-subject and principle as a between-subject variable). Error responses were excluded from the analysis of reaction times. If the standard score for a reaction time was outside the range of +/- 2.5 for a given participant, the reaction time was adjusted to that equivalent to a standard score of +/- 2.5. As the error rates were extremely low (2.78% for Exp. 1 and 2.70% for Exp. 2), with even distributions of errors across the conditions, no error analyses are reported here. Reaction times for Experiment 1 and for Experiment 2 are shown in Figure 3.

Reaction Times Analysis

Experiment 1

In the subject analysis, there were significant main effects of prime, F_1 (2, 88) = 9.13, $p <$ .0001, and of word-formation principle, F_1 (3, 132) = 7.38, $p < .0001$, although there was no significant interaction, $p = .822$. In the item analysis, there was only a significant main effect of prime, F_2 (2, 232) = 8.51, $p < .0001$. Planned comparisons using Tukey's Honestly Significant Difference (HSD) test (all significant HSD differences reported in this paper are at the 0.05 level) for the main effect of prime in the subject analysis revealed that reaction times in the first constituent condition were significantly faster than both the second constituent condition and the unprimed condition for all word-formation principle conditions, with HSD = 14.79 in all cases. Planned comparisons for the main effect of prime in the item analysis revealed that reaction times in the first constituent condition were significantly faster than both the second constituent condition and the unprimed condition for all word-formation principle conditions, with HSD = 15.25 in all cases. Planned comparisons for the main effect of word-formation principle in the subject analysis revealed that reaction times in the modifier + modified condition were significantly slower than in the verb + complement and the synonymous pairs, with HSD = 13.56 in all cases.

Experiment 2

In the subject analysis, there were significant main effects of prime, F_1 (2, 88) = 31.80, $p < .0001$, and of word-formation principle, F_1 (3, 132) = 5.91, $p < .005$, although there was no significant interaction, $p = .189$. In the item analysis, there was only a significant main effect of prime, F_2 (2, 232) = 21.74, $p < .0001$. Planned comparisons for the main effect of prime in the subject analysis revealed that reaction times in both the first constituent and the second constituent conditions were significantly faster than in the unprimed condition for all conditions. However, reaction times in the first constituent were also significantly faster than in the second constituent condition for all word-formation conditions, with HSD = 12.64 in all cases. Similarly, planned comparisons for the main effect of prime in the item analysis revealed that reaction times in both the first constituent and the second constituent conditions were significantly faster than in the unprimed condition for all conditions. However, reaction times in the first constituent were also significantly faster than in the second constituent condition for all word-formation conditions, with HSD = 14.87 in all cases. Planned comparisons for the main effect of word-formation principle in the subject analysis revealed

that reaction times in the synonymous pairs were faster than in both the modifier + modified and the complement + verb conditions, with HSD = 11.79 in all cases.[13]

2.1.3 Discussion

In Experiment 1, although there were slightly faster reaction times in the second constituent condition, significant differences compared to the unprimed condition were only found in the first-constituent condition. This finding suggests that phonological information was activated by the kana orthography on-reading primes. This phonological information would be associated with many kanji and, from those kanji, to even more compound words as a component element of their pronunciation. However, the heightened levels of activation only facilitated lexical decision reactions to the target two-kanji compound words when the on-reading was shared by the first-constituent. This may be explained by the following assumptions. In the first-constituent condition, because the phonological information activated by the prime would be consistent with the initial part of the pronunciation for the two-kanji compound word, it is relatively easy for the participants to extend that pronunciation, when the phonological information is activated again in processing the two-kanji compound word target, by adding the pronunciation of the second-constituent that would also become available in processing the target compound word. However, in the second-constituent, the situation would be more complex. Here, the activated phonological information from the on-reading prime first needs to be mentally shifted to a word final position before inserting the on-reading for the first-constituent of the two-kanji compound word as it becomes available from the processing of the target stimulus item.[14]

In Experiment 2, there was significant priming in both the first constituent and second constituent conditions compared to the unprimed condition, although there were also significant differences between the first constituent and the second constituent conditions.[15] In contrast to Experiment 1 with on-reading primes, it is not possible to explain the priming in both constituent conditions in Experiment 2 in terms of phonological information, because there is no phonological overlap between the kun-reading primes of the constituent kanji of the two-kanji compound words and the compound-word targets. Rather, because relatively few kun-readings are shared by kanji, the presentation of kana orthography kun-reading primes would, in most cases, activate only the appropriate lemma unit representations, which would, in turn, activate relevant semantic representations and lemma units for morphologically-related compound words. The differences in the patterns of facilitation between Experiment 1 and Experiment 2 may, therefore, be taken as a reflection of the differences in the characteristics of on-readings and kun-readings, noted earlier. These

---

[13] Although the analyses of these experiments indicate effects of word-formation principle, the results in Joyce (1999, 2002a) may caution against thinking that word formation principle per se influences lexical retrieval. Even though effects of word-formation condition were also indicted in Joyce's (1999, 2002a) experiments, there was no consistent pattern over the two experiments (or even with the present experiments) which would be expected if word formation principle were indeed influencing retrieval.

[14] These results seem to offer a possible clue as to why Hirose (1992) obtained an advantage for the first constituent. In his Experiment 1, the participants responded to the prime kanji as well by naming it. The bias towards the activation of phonological information, which this would have evoked, may have been the cause of the first constituent advantage Hirose found, when the participants were processing the target compound words.

[15] Although this pattern of facilitation matches that of Hirose's (1992) Experiment 1, it should be remembered that it is extremely difficult to explain the facilitation in the second constituent condition compared to the baseline unprimed condition with his notions of retrieval cues and clustered representations based on a shared first kanji.

characteristics are modeled within the Japanese lemma unit model in terms of the differences in the number of connections to lemma units that exist for the two kinds of reading, where on-readings are shared by many more kanji than kun-readings, and by the strengths of these connections, which would be the product of frequency.

The presence of cross-script priming in both Experiment 1 with on-reading primes and in Experiment 2 with kun-reading primes provides evidence for the existence of connections from the kana representations of morphemes to the lemma units representing compound words normally represented orthographically as two-kanji compound words. Thus, the results of these two cross-script priming experiments are consistent with Joyce's (1999, 2002a) suggestion that the incorporation of lemma unit representations within the Japanese lemma unit model provides an effective way of integrating both kanji and kana processing within a single model of the Japanese model lexicon.

2.2 Cross-Modal Priming: Experiments 3 and 4

In addition to providing an interesting approach to unifying both kanji and kana processing within a single integrated model, as discussed in the previous section, the lemma unit model is capable of accounting for the complexity in terms of phonological representations due to the Japanese writing system's dual-reading system of on-readings and kun-readings. Continuing to investigate this aspect of the Japanese lemma unit model, this section describes two cross-modal versions of the cross-script experiments.

Although Bowers and Michita (1998) reported finding only non-significant priming in a long-term cross-modal condition, the cross-modal priming paradigm has been widely used in exploring the representation of morphologically related words (Allen & Badecker, 2002; Feldman & Laralee, 2001; Feldman, & Soltano, 1999; Frost, Deutsch, Gilboa, Tannenbaum, & Marslen-Wilson, 2000; Marslen-Wilson, Tyler, Waksler, & Older, 1994; Zhou & Marslen-Wilson, 2000). For instance, Marslen-Wilson, Tyler, Waksler, and Older, (1994) have investigated English derivative words, finding evidence for morphological decomposition of semantically transparent forms. More recently, Allen and Badecker (2002) have argued for a two-level model of the lexicon which distinguishes lemmas and lexemes, in their study of an asymmetry in priming for regular/irregular English verbs proposing form-based inhibition for orthographically similar verb forms (e.g., *gave-give*). Zhou and Marslen-Wilson (2000) present results of cross-modal studies for English and Chinese compound words which contrasted the degree of their semantic transparency. The framework for the lexical representation of compound words that Zhou and Marslen-Wilson propose—consisting of orthographic, phonological and semantic representations—is similar in its basic structure to the Japanese lemma unit model. Although their framework does not include lemma units to mediate connections, in common with the Japanese lemma model, the level of semantic transparency in compound words is seen as a reflection of the degree of semantic overlap between compounds and their constituents morphemes.

The purpose of the cross-modal experiments is to investigate whether the auditorily-presented on-readings (Experiment 3) and kun-readings (Experiment 4) for the constituent morphemes of two-kanji compound words will facilitate the recognition of the two-kanji compound words when presented visually. Such a finding would further support the existence

of connections between phonological access representations and lemma units within the lemma unit model.

2.2.1 Method

Participants

Ninety native Japanese students (45 in Exp. 3 and 45 in Exp. 4, average age 19.9, SD = 2.9) of the University of Tsukuba participated in the experiments as volunteers. None of the participants took part in the previous experiments.

Design

The 3 x 4 two-factor design, with both factors as within-subject variables, for the cross-modal experiments is identical to that of the cross-script experiments.

Experiment 3 Materials

The two-kanji compound word targets used in Experiment 3 are the same as those used in Experiment 1. The counterbalancing of the combination of the six sets of presentation lists was also the same, to ensure that each participant would, at most, only hear a prime repeated twice, once in each block of trials. Wave files for the auditory primes were created from the sound files in a database on the lexical properties of Japanese (Amano & Kondō, 1999). These files were edited to remove the silence both before and after the readings with sound editing software (Sound Forge XP, version 4.5c).[16] As the average duration of these sound files 342.9 ms (SD = 70.3, range 158 ms to 519 ms), a sound file with 340 ms of silence was prepared for the unprimed condition.

Experiment 4 Materials

The two-kanji compound word targets used in Experiment 4 are the same as those used in Experiment 2. In a similar fashion to Experiment 3, wave files were created for the auditory primes. As the average duration of these sound files 497 ms (SD = 131, range 220 ms to 808 ms), a sound file with 500 ms of silence was prepared for the unprimed condition.

Apparatus

The apparatus used in the cross-modal experiments is the same as that used in the cross-script experiments, except that the auditory primes were presented through headphones connected to the personal computer.

Procedure

The presentation procedure for the two experiments was almost identical; the primary difference was that on-reading primes were presented in Experiment 3 and kun-reading primes were presented in Experiment 4.

At the start of a trial, a plus symbol (+) appeared in the center of the screen as a fixation point for 250 ms. This remained on the screen as the prime was presented auditorily through headphones. Presentation of the target stimulus item at the center of the computer screen

[16] As there are no single sound files in the database for the on-readings /bo/, /da/, /getsu/, /gyo/, /hoku/, /min/, and /zetsu/, these readings were edited from a number of longer sound files. Ten native speakers were asked to listen to the edited versions and judge the most appropriate for the target on-reading.

immediately followed offset of the auditory prime. This remained on the screen until the participant pressed a button on the response box for the lexical decision. There was a 1,500 ms inter-trial interval. Participants were instructed to press a green button for a compound word and a red button for a non-word as quickly and as accurately as possible. The instructions to the participants were printed on a card, which they were asked to read. After briefly repeating the instructions verbally to confirm that they had been understood, there was practice session of ten trials, and the whole experiment took between 15 to 20 minutes to complete.

2.2.2 Results

ANOVAs were carried out for the effects of prime and word-formation principle both by subject (both factors as within-subject variables) and item (prime as a within-subject and principle as a between-subject variable). The treatment of error responses and of data points outside the range of a standard score +/- 2.5 was identical to that applied in the cross-script experiments. As the error rates were extremely low (2.63% for Exp. 3 and 3.13% for Exp. 4), with even distributions of errors across the conditions, no error analyses are reported here. Reaction times for Experiment 3 and for Experiment 4 are shown in Figure 4.

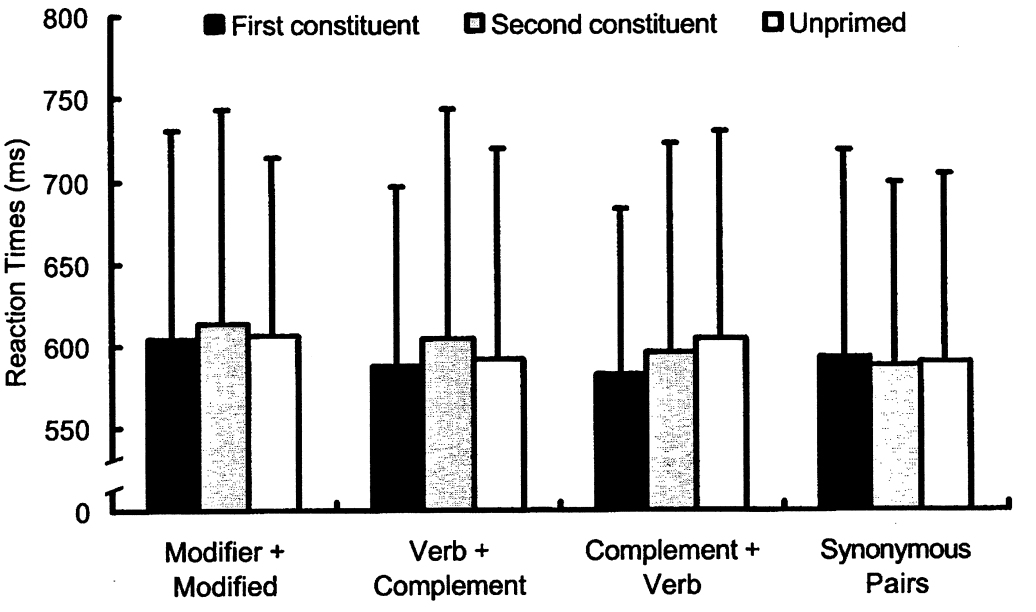
Reaction Times Analysis

Experiment 3

In the subject analysis, although there were no significant effects for prime, $p = .159$, there was a significant effect of word-formation principle, F_1 (3, 132) = 6.915, $p < .0001$, with no significant interaction. In the item analysis, no significant effects were found. Planned comparisons for the main effect of word-formation principle in the subject analysis revealed that reaction times in the modifier + modified condition were slower than in other conditions, with HSD = 10.9 in all cases.

Experiment 4

In the subject analysis, there were significant main effects of prime, F_1 (2, 88) = 9.22, $p < .0001$, and of word-formation principle, F_1 (3, 132) = 15.57, $p < .0001$, although there was no significant interaction, $p = .491$. In the item analysis, there was only a significant main effect of prime, F_2 (2, 232) = 5.96, $p < .005$. Planned comparisons for the main effect of prime in the subject analysis revealed that reaction times in both the first constituent and the second constituent conditions were significantly faster than in the unprimed condition for all conditions. However, there were no significant differences in reaction times between the first constituent and the constituent conditions, with HSD = 11.34 in all cases. Similarly, planned comparisons for the main effect of prime in the item analysis revealed that reaction times in both the first constituent and the second constituent conditions were significantly faster than in the unprimed condition for all conditions. However, there were no significant differences in reaction times between the first constituent and the constituent conditions, with HSD = 12.26 in all cases. Planned comparisons for the main effect of word-formation principle in the subject analysis revealed that reaction times in the synonymous pairs were faster than in all other conditions, with HSD = 10.98 in all cases.



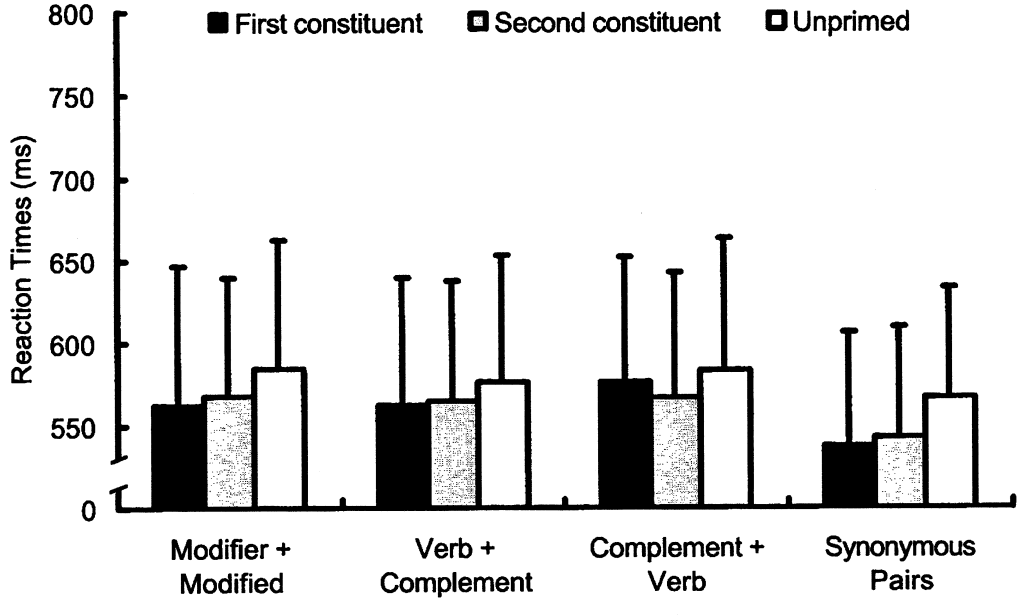


Figure 4. Mean Reaction Times (in Milliseconds) as a Function of Word Formation Principle by Prime Condition in Experiment 3 and Experiment 4

2.2.3 Discussion

In Experiment 3, no priming effects were found in either the first constituent or second constituent conditions. That is, the auditory presentation of an on-reading for a constituent morpheme of a two-kanji compound word is clearly not effective in priming the visual word recognition of the two-kanji compound word target. While this finding would seem to be consistent with Bowers & Michita's (1998) claim for modality-specificity of phonological information, the lack of priming in this experiment with on-reading primes should rather be interpreted in terms of the characteristics of on-readings—being shared by many kanji, the auditory presentation of the on-reading for the constituent morpheme of a two-kanji compound word target simply provides insufficient information to facilitate responses to visually-presented targets.

In Experiment 4, significant priming was found in both the first constituent and second constituent conditions compared to the unprimed condition; more consistent with the results reported by Joyce (1999, 2002a), the priming from the first-constituent and second-constituent conditions were at similar levels. Similar to the results of Experiment 2 with cross-script kun-readings, it is not possible to explain the priming in both constituent conditions in terms of phonological information, because there is no phonological overlap between the kun-reading primes of the constituent morphemes of the two-kanji compound words and the compound word targets. Again, because relatively few kun-readings are shared by kanji, the auditory presentation of the kun-readings would activate, in most cases, appropriate lemma unit representations. These representations would, in turn, pass on activation to relevant semantic representations and lemma units for morphologically-related compound words. The fact that priming was not found for the auditory presentation of on-readings but was found for both constituent conditions in cross-modal kun-reading experiment is consistent with the pattern of results for the two cross-script priming experiments. This further underscores the differences in the characteristics of on-readings and kun-readings.

The presence of cross-modal priming in Experiment 4 is consistent with the Japanese lemma unit model, where phonological access codes are assumed to link to lemma units.

3. MODELING CONSIDERATIONS

3.1 The Japanese Lemma Unit Model

As discussed in Part 1, the Japanese lemma unit model (Joyce, 1999, 2002a), incorporating lemma unit representations to mediate the connections between access representations and semantic representations, provides an attractive means of accounting for the diversity in the morphology of two-kanji compounds.

In proposing the lemma unit model, Joyce (1999, 2002a) argued that the inclusion of lemma units also provides a promising approach to handling, on the one hand, orthographic issues arising from the multi-script nature of the Japanese system and, on the other hand, phonological issues due to its dual-reading system. Part 2 of this paper has presented a series of cross-script and cross-modal versions of constituent-morpheme priming that provide support for the existence of links from kana orthography representations to lemma units and

for quantitative differences in the number of links to lemma units from on-reading and kun-reading phonological representations.

As noted earlier, the lemma units are connections or way-stations that develop when semantic information regularly co-occurs with form information (Schreuder & Baayen, 1995; Taft, Liu, & Zhu, 1999). The regular co-occurrence of the meaning 'mountain' with the orthographic forms 山 and やま and the phonological form /yama/ will lead to the development of a lemma unit to mediate these relationships. Similarly, the co-occurrences of the meaning 'mountain climbing,' with both the Native Japanese word 山登り /yamanobori/ and the Sino-Japanese word 登山 /tozan/ would lead to the development of two separate lemma units linking the meaning to the appropriate orthographic and phonological forms, as depicted in Figure 1.

When multiple units at the same level link to a unit at another level, it is necessary to specify the order of the links (Taft, Liu, & Zhu, 1999). Although this is indicated with numbered connections in Figure 1, the ordering of elements is seen as a function of the lemma units rather than an artifact of the connections.[17] Arguing that the central role of morphology is in computing meaning, Schreuder and Baayen (1995) posit lemma units (concept nodes) as part of a lexical representation interfacing between external access representation and internal syntactic and semantic representations, which specify combinatory properties, such as sub-categorization and word class. The lexical representation also includes mechanisms of licensing and composition, which operate on activated concept nodes. These mechanisms of licensing and composition appear to be related to the morphological awareness that Hatano (1995) refers to as compounding schemata—the implicit awareness of how kanji can be combined in compound words—acquired through learning to use the Japanese writing system. Discussing the Japanese writing system within the context of the practice account of literacy (Scribner & Cole, 1981), Hatano (1995) argues that there are cognitive consequences on modes of communication from acquiring and engaging in Japanese literacy practices. Specifically, he claims that literacy in Japanese includes the acquisition of compounding schemata, by which new words are made, and cognitive skills, that are required to solve homonymic ambiguity and to infer the meanings of unfamiliar words (Hatano, Kuhara, & Akiyama, 1981). In his discussions, however, Hatano (1995) suggested that experienced readers of Japanese have, in addition to the usual mental lexicon of words, a mental lexicon of kanji, or rather the corresponding morphemes, as building blocks for compound words. Although the notion of a separate lexicon for kanji is undoubtedly prompted by the bound nature of on-readings, the idea of two separate lexicons is far from appealing. It is also completely unnecessary, because the relationships between semantic representations for morphemes, orthographic representations for kanji, and phonological representations for both on- and kun-readings can be adequately explained through the mediation of lemma units.

3.2 Connectionist Modeling

Developing out of the multilevel interactive-activation framework that Taft (1991, 1994) has advocated and developed to account for the lexical retrieval of Chinese words (Taft, Liu,

& Zhu, 1999; Taft & Zhu, 1995, 1997a, 1997b; Taft, Zhu, Peng, 1999), the Japanese lemma unit model is a localist connectionist model. While not entering here into a discussion of the various merits and demerits of localist modeling over distributed connectionist modeling (see McClelland, Rumelhart, & the PDP Research Group, 1986; Rumelhart, McClelland, & the PDP Research Group, 1986; Seidenberg & McClelland, 1989; see also Page, 2000 for arguments in favor of the localist position),[18] in concluding this paper, a few comments would seem appropriate about the mediating or interfacing nature of the lemma unit representations.

While differing in conception, there are resemblances between the lemma unit representations and the layer of hidden units that Ijuin, Fushimi, Patterson, and Tatsumi (1999) have implemented in their connectionist model of word naming for two-kanji compound words. Not only do they both mediate the connections between orthographic and phonological access representations, but also they both deal with the dual-reading system in differentiating on- and kun-readings. Ijuin, et al. (1999) report that their simulations showed frequency and consistency effects, as well as an interaction between these, that are generally comparable to those observed by Fushimi, Ijuin, Patterson and Tatsumi (1999) for Japanese skilled readers. Inconsistencies, or irregularities, in the pronunciation of two-kanji compound words (Sakuma, Sasanuma, Tatsumi, & Masaki, 1998; Wydell, Patterson, & Humphreys, 1993) are often the result of hybrid compounding, where elements from different lexical stratums are combined. Given this fact, the kind of morphological information concerning allomorphic realizations as on-readings and kun-readings that the lemma units provides would seem to be important for accounting for the skilled Japanese reader's ability to cope with the inconsistencies in the pronunciation of some two-kanji compound words.

The Japanese lemma unit model has been primarily proposed as a model of visual word recognition that is capable of providing a satisfactory account of the lexical retrieval and representation of two-kanji compound words in the Japanese mental lexicon from a morphological perspective. Accordingly, issues relating to how the balance between morphological, semantic, and syntactic information may be implemented within the Japanese lemma unit model have not been discussed in detail. The notion of morphology that is incorporated in the Japanese lemma unit model comes to resemble Anderson's (1992) interpretation of morphology as the study of the relations between words rather than the study of the minimal signs combined in complex words, which in the model is realized by the links mediated by the lemma units. However, one major challenge for mental lexicon models seeking to account for the higher cognitive processes involved in comprehension is to implement more principled divisions of semantic and syntactic information.

As already noted, the lemma unit representations in the Japanese lemma unit model (Joyce, 1999, 2002a) and the concept nodes in the model of morphological processing proposed by Schreuder and Baayen (1995) are quite similar in function. Both are posited as mediating links, which interface between external access representations, and internal semantic and syntactic representations. The general form of a lexical representation presented

[17] However, it is not yet clear whether the lemma units can also serve as the locus for the positional sensitivity effects for verbal constituents with high positional ratios that Joyce (2003a, 2003b) reports.

[18] Seidenberg and McClelland (1989) comment that the notion of lexical access has no place in their distributed activation model modeling. While acknowledging their basic insight about the kinds of theoretical questions that attend to certain metaphors, and agreeing that the concept of parallel processing renders talk of routes and multiple lexicons rather meaningless, still basic theoretical questions concerning how visual words are recognized and how words are represented and structured remain to be answered whatever the approach.

by Schreuder and Baayen is also fairly similar to the notion of a lexical entry suggested by Levelt (1989). In his work on speech production, Levelt has depicted the make-up of a lexical entry as the four segments of a circle (p. 182 and p. 188); the lower half, referred to as the morpho-phonological form, consists of morphological and phonological information, while the upper half of the circle, which Levelt refers to as the lemma, consists of meaning and syntactic information. In contrast to Figure 1, the schematic representation in Figure 5 illustrates how syntactic information could also be expressly incorporated within the local connectionist framework of the Japanese lemma unit model. In this schematic representation, the four main groupings of information constituting a lexical entry—orthographic (kanji, kana, and alphabet), phonological (on- and kun-readings), semantic and syntactic—are linked by the lemma unit representations at the center.[19]

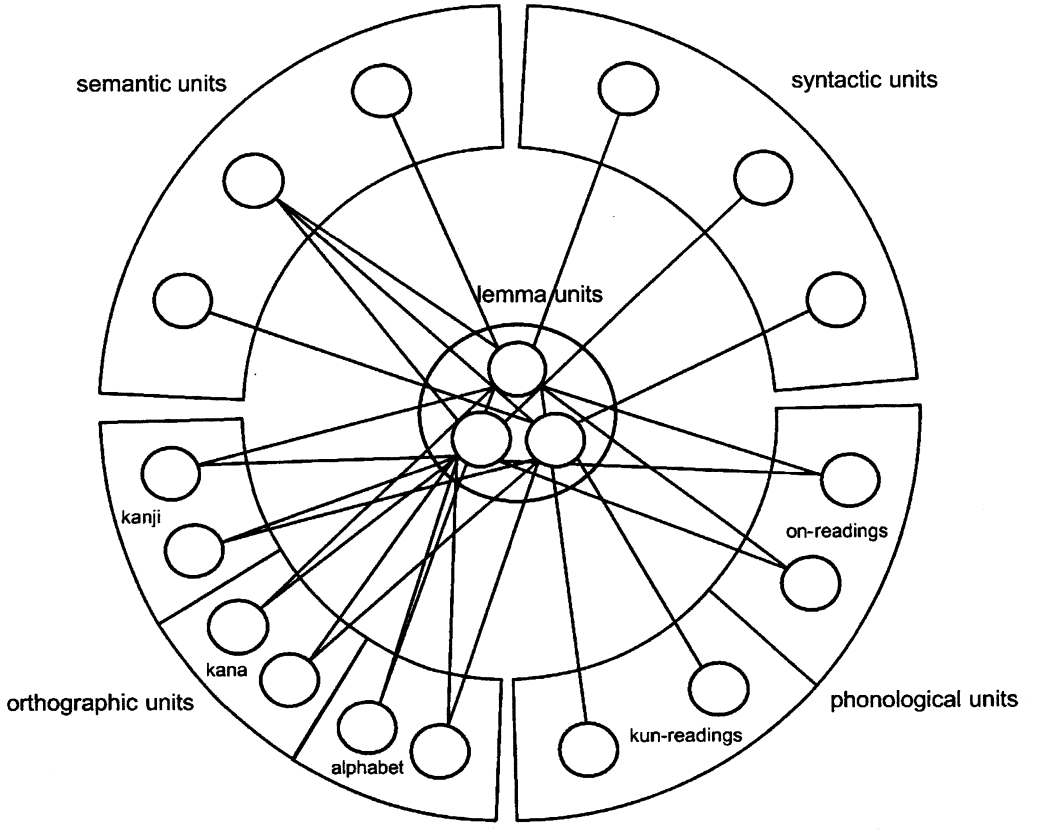


Figure 5. Schematic Representation of Japanese Lemma Unit Model, with Lemma Unit Representations Mediating the Connections between Orthographic (Kanji, Kana, and Alphabet) and Phonological (On- and Kun-Readings) Access Representations, Semantic Representations and Syntactic Representations

[19] Syntactic information is also given separate representation in the parallel distributed processing model proposed by Tsuzuki, Kawamoto, and Yukihiro (1999) which is concerned with the resolution of lexical ambiguity in homophone Japanese nouns.

In summary, this paper has considered important morphological, orthographic and phonological issues related to the Japanese language and the complexity of the Japanese writing system. These issues have been examined within the context of the Japanese lemma unit model advocated by Joyce (1999, 2002a), which is a modified adaptation to Japanese of the multi-level interactive-activation framework (Taft, 1991, 1994). The special feature of this model is the incorporation of lemma unit representations, as connections mediating the links between access representations, both orthographic and phonological, and semantic representations. While there are still aspects of the model that need to be considered, such as how to account for positional sensitivity of verbal constituents with high positional ratios and the inclusion of syntactic information, the Japanese lemma unit model provides an appealing way to account for the diversity in the morphology of two-kanji compound words. Moreover, the inclusion of lemma unit representations also provides a promising approach to handling, on the one hand, orthographic issues arising from the multi-script nature of the Japanese system and, on the other hand, phonological issues due to its dual-reading system, as supported by the results of a series of cross-script and cross-modal versions of constituent-morpheme priming experiments.

REFERENCES

Allen, M., & Badecker, W. (2002). Inflectional regularity: Probing the nature of lexical representation in a cross-modal priming task, *Journal of Memory and Language, 46,* 705-722.

Amano, S., & Kondō, T. (Eds.). (1999). *Nihongo no goitokusei* [Lexical properties of Japanese] NTT database series. Tokyo: Sanseidō.

Anderson, S. R. (1992) *A-morphous morphology.* Cambridge, England: Cambridge University Press.

Backhouse, A. E. (1984). Aspects of the graphological structure of Japanese. *Visible Language, 18,* 219-228.

Besner, D. & Hildebrandt, N. (1987). Orthographic and phonological codes in the oral reading of Japanese kana. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 335-343.

Bowers, J. S., & Michita, Y. (1998). An investigation into the structure and acquisition of orthographic knowledge: Evidence from cross-script Kanji-Hiragana priming. *Psychonomic Bulletin and Review, 5,* 259-264.

Butterworth, B. (1983). Lexical representation. In B. Butterworth, (Ed.), *Language production: Volume 2 Development, writing and other language processes* (pp. 257-294). London, England: Academic Press.

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

Chen, H.-C. (Ed.). (1997). *Cognitive processing of Chinese and related Asian languages.* Sha Tin, N.T., Hong Kong: Chinese University Press.

Chen, H.-C., & Zhou, X. (1999). Processing East Asian languages: An introduction. *Language and Cognitive Processes, 14,* 425-428.

Coulmas, F. (1989). *The writing systems of the world.* Oxford, England: Basil Blackwell.

Coulmas, F. (1996). *The Blackwell encyclopedia of writing systems*. Oxford, England: Blackwell.

Daniels, P. T. (1996). The study of writing systems. In P. T. Daniels & W. Bright (Eds.). *The world's writing systems* (pp. 3-17). New York, NY: Oxford University Press.

Daniels, P. T. (2001). Writing systems. In M. Aronoff & J. Rees-Miller (Eds.). *The handbook of linguistics* (pp. 43-80). Oxford, England: Blackwell.

De Groot, A. M., Thomassen, A. J, & Hudson, P. T. (1982). Associative facilitation of word recognition as measured from a neutral prime. *Memory & Cognition, 10,* 358-370.

DeFrancis, J. (1989). *Visible speech: The diverse oneness of writing systems*. Honolulu, HI: University of Hawaii Press.

Feldman, L. B. (Ed.). (1995). *Morphological aspects of language processing*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Feldman, L. B., & Larabee, J. (2001). Morphological facilitation following prefixed but not suffixed primes: Lexical architecture or modality-specific processes? *Journal of Experimental Psychology: Human Perception and Performance, 27,* 680-691.

Feldman, L. B., & Soltano, E. G. (1999). Morphological priming: The role of prime duration, semantic transparency, and affix position. *Brain and Language, 68,* 33-39.

Fischer, S. R. (2001). *A history of writing*. London: Reaktion Books.

Flores d'Arcais, G. B. (1992). Graphemic, phonological and semantic activation processes during the recognition of Chinese characters. In H. -C. Chen, & O. J. L. Tzeng (Eds.), *Language processing in Chinese* (pp. 37-66). Amsterdam: North-Holland.

Forster, K. I. (1976). Accessing the mental lexicon. In R. J. Wales, & E. Walker (Eds.). *New approaches to language mechanisms: A collection of psycholinguistic studies* (pp. 257-287). Amsterdam: North-Holland.

Forster, K. I. (1989). Basic issues in lexical processing. In W. Marslen-Wilson, (Ed.), *Lexical representation and process* (pp. 75-107). Cambridge, MA: MIT Press.

Frost, R., Deutsch, A., Gilboa, O., Tannenbaum, M., & Marslen-Wilson, W. (2000). Morphological priming: Dissociation of phonological, semantic, and morphological factors. *Memory & Cognition, 28,* 1277-1288.

Fujii, T. (1996). Goi [Vocabulary]. In Saji, K. & Sanada, S (Eds.), *Nihongo kyōshi yōsei shiriizu: Onsei, goi, moji hyōki* (pp. 85-104). Tokyo: Tokyo Hōmei Shuppan.

Fushimi, T., Ijuin, M., Patterson, K., & Tatsumi, I. F. (1999). Consistency, frequency, and lexicality effects in naming Japanese kanji. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 382-407.

Hatano, G. (1995). The psychology of Japanese literacy: Expanding 'the practice account'. In L. M. W. Martin, K. Nelson, & E. Tobach (Eds.), *Sociocultural psychology: Theory and practice of doing and knowing* (pp. 250-275). Cambridge, England: Cambridge University Press.

Hatano, G., Kuhara, K., & Akiyama, M., (1981). Kanji help readers of Japanese infer the meaning of unfamiliar words. *The Quarterly Newsletter of the Laboratory of Comparative Human Cognition, 3,* 30-33.

Hatta, T., & Saito, H. (Eds.) (1999). Cognitive processing of the Japanese and Chinese languages I [Special issue]. *Psychologia, 42(4).*

Hatta, T., & Saito, H. (Eds.) (2000). Cognitive processing of the Japanese and Chinese languages II [Special issue]. *Psychologia, 43(1).*

Hirose, H. (1992). Jukugo no nichi katei ni kan suru kenkyū: Puraimingu hō ni yoru kentō [An investigation of the recognition process for jukugo by use of priming paradigms]. *Shinrigaku Kenkyū, 63,* 303-309.

Hirose, H. (1994). Kanji no kioku kōzō ni kansuru kenkyū: 1ji kanji to jukugo no kanren ni tsuite [An investigation of the memory structure of Kanji: The relation of Kanji and Jukugo]. *The Japanese Journal of Psychonomic Science, 12,* 71-76.

Hirose, H. (1996). Functions of graphemic components of kanji on recognition of jukugo. *Hiroshima Forum for Psychology, 17,* 49-55.

Ijuin, M., Fushimi, T., Patterson, K., & Tatsumi, I. (1999). A connectionist approach to Japanese kanji word naming. *Psychologia, 42,* 267-280.

Jarema, G., Kehayia, E., & Libben, G. (Eds.). (1999) Mental lexicon [Special issue]. *Brain and Language, 68(1/2).*

Joyce, T. (1999). Lexical access and the mental lexicon for two-kanji compound words: A priming paradigm study. *Proceedings of the 2nd International Conference on Cognitive Sciences and 16th Annual Meeting of the Japanese Cognitive Science Society Joint Conference*, 27-30 July, Tokyo, Japan, 511-514.

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

Joyce, T. (2002b). *The Japanese mental Lexicon: The lexical retrieval and representation of two-kanji compound words from a morphological perspective.* Unpublished doctoral thesis. University of Tsukuba, Japan.

Joyce, T. (2003a). *Frequency and verb-morphology effects for constituents of two-kanji compound words.* Poster session presented at the 4th Tsukuba International Conference on Memory (Human Learning and Memory: Advanced in Theory and Application), 11-13 January, Tsukuba, Japan.

Joyce, T. (2003b). Kanji niji jukugo ni okeru dōshi kōzō yōso no ichiteki hand [Positional frequency of verbal constituents within two-kanji compound words] *Proceedings of the 67th Meeting of the Japanese Psychological Association*, Tokyo University, Tokyo, Japan, p. 590.

Joyce, T. & Ohta, N. (1999). The morphology of two-kanji compound words: Data from a word-formation classification survey. *Tsukuba Psychological Research, 22,* 45-61.

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

Kageyama, T. (1982). Word formation in Japanese. *Lingua, 57,* 215-258.

Kaiho, H., & Nomura, Y. (1983) *Kanji jōhō shori no shinrigaku* [The psychology of kanji information processing]. Tokyo: Kyōiku Shuppan.

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

Leong, C. K., & Tamaoka, K. (Eds.). (1998). Cognitive processing of the Chinese and Japanese languages [Special issue]. *Reading and Writing, 10(3-5).*

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* Cambridge, MA: MIT Press.

Marslen-Wilson, W., Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review, 101,* 2-33.

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,* 375-407.

McClelland, J. L., Rumelhart, D. E., and the PDP Research Group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 2. Psychological and biological models*. Cambridge, MA: MIT Press.

Monsell, S. (1985). Repetition and the lexicon. In A. W. Ellis, (Ed.), *Progress in the psychology of language: Vol. 2* (pp. 147-195). London: Lawrence Erlbaum Associates.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106,* 226-254.

Nomura, M. (1988). Niji kango no kōzō [The structure of 2 kanji Sino-Japanese words]. *Nihongogaku, 7*(5), 44-55.

Ozaki, Y., Todome, H., Nishioka, H., Yamada, K., & Yamada, T. (Eds.). (1992) *Daijigen* [Large dictionary of etymology] Tokyo: Kadokawa Shoten.

Page, M. (2000). Connectionist modelling in psychology: A localist manifesto. *Behavioral and Brain Sciences, 23,* 443-512.

Paradis, M., Hagiwara, H., & Hildebrandt, N. (1985). *Neurolinguistic aspects of the Japanese writing system*. Orlando, FL: Academic Press.

Rubin, G. S., & Becker, C. A., Freeman, R. H. (1979). Morphological structure and its effect on visual word recognition. *Journal of Verbal Learning and Verbal Behavior, 18,* 757-767.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review, 89,* 60-94.

Rumelhart, D. E., McClelland, J. L., and the PDP Research Group. (Eds.). (1986). Parallel distributed processing: *Explorations in the microstructure of cognition. Vol. 1. Foundations*. Cambridge, MA: MIT Press.

Saito, H. (1997). Shinteki jisho [Mental lexicon]. In Y. Ohtsu, & T. Gunji (Eds.), *Linguistic Sciences* Vol. 3 (pp. 94-153). Tokyo: Iwanami Shoten.

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

Sakuma, N., Sasanuma, S., Tatsumi, I. F., & Masaki, S. (1998). Orthography and phonology in reading Japanese kanji words: Evidence from the semantic decision task with homophones. *Memory & Cognition, 26,* 75-87.

Sandra, D. (1994). The morphology of the mental lexicon: Internal word structure viewed from a psycholinguistic perspective. *Language and Cognitive Processes, 9,* 227-269.

Sandra, D., & Taft, M. (Eds.) (1994). Morphological structure, lexical representation and lexical access [Special issue] *Language and Cognitive Processes, 9(3).*

Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. In L. B. Feldman, (Ed.), *Morphological aspects of language processing* (pp. 131-154). Hillsdale, NJ: Lawrence Erlbaum Associates.

Scribner, S., & Cole, M. (1981). *The psychology of literacy*. Cambridge, MA: Harvard University Press.

Shafiullah, M., & Monsell, S. (1999). The cost of switching between kanji and kana while reading Japanese. *Language and Cognitive Processes, 14,* 567-607.

Shibatani, M. (1990). *The languages of Japan.* Cambridge, England: Cambridge University Press.

Shimamura, A. P. (1987). Word comprehension and naming: An analysis of English and Japanese orthographies. *American Journal of Psychology, 100,* 15-40.

Smith, J. S. (1996). Japanese writing. In P. T. Daniels & W. Bright (Eds.). *The world's writing systems* (pp. 209-217). New York: Oxford University Press.

Sproat, R. (2000). *A computational theory of writing systems.* Stanford, CA: Cambridge University Press.

Sugishita M., Otomo K., Kabe S., & Yunoki K. (1992). A critical appraisal of neuropsychological correlates of Japanese ideogram (kanji) and phonogram (kana) reading. *Brain, 115,* 1563-1585.

Taft, M. (1991). *Reading and the mental lexicon.* Hillsdale, NJ: Lawrence Erlbaum Associates.

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

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., Liu, Y., & Zhu, X. (1999). Morphemic processing in reading Chinese. In J. Wang, A. Inhoff, & H.-C. Chen, (Eds.), *Reading Chinese script: A cognitive analysis* (pp 91-113). Hillsdale, NJ: Lawrence Erlbaum Associates.

Taft, M., & Zhu, X. (1995). The representation of bound morphemes in the lexicon: A Chinese study. In L. B. Feldman, (Ed.), *Morphological aspects of language processing* (pp. 293-316). Hove, England: Lawrence Erlbaum Associates.

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

Taft, M., & Zhu, X. (1997b). Using masked priming to examine lexical storage of Chinese compound words. In H.-C. Chen (Ed.), *Cognitive Processing of Chinese and Related Asian Languages* (pp. 233-241). Sha Tin, Hong Kong: The Chinese University Press.

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

Tamamura, F. (1985). Go no kōsei to zōgohō [The structure of words and principles of word formation] In Kokuritsu Kokugo Kenkyūjo. *Go no kenkyū to kyōiku (shita) Nihongo kyōiku shidō sankōsho 13* (pp. 1-78). Tokyo: Ōkurashō Insatsukyoku.

Tamaoka, K. (1991). Psycholinguistic nature of Japanese orthography. *Studies in Language and Literature, 11,* 49-82.

Tamaoka, K. (1994). A review of psychological studies of kana and kanji: A single phonological route to a multiple interactive activation. *Studies in Language and Literature, 13,* 1-18.

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

Tsuzuki, T., Kawamoto, A. H., & Yukihiro, R. (1999). Goiteki tagisei no shori ni kansuru heiretsu bunsan shori moderu: Bunmyaku to tomo ni teiji sareta tagigo no ninchi ni kansuru jikken dēta no rironteki tōgō [A parallel distributed processing model of lexical

ambiguity resolution: The theoretical integration of empirical data for processing ambiguous Japanese words presented with contexts]. *Ninchi Kagaku, 6,* 91-104.

Wydell, T. N., Patterson, K. E., & Humphreys, G. W. (1993). Phonologically mediated access to meaning for kanji: Is a rows still a rose in Japanese kanji? *Journal of Experimental Psychology Learning, Memory, and Cognition, 19,* 491-514.

Yamada, T. (1967). Japan - The writing system: Historical research and modern development. In T. A. Sebeok (Ed.). *Current trends in linguistics Vol. 2* (pp. 693-734), The Hague, Mouton.

Yokosawa, K. & Umeda, M. (1988). Processes in human kanji-word recognition. *Proceedings of the 1988 IEEE international conference on systems, man, and cybernetics* (pp. 377-380). August 8-12, 1988, Beijing and Shenyang, China.

Yokoyama, S., Sasahara, H., Nozaki, K., & Long. E., (1998). *Shinbun denshi media no kanji: Asahi Shinbun CD-ROM ni yoru kanji hindohyō* [Electronic newspaper media kanji: Kanji frequency lists based on Asahi Newspaper CD-ROM]. Tokyo: Sanseidō.

Zhou, X., & Marslen-Wilson, W. (2000). Lexical representation of compound words: Cross-linguistic evidence. *Psychologia, 43,* 47-66.

AUTHOR NOTE

This research has been supported first by a scholarship from the Japanese Ministry of Education, Culture, Sports, Science and Technology, and then by a Japanese Society for the Promotion of Science (JSPS) postdoctoral fellowship for foreign researchers. Gratitude is extended to Nobuo Ohta for supervising my research at the University of Tsukuba and to Junichi Oda for acting as host at the Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies. Part of this research was presented at the 65th Meeting of the Japanese Psychological Association, 7-9 November 2001, Tsukuba, Japan.