30 Morphological representation and processing of Sino-Korean words

Kwangoh Yi

Abstract

Sino-Korean words are Chinese words that have been borrowed into the Korean language over a long period of history. Currently representing over 70 percent of the entire Korean vocabulary, Sino-Korean words are typically two syllables in length and morphologically compound in structure. Behavioral studies have shown that although morphologically related Sino-Korean primes neither facilitate nor inhibit the processing of a target Sino-Korean word, orthographically related Sino-Korean primes tend to inhibit the recognition of a target word. Such findings are inconsistent with existing models for English and Chinese. Based on these Korean research findings, it is proposed that a level of morphological representations exists between word-form and semantic representation, and that there are strong inhibitory connections between word-form representations that share a given syllable in a particular word position. A number of related issues are also discussed, such as the representation of morphology within the mental lexicon, individual differences in terms of reading proficiency, semantic versus morphological relatedness, and the position of critical constituents.

Introduction

The Korean alphabet, or Hangul, consists of fourteen consonant and ten vowel letters. While each letter of Hangul represents a phoneme, it is more accurate to characterize the present Korean writing system as being morphophonemic in nature, rather than purely phonemic. In order to correctly spell a Korean word, it is necessary to consider both its phonological and morphological characteristics. According to spelling regulations publicized by the Korean government in 1948, the Korean language should be written with Hangul in order to both precisely transcribe the pronunciation of words and to clarify their meanings. These regulations are the product of a long struggle by the Korean people to develop effective orthographic uses of their essentially phonemic script. When Hangul was first invented by King Sejong, great emphasis was placed on

accurately spelling a word according to its sound. However, as time passed, it gradually became evident that a system of spelling based solely on pronunciation was not completely effective. Accordingly, a movement emerged calling for a system of spelling that reflects both the phonological and morphological characteristics of words. Thus, modern Korean orthography may be seen as striking a compromise between the need to represent both phonology and morphology. In this sense, modern Korean orthography is very similar to English orthography.

Although Hangul is an alphabetic script, Hangul letters are not arranged linearly from left to right, as in English, for example. Rather, the arrangement is more complex involving both left-to-right and top-to-bottom directions in order to form character-like syllabic blocks. The block arrangement of Hangul letters serves to make the syllable particularly salient, and so the syllable is easy to identify in Korean texts. Reflecting this, the Hangul writing system has sometimes been referred to as an alphabetic syllabary (Taylor, 1980). Given the fact that the syllable typically corresponds to a morpheme, and that, as already noted, spelling conventions also emphasize the morphology of words, the morpheme is also a very salient unit in Korean orthography. Thus, it would seem that orthographic and morphological processing are both fundamental to the recognition of Korean words.

It should, however, be noted that the correspondence between the syllable and the morpheme varies according to word type. The correspondence is especially high in the case of Sino-Korean vocabulary, compared to native Korean and Western loanwords. Sino-Korean words (hereafter abbreviated as S-K) are mostly compound words, composed of two or more constituent morphemes. Assimilated into the Korean language from Chinese over a long period of history, S-K words presently represent over 70 percent of all Korean words. Each syllable of an S-K word was traditionally represented orthographically by a Chinese character, and in terms of its meaning and pronunciation, the syllables still correspond to the original Chinese meanings and pronunciations. Due to the huge numbers of S-K words in Korean, clearly essential tasks for researchers of Korean word recognition are to understand how S-K words are represented in the Korean mental lexicon and how they are processed in visual and spoken word recognition.

The remainder of this paper will discuss some important findings concerning the morphological representation and processing of S-K words, and will draw some crosslinguistic comparisons with other languages, such as English, Chinese, and Japanese. As a little background to the discussions, the next section continues by briefly summarizing some studies of morphological processing. The Chinese and Japanese studies mentioned are of particular relevance, because many S-K compound words are also common to those languages.

Morphological priming in languages other than Korean

Researchers investigating word recognition have focused on the morpheme as an important sub-lexical unit. For instance, a number of studies employing the priming paradigm have reported morphological facilitation and orthographic inhibition. Morphological facilitation refers to the robust finding that lexical decision responses to a target word are faster when it has been preceded by a morphologically related word compared to when preceded by an unrelated word. In contrast, orthographic inhibition refers to the finding that responses are delayed when the prime and target share an identical string of letters relative to when there is no orthographic overlap.

Stolz and Feldman (1995) have reported both morphological facilitation and orthographic inhibition for English. In their experiment, they manipulated the lag – in terms of the number of intervening items – between a prime and a target. In the short-lag condition, the target immediately followed the prime with no intervening items, but in the long-lag condition, on average ten unrelated words were presented between the prime and its target. Consistent with their hypothesis that orthographic overlap would not influence responses in the long-lag condition, Stolz and Feldman found no orthographic effects for any prime–target pairs, but they did observe significant morphological priming for both opaque and transparent prime–target pairs. However, in the short-lag condition, both morphological facilitation and orthographic inhibition were observed. The fact that orthographic pairs (e.g., *market–mark*) only inhibited each other in the short-lag condition implies that orthographic activation is short-lived.

In addition to studies in European languages, like English, there has also been research into the morphological processing of Asian languages, such as Chinese and Japanese, where morphemes have been shown to play a significant role in the processing of compound words. The results from Chinese studies are especially interesting. Most studies have shown facilitation when a prime and target share a morpheme. In addition, facilitatory effects, somewhat weaker than morphological facilitation, have also been reported for homographic–homophonic pairs which are not morphologically related (Zhou, Marslen-Wilson, Taft and Shu, 1999). It should be noted that orthographic facilitation, not orthographic inhibition like English, is found in Chinese.

In the case of Japanese, Hirose (1992) has also reported morphological facilitation in a constituent-morpheme priming study, although the effects interacted with the position of the critical morpheme within the compound word target. Because the priming effects were greater when the morpheme prime appeared in word-initial position within the target words, Hirose suggested that kanji compound words are arranged according to the initial kanji character. However, the results from more recent studies by Joyce (1999, 2002) clearly challenge this interpretation. Contrasting a number of word-formation

principles, or morphological structures for compound words, Joyce observed morphological facilitation in all morphological structure conditions from both constituent prime conditions, and in the majority of cases at a very similar level.

Thus, while the Chinese and Japanese studies have also yielded considerable evidence of morphological facilitation, which is consistent with English studies, they have failed to provide any evidence of orthographic or phonological inhibition. In that respect, they differ from the English data.

Experiments in Korean

The role of morphology within the recognition of Korean words has been an important issue for researchers who emphasize the language-dependent nature of lexical processing. For example, the series of studies by Jung and Yi (2004), Yi (1999a, 1999b), and Yi and Yi (1999) have explored the representation and processing of S-K words composed from two constituent morphemes. Yi and Yi (1999) first addressed representational issues in a lexical decision experiment employing a long-lag presentation similar to that used in Stolz and Feldman (1995). Taking the target stimulus "pan-hang" (meaning *resistance*) as an example, Table 30.1 illustrates the three kinds of prime–target relations in their experiment. The first type of prime was a morphologically related word sharing a constituent morpheme with the target word in the word-initial position (i.e."pan-chik" meaning *violation*). The second type of prime was an orthographically related word sharing a syllable in the word-initial position (i.e. "pan-jang" meaning *captain*). The third type was a neutral prime, or control, that was both morphologically and orthographically unrelated to the target (i.e. "kong-kai" meaning *open*). In the long-lag condition, the results were similar to those reported by Stolz and Feldman. As shown in Figure 30.1, morphological facilitation was found, but orthographic inhibition was not observed. This result is consistent with the view that lexical entries for S-K words can be contacted through the activation of their constituent morphemes. In sharp contrast, however, a very different pattern of results was produced in the short-lag condition, which included four stimulus onset asynchronicity (SOA) conditions of 60, 200, 300, and 1,000 msec. Although

Table 30.1. *The prime–target relations of Sino-Korean word stimuli in Yi and Yi (1999)*

	Morphological (MORPH)	Orthographic (ORTHO)	Control (CON)
Prime	반칙{反則} "pan-chik" *violation*	반장{班長} "pan-jang" *captain*	공개{公開} "kong-kai" *open*
Target		반항{反抗} "pan-hang" *resistance*	

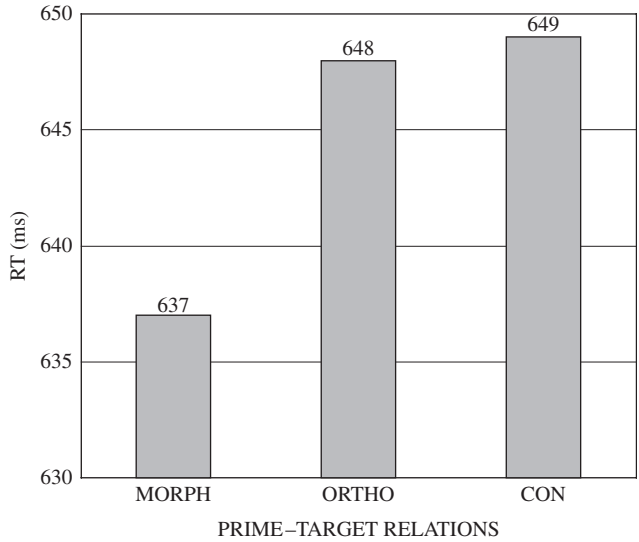

Figure 30.1. Mean lexical decision times (LDTs) for morphologically related, orthographically related, and control pairs at long lags

there was orthographic inhibition in the orthographically related condition, morphological facilitation was not observed in the morphologically related prime condition, as shown in Figure 30.2. This pattern of results was identical across the various SOA conditions from the shortest of 60 msec to the longest of 1 sec. This lack of morphological facilitation is particularly remarkable because such results have not been reported for any other language before.

How can this lack of morphological facilitation for morphologically related words be explained? Do these results compel us to believe that morphology does not have a role in lexical retrieval? Perhaps not. If we accept that morphologically related words facilitate each other and that orthographically related words inhibit each other, then one possible interpretation of these results is that these two forms of priming acted to effectively cancel each other out. That is, the morphological facilitation for the morphologically related words is negated by orthographic inhibition. This interpretation would be consistent with the results of a study by Jung and Yi (2004) that investigated the effects of morpheme position with S-K words. It has been claimed that word-initial morphemes play a more pivotal role than word-final morphemes for English compounds (Taft & Forster, 1976). The basic suggestion is that the organization of lexical entries is based on their word-initial constituents. A similar proposal was made by Hirose (1992) concerning Japanese two-kanji compound words, although as already noted, the results from Joyce (1999) of generally similar

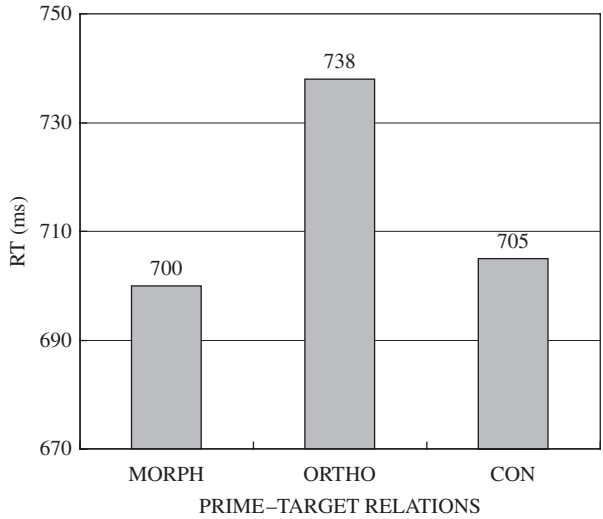

Figure 30.2. Mean LDTs to targets at short lags (SOA = 300)

levels of morphological facilitation from both word-initial and word-final constituent morphemes are clearly not compatible with the notion of an advantage for word-initial morphemes. As shown in Figure 30.3, the results from Jung and Yi's study suggest that morphological facilitation and orthographic inhibition interact with the position of the critical constituents. Morphological priming was greater when the shared morpheme was in different positions than when it was in the same position. However, orthographic inhibition was stronger when the shared constituent was in the same position in the primes and targets – either word-initial for both or word-final for both.

Another issue for research concerned with the effects of morphological relatedness that is frequently noted is the possible confounding effect of semantic priming. In order to tease apart semantic priming and morphological priming a number of studies have included a semantically related condition to compare with the morphologically related condition. For example, Zhou et al. (1999), using Chinese compound words, reported even greater levels of morphological facilitation compared to semantic facilitation. However, Yi (1999) ran an experiment with S-K compounds and found that the level of morphological facilitation was considerably lower than the level of semantic facilitation, as shown in Figure 30.4. This pattern of results for S-K words is very different from that reported by Zhou and Marslen-Wilson. Yi's results would suggest that inter-lexical connections are stronger than morpho-lexical connections in the S-K lexical system.

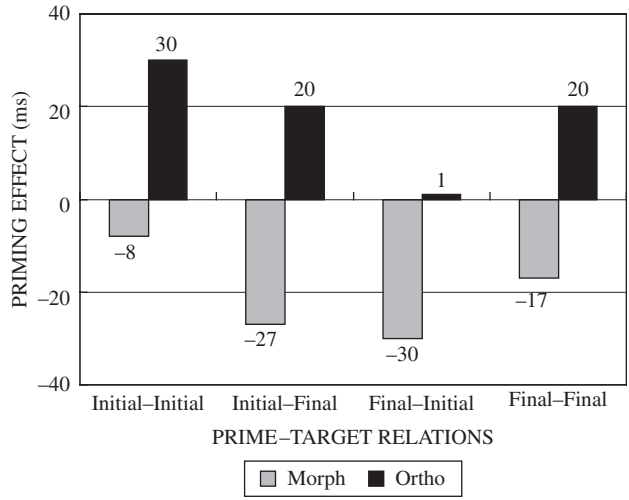

Figure 30.3. Positional effects of critical morphemes (SOA = 400)

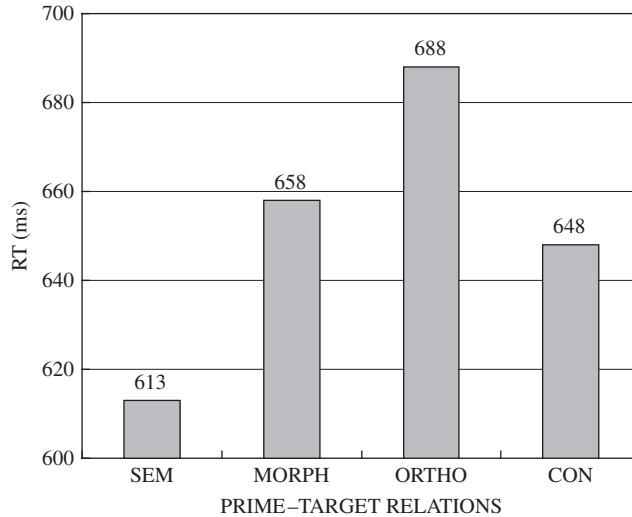

Figure 30.4. Mean LDTs (ms) for semantically, morphologically, and orthographically related pairs (SOA = 200)

Another fascinating – albeit extremely elusive – issue for word recognition research is individual differences in terms of reading performance. Of particular relevance in this context are the results reported by Yi (1999), which are clearly supportive of the notion that the morphological representation and

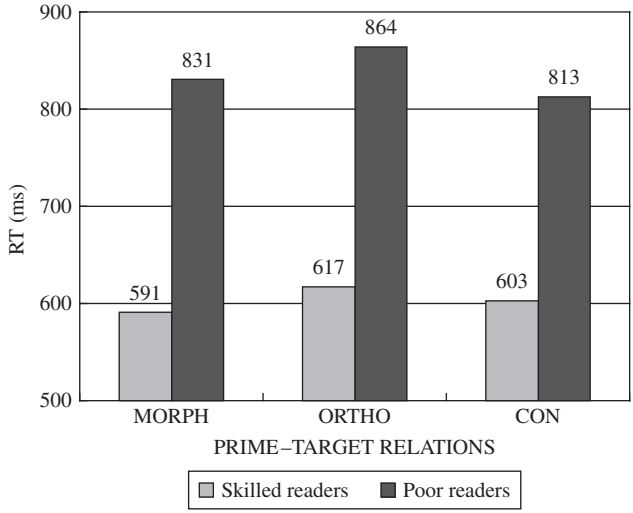

Figure 30.5. Mean LDTs as a function of reading performance (SOA = 300)

processing of S-K words varies as a function of individual reading skill. By categorizing participants into skilled and poor readers according to mean lexical decision reaction times, Yi discovered that the patterns of priming for the two groups differed across a range of SOA conditions, as shown in Figure 30.5. The main characteristic in the priming pattern for the skilled readers was that morphological facilitation was relatively large for morphologically related words, but the levels of orthographic inhibition were very low, failing to reach statistical significance. In contrast, the poor readers were only found to show orthographic inhibition for the orthographically related pairs. No morphological facilitation for morphologically related words was observed with the poor readers.

Morphological level representations in a mental lexicon model

As briefly reviewed in the previous section, the findings from the experimental studies investigating the lexical processing of S-K words would seem to indicate that S-K morphemes are represented in the mental lexicon, but that morphological activation is not an obligatory stage or route in accessing the lexical entries of S-K words. This is clearly at odds with general notions of morphological involvement in the organization of the mental lexicon that posit a level of morphological representation between the orthographic and lexical levels and thus regard morphological activation as an intermediate and mandatory stage.

Most models of morphological representation and processing that have been proposed for East Asian languages take for granted the existence of three levels of representations; character, morpheme, and word. These layers of representation are usually organized hierarchically with character representations at the lowest level and lexical representations at the highest level, positing a layer of morphological representations between these, which is seen as being logically necessary in order to account for morphological facilitation between morphologically related words at short lags. However, the assumption that the layer of morpheme representation is sublexical is not consistent with the experimental results reviewed above. Morphological facilitation was not observed for S-K compound words sharing a common morpheme at short lags. Rather, it would seem that the morphological representations of S-K words cannot be activated prior to the activation of whole word representation because the mapping between characters and morphemes would be too diffuse. The orthographic syllables of S-K words can usually be mapped onto tens of morphemes. However, an alternative arrangement of representational layers also seems feasible; one in which morpheme representations exist above whole-word representations. Since such an arrangement would imply that morphological facilitation arises due to the activation of shared morpheme units following whole-word unit activation, it would clearly be more consistent with the reviewed Korean results. According to this supralexical perspective, when primes and targets share a common morpheme, then morphological activation due to the presentation of the prime will facilitate the processing of the target word via the excitatory connections between the relevant morphemic representations and the appropriate whole-word representations (Giraudo & Grainger, 2000).

General discussion

In summary, the results reviewed here from a series of Korean studies challenge some of the underlying assumptions in models of morphological representation and processing based on other languages. Although S-K words are compounds, consisting of two or more morphemes, the results provide little evidence of morphological facilitation for S-K words across a variety of SOAs. Such a finding may be explained by assuming that orthographic inhibition is due to inhibitory connections between whole-word units at the same level, while morphological facilitation arises from the connections between units at both the morphemic and lexical levels. Moreover, assuming that orthographic inhibition and morphological facilitation operate to cancel each other out in the case of morphologically related pairs, because no morphological facilitation occurs for orthographically related pairs, only orthographic inhibition is observed. The Korean data also suggest that positional effects are stronger when the shared constituent is purely orthographic in nature compared to the weaker priming for

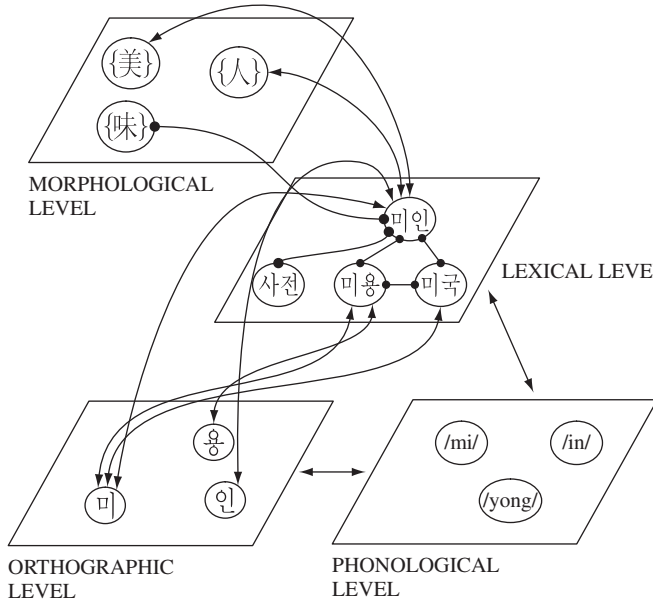

Figure 30.6. A mental lexicon model for the representation and processing of Sino-Korean words

shared morpheme constituents. The fact that words sharing a constituent at either word-initial or word-final position strongly inhibit each other implies that this is an important characteristic relating to the arrangement of lexical-level units. Assuming that competition is stronger between lexical units that share a syllable in the same position, inhibition will be stronger between primes and targets that share a constituent at either the word-initial or the word-final positions compared to when the common morpheme is in a different position.

Interestingly, greater levels of morphological facilitation and lower levels of orthographic inhibition were observed with skilled readers compared to poor readers, where a reversed pattern of priming was seen. This suggests that while lexical processing relies primarily on the activation of lexical units in the case of poor readers, the activation of sublexical units, such as morphemes, is more important for skilled readers.

Figure 30.6 presents a mental lexicon model for the representation and processing of S-K words. The first major feature of the model is the positioning of the layer of morphological representations above the layer of lexical units. The second important feature is the greater weighting of inhibition connections between lexical-level representations that have the same syllable in the same position. Individual differences in terms of lexical processing for S-K words can

be accounted for by the connection weights between whole-word units and morpheme units. Exposure to more S-K words, or Chinese characters, would help to establish and strengthen the links between words and their constituent morphemes.

Naturally, there are still many other issues that need to be addressed in investigating the morphological representation and processing of S-K words, such as the effects of Hanja proficiency, for example, on the processing of S-K words, which is likely to have serious theoretical and practical implications. The results from cross-psycholinguistic studies using Chinese, Japanese and other Asian languages are expected to shed light on such questions.