Acoustic Cues of Korean Nuclear Tones

Ho-Young Lee

The purpose of this paper is to discuss what acoustic cues are significant in distinguishing Korean nuclear tones acoustically. It is shown that target frequencies, direction of target frequencies, F0 of the beginning and ending points, timing of peaks and valleys, duration of the rising and falling phases, duration of the whole tone, and F0 of the preceding syllable function as acoustic cues.

Peak frequency was crucial in distinguishing High Fall from Low Fall and Full Rise from Low Rise. F0 of the beginning point was significant in distinguishing High Fall from Low Fall, Low Rise from Mid Level, and High Level from Full Rise after a voiceless consonant. And F0 of the ending point was important in distinguishing High Fall from Low Fall. Direction of target frequencies was significant in distinguishing High Fall from Full Rise and Low Rise from Low Fall.

Duration of the whole tone was significant in distinguishing Low Fall from Rise Fall and High Level from Full Rise. The former pair of nuclear tones was also distinguished by Peak timing and the latter pair by valley timing. The F0 of the penultimate syllable was crucial in distinguishing Low Fall from Low Level.

1. Introduction

The greater part of the intonational meaning of a tune is conveyed by the intonation pattern imposed on the final syllable of the intonation phrase in Korean. I call this intonation pattern the **nuclear tone** (Lee 1990, 1991, 1997, 1999)[1], following the British tradition of intonational analysis (O'Connor and Arnold 1973, Cruttenden 1997). On the other hand, Jun (1993, 1999) and Jun and Oh (1996) call this intonation pattern the **boundary tone**, following the Pierrehumbertian intonational phonology (Pierrehumbert 1980,

[1] In Lee (1990, 1991), I called the intonation pattern of the intonation phrase final syllable the boundary tone, but changed its name in subsequent papers and books.

Beckman and Pierrehumbert 1986, Pierrehumbert and Beckman 1988, Silverman et al. 1992).

In Lee (1990, 1991), I established 9 nuclear tones — Low Level, Mid Level, High Level, High Fall, Low Fall, Full Rise, Low Rise, Rise Fall and Fall Rise — based on my auditory impression. I illustrated the acoustic characteristics of these nuclear tones in Lee (1997).

Jun (1993) and Jun and Oh (1996) set up 6 boundary tones — L%, H%, LH%, HL%, LHL% and HLH% — and briefly discussed acoustic characteristics of these boundary tones. Three more boundary tones — HLHL%, LHLH% and LHLHL% — were added in Jun (1999).

Lee (1999) was the first attempt to discuss the nuclear tone from the point of view of experimental phonetics. I measured major target frequencies of each nuclear tone and compared similar nuclear tones based on quantitative data.

Although major target frequencies and direction of pitch targets certainly are the most important acoustic cues of Korean nuclear tones, there seem to be other important acoustic cues such as timing of peaks and valleys, duration of the whole tone, F0 of the preceding syllable, and so on. Hence this paper aims to find out important acoustic cues of Korean nuclear tones based on quantitative acoustic data.

2. Experiment

2.1. Material

For this experiment, I designed two declarative sentences, two yes-no questions, and two wh-questions. Since the F0 curves of the same nuclear tone may look different depending on whether the final syllable begins with a voiceless consonant or not, one of each pair of sentences is designed to contain "-wasso" and the other "-oyo" at the end.

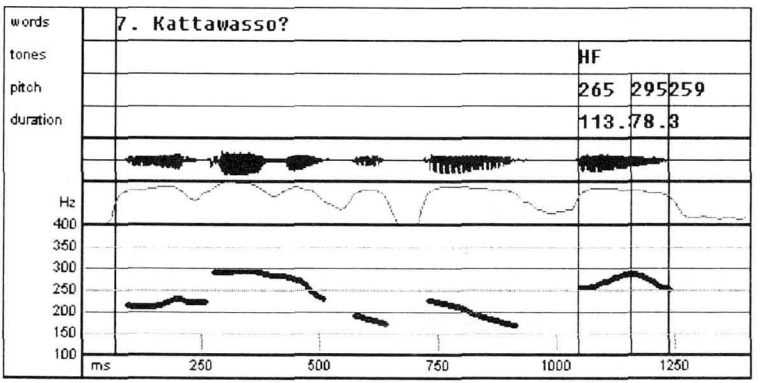
In order to elicit various nuclear tones, subjects were asked to produce the declarative sentences with three different attitudes — indifferent/hostile, friendly/intimate, and annoyed, the yes-no questions with five different attitudes — friendly/intimate, interested, surprised/dubious, hostile, and sarcastic, and the wh-questions with four different attitudes — indifferent/hostile, friendly/intimate, annoyed, and anxious. Appropriate contexts are

also given as in (1).

(1) "No chinan chumare odi kattawanni?" (Where have you been last weekend?)

1. Seoule kattawasso. (indifferent/hostile, Why do you want to know that?) (I've been to Seoul.)
2. Seoule kattawasso. (friendly/intimate, Are you curious?)
3. Seoule kattawasso. (annoyed, Why do you keep asking the same question?)

(a)



(b)

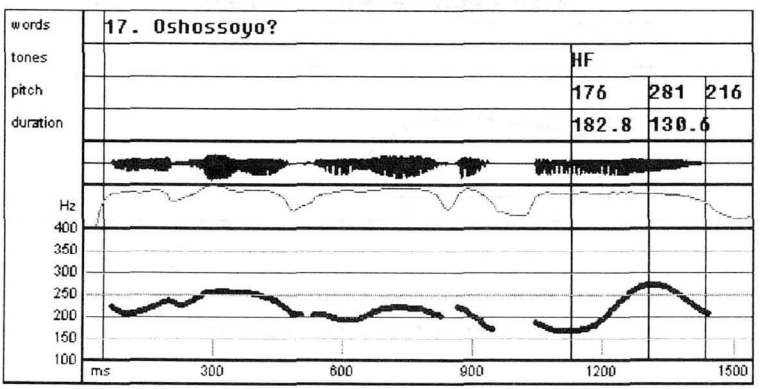


Figure 1. Voiceless consonants are manifested as blanks in F0 curves. A blank appears in (a) before High Fall but not in (b).

2.2. Procedure

Five female Seoul speakers (MK, SJ, YK, KH, and YS) in their twenties were selected as subjects. A total of twenty four test sentences (2 declarative sentences x 3 attitudes, 2 yes-no questions x 5 attitudes, 2 wh-questions x 4 attitudes) were read six times. To elicit as many nuclear tones as possible, I did not randomize the test sentences. I asked the subjects to try to read each test sentence with a different intonation pattern. The subjects' utterances were recorded and digitized with Kay's CSL and analyzed with Scicon's PitchWorks. T-tests were conducted to prove that a certain acoustic cue was statistically significant in distinguishing a certain pair of nuclear tones.

2.3. Measurements

In Lee (1999), I measured only the target frequencies of Korean nuclear tones. But in this experiment, I also measured F0 of the beginning point in High Level, Mid Level, Low Level, High Fall, and Low Fall to see if this F0 functions as an acoustic cue.

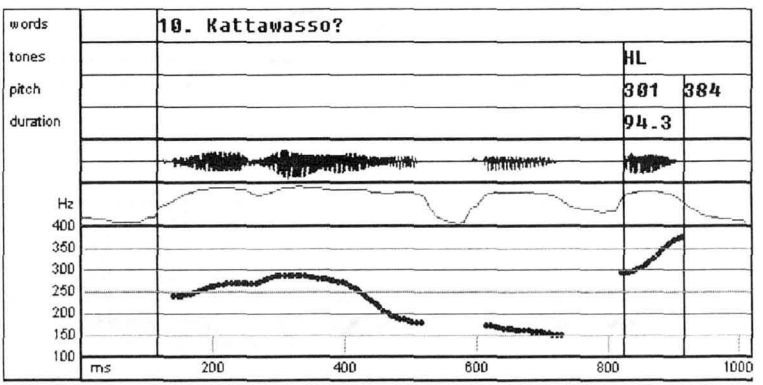
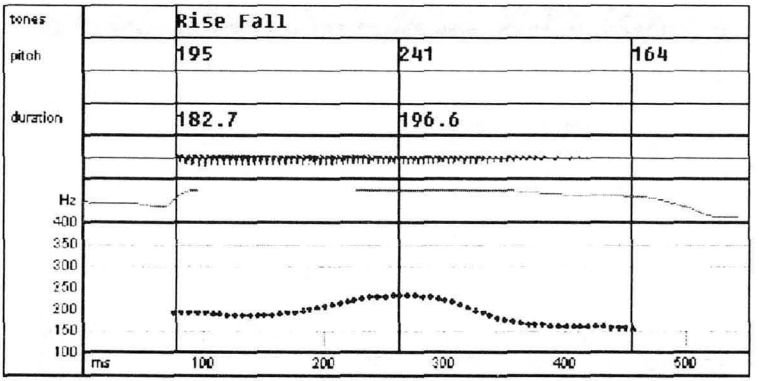


Figure 2. Not only the highest F0 but also F0 of the beginning point in High Level was measured. The duration of the tone was measured from the beginning of the last vowel.

In this experiment, I also measured timing of peaks and valleys, duration of falling and rising phases, and duration of the whole tone. In sentences ending in "-wasso", I measured the timing of peaks and valleys and the duration of the whole tone from the beginning of the last vowel "o".

(a)



(b)

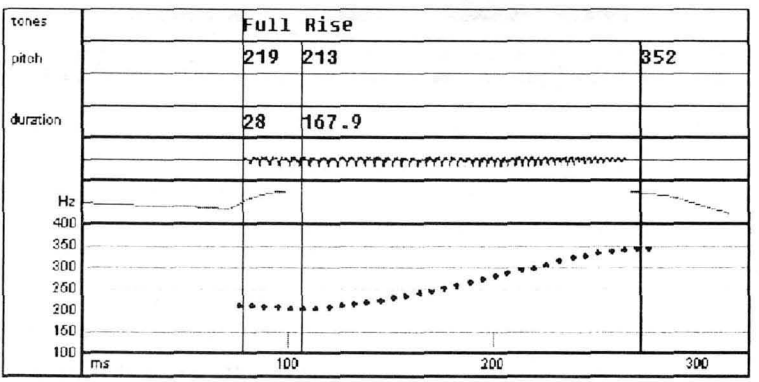


Figure 3. Peak timing (182.7ms) in (a) and valley timing (28ms) in (b). The duration of the falling phase (196.6ms) in (a) and that of the rising phase (167.9ms) in (b) were also measured. The duration of the whole tone equals to the sum of the peak or valley timing and the duration of the falling or rising phase.

But in sentences ending in "-oyo", it was very difficult to find the syllable boundary. Jun and Oh (1996) placed the syllable boundary where the F2 of the semi-vowel /j/ reaches its peak. As in figure 4, however, it was often impossible to find the F2 peak of the semi-vowel /j/. Hence I neglected the timing and duration of nuclear tones which appeared in sentences ending in "-oyo". But I did not have any problem in measuring F0 of the beginning point of the nuclear tones occurring in these sentences.

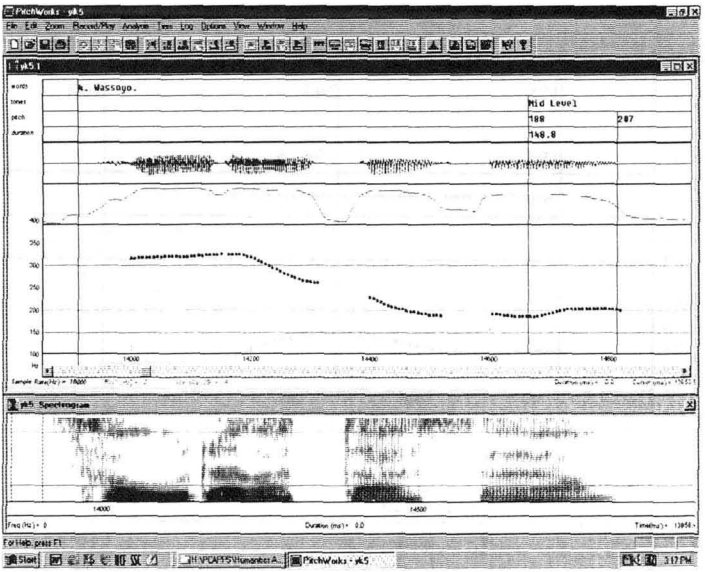
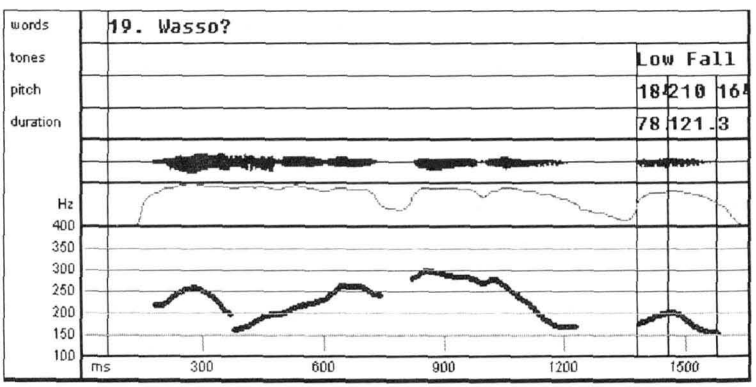


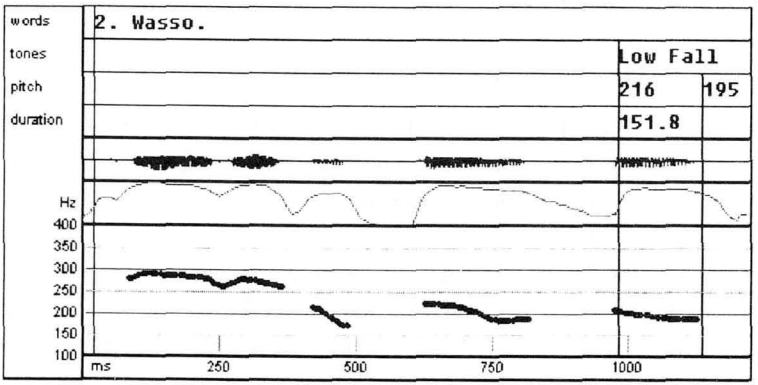
Figure 4. The whole duration of the Mid Level in this figure was neglected due to the difficulty in placing the syllable boundary. But the F0 of the beginning point was taken into account in this experiment.

Most High Falls and Low Falls have certain peak timing. But there were examples where I could not find the peak timing. Hence I classified High Falls and Low Falls into two groups — those with the peak timing (HF1 and LF1) and those without it (HF2 and LF2).

(a)



(b)



(c)

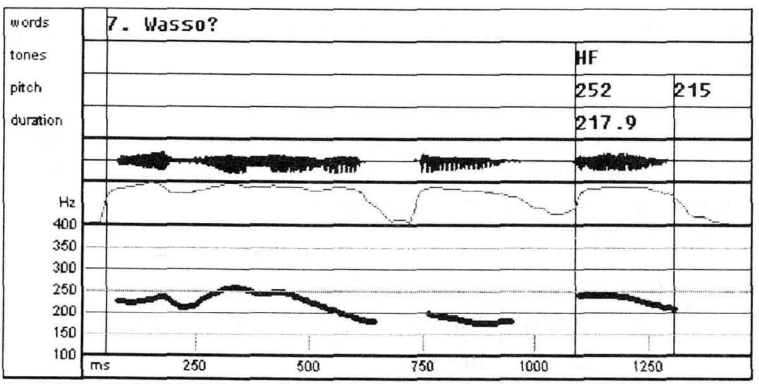


Figure 5. Low Fall with peak timing (a) and Low Fall without peak timing (b) High Fall without peak timing (c)

Average target frequencies, timing and duration of each nuclear tone might be slightly different from speaker to speaker. Since the purpose of this experiment was to find out the acoustic cues of Korean nuclear tones produced by female speakers, I did not take individual differences into account. Since all the subjects were females, I could not find any evidence that individual differences were statistically significant in the discussion of the acoustic cues of Korean nuclear tones.

3. Results and Discussion

3.1. F0 levels and direction of F0 levels

The terms "High Fall, Low Fall, Low Rise, Full Rise, Rise Fall, High Level, etc." imply that major target frequencies and their direction are the most important acoustic cues of Korean nuclear tones. Table 1 shows the F0s of targets and beginning points of Korean nuclear tones. In this table, L1 represents the F0 of the beginning point in the case of High Fall, Low Fall, High Level and Mid Level, and the lowest F0 of the rising phase in the case of Rise Fall, Full Rise and Low Rise. H represents the peak frequency and L2 the lowest F0 of the falling phase. Target frequencies are written in bold numbers. HF1 and LF1 have peak timing whereas HF2 and LF2 do not. The F0 values in the left part of the table were measured from sentences ending in "-wasso" and those in the right part from sentences ending in "-oyo".

Table 1. Average target frequencies and average F0s of beginning points. Standard deviations are given in parentheses. Low Rise occurred only once in sentences ending in "-oyo".

Sentences ending in "-wasso"				Sentences ending in "-oyo"			
Tone	L1	H	L2	Tone	L1	H	L2
HF1	265(30)	**294(38)**	**234(51)**	HF1	183(22)	**278(38)**	**218(52)**
HF2		**261(17)**	**209(25)**				
LF1	187(18)	**203(17)**	**167(20)**	LF1	180(18)	**203(19)**	**166(16)**
LF2		**214(20)**	**173(20)**				
RF	**192(21)**	**218(26)**	**164(16)**	RF	**197(17)**	**218(17)**	**169(13)**
HL	269(50)	**360(61)**		HL	172(13)	**324(52)**	
FR	**189(13)**	**312(31)**		FR	**171(10)**	**286(15)**	
LR	**174(13)**	**211(27)**		LR	**187**	**205**	
ML	206(23)	**221(23)**		ML	174(10)	**198(12)**	
LL		200(17)	**168(8)**				

The peak frequency and the lowest F0 of the falling phase (L2) turned out to be significant acoustic cues in distinguishing High Fall from Low Fall. The peak frequency of High Fall was significantly higher than that of Low Fall ($t=20.86$, $p<0.01$). And High Fall ended in a significantly higher

pitch compared with Low Fall ($t=11.31$, $p<0.01$). The F0 of the beginning point was also an important acoustic cue in sentences ending in "-wasso". High Fall began higher than Low Fall ($t=7.24$, $p<0.01$). Low Fall was never used in Yes-No questions, nor did High Fall in declarative sentences except one case. Both High Fall and Low Fall appeared in wh-questions.

The F0 curves of High Levels and Full Rises looked very similar especially in sentences ending in "-oyo". But the F0 of the beginning point was a significant acoustic cue in sentences ending in "-wasso". As can be seen in figure 2 and 3b, High Level began significantly higher than Full Rise ($p=-6.77$, $p<0.01$). Both High Level and Full Rise were used in yes-no questions and wh-questions, but not in declarative sentences.

The F0 of the beginning point also played an important role in distinguishing Low Rise from Mid Level in sentences ending in "-wasso" ($t=-3.03$, $p<0.01$). But the F0 of the beginning point was meaningless in sentences ending in "-oyo".

Since Low Fall and Rise Fall have similar target frequencies, the F0 contours of these two tones were not markedly different especially in sentences ending in "-oyo". In sentences ending in "-wasso", a valley in F0 curve was often observed in the beginning of Rise Fall. As will be discussed later, the important acoustic cues that distinguish Rise Fall from Low Fall were the peak timing and the duration of the whole tone.

Low Rise ends in a significantly lower pitch compared not only with Full Rise ($t=7.24$, $p<0.01$) but also with High Level ($t=5.36$, $p<0.01$). Low Rise was used in declarative sentences and wh-questions, but not in yes-no questions. High level and Full Rise were not used in declarative sentences except one case.

Low Fall and Low Level fall from a similar pitch and end in a similar pitch. These two tones are not distinguished by the F0 of the beginning or ending point. As will be discussed in section 3.3., these two tones are distinguished by the pitch movement from the preceding syllable.

Direction of target frequencies is also important in distinguishing nuclear tones which have more than two target frequencies. High Fall and Full Rise have different direction of target frequencies. Low Fall and Low Rise also have different direction of target frequencies.

3.2. Timing and duration

Timing and duration also played an important role in distinguishing

nuclear tones acoustically. Table 2 shows information about peak timing of High Fall, Low Fall, and Rise Fall, duration of the rising phase occurring in Rise Fall, High Level, Full Rise, Low Rise, and Mid Level, duration of the falling phase, and duration of the whole tone. As in table 1, HF1 and LF1 have peak timing whereas HF2 and LF2 do not. Since it was almost impossible to accurately measure the timing and duration values in sentences ending in "-oyo", the timing and duration values in table 2 were measured from sentences ending in "-wasso".

Table 2. Peak timing or duration of the rising phase (T1), duration of the falling phase (T2), and duration of the whole tone (T1+T2). Standard deviations are given in parentheses.

Tone	T1	T2	T1+T2
HF1	89(23)	119(59)	208
HF2		176(22)	176
LF1	101(39)	124(23)	225
LF2		180(50)	180
RF	162(35)	191(48)	353
HL	133(37)		133
FR	185(21)		185
LR	209(51)		209
ML	166(44)		166
LL		238(50)	238

Timing and duration functioned as important acoustic cues in distinguishing Low Fall from Rise Fall. Rise Fall had significantly longer peak timing ($t=-4.72$, $p<0.01$) and significantly longer total duration ($t=-8.75$, $p<0.01$) compared with Low Fall.

Timing and duration also played an important role in distinguishing High Level from Full Rise. Jun and Oh (1996) claimed that LH%, corresponding to Full Rise, has a low plateau in the beginning whereas H%, corresponding to High Level, does not. The duration of the low plateau corresponds to valley timing in sentences ending in "-wasso". The average valley timing of High Levels was 7ms and that of Low Levels 29ms. And the standard deviation was 9ms in the case of High Levels and 14ms in the case of Full Rises. These differences were statistically significant ($t=-7.80$, $p<0.01$). Full Rise also had significantly longer total duration compared with High Level

($t=-5.76$, $p<0.01$).

3.3. F0 of the penultimate syllable

The F0 of the penultimate syllable of an intonation phrase normally reaches the bottom of the pitch range before High level, Mid Level, High Fall, or Full Rise begins whereas the penultimate syllable ends higher than the beginning point of Low Level, Low Rise, and Rise Fall.

Table 3. The relationship between the nuclear tone and the F0 of the penultimate syllable. This table shows target frequencies and standard deviations. These values are adopted from Lee (1999).

Tone \ value	F0 of the penultimate syllable	L1	H	L2
HF	187 (25)		302 (48)	211 (58)
LF	202 (29)		230 (19)	177 (16)
RF	226 (35)	198 (23)	217 (29)	163 (15)
HL	180 (17)		380 (56)	
FR	174 (16)	178 (12)	306 (22)	
LR	223 (29)	184 (13)	208 (16)	
ML	172 (16)		231 (19)	
LL	189 (24)			161 (17)

As can be seen in Table 1 and 3, Low Fall and Low Level have almost the same peak and low target frequencies. And these two tones also have similar duration as in Table 2. Hence the pitch movement from the penultimate syllable is the only significant acoustic cue in distinguishing these tones. A pitch rise from the penultimate syllable is observed in Low Falls whereas this pitch rise is not observed in Low Levels.

4. Conclusion

So far, I have discussed what acoustic cues are significant in distinguishing Korean nuclear tones acoustically. The results of this experiment show that target frequencies, direction of target frequencies, F0

of the beginning and ending points, timing of peaks and valleys, duration of the rising and falling phases, duration of the whole tone, and F0 of the penultimate syllable function as acoustic cues, and that each acoustic cue is significant in distinguishing a certain pair of nuclear tones.

Peak frequency was crucial in distinguishing High Fall from Low Fall and Full Rise from Low Rise. The F0 of the beginning point was significant in distinguishing High Fall from Low Fall, Low Rise from Mid Level, and High Level from Full Rise in sentences ending in "-wasso". And the F0 of the ending point was important in distinguishing High Fall from Low Fall. Direction of target frequencies was crucial in distinguishing High Fall from Full Rise and Low Rise from Low Fall.

Duration of the whole tone was significant in distinguishing Low Fall from Rise Fall and High Level from Full Rise. The former pair of nuclear tones was also distinguished by Peak timing and the latter pair by valley timing. The F0 of the penultimate syllable was crucial in distinguishing Low Fall from Low Level.

It seems to me that not all the above-mentioned acoustic cues are significant in perception and that certain acoustic cues may be more important than others in perception. I hope that these assumptions will be examined in future studies.

References

Beckman, M. and J. Pierrehumbert (1986) Intonational Structure in English and Japanese, *Phonology Yearbook* 3.

Cruttenden, A (1997) *Intonation*, 2nd ed., Cambridge, Cambridge University Press.

Jun, S.A. (1993) *The phonetics and phonology of Korean prosody*, PhD dissertation, Ohio State University.

Jun, S.A. (1999) K-ToBI (Korean ToBI) labelling conventions, version 3.0, ms., UCLA.

Jun, S.A. and M. Oh (1996) A prosodic analysis of three types of wh-questions in Korean, *Language and Speech* 39.1.

Lee, H.Y. (1990) *The structure of Korean prosody*, PhD thesis, University College London, [Published by Hanshin Publishing Co.].

Lee, H.Y. (1991) The intonation system of Korean (in Korean), *Onohak* 13, The Linguistic Society of Korea.

Lee, H.Y. (1997) *Korean Prosody* (in Korean), Seoul, The Korean Research Center.

Lee, H.Y. (1999) An Acoustic Phonetic Study of Korean Nuclear Tones (in Korean), *Malsori* 38. The Phonetic Society of Korea.

O'Connor, J. D. and G. F. Arnold (1973) *Intonation of colloquial English*, London, Longman.

Pierrehumbert, J. (1980) *The phonology and phonetics of English intonation*, unpublished PhD thesis, MIT.

Pierrehumbert, J. and M. Beckman (1988) *Japanese Tone Structure*, Cambridge, MA, MIT Press.

Silverman, K. et al. (1992) ToBI: a Standard for Labeling English Prosody, *Proceedings, Second International Conference on Spoken Language Processing* 2, Banff, Canada.

Appendix: List of Test Sentences

"No chinan chumare odi kattawanni?" (Where have you been last weekend?)

1. Seoule kattawasso. (indifferent/hostile, Why do you want to know that?) (I've been to Seoul.)
2. Seoule kattawasso. (friendly/intimate, Are you curious?)
3. Seoule kattawasso. (annoyed, Why do you keep asking the same question?)

"No chinan chumare odi kattawanni?" (Where have you been last weekend?)

4. Seoule kattawassoyo. (indifferent/hostile, Why do you want to know that?) (I've been to Seoul.)
5. Seoule kattawassoyo. (friendly/intimate, Are you curious?)
6. Seoule kattawassoyo. (annoyed, Why do you keep asking the same question?)

"Chinan chumare ilbone kattawassoyo." (I've been to Japan last weekend.)

7. Ilbone kattawasso? (friendly/intimate, You've achieved what you've been dreaming.) (You've been to Japan?)
8. Ilbone kattawasso? (interested, You told me that you would go to America.)
9. Ilbone kattawasso? (surprised/dubious, You told me you'd never go to Japan.)
10. Ilbone kattawasso? (hostile, I warned you not to go to Japan. I won't

forgive you.)

11. Ilbone kattawasso? (sarcastic, You've been to Japan? You're lying again.)

"Chinan chumare ilbone kattawasso." (I've been to Japan last weekend.)

12. Ilbone tanyooshossoyo? (friendly/intimate, You've achieved what you've been dreaming.) (You've been to Japan?)
13. Ilbone tanyooshossoyo? (interested, You told me that you would go to America.)
14. Ilbone tanyooshossoyo? (surprised/dubious, You told me you'd never go to Japan.)
15. Ilbone tanyooshossoyo? (hostile, I warned you not to go to Japan. I won't forgive you.)
16. Ilbone tanyooshossoyo? (sarcastic, You've been to Japan? You're lying again.)

17. No onje Seoule wasso? (indifferent/hostile, Why have you come? You shouldn't.) (When have you come to Seoul?)
18. No onje Seoule wasso? (friendly/intimate, I'm happy because you've come.)
19. No onje Seoule wasso? (annoyed, Why do you make me repeat the same question?)
20. No onje Seoule wasso? (anxious, Your parents are worrying about you.)

21. Onje Seoule oshossoyo? (indifferent/hostile, Why have you come? You shouldn't.) (When have you come to Seoul?)
22. Onje Seoule oshossoyo? (friendly/intimate, I'm happy to meet you in Seoul.)
23. Onje Seoule oshossoyo? (annoyed, Why do you make me repeat the same question?)
24. Onje Seoule oshossoyo? (anxious, Your family are worrying about you.)

Department of Korean Language and Literature
Pukyong National University
599-1 Daeyon-dong, Nam-ku
Pusan 608-737, Korea
E-mail : leehy@mail.pknu.ac.kr