The Handbook of East Asian Psycholinguistics

Volume III Korean

Edited by
Chungmin Lee
Greg Simpson
and
Youngjin Kim

General Editor
Ping Li

CAMBRIDGE UNIVERSITY PRESS
and Kang (1994) may have been due to differences in the frequency of the syllables that the words comprised. If the syllable is a processing unit in Hangul, then a word-level variable such as frequency would not be expected to have a large effect.

In order to confirm that the syllable enjoys an independent status as a processing unit in Hangul that is separate from both lexical and subsyllabic (i.e. graphemic) processes, we conducted two further experiments. In the first (Simpson & Kang, 2004, Experiment 3), we compared naming times for bound syllables and pseudo-syllables, controlled for the subsyllabic variables of bigram frequency, number of graphemes per syllable, and initial phoneme. In the final experiment (Experiment 4), we compared high- and low-frequency bound syllables. The results of both experiments confirmed an independent role for the syllable. Bound syllables were named significantly faster than were pseudo-syllables and high-frequency bound syllables were named more rapidly than their low-frequency counterparts.

Taken together, these experiments suggested that the syllable can be an independent processing unit in Hangul word recognition. The degree to which this characteristic may generalize to other languages remains to be seen. A small number of studies have investigated syllable processing in English (Prinzmetal, Treiman & Rho, 1986) and other languages, such as French (Cutler, Mehler, Norris & Segui, 1986) and Spanish (Carreira, Alvarez & de Vega, 1993). Because languages differ both in the clarity of syllabification in speech, and (with Hangul being a prime example) in their written form, considerably more work will be needed to test the generality of syllable-based processing in word recognition.

Summary and conclusions

We have suggested that crosslinguistic analyses are critical to a complete understanding of word-recognition processes. The ultimate goal of research in this area is to gain an understanding of the architecture of the word-recognition system. One component of this goal is the identification of those processes that are universal among readers of any language and orthography, and those that vary with characteristics of the language. Underlying all of our work is the question of how phonological information is used in word recognition, and how it interacts with other components of the word-recognition system. We hope that our research helps to guide future work by suggesting some aspects of word recognition worthy of such direct crosslinguistic comparison.

33 Prosody in sentence processing

Sun-Ah Jun

Introduction

Prosody refers to a grouping within an utterance and the prominence relations between the members within the group. Groupings within an utterance, called prosodic units, are hierarchically organized so that a prosodic unit can include one or more smaller prosodic units. Since the grouping and the prominence relations between the members are often marked by intonation, the terms 'intonation' and 'prosody' are often used interchangeably. Intonation, though traditionally defined as the global changes in pitch over the course of a sentence or a phrase, has an internal structure. Some pitch events mark the boundaries of groupings, either small or large, while others mark the prominent members within a group. In this way, an intonation contour marks a hierarchy of groupings and reflects the metrical structure of the group. The pitch events marking the internal structure of intonation can be represented by two distinct pitch levels, High (H) or Low (L) and their combinations (e.g. HL for falling and LH for rising). This view of intonation is known as an autosegmental-metrical model of intonation or intonational phonology, started in the late 1970s and early 1980s through the seminal works of Bruce (1977) on Swedish intonation and Pierrehumbert and her colleagues on English intonation (e.g. Pierrehumbert, 1980; Beckman & Pierrehumbert 1986; Liberman & Pierrehumbert, 1986; Pierrehumbert & Hirschberg, 1990).

This model of intonation has been applied to Japanese (Pierrehumbert & Beckman, 1988) and Korean (Jun, 1993), and has been expanded to many other languages including German (Grice & Benzmüller, 1992) and Greek (Arvaniti & Baltazani, 2005; see S.-A. Jun, 2005 for a similar analysis of eight other languages). As a phonological model, this model specifies only distinctive tonal events which are specific to each language or a dialect. Non-distinctive, and thus predictable, tones are not specified. Syllables with no tonal

---

1 For most of these languages, a prosodic transcription system known as ToBI (Tones and Break Indices) has been developed based on the intonational phonology (i.e. tones) of each language and the prosodic groupings defined by the degree of juncture between words (i.e. break indices) (see Jun, 2005, Chapter 2 for the history and the principles of ToBI).
Intonation of Korean and English

Intonation of Korean

The intonational phonology of Korean proposed in S.-A. Jun (1993, 1998) and the Korean ToBI (Tones and Break Indices) model, and the transcription system of intonation and phrasing reported in S.-A. Jun (2000)\(^2\) posit two prosodic units above the word: an Intonation Phrase (IP) and an Accenthual Phrase (AP). An IP can have one or more APs which can in turn have one or more words. An

\(^2\) The manual of Korean-ToBI conventions, and associated sound files are accessible in www.linguistics.ucl.ac.uk/people/jun/sun-ah.htm.

IP is defined by phrase-final lengthening and a boundary tone (e.g. L%, H%, LH%, HL%, LHL%, HLH%, LHLH%) realized on the last syllable of the phrase. It is optionally followed by a pause. An AP is defined by a prosodic tone (LHLH or HLHHL) marking the beginning and the end of the phrase (in Korean ToBI, the AP final tone is transcribed with a diacritic 'u' (e.g. Ha), reflecting the function of the AP boundary marker). An AP has no phrase-final lengthening and is not followed by a pause. The end of an AP is marked by a rising tone (LHa), realized on the last two syllables of the phrase (L on the penultimate and Ha on the final syllable). The beginning of an AP is marked by either a rising tone (LH) or a high plateau (HH) on the two phrase-initial syllables. The tone on the phrase-initial syllable is H when the syllable begins with a tense or aspirated consonant, /h/ or /s/, and L otherwise. The H tones on the AP-initial syllables (the first two Hs in HHLHa) are realized much higher than the H tone after an L tone (the first H in HHLHa), as discussed in Lee (1999). These tone patterns are fully realized when an AP has four or more syllables. However, when an AP has three or fewer syllables, the two medial tones (HL) may be undershot. For L-initial APs, this results in a simple rise (LHa), an early rise (LHLHa), or a late rise (LLHa) pattern. For H-initial APs, the result is a high plateau (HHa) or a fall-rise (HLHa) pattern. The AP final tone is in general High, but is sometimes (11%; data from S. Kim, 2004) realized as Low before an H-initial AP or before an IP-final AP with a L% boundary tone (see below), resulting in a falling pattern (HLa, HLLa, or HHLa), a low plateau (LLa), or a rise-fall (LHLa) AP pattern.

The tones marking an AP are phraseal tones and are not linked to words within a phrase. Thus, the tonal shape of a word changes depending on its location within an AP. For the same reason, a word-initial segment affects the AP initial tone only when the word comes at the beginning of an AP, but not when it comes in the middle of an AP. In general, an AP contains three to four syllables, and when it has more than six syllables forming two words, it splits into two APs (Jun, 2003b; S. Kim, 2004). Thus the most common AP contains only one word (Schafer & Jun, 2000, 2002; S. Kim, 2004).

When a word is contrastively focused, AP boundaries are often deleted between words after focus. When this post-focal dephrasing occurs, an AP can contain multiple words. However, the degree of juncture before the focused word is larger than a default AP boundary and smaller than a default IP boundary. The pitch range of the focused phrase is much larger than that of a default AP, and the phonetic realization of the initial segment of the focused phrase is stronger than that of a default AP, reflecting the hierarchy of prosodic units based on the degree of phrase-initial strengthening (Jun, 1993, Fougeron & Keating, 1997, T. Cho & Keating, 2001). Because of this, and based on data from sentence processing (Jun and Kim, 2004), S.-A. Jun (2004) revised the earlier model and proposed a prosodic unit between an IP and an AP, called an Intermediate Phrase (ip).
An IP in general contains two or three APs, and is defined by either a higher AP-final boundary tone or by a pitch reset between APs, or both. It shows a small degree of phrase-final lengthening. In the case of a syntactically heavy constituent such as a small clause or a heavy XP (e.g. NP, VP) is often marked by an IP boundary, and a large clause boundary is more often marked by an IP boundary. APs within an IP show a downstep-like relationship. That is, the f0 peak of an AP is lowered compared to that of the preceding AP, and the downstep chain is broken (i.e. pitch is reset) at the beginning of a new IP. Note that this downstep is observed only when all the APs within an IP begin with the same type of tone, either H or L, as triggered by the segment type. Further research is needed to define a more general cue to the presence of IP boundaries.

**Intonation of English**

The intonational phonology of English proposed in Beckman and Pierrehumbert (1986) and the English ToBI transcription system summarized in Beckman and Ayers-Elam (1994) posits two prosodic units above the word: An Intonation Phrase (IP) and an Intermediate Phrase (ip). An IP is the highest prosodic unit defined by intonation and can contain one or more Intermediate Phrases. The intonation structure of English is shown in Figure 33.1. An IP is marked by a boundary tone (L% or H%; T% in Figure 33.1; T meaning some tone, L or H), realized on the phrase-final syllable, and an optional High tone at the beginning (%H), realized on the phrase-initial syllable. It is also marked by phrase-final lengthening and is optionally followed by a pause. An IP must contain at least one pitch accent (T*), a pitch excursion realized on a prominent, stressed syllable, and is marked by a phrase accent (T), which is realized on the syllables between the last pitch-accented word and the end of the IP.

There are five pitch accent types proposed in English ToBI: L*, H*, L*+H, L+H*, H+H* (plus versions with down-stepped High tones: !H*, L+!H*, L*+!H). Any of the five pitch accents can come at the beginning of an IP, although the three variants with down-stepped H (H*, L+H*, L*+H) cannot. The starred tone is realized on the stressed syllable of a word, σ in (Figure 33.1), and the tone preceding or following the starred tone (L in L+H* or H in L*+H) is realized immediately preceding or following the stressed syllable. Therefore, the f0 peak of an L+H* accent is realized earlier than the f0 peak of an L*+H accent.

Pitch accents are associated with the stressed syllables of the semantically and pragmatically prominent words in a sentence, and the type of pitch accent conveys the informational status of the pitch-accented item in the discourse (Pierrehumbert & Hirschberg, 1990). Though every word has stress, not every word receives pitch accent. Whether a word receives pitch accent or not is determined postlexically based on the meaning of the utterance. This is different from pitch accent in Tokyo Japanese, where there is only one type of pitch accent (H*+L) and pitch accentedness is a lexical property. Since not every word in English receives pitch accent, words without pitch accent are not specified with a tone, and their pitch values are determined by rule-governed interpolation between the tonal targets of surrounding words.

In English, the last pitch accent of an IP is the most prominent pitch accent within that IP, and is called the nuclear pitch accent (NPA). That is, the IP is the domain headed by the NPA. An IP is also the domain of the NPA derived from focus. When a word is contrastively focused, the word receives an NPA and any pitch accent that might have appeared on a post-focus word in the neutral production of the utterance is deleted, this is known as deaccenting. The words preceding the focused word also show signs of reduced prominence. They either lose their pitch accents or are produced in a reduced pitch range. The focused word is produced with an expanded pitch range, higher amplitude, and longer duration. It may also be separated from surrounding material by a pause before and/or after. Finally, an IP is the domain of downstep. That is, pitch range is reset across IP boundaries.

**Similarities**

The prosodic systems of Korean and English are similar in a few respects. Both languages have at least two prosodic units above the word, and these are marked with...
by intonation. The IP in both languages is marked similarly, by phrase-final lengthening, an obligatory boundary tone, and an optional pause after. Though the number of boundary tones is far fewer in English, some sentence types are marked by the same boundary tones in the two languages. For example, yes/no questions are marked by a high boundary tone while declaratives and imperatives are marked by a low boundary tone.

Though the Korean AP is a prosodic unit larger than a word, its function of marking new/old information is similar to that of the English pitch accent. In Korean, a word with new information comes at the beginning of an AP and a word with old information tends to come in later in an AP (Kang, 1996). In English, a word with new information receives pitch accent and a word with old information tends not to receive pitch accent.

The realization of focus is also similar in both languages. Pitch range is expanded during the focused word and reduced after focus. In Korean, AP boundaries tend to be deleted after focus (dephrasing), and in English, pitch accent tends to be deleted after focus (deaccenting). The domain of dephrasing or deaccenting is an Intermediate Phrase in both languages (assuming the revised model of Korean intonation in Jun, 2004).

**Differences**

One of the biggest differences between English and Korean prosody is that English is a lexical stress language and Korean is not. In English, the prominence of a word is cued by pitch accent which is associated with the stressed syllable of the prominent word. In Korean, the prominence of a word is achieved by placing the word at the beginning of a phrase. Thus, English is known as a 'head' prominence language and Korean an 'edge' prominence language (Jun, 2005, Ch. 16).

Though the Intonation Phrase is defined similarly in English and Korean, the smaller phrases are not. The IP in English has phrase-final lengthening, though not as much as the IP. The ip or AP in Korean has no substantial phrase-final lengthening. The IP in English is marked by a phrase accent whose realization is not localized on the phrase-final syllable, but covers any syllables between the last pitch-accented word and the end of the phrase. However, it is not clear if there is any tone specific to an IP in Korean when it is defined only by pitch reset. The AP in Korean is defined by phrasal tones marking both the beginning and the end of the phrase. Having two tones (H or L) at the beginning of an AP depending on the phrase-initial segment type is unique to Korean intonation. Since most words form an AP by themselves in Korean (Schafer & Jun, 2002; S. Kim, 2004), the association of a tone with a word-initial segment is perceptually very salient (T. Cho, 1996; M.-R. Kim et al., 2002).

---

**Prosysy in sentence processing**

Finally, the pragmatic meaning of a sentence is delivered by the IP boundary tone realized on the phrase final syllable in Korean but by the whole intonation contour (from the combined meaning of pitch accent, phrase accent, and boundary tone) in English. For example, one of the functions delivered by a LHL% boundary tone in Korean is annoyance or irritation. In English, this meaning is achieved by a sequence of L* pitch accent, H* pitch accent, and L-L% boundary tones.

**The role of prosody in sentence processing**

Given the similarities and differences in the prosody of English and Korean, there are similarities and differences in the way prosody influences sentence processing in the two languages. For many spoken sentences in each language, prosodic structure helps to resolve ambiguity at other levels of linguistic analysis. For example, the English sentence in (1) can mean either (a) or (b) depending on the prosodic phrasing of the utterance: a prosodic boundary comes after the girl in (1a), but before the girl in (1b). Similarly, the Korean sentence in (2) means (2a) if a prosodic boundary comes between Soyengi ‘Soyeng-NOM’ and pap ‘a meal’ but means (2b) if there is no boundary there.

1. The hostess greeted the girl with a smile (Lehiste, 1973)
   a. The hostess greeted // with a smile → The hostess smiled
   b. The hostess greeted // the girl with a smile → The girl smiled

2. Soyengi pap mekessin? Soyeng-Nom a meal interrogative ending
   a. Soyengi // pap // mekessin ‘Soyeng, did you eat your meal?’
   b. Soyengi // pap // mekessin ‘Did you eat Soyeng’s meal?’

As shown in (1) and (2) above, an intended syntactic and semantic structure in each language is cued by the prosodic phrasing of the sentence. Accordingly, it has been found that when the boundary of a prosodic unit comes at a place corresponding to a syntactic/semantic group, native speakers of each language take less time in processing the sentence/phrase compared to the case where the prosodic boundary does not match the syntactic/semantic boundary (e.g. Warren et al., 1995; Schafer, 1997; Kjelgaard & Speer, 1999; Speer et al., 1999; Schafer et al., 2000 for English; Schafer & Jun, 2000, 2002; Kang & Speer, 2003; H.-S. Kim, 2004, for Korean). For example, in a cross-modal naming task where subjects complete a sentence after hearing a sentence fragment and seeing a target word (the word immediately following the sentence fragment) on a computer screen, Kjelgaard and Speer (1999) found that, when the target word is is, English speakers complete the sentence fragment shown in (3) much faster, i.e. process faster, when an Intonation Phrase boundary comes after the verb leaves than after the noun the house.
When Roger leaves the house

Similarly, in a cross-modal naming task, Schafer and Jun (2000, 2002) found that native speakers of Korean process a noun phrase (Adjective+NP1 +NP2; e.g. hyoomnyeonghan aktiy appa ‘wise baby’s daddy’) faster when the accentual phrasing of the noun phrase and the semantic/pragmatic meaning of the phrase (e.g. wise // baby’s daddy) match than when they do not (e.g. wise baby’s // daddy). This study shows that Korean speakers are sensitive to the existence of an AP boundary in sentence processing even though, unlike the Intonation Phrase (or the Intermediate Phrase) in English, the Korean AP has no consistent final lengthening. As found in H.-S. Kim and Lee (2004), prosodic phrases realized with strong acoustic cues such as an Intonation Phrase exert more influence on sentence parsing than those marked by weaker acoustic cues. Thus, the English IP and IP, whose boundaries are marked by phrase-final lengthening in addition to the tonal cues, behave similarly in sentence processing, but the Korean AP and IP do not: only the IP is marked by phrase-final lengthening, and IPs have stronger influence than APs on sentence processing.

Schafer and Jun's finding that prosodic phrasing in Korean disambiguates a syntactically ambiguous string has been confirmed by more recent studies in which the domain of ambiguity was larger. Kang and Speer (2003) showed that a globally ambiguous sentence like (4) is disambiguated by prosodic phrasing. Two possible meanings are shown in (4a) and (4b).

(4) Cinwen-ika ywupaytanga-n sancang-ul kongkyekhayssye.
    'Cinwen-NOM got.exiled-REL villa-ACC attacked'

a. ‘Cinwen attacked the mountain villa where (pro) got exiled.’
   b. ‘(pro) attacked the mountain villa where Cinwen got exiled.’

When there was no Intonation Phrase boundary after the subject NP Cinwenika (thus the subject NP and the following verb formed one prosodic unit), speakers overwhelmingly chose the subject NP (82.6%) as the subject of the verb ywupaytanghan (got.exiled), thus interpreting the sentence as (4b).

H.-S. Kim (2004) also showed the role of prosodic phrasing in disambiguation of a larger phrase. She investigated whether a temporarily ambiguous string shown in (5), nanun kkochul coahanun mincleykey, is produced differently depending on the two different structures/meanings of the sentence, (5a) and (5b), and whether speakers use prosody to disambiguate between the different interpretations. In (5a), the second word, kkochul ‘a flower-ACC’, is the object of the relative clause; in (5b) the second word is the object of the main clause. Kim found that the ambiguous string is disambiguated by prosodic phrasing: an IP break after ‘I-top’ and an AP break after ‘a flower’ for (5a), and an AP break after ‘I-top’ and an IP break after ‘a flower’ for (5b). Subjects processed the sentence significantly faster when the prosodic phrasing matched the intended meaning of the structure than when it did not.

(5) nanun kkochul coahanun mincleykey (cangmilil)\textsuperscript{4} cwoessta\textsuperscript{4} ‘I-top flower-ACC like-rel Minc-DAT a rose gave’

a. SR reading: I gave the rose [to Minci who likes flowers]
   b. SOR reading: I gave [the flowers] [to Minci whom ɛ like]

Thus, English and Korean speakers are similar in manipulating prosodic phrasing to deliver the intended meaning and structure of a sentence and in paying attention to prosodic phrasing cues in processing the different meanings.

On the other hand, English and Korean differ in the relation between prosodic features and the meaning in the case of focused wh-pronouns. Schafer et al. (2000) found that when a wh-word is focused (nuclear pitch accentuated), as in (6a), the wh-word is interpreted as an embedded question; the absence of pitch accent on the wh-word, as in (6b), biased listeners to a relative clause reading.

(6) a. I asked the pretty little girl WHO is cold
   b. I asked the pretty little girl who is COLD

The same meaning as (6a) is achieved in Korean by placing the wh-word at the beginning of an Accental Phrase and erasing the Accental Phrase boundary after the wh-word, as shown in (7a). If the AP boundary after the wh-word is maintained as in (7b), the wi-word is interpreted as an indefinite pronoun as described in Jun & Oh (1996). The meaning of (6b) cannot be conveyed in Korean without changing word order and structure.

(7) nanun yeypun sonyeeykey nuka chwupnyako mulessta.
    ‘I-TOP pretty little girl who cold asked’

a. {nanun} {yeypun} {sonyeeykey} {nuka chwupnyako} {mulessta}.
   → ‘I asked the pretty little girl WHO is cold’
   b. {nanun} {yeypun} {sonyeeykey} {nuka} {chwupnyako} {mulessta}.
   → ‘I asked the pretty little girl if there is anyone who is cold’

Future research

Studies examining the role of prosody in sentence processing are relatively new in the field of psycholinguistics. To better understand the mechanism of auditory sentence processing both in general and specifically in Korean, the prosody of more syntactic structures and more varied pragmatic conditions should be

\textsuperscript{4} The word 'cangmilil' is not included for the meaning of (5b).
examined using a range of methodologies. Further, we need to refine the intonational model of Korean to describe the prosodic phrasing and the tonal patterns of utterances in production experiments and to allow us to manipulate prosodic features for processing experiments. This would also help us to pursue syntax–prosody interface studies (e.g. Jun, 1998).

Studies on auditory sentence processing have investigated the effect of prosody on processing by adult native speakers of Korean. A natural extension of this is to investigate the acquisition of Korean prosody by children (as well as by second-language learners) and to study how prosody guides parsing of word and sentence boundaries in Korean (see Jun & Oh, 2000; Y. Choi, 2003; S. Kim 2004; cf. Jusczyk, Cutler & Redanz, 1993; Jusczyk, 1999). Most studies on first-language acquisition have focused on the areas of morphosyntax and phonology; prosody is rarely the focus of such studies. It is known that children acquire intonation before they acquire words (Lewis, 1951; Crystal, 1979). Acquiring a different boundary tone for a different sentence type (e.g. L tone for a declarative or H tone for an interrogative) can be easily observed (e.g. Jun, 2006), but careful study is needed to investigate the pragmatic meaning of different boundary tones (e.g. M.-J. Park, 2003) and the acquisition of prosodic phrasing (e.g. Y. Choi, 2003; Y. Choi & Mazuka, 2003, this volume).

In sum, prosody plays an important role in the area of psycholinguistics as well as other sub-areas of linguistics, and analyzing intonation in the framework of intonational phonology provides a useful tool for comparing prosodic features across languages. It can also help experimenters to manipulate the prosodic features of a language in their experimental design and materials. Further research is needed to explore both the processing of prosody and the role of prosody in processing.

34 Korean sentence processing

Youngjin Kim and Kwangill Choi

One of the most important issues in the study of language processing concerns the universality of sentence processing strategies across different languages (e.g. Fillart, 1998; de Vincenzi, 2000). As all researchers agree, if the goal of psycholinguistics is to study the human sentence processor and not the sentence processing mechanisms of a specific language, the test of universality is not an option but a necessary condition for evaluating various models of sentence processing.

Crosslinguistic comparisons of sentence processing can be attempted from two different perspectives. One is the universality hypothesis. It argues that there are universal processing strategies that apply to all languages, because processing strategies are independent of specific languages and are based on cognitive universals. The strategies or principles that could be considered as being universal are minimal attachment and late closure (Frazier, 1987) as well as the minimal chain principle (de Vincenzi, 1991). These principles are assumed to operate in all languages, with only the ‘vocabulary,’ i.e. lexical items and specific grammar, differing across languages.

Another reason to do crosslinguistic studies might be the expectation that different languages show different processing strategies, under an assumption that parsing strategies are a reflection of language-specific characteristics and a by-product of exposure to a given language. Some processing strategies may not be universal but are instead language specific or parameterized (e.g. Mazuka, 1998). So, there is no need, at least in principle, to test the same strategy in different languages. From this perspective, first of all we should try to find important grammatical aspects or structures that are unique to a language and then examine their sentence processing implications. For example, in English, word order is most typically used for the basic syntactic function of indicating ‘who did what to whom.’ In Korean, on the other hand, case markers on words indicate the syntactic function, and word order is used mostly for highlighting and backgrounding information semantically or pragmatically. Therefore, the role of word order information in syntactic parsing procedures may be different from language to language. Our job may be to combine all the findings from studies of different structures of