

Prosody-Based Word Recognition for L2 Speakers

GWANHI YUN
Daegu University

0. Introduction

Previous studies in word recognition have attempted to explain how adults can recognize words or discriminate a word from other lexical alternatives quickly and accurately from speech signals (Grosjean 1980, Marslen-Wilson and Tyler 1980, Warren and Marslen-Wilson 1987). One of the most prevalent models, i.e., the "cohort model", proposes that word recognition involves two steps. First, when L1 listeners hear phonetic input through some duration of time, multiple words similar to such signals are activated. Next, as more segmental input induces listeners to select the appropriate word out of the candidate words. Under this model, left-to-right evaluation of segmental information is crucial to recognize words (Marslen-Wilson 1987). Recently, however, such segment-based lexical processing turned out to be insufficient for L1 adults' word differentiation. Thus an increasing number of studies have focused on the role of prosody in word recognition for L1 speakers.

As is emphasized under two major models on word recognition, i.e., the "cohort model" and the "neighborhood activation model", two controversial issues in the role of prosody in word recognition are: what prosodic cues are crucial for segmentation? And what the unit of perception is between syllables and phonemes? First, let us turn to the former issue. Listeners decode speech signals with reference to a variety of prosodic cues to detect word boundaries. For instance, listeners rely on intonation patterns to locate word boundaries (Taylor et al. 1998). Furthermore, the longer the portion of the word listeners hear, the more accurately they are able to segment it (Lee 1999, Hudson and Bergman 1985). Recently, a stress-based model has been in the spotlight for word identification (Cutler 1994, Cutler and Butterfield 1992, Cutler and Norris 1988). Nakatani and Schaffer (1978) show that English adult listeners assign word boundaries before stressed syllables. Cutler and colleagues show that even infant listeners segment words on the basis of stress cues.

Another issue as to the role of prosodic elements in word recognition centers on the unit of perception. That is, what is the basic unit when listeners recognize words? This tackles the question of whether the unit for lexical representation is a syllable or a phoneme. Cutler et al. (1986) report that French listeners are sensitive to syllables. Additionally, Pitt and Samuel's (1990) experiment reveals that

English listeners recognize the location of phonemes as well as stressed syllables within words. Pallier (1997) shows that even French listeners can recognize the exact location of a specific phoneme like English listeners. Combined with previous studies, this evidence suggests that L1 speakers' mental lexicon is equipped with syllabically-structured representations based on phonemes.

Compared to ample studies on word recognition by L1 speakers of English or French, there has been comparatively little focus on L2 listeners' sensitivity to stress for word identification in other target languages. Hence the current study explores whether Korean L2 (English) listeners can effectively differentiate or retrieve the original words on the basis of the initial syllable or foot containing primary or secondary stressed vowels. In effect, this study tests whether a bigger prosodic size than syllables can be the unit of word recognition.

1. Goals

The present study investigates whether and how Korean L2 learners of English recognize or differentiate a word from others on the basis of subtle stress differences in English. It also tests whether the principle of a cohort model applies to L2 listeners' word recognition.

The focus of this study is essentially to answer two questions. First, how do Korean L2 speakers produce primary- and secondary-stressed vowels in English? This study examines whether Korean speakers use three major phonetic correlates of stress, i.e., vowel duration, amplitude, and F0 to distinguish the delicate stress differences. If they produce different primary and secondary stressed vowels, it is highly likely that they might also be sensitive to acoustic signals enough to perceive such differences on the perception side. Furthermore, it might induce listeners to activate the source word solely on the basis of stress information. The present study was partially stimulated by Mattys' (2000) study conducted for native speakers of English. We would like to see if Korean speakers replicate similar results with an almost identical set of English stimuli for Korean L2 speakers.

The second question to be addressed is whether Korean L2 listeners restore or guess the original source words on the basis of stress differences between primary and secondary stressed syllables. This study aims to explore whether two effects found for native speakers of English are evident for Korean speakers of English (Mattys 2000): fragment size effect and stress degree effect. As for the fragment size effect, it is predicted that syllable length might contribute to the degree of restoration of English words since more segmental information coupled with foot information can lead to more accurate activation of words (e.g. [prá] vs. 'prose' [[prásɪ] in 'prosecution' or 'prosecutor'). That is, the longer the fragment they hear, the faster and the more accurately they will identify the source words. Next, as for stress degree effect, it can be expected that Korean listeners have difficulty in accurately differentiating the source words solely on the basis of subtle stress cues

without reference to full segmental information, probably due to lack of sensitivity to the stress system of English. Of course, the possibility cannot be excluded that their perceptual sensitivity might vary depending on individuals' stress knowledge.

2. Production and Perception Experiments

Subjects: Thirty-nine English-learning Korean speakers participated in the perception experiment. Twenty-four of thirty-nine participated in the production experiment. All were junior or senior undergraduates recruited at Daegu University with either an English major or minor. Twenty-four were females and fifteen were males. They ranged in age from 21 to 26 years, with an average age of 24. Their mean TOEIC scores were 665 points, and their self-rating of their own English proficiency was 5.5 out of 10 scales. On this basis, they were of approximately intermediate level in English. Before the experiment, they were asked to mark primary stress on the stimuli English words. The average score for females was 68% and for males 55%. They had an average 12 years of experience learning English with no known hearing or production problems. All the subjects were paid for participation.

Stimulus Materials: Because this study examines whether Korean L2 learners of English show native-like patterns of word recognition, test materials used are identical to those employed in Mattys (2000). Test stimuli consisted of 24 pairs with 4 syllables. One group of words has primary stress on the initial syllable and secondary stress on the third syllable as illustrated in (1a). The other group of words has secondary stress on the initial syllable and primary stress on the third syllable as exemplified in (1b). Both groups of words consisted of the identical segments within the first foot or the first two syllables, but they had different stress patterns. The first or second syllable was (C)(C)V(C) (e.g., *prosecutor*, *delegating*, *consequently*, *aggravating*, *replicating*, *category*). The vowels in the initial syllable were /ɪ, ε, æ, α/. Additional twenty filler words were interpolated with target words. 10 pairs of words had primary stresses on the first syllable and vowel qualities were identical (e.g., *móment/mómentary*, *célibate/célibacy*, etc). A total of 696 tokens (24 pairs x 29 subjects) were taken for production and perception analyses.

(1) Test stimuli

a. Initial-Primary	b. Initial-Secondary
prósecùtor	pròsecútion
délegàting	délegátion
présidèncy	prèsidéntial
cátègòry	càtegórical
cónsequèntly	cònsequéntial
návigàtor	nàvigátion
víndicàting	víndicátion

Gwanhi Yun

fábricáting	fàbricación
ségregàting	sègregación
réplicàting	rèplicación
ágravàting	àgravación
céremòny	cèremónial

Procedures: First, for a production experiment, each of the target tokens was placed in a carrier sentence "Please say _____" designed to elicit natural production of the real words. First, in order to see whether/how Korean speakers realize stress differences like native speakers of English, twenty four Korean learners were asked to read the carrier sentences with target words. To compare Korean speakers' production with English speakers' and to replicate previous findings, five native speakers of English also participated in the identical production experiment. The stimuli were presented in a random order. Subjects were recorded with PRAAT, and the sentences were digitized at 44kHz sampling rate. In order to explore usage of phonetic correlates of stress, three acoustic parameters were measured: vowel duration, F0 of stressed vowels, and amplitude of vowels. Vowels were measured from the beginning of the first formant to the end of the second formant in the spectrogram. F0 and the amplitude were measured roughly at the mid-point of the whole vowel in time dimension. Since stress differences were examined within subjects, durational and spectral measures were subjected to a one-way repeated-measures ANOVA through PAWS Statistics (SPSS) 18.

Second, the perception experiment proceeded in two blocks. To obtain listening stimuli, one native speaker's production of the stimuli in (1) were recorded. The listening stimuli were created by excising the first syllable ((C)(C)V(C)) or two syllables from the original source words. In one block, thirty nine Korean listeners were asked to listen to one syllable with primary or secondary stress (e.g., [prá], [prà], [pré], [prè], etc). Next, they saw two forced choices on the computer screen (e.g., ① prosecutor ② prosecution), and then they had to guess the original source and press 1 or 2 on the keyboard. With the forced choices, the items in column number one were the words with initial primary stress while those in column number two were those with initial secondary stress. In the second block, in order to see the effect of fragment size on word identification, the same subjects heard the first two syllables consisting of the primary stress and unstressed syllables or of the secondary stress and unstressed syllables (e.g., [prásɪ], [pràsɪ], etc.). On the basis of the first foot, they had to perform the same guessing task as in block I. They were instructed to press 1 or 2 when they saw the multiple choice (e.g., ① prosecutor ② prosecution).

Listening stimuli and randomizing stimuli were run using the software E-Prime. In both the blocks, stimuli were presented binaurally over headphones at a comfortable listening level. They were instructed to make fast and accurate

responses by pressing the appropriate keys as soon as they had made a decision. A pause of 2000ms timeout after stimulus offset with a word "Next" on the screen. The test session began after 5 trials. Thus a total of 3744 responses were obtained and analyzed within these two blocks (24 pairs \times 2 blocks \times 39 subjects).

Accuracy and reaction times were measured and obtained through E-Prime, in order to see how accurately and how fast Korean listeners activate the source words on the basis of short or long fragments without the full segmental information. Those two parameters were evaluated within each block. Furthermore, since those were examined within subjects, accuracy and reaction times were subjected to repeated-measures ANOVA.

Statistical analyses examined two effects: (i) the effect of stress difference on the accuracy and reaction times, and, (ii) the effect of fragment size on the accuracy and reaction times.

3. Results

3.1. Korean and English speakers' production of stress in English

Table 1 shows average values of three phonetic correlates, i.e., duration, F0, and intensity of primary stressed and secondary stressed vowels produced by Korean learners. A one-way repeated-measures ANOVA with a factor of stress (primary vs. secondary) was performed. First, primary stressed vowels in word initial syllables were on average 9 ms longer than secondary stressed ones (e.g., [prá] in 'prosecutor' vs. [prá], 89 ms vs. 78 ms). The difference statistically reached marginal significance ($F(1,24)=3.6$, $p=.06$). However, of interest is the finding that primary stressed vowels were 17 ms shorter than secondary stressed vowels in the third syllables as well (e.g., [kjù] in 'prosecutor' vs. [kjù] in 'prosecution', 134 ms vs. 146 ms), and the difference between the durations of primary and secondary stressed vowels was also marginally significant ($F(1,24)=3.69$, $p=.06$). Furthermore, both the initial and the third syllables had the equal F0 between primary and secondary stressed vowels (initial, 178 Hz vs. 178 Hz, $F(1,24)=0.01$, $p>0.05$; third, 164 Hz vs. 162 Hz, $F(1,24)=0.6$, $p>0.05$). Finally, primary stressed vowels were as loud as secondary stressed ones in word-initial syllable (73 dB vs. 73 dB, $F(1,24)=0.7$, $p>0.05$), but the difference in the third syllable was highly significant (72 dB vs. 71 dB, $F(1,24)=14.6$, $p=0.001$).

Table 1. Korean speakers' production of stress differences

	1 st syllable			3 rd syllable		
	Pri.strs	Sec.strs	F & p-val	Pri.strs	Sec.strs	F & p-val
Dur. (ms)	89	78	F(1,24)=3.6 P=.06*	134	146	F(1,24)=3.6 P=.06*
F0 (Hz)	178	178	F(1,24)=.01 P>.05	162	164	F(1,24)=.6 p>.05
Int.(dB)	73	73	F(1,24)=.7 p>.05	71	72	F(1,24)=1.6 P=.001***

In summary, in word-initial syllables, difference between primary and secondary stressed vowels was significant only in duration of the vowels, while the F0 and intensity differences were neutralized. Furthermore, primary and secondary stressed vowels were different in duration and intensity in the third syllables even though the latter had longer and higher intensity than the former unlike native speakers of English. These results indicate that durational difference was kept as a major stress cue for Korean learners unlike native speakers who use all three phonetic cues to differentiate the degrees of stress. They also seem to suggest that temporal properties like duration might be less marked for acquisition of English stress than spectral cues like F0 or amplitude. Mattys' (2000) preliminary production study shows that initial primary stressed vowels are higher in F0, longer and louder than initial secondary stressed vowels for native speakers of English. These differences have long been found to hold true between stressed vowels and unstressed vowels for native speakers of English (Beckman 1986, Fry 1958, Lehiste 1970). In a nutshell, it seems that Korean speakers use fewer cues than native speakers of English to differentiate subtle degree of stress in English on the production side.

In order to confirm and replicate previous results found for native speakers, the present study obtained 5 English native speakers' production. Table 2 shows mean durations, F0s, and intensity for primary and secondary stressed vowels. As illustrated, effects of all three cues are evident only in the initial syllable, but not in the third syllables. Specifically, primary stressed vowels were on average 5 ms longer than secondary stressed ones ($F(1,5)=8.8$, $p<0.05$) and were 5 Hz higher than secondary stressed one ($F(1,5)=14$, $p<0.05$). Furthermore, there was a significant difference in loudness ($F(1,5)=1.9$, $p<0.05$). This finding confirms stress-sensitive acoustic differences for native speakers of English.

Table 2. Vowel duration, F0, and intensity for native speakers of English

	1 st syllable		3 rd syllable	
	Pri.strs	Sec.strs	Pri.strs	Sec.strs
Dur. (ms)	91	86	118	123
F0 (Hz)	146	141	138	141
Int.(dB)	80	79	75	77

Putting together the results in Table 1 and 2, we see that vowel duration, F0, and intensity are crucial stress cues enough to judge whether a syllable is primary or secondary stress for native speakers of English, especially for the initial position, while primary vowel duration is employed as a potential cue for Korean speakers.

3.2. Korean Listeners' Word Restoration

Table 3 exhibits accuracy of restoration of the source words as a function of stress (primary vs. secondary) and size of auditory stimuli (mono- vs. di-syllables) obtained for thirty eight Korean L2 listeners. In order to determine whether stress and stimuli size on affects the accuracy of word restoration, a two-way repeated-measures ANOVA was conducted. Analysis found significant main effects of stress ($F(1,37)=7.826$, $p<.001$) and fragment size ($F(1,37)=5.643$, $p=0.02$), but interaction of stress and stimuli size reached no significance ($F(1,37)=0.846$, $p>0.05$). Specifically, as illustrated in Table 3, Korean L2 listeners identified the source words more accurately when they heard initial primary-stressed (one or two) syllables than when they heard secondary-stressed counterparts (55% vs 44%). This finding suggests that Korean listeners were sensitive to stress difference to some extent to distinguish primary and secondary stress and to restore words differently on the basis of stress cues. Next, accuracy of word identification was substantially higher when the auditory stimuli were initial two syllable feet than monosyllables (51% vs. 48%). This indicates that longer fragments might have contributed to the restoration of source words slightly more than shorter ones.

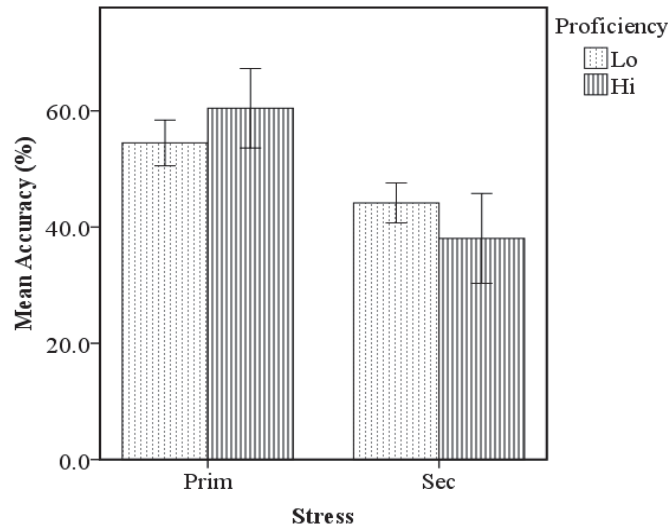
Table 3. Accuracy of identification of the source words as a function of stress pattern and stimuli size (%)

Stress	Initial-primary stress		Initial-secondary stress	
	1 st syl.	1 st & 2 nd sylls.	1 st syl.	1 st & 2 nd sylls.
Females	56.2	54.8	43.1	48.9
Males	53.2	56.4	40.8	44.2
All	55.2	55.2	42.5	47.7

However, within initial primary stress condition, fragment size effects were not found ($F(1,37)=0.07$, $p>0.05$), whereas such effects were significantly marginal within secondary stress condition (42.5% vs. 47.7%, $F(1,37)=2.9$, $p=0.09$) as shown in Table 3.

In order to see whether L2 proficiency is related to Korean listeners' sensitivity to stress and/or word restoration, accuracy results were classified by dividing 30 participants into two groups (high-proficiency group with TOEIC scores higher than 700; low-proficiency group with less than 700). A two-way (proficiency \times stress) repeated-measures ANOVA was performed. Analysis displayed no main effect of proficiency (49% vs. 49%, $F(1,14)=0.003$, $p>0.05$), but stress effect approached significance (57% vs. 41%, $F(1,14)=17.13$, $p=0.001$). Interaction of proficiency and stress reached marginal significance ($F(1,14)=3.324$, $p=0.09$). This result indicates that Korean learners' restorability of the source words was not affected by their L2 proficiency level characterized by TOEIC scores. In other words, whether more or less advanced learners, they were sensitive to stress difference and were able to identify the source words by way of primary stress cues rather than secondary stress cues. Figure 1 depicts the interaction between proficiency and stress.

Figure 1. Interaction between proficiency and stress



Finally, in order to see whether Korean listeners' restorability of the source words is affected by vowel quality of the initial primary- or secondary-stressed syllables, a one-way ANOVA was conducted. Analysis displayed no significant effect of vowel quality on the accuracy of word restoration ($F(3,155)=0.38$, $p>0.05$). Specifically, whether the stressed syllables varied among /ɪ, ɛ, æ, ɑ/ (e.g., [vɪndɪ] 'vindication', [dɛlɔ] 'delegating', [nævɪ] 'navigator', [kɑnsɪ] 'consequently'), accuracy of word detection did not vary.

Table 4 exhibits reaction time when it took to judge the stimuli or restore the source words on the basis of initial one or two syllables by stress and stimuli size. To examine whether stress difference and stimuli size affect reaction time, a two-

way repeated-measures ANOVA was performed. Analysis showed that there were no main effects of stress ($F(1,36)=0.0$, $p>0.05$) and no interaction of stress and fragment length ($F(1,36)=1.29$, $p>0.05$). However, the effect of fragment size reached significance ($F(1,36)=89.7$, $p<0.001$). First, these findings indicate that stress difference was not conducive to rapid judgment. Specifically, it took almost the same amount of time to identify the source words whether the stimuli contained primary- or secondary-stressed vowels. Second, they suggest that the longer the fragments were, the faster the source words were identified although fast reaction time did not contribute to more accuracy (e.g., [rɛ̃plɪ] vs. [rɛ̃] 'replicating' or 'replication', 2.9 sec vs. 3.7 sec).

Table 4. Reaction time for identification of the source word as a function of stress pattern and stimuli size

Stress	Initial-primary stress		Initial-secondary stress	
	1 st syl.	1 st & 2 nd sylls.	1 st syl.	1 st & 2 nd sylls.
Females	3.8	3.8	3.7	3.0
Males	3.4	2.6	3.5	2.8
All	3.7	2.8	3.6	3.0

Such effect of fragment size on reaction time was found for stimuli with initial primary stress (e.g., [nævɪ] vs. [næ] 'navigator', 2.9 sec. vs. 3.7 sec., $F(1,36)=82.32$, $p=0.001$). It also took faster to judge longer fragments, i.e., initial foot with secondary-stressed vowels than shorter ones, i.e., initial monosyllables (e.g., [nævɪ] vs. [næ] for 'navigation', 2.9 sec. vs. 3.7 sec., $F(1,36)=64.22$, $p<0.001$). Figure 3 clearly illustrates no interaction between stress and fragment size. The main effect of fragment size, no effect of stress and no interaction between those factors were found for both females and males (females, stress, $F(1,22)=0.11$, $p>0.05$; fragment size, 2.9 sec. vs. 3.7 sec., $F(1,22)=57$, $p<0.001$; interaction, $F(1,22)=0.94$, $p>0.05$; males, stress, $F(1,13)=0.39$, $p>0.05$; fragment size, 2.8 sec. vs. 3.6 sec., $F(1,13)=30.46$, $p<0.001$; interaction, $F(1,13)=0.31$, $p>0.05$).

In summary, reaction time was significantly affected by fragment size, not by stress for Korean L2 listeners; however faster reaction times did not guarantee higher accuracy of source words.

3.3. Native English Speakers' Perception of stress

In order to compare the effects of stress and fragment size on accuracy of the source words and reaction time obtained for Korean L2 listeners with those of native English speakers, a two-way repeated-measures ANOVA was used. Analysis displayed a marginally significant main effect of stress ($F(1,3)=6.5$, $p=0.08$) and no main effect of fragment size ($F(1,3)=2.19$, $p>0.05$). Furthermore, interaction of stress \times fragment size did not approach significance ($F(1,3)=0.53$,

$p > 0.05$). As illustrated in Table 5, stimuli with primary stress were restored to the source words more than those with secondary stress (77.7% vs. 27.7%). This shows that native English speakers tend to resort to primary stress more than secondary stress. However, accuracy was not significantly different regardless of shorter or longer fragments (50.6% vs. 54.8%).

Table 5. Mean accuracy for identification of the source words as a function of stress pattern and stimuli size (%)

Stress	Initial-primary stress		Initial-secondary stress	
	1 st syl.	1 st & 2 nd sylls.	1 st syl.	1 st & 2 nd sylls.
Fragment size	74.8	80.8	26.4	29
Average	77.7		27.7	

Next, in order to explore subject variation, accuracy was submitted to one-way ANOVA for stress or fragment size effect within each subject condition. As clearly demonstrated in Table 6, the initial one or two syllables with primary stress played a role as more reliable cues in restoring the source words than those with secondary stress for three out of four native speakers, i.e., E1, E2, and E3. As for E4, there was no significant effect of stress on accuracy. In addition, there were no significant differences in reaction time for fragment size for all four native English listeners. That is, longer stimuli did not contribute to higher accuracy of word detection. This finding suggests that the second syllable following the initial primary- or secondary-stress did not contain pivotal information enough to induce listeners to identify words and more segmental information alone does not guarantee more accurate word restoration.

Table 6. Mean accuracy for identification of the source words as a function of stress pattern and stimuli size (sec.)

speakers	stress		Fragment size	
	primary	secondary	1 syllable	2 syllables
E1	82	35	77	77
	F(1,32)=45.3, $p < 0.001$ ***		F(1,57)=0.0, $p > 0.05$	
E2	71	44	65	74
	F(1,34)=13, $p = 0.001$ **		F(1,72)=2.2, $p > 0.05$	
E3	97	25	96	96
	F(1,24)=196, $p < 0.001$ **		F(1,47)=0.0, $p > 0.05$	
E4	61	58	66	73
	F(1,44)=0.16, $p > 0.05$		F(1,75)=1.4, $p > 0.05$	

Next, in order to explore the effects of stress and fragments size on reaction time that took to identify the source words, a two-way repeated-measures ANOVA was performed. Analysis revealed that there were no main effect of stress ($F(1,3)=0.68$, $p > 0.05$) and no interaction of stress and fragment size ($F(1,3)=1.34$, $p > 0.05$). However, main effect of fragment size reached a marginal-

ly significant effect ($F(1,3)=7.67$, $p=0.70$). Specifically, although stress differences affect the accuracy, they did not influence how long it took to restore the source words for native speakers of English.

Effects of stress and fragment size on reaction time within each subject were explored by a two-way ANOVA. Analysis showed that only fragment size affected reaction time for three native speakers of English, i.e., E1, E2, and E3. Stress effect and interaction effect were not found for any native listeners. Specifically, for E1, E2, and E3, it took faster to judge longer fragments (i.e., initial two syllables, [sɛrɪ] in 'ceremony') than shorter ones (i.e., [sɛ] in 'ceremony').

In summary, despite a relatively small number of native speakers of English, overall, stress but not fragment size played a crucial role for word detection. That is, initial syllables or feet with primary stress were restored to the source word substantially more accurately than those with secondary stress (77.7% vs. 27.7%). On the other hand, not stress but fragment size alone influenced reaction time which took to judge the auditory stimuli. Specifically, it took faster to restore the source words from the longer fragments than from the shorter ones although the former did not guarantee higher accuracy over the latter (2.7 sec. vs. 3.4 sec).

Compared with the results obtained for Korean L2 listeners, native English listeners restored the source words even more accurately on the basis of initial primary-stressed signals than secondary-stressed ones, implying that they have perceptually greater sensitivity to stress differences than Koreans (77.7% vs. 55.2%). Similar to Korean listeners, native speakers also made judgments on the longer fragments faster than on the shorter ones (English 2.7 sec. vs. 3.4 sec; Korean 2.9 sec. vs. 3.7 sec.). Overall, stress effect was found for accuracy, and fragment size effect obtained for reaction time for both native speakers of English and Korean L2 listeners.

4. Discussion

This study investigated whether Korean L2 learners of English produce subtle differences between primary and secondary stressed vowels found for native speakers of English and whether they can restore the source English words solely on the basis of one or two-syllable fragments containing primary- and secondary-stressed vowels without reference to the full segmental information.

First, Korean L2 speakers produced primary-stressed vowels with longer duration than secondary ones, particularly in the first syllable but not in the third syllable. However, F0 and amplitude were differently realized enough to distinguish subtle stress differences. On the other hand, results obtained for five native speakers of English in the current study replicated previous findings (Beckman 1986, de Jong and Jawaydeh 2002, Lehiste 1970, Mattys 2000). That is, so-called three major phonetic correlates of stress, i.e., vowel duration, F0, and amplitude were all significantly different between primary- and secondary-stressed vowels, particularly in the first syllable but not in the third syllable. Furthermore, another difference is that difference in vowel duration was 5 ms for Korean speakers,

while that was 11 ms for native English speakers. These findings provide several implications on the acquisition of English stress by Korean L2 speakers. First, from the fact that vowel duration is only a reliable indicator of stress degree for Korean speakers, it might be the case that temporal properties like duration is less marked for L2 acquisition than the other spectral cues like F0 and amplitude. Second, it implies that L2 learners seem to be at a stage of acquiring vowel duration as a phonetic cue of stress which does not exist as a phonemic entity in their L1 Korean, supporting Kuhl and Iverson (1995). Such differences in availability of phonetic cues of stress might account for differences in perceptual sensitivity to stress between Korean L2 and native English listeners.

Second, results in the current experiment showed that accuracy of restoration of the source words was affected by stress like native speakers of English employed in the current study as well as those in Mattys (2000). Specifically, the source English words were identified more accurately on the basis of initial primary-stressed than secondary-stressed syllables. However, in the present study, reliability of primary stress was much greater for native speakers of English than Korean L2 listeners. Interestingly, the accuracy for Korean listeners is similar to that for English listeners in Mattys' (2000) study. Comparing the current study with Mattys (2000), even native speakers did not show consistent patterns of perceptual sensitivity with regard to primary- and secondary-stressed syllables.

In any case, the results in our study provide interesting implications as to L2 word recognition. First, it seems that the first syllable or foot with primary stressed vowels plays a more crucial role in activating the source words than that of secondary stressed vowels even for L2 speakers, although their perceptual sensitivity is relatively lower than native speakers of English. Second, it suggests L2 speakers' mental lexicons contain representations with prosodic as well as segmental information. Finally, our results lend additional support to prosody-based word recognition even for L2 listeners in that they can differentiate the subtle nuance in English stress and identify the source words on the basis of such clues, being segmental information available equal (Mattys and Samuel 1997, Vroomen and de Gelder 1997). This is evidenced in that accuracy was significantly higher for primary-stressed than secondary-stressed signals, whereas longer fragments did not elevate the accuracy over shorter ones. Of course, longer fragments definitely assisted listeners to judge the stimuli and guess the source words faster than shorter ones. Nevertheless, more segmental information did not guarantee higher accuracy of restoration of the source words, implying that segmental information alone is not sufficient enough to activate the appropriate words for L2 listeners as well as L1 listeners. For that reason, it makes sense to infer that Korean L2 learners seem to be equipped with lexical representation containing stress as well as segmental information.

Third, it is worthwhile to interpret our results in terms of the cohort model for L2 listeners. As previously mentioned, under the cohort model (Grosjean 1980, Marslen-Wilson and Tyler 1980, Marslen-Wilson 1987), words can be recognized from partial acoustic input before all the segments of a word are heard. Listeners

can discriminate a word from others as soon as available information is given. When listeners hear a portion of word in a left-to-right fashion, they activate multiple words on the basis of what they hear. As they hear more and more segments, they identify the words. This model for L1 listeners seems to accord with the results obtained for Korean L2 listeners. They can discriminate one word from another (e.g., 'navigator' vs. 'navigation'). Furthermore, the accuracy of identification increased as they heard longer fragments (e.g., for Koreans: [næ]'na' vs. [nævɪ] 'navi' => 48% vs. 51%). However, for native speakers of English, mere longer fragments did not significantly contribute to higher accuracy (50% vs. 54%). These findings suggest that the segmentally-based word recognition model, i.e., cohort model does not present a perfect account for word recognition for L1, but instead does more for L2 listeners. Thus, following Walley (1993), we adopt a revised version of the cohort model where listeners' lexical representation and processing may not be exclusively segmentally-based but rather holistic to handle both L1 and L2 listeners. That is, it might be the case that listeners incorporate segmental and prosodic information to identify and restore the speakers' intended words.

5. Conclusion

The current study showed that on the production side, only vowel duration was significantly different between primary- and secondary-stressed syllables for Korean L2 speakers. On the other hand, it replicated the previous findings for native speakers of English by showing that vowel duration, F0, and amplitude are reliable indicators of stress degrees for native speakers of English. On the perception side, subtle differences in stress played a crucial role as a perceptual cue to differentiate one word from another word with the same segmental information for Korean L2 listeners although their perceptual sensitivity was lower than native speakers of English. However, fragment size did not affect the accuracy of identification of the source words. Regardless of stress differences, fragment size affected the reaction time it took to judge the stimuli, showing that it took faster to judge the longer fragments than the shorter ones for both Korean and English speakers. These results suggest that a prosody-based word recognition model holds true for L2 learners as well as for L1 listeners. In addition, it is likely that L2 listeners activate words on the basis of segmental information coupled with prosodic information.

References

- Beckman, Mary. 1986. *Stress and Non-stress Accent*. Dordrecht: Floris.
Cutler, Anne. 1994. Segmentation Problems, Rhythmic Solution. *Lingua* 92: 81-104.

- Cutler, Anne, and David Carter. 1987. The Predominance of Strong Initial Syllables in the English Vocabulary. *Computer Speech and Language* 2: 133-142.
- Cutler, Anne, and Dennis Norris. 1988. The Role of Strong Syllables in Segmentation for Lexical Access. *Journal of Experimental Psychology: Human Perception and Performance* 14: 113-121.
- Cutler, Anne, and Sally Butterfield. 1992. Rhythmic Cues to Speech Segmentation: Evidence from Juncture Misperception. *Journal of Memory and Language* 31: 218-236.
- Cutler, Anne, Jacques Mehler, Dennis Norris, and Juan Segui. 1986. The Syllable's Differing Role in the Segmentation of French and English. *Journal of Memory and Language* 25: 385-400.
- de Jong, Kenneth and Bushra Zawaydeh. 2002. Comparing Stress, Lexical Focus, and Segmental Focus: Patterns of Variation in Arabic Vowel Duration. *Journal of Phonetics* 30: 53-75.
- Fry, Butler. 1955. Duration and Intensity as Physical Correlates of Linguistic Stress. *Journal of the Acoustical Society of America* 27: 765-768.
- Grosjean, Francois. 1980. Spoken Word Recognition and the Gating Paradigm. *Perception & Psychophysics* 28: 267-283.
- Hudson, Patrick, and Marijke Bergamn. 1985. Lexical Knowledge in Word Recognition: Word Length and Word Frequency in Naming and Lexical Decision Tasks. *Journal of Memory and Language* 24: 46-58.
- Kuhl, Patricia, and Paul Iverson. 1995. Linguistic Experience and the "Perceptual Magnet Effect." In Winifred Strange ed., *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-Language Speech Research*, 121-154. Baltimore: York.
- Lee, Chang Hee. 1999. A Locus of the Word-length Effect on Word Recognition. *Journal of Reading Psychology* 20: 129-150.
- Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, MA: MIT Press.
- Marslen-Wilson, William. 1987. Functional Parallelism in Spoken Word Recognition. In U. H. Frauenfelder & L. K. Tyler (Eds.), *Spoken word recognition*. Cambridge, MA: MIT Press.
- Marslen-Wilson, William, and Lorraine Tyler. 1980. The Temporal Structure of Spoken Language Understanding. *Cognition* 8: 1-71.
- Mattys, Sven L. 2000. The Perception of Primary and Secondary Stress in English. *Perception and Psychophysics* 62(2): 253-265.
- Mattys, Sven L, and Arthur Samuel. 1997. How Lexical Stress Affects Speech Segmentation and Interactivity: Evidence from the Migration Paradigm. *Journal of Memory & Language* 36: 465-494.
- Nakatani, Lloyd, and Judith Schaffer. 1978. Hearing 'Words' Without Words: Prosodic Cues for Word Perception. *Journal of the Acoustical Society of America* 63: 234-245.
- Pallier, Christophe. 1997. Phonemes and Syllables in Speech Perception: Size of Attentional Focus in French. In *Proceedings of Eurospeech '97* (4): 2159-2162.

Prosody-Based Word Recognition for L2 Speakers

- Pitt, Mark, and Arthur Samuel. 1990. The Use of Rhythm in Attending to Speech. *Journal of Experimental Psychology: Human Perception and Performance* 16: 564-573.
- Taylor, Paul, Simon King, Stephen Isard, and Helen Wright. 1998. Intonation and Dialog Context as Constraints for Speech Recognition. *Language and Speech* 41(3-4): 493-512.
- Vroomen, Jean, and Beatrice de Gelder. 1997. Trochaic Rhythm in Speech Segmentation. Paper presented at the 38th Meeting of the Psychonomic Society, Philadelphia.
- Warren, Paul, and William Marslen-Wilson. 1987. Continuous Uptake of Acoustic Cues in Spoken Word Recognition. *Perception & Psychophysics* 41: 262-275.
- Walley, Amanda. 1993. The Role of Vocabulary Development in Children's Spoken Word Recognition and Segmentation Ability. *Developmental Review* 13: 286-350.

Department of English Language and Literature
Daegu University
201 Daegudaero, Jillyang, Gyeongsan, Gyeongbuk
Korea 712-714

ghyun@daegu.ac.kr