



## A PROSODIC ANALYSIS OF THREE SENTENCE TYPES WITH "WH" WORDS IN KOREAN

Sun-Ah Jun\* and Mira Oh\*\*

\*Linguistics Department, UCLA, Los Angeles, California 90024-1543, U.S.A.

\*\*English Department, Yeoo Technical College, Kyunggi-do, 469-800, Korea

### ABSTRACT

"Wh" words in Korean have two interpretations: either as wh-pronouns as in a wh-question or as indefinite pronouns as in a yes/no-question. The lexical ambiguity of a wh-word has been impressionistically noted to be differentiated by prosodic features. In this paper, we investigate which prosodic features, such as pitch range and accentual phrasing, differentiate these two usages, both in production and perception. We additionally investigate a third usage involving incredulity. Production data show that the three question types are distinguished by boundary tones, pitch ranges and accentual phrasing. Yes/no questions and other types are distinguished by different accentual phrasing. Incredulity and wh-questions are in general distinguished by different peak amplitudes and pitch ranges, in addition to a boundary tone. Interestingly, not all speakers use the same strategy to distinguish incredulity and wh-questions. The results of the perception test show that subjects often confused incredulity questions with wh-questions, suggesting that accentual phrasing is a stronger perceptual cue than pitch range or amplitude difference.

### I. INTRODUCTION

"Wh" words in Korean have two functions: either as a wh-pronoun as in a wh-question, or as an indefinite pronoun as in a yes/no-question. For example, since /nuka/ has two possible meanings, 'who' and 'anyone', /nuka wajo/? (wa+jo 'to come+interrogative ending') can mean 'Who is coming?' or 'Is there anyone coming?'. It has often been mentioned in the literature that the two types are distinguished by having different prosodic features such as a pitch accent [2, 3, 10] or boundary tones [10]. But to our knowledge, no instrumental phonetic study has been done in this regard. Maekawa [13] also notes a similar phenomenon in Tokyo Japanese and shows the intonational difference between wh and non-wh questions using perception test of synthetic intonation.

In this study, we added a third type of wh-question: an incredulity-question. This question type is employed when the speaker cannot believe what he or she just heard [12, 14]. This type differs from the wh-question or echo-question in that the speaker uses the wh-word even though he or she knows the answer to the question, but it also differs from the non-wh or yes/no question in that the wh-word is not indefinite. Thus, /nuka wajo/? can also mean 'WHO did you say is coming? (= I can't believe that X is coming.)'. This type again differs from an echo-question in that the incredulity question

implies surprise while the echo-question does not. It has been claimed that an "incredulity" reading in English is associated with shorter duration, greater amplitude, and a larger pitch range in production compared to an "uncertainty" reading, but only a large pitch range has been found to be perceptually salient in the perception of "incredulity" [12, 5].

As shown in previous studies of intonational meaning [9, 1], particular prosodic features contribute to utterance interpretation. To find out what prosodic features are used to distinguish the three types of wh-question and which feature is perceptually most salient, we examined both production and perception data. The production results show that the three types are distinguished by different boundary tones, accentual phrasing [4, 6, 7, 8, 11], pitch range, and peak amplitude. We also found that the incredulity question is not always produced with a higher pitch in all parts of an utterance, and the way of realizing the different pitch range is not the same across speakers. The results of the perception test suggest that accentual phrasing is a more important perceptual cue than pitch range or amplitude difference.

### II. METHOD

#### II.1 Production

*Subjects* : Four Seoul speakers, two males (speaker C and speaker L) and two females (speaker O and speaker Y), in their early thirties. Speakers C, L, and Y are graduate students at the Ohio State University, Columbus, Ohio. Speakers C and Y have been in the States about four years, and speaker L three years. Speaker O is the second author of this paper.

*Material* : Twenty sentences were uttered employing three question types, wh-question, incredulity question, and yes/no question. Each sentence had a wh-word initial phrase preceded by an adverbial phrase. We added an adverbial phrase to examine the pitch range difference outside the wh-word phrase. Every sentence ended in a question particle /-jo/, so that a boundary tone was not specific to the particle, but to the question type. The wh-word initial phrase contained one of four wh-words, /nuka/ 'who', /mwə/ 'what', /əntʃe/ 'when', and /əti/ 'where', and two wh-words after the postposition, /nukular/ and /nukuhako/, both meaning 'with whom'. The adverbial clause was either /onil tʃənjəke/ 'tonight' or /atʃuməni-ka(nin, lil)/ 'madam-case marker'. (1) shows representative sentences.

- (1) a. /onil tʃənjəke nuka wajo/ 'Tonight, who is coming (or 'does anyone come', or WHO is coming)?'  
 b. /atʃuməninin əntʃe ətʃiləwəjo/ 'Madam, when do you feel dizzy (or 'is there anytime that you feel dizzy', or 'WHEN do you feel dizzy')?'

**Procedure :** We wrote a short dialogue (question and answer pair) for each type of wh-question on a card (5.5 inch x 8.5 inch), and 60 (3 types x 20 sentences) dialogues were pseudo-randomized so that the same type of question did not appear three times in a row. Speakers repeated the 60 dialogues three times. For every recording, the second author participated as a partner, helping the speaker to respond appropriately and naturally. The recording was made at the Linguistics department of the Ohio State University. Using the xwaves program in the Sun system, utterances were digitized and pitch tracks and amplitude of each sentence were displayed.

We measured f0 values and time (reference to the beginning of an utterance) of seven points on the pitch track, as shown in Figure 1; the lowest f0 (=A) and the highest f0 (=B) of the adverbial phrase, the lowest f0 after 'B' (=C), the highest f0 of the wh-phrase (=D), the lowest or the beginning of low plateau after 'D' (=E), the beginning of the question particle 'jo' (=F), and the utterance final f0 (=G). We decided the point 'F' as the highest F2 of [j] in the spectrogram. The spectrogram was aligned with the waveform and the pitch tract. In addition, we measured the peak rms amplitude of two points of an utterance, the wh-word and the question particle at the end of an utterance.

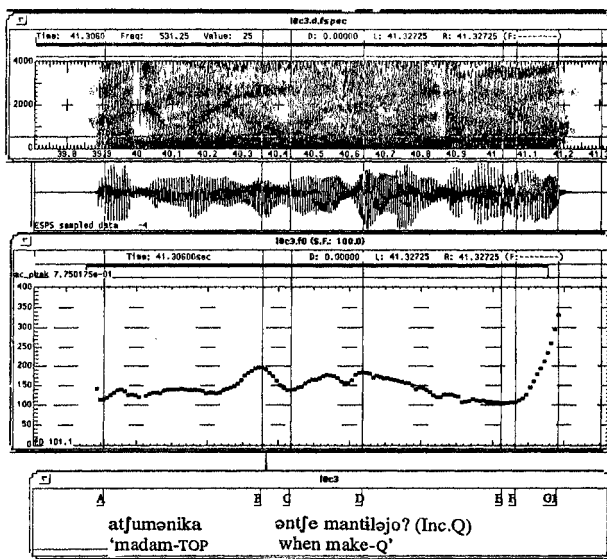


Figure 1. Measurement points: the lowest f0 of the adverbial phrase (=A), the highest f0 of the adverbial phrase (=B), the lowest f0 after 'B' (=C), the highest f0 of the wh-phrase (=D), the lowest f0 after 'D' (=E), the beginning of the question particle 'jo' (=F), and the utterance final f0 (=G). (G1 indicates a 'H%'.)

## II.2 Perception

**Subjects :** 30 college students in their late teens or early twenties, 24 males and 6 females. All subjects were native Seoul speakers.

**Material :** 45 sentences (3 wh types x 15 sentence types) for each speaker taken from the first reading of the

production data. Sentences were randomized separately for 4 speakers and recorded on to a tape. The perception tape contains instructions in Korean with 6 example sentences and the main test of 180 sentences, which is blocked by a speaker. Within data from each speaker, a beep was placed every five sentences; each sentence was followed by 4-seconds of silence. Data from different speakers were separated by two beeps. The whole test took 30 minutes.

**Procedure :** The test was run twice in a classroom of Yeojoo College, Kyunggi-do, Korea, once with 21 subjects wearing headphones, and a second time with 9 subjects not wearing headphones. Subjects were asked to choose one of three possible answers for each stimulus by placing a mark on the answer sheet. Each answer sheet had 45 rows and 4 columns (i.e., stimulus number, wh-question, yes/no question, incredulity question), containing data from one speaker. Every five rows were grouped by a double line. Instructions were presented in both oral and written form.

## III. RESULTS AND DISCUSSION

The three question types are distinguished by the combination of boundary tones and accentual phrasing. Yes/no-questions and incredulity-questions very often have a H boundary tone (H%) and wh-questions have various boundary tone types, most commonly a LH%, except for speaker C. Figure 2 shows the percentage of each boundary tone used in each type of question. Speaker C used four kinds of boundary tones for wh-questions.

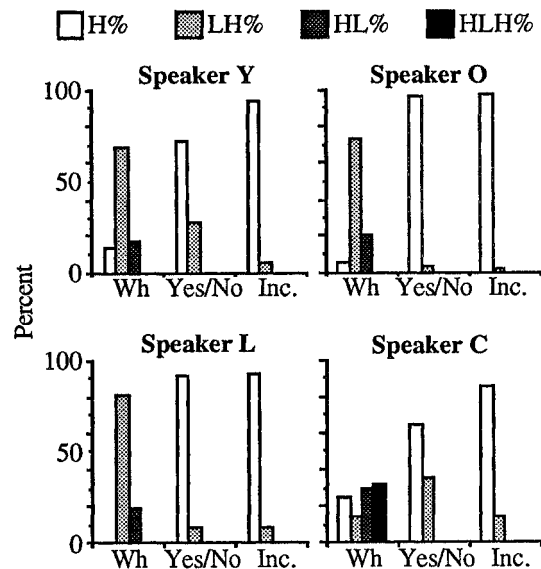


Figure 2. Types of boundary tones per each wh-question and percent of each boundary tone for four speakers.

Within a given category of boundary tone, the pitch was in general higher in incredulity-questions than that in yes/no-questions which were again higher than that in wh-questions. This difference is very prominent between wh-questions and incredulity questions, but less so between incredulity questions and yes/no questions. The main difference between the yes/no question and the other two types is found to be in accentual phrasing. All speakers produced the wh-phrase in wh-questions and incredulity questions in one Accentual Phrase. But they produced the

wh-phrase in yes/no questions in two Accentual Phrases, one for the wh-word and the other for the verb phrase.

According to Jun [7, 8] and others [4, 11], an Accentual Phrase in the Seoul dialect starts with a low tone and ends with a high tone, with an additional initial high peak around the second syllable of a phrase longer than four syllables. Thus, if the wh-word forms one Accentual Phrase, the word would end with a high f0 and the following lowest f0 point would be the beginning of the following Accentual Phrase, the beginning of the verb. On the other hand, if the wh-word and the following verb phrase form one Accentual Phrase, the peak f0 of the wh-phrase would be realized around the second syllable of the wh-word, but the following lowest f0 would be toward the end of the verb phrase. Thus the difference between 'D' and 'E' would be small in yes/no questions, but large in the other two types. Table 1 shows the difference in mean time between 'D' (=peak f0 of the wh-phrase) and 'E' (=the lowest f0 after the peak) in three types of wh-questions.

Table 1. The mean time difference (ms) between 'D' (=peak f0 of the wh-phrase) and 'E' (=the lowest f0 after the peak) of three wh-question types for four speakers. (One standard error in the parenthesis)

| Q type | wh-Q          | yes/no Q      | incredulity Q |
|--------|---------------|---------------|---------------|
| spk. Y | 682.40 (16.4) | 370.58 (12.9) | 654.86 (17.3) |
| spk. O | 686.66 (20.5) | 418.31 (13.6) | 693.64 (21.6) |
| spk. L | 602.02 (16.9) | 305.82 (9.3)  | 602.40 (19.9) |
| spk. C | 598.60 (17.5) | 324.52 (11.4) | 682.64 (22.2) |

Multiple t-tests show that all speakers show the time difference ('D-E') in yes/no questions is significantly shorter than that in other two types ( $p < 0.01$ ). The time difference between wh-questions and incredulity questions is not significant for all speakers at  $p = 0.01$  level.

The wh-question and the incredulity question differ in their pitch ranges and peak amplitudes. For all speakers, the incredulity-question generally showed a larger pitch range than that of the wh-question, and all speakers except for speaker C showed a greater peak amplitude than that for wh-questions [5]. However, interestingly, the large pitch range of incredulity questions is not always apparent for every speaker. Speaker Y produced similar pitch ranges for wh-questions and incredulity questions, but consistently produced the pitch range of the incredulity question lower than that of wh-question. She showed an expanded pitch range only for the boundary tone, which was about 120 Hz higher for the incredulity than for the wh-question. The other speakers employed larger pitch range for the wh-phrase of incredulity questions, but they used different strategies in expanding their pitch range. For example, speaker L and speaker C achieved a wider pitch range for the incredulity reading by raising the pitch of the H tone but keeping the same L tone pitch as in wh-questions, while speaker O achieved a larger pitch range by lowering the pitch of the L tone but keeping the same H tone pitch as in wh-questions. Figure 3 shows f0 values and times of three points, B-C-D, for four subjects; time of 'C' and 'D' is plotted relative to the end of an adverbial phrase, 'B'.

The same pitch range relationship was found within an adverbial phrase, between 'A' and 'B'. Figure 3 also shows that, in general, it takes longer to realize the peak f0 of wh-phrases (=D) for incredulity question than for wh-question. This is due to the lengthening of the first syllable of the wh-word in incredulity questions. But the total duration of incredulity questions is in general similar to that of the non-surprised version, different from what is

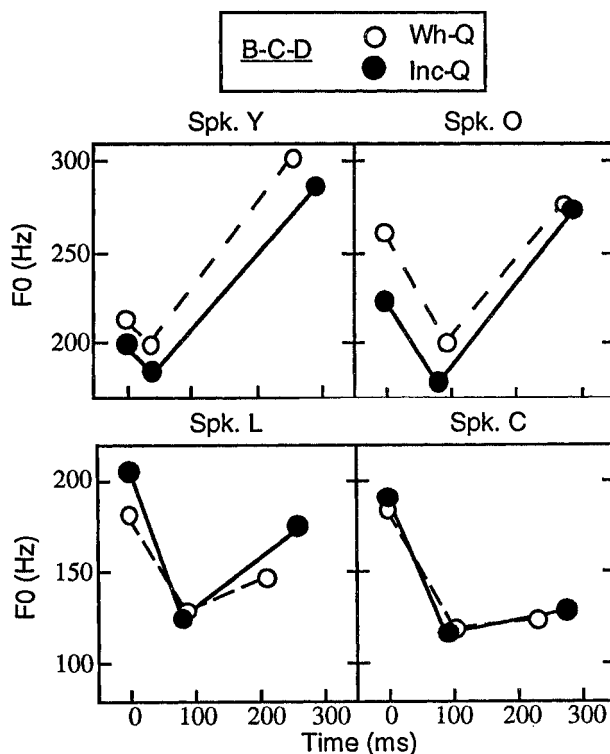


Figure 3. F0 and mean time of B-C-D of two question types: 'B' (the end of an adverbial phrase), 'C' (=lowest f0 after 'B' before a peak) and 'D' (=peak f0 of wh-phrase). Time of 'C' and 'D' is plotted by a reference to 'B'. X-axis refers to time in ms and Y-axis refers to a fundamental frequency in Hz.

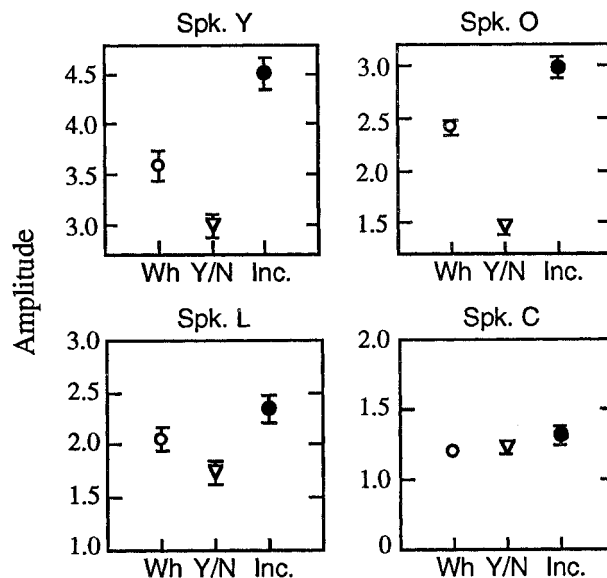


Figure 4. Rms amplitude (in arbitrary scale) of wh-word in three question types for each speaker. (An error bar represents one Standard Error.)

found in English [5].

The amplitude of the wh-word and question particle 'jo' was consistently greatest for incredulity

questions. But speaker C produced no significantly different amplitude for wh-word for all question types. The result of the perception test did not show any difference for this speaker, suggesting that amplitude is not the main perceptual cue. Amplitude (rms in arbitrary scale) of wh-words is shown in Figure 4.

Table 2 summarizes the differences between speakers for different prosodic features.

Table 2.

| spk | Inc.Q against wh-Q |        |       | % tones         |        | Acc.Phr. |        |
|-----|--------------------|--------|-------|-----------------|--------|----------|--------|
|     | f0 (C)             | f0 (D) | ampl. | wh-Q            | others | y/n Q    | others |
| Y   | low                | low    | high  | H,LH,HL         | H,LH   | two      | one    |
| O   | low                | same   | high  | H,LH,HL         | H,LH   | two      | one    |
| L   | same               | high   | high  | LH, HL          | H,LH   | two      | one    |
| C   | same               | high   | same  | H,LH,HL,<br>HLH | H,LH   | two      | one    |

The results of the perception test show that subjects were in general good at distinguishing three types of wh-questions, and the response pattern was very similar across four speakers. This indicates that subjects are good at interpreting the different strategy employed by different speakers. The perception responses combining all four speakers are shown in Figure 5; 'A' stimuli is the wh-question, 'B' stimuli is the yes/no question, and 'C' stimuli is the incredulity question. The percentage is based on 450 tokens (15 sentences x 30 subjects) per each type of question. In general, 'A' stimuli was easiest and 'C' stimuli was hardest for subjects to answer. We think this is due to insufficient information related to pitch range or amplitude of the context. The utterance-initial adverbial phrase did not help markedly the normalization of pitch range because the adverbial clause itself showed the same degree of pitch range as that of the wh-phrase. For all speakers, the incredulity question was most often misperceived as a wh-question. Yes/no question types were equally misperceived either as wh-questions or incredulity questions. This indicates that accentual phrasing is a more salient cue than pitch range and amplitude difference. We believe that subjects are more sensitive to an accentual phrasing because a different accentual phrasing determines the lexical category of the wh-word, i.e. a wh-pronoun or an indefinite pronoun, while a different pitch range and/or amplitude changes a pragmatic meaning, which often depends on the context of a dialogue, thus hard to tell by listening to only one sentence in isolation.

#### IV. CONCLUSION

Three types of wh-questions due to a lexically or pragmatically ambiguous wh-word in Korean have been shown to have different prosodic features. In yes/no questions, the wh-word is produced in one Accentual Phrase separated from the verb phrase, while the wh-word and the following verb phrase together form one Accentual Phrase in incredulity and wh-questions. Incredulity questions are distinguished from wh-questions by a larger pitch range and higher amplitude, as found in [5], and also by a boundary tone. However, not all speakers use the same strategy in distinguishing incredulity and wh-questions. But the perception test shows that the different strategies are equally accepted by subjects. The fact that subjects often confused incredulity questions with wh-question suggests that accentual phrasing is a stronger perceptual cue than pitch range or amplitude difference.

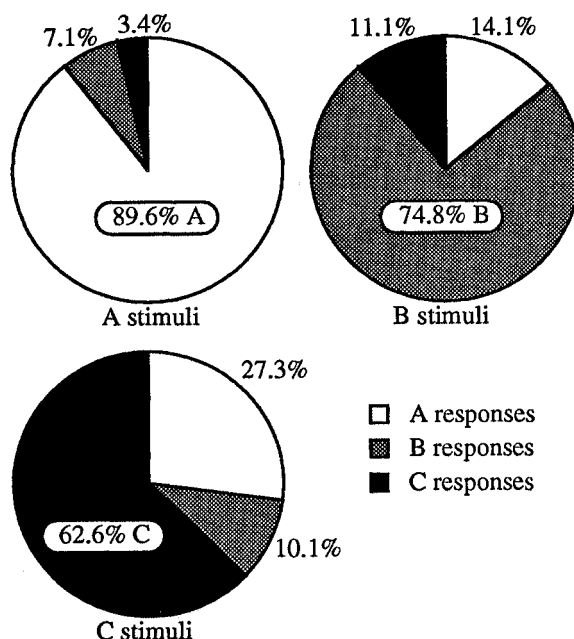


Figure 5. Percentage of response to each type of wh-question. Data combining 4 speakers.

#### REFERENCES

- [1] D. Bolinger, *Intonation and its uses: melody in grammar and discourse*. London: Edward Arnold. 1989.
- [2] S. J. Chang, "A Generative Study of Discourse: Pragmatic Aspects of Korean with reference to English," *Language Research* 9.2. (Supplement) 1973.
- [3] J.-W. Choe, "Pitch-accent and q/wh words in Korean," S. Kuno et al. (eds.) *Harvard Studies in Korean Linguistics*. Harvard Univ.:Cambridge, 1985.
- [4] K. deJong, "Initial Tones and Prominence in Seoul Korean", paper presented at the 117th meeting of the Acoustical Society of America. Syracuse, N.Y. 1989.
- [5] J. Hirschberg & G. Ward, "The influence of pitch range, duration, amplitude and spectral features on the interpretation of the rise-fall-rise intonation contour in English," *Journal of Phonetics* 20: 241-251. 1992.
- [6] S.-A. Jun, "The accentual pattern and prosody of Chonnam Dialect of Korean," In Kuno et al., eds., *Harvard Studies in Korean Linguistics* III, 1989.
- [7] S.-A. Jun, "The Prosodic Structure of Korean - in terms of voicing," in E.-J. Baek (ed.) *Proceedings of the Seventh International Conference on Korean Linguistics*, Vol. 7., Univ. of Toronto Press. 1990.
- [8] S.-A. Jun, *The Phonetics and Phonology of Korean Prosody*. Ph.D. dissertation, the Ohio State Univ. 1993.
- [9] D. R. Ladd, *The structure of intonational meaning*. Bloomington: Indiana University Press. 1980.
- [10] H.-Y. Lee, *The structure of Korean prosody*. Ph.D. dissertation. University College London. 1990.
- [11] S. Lee, "Intonational Domains of the Seoul Dialect of Korean," paper presented at 117th ASA meeting. 1989.
- [12] M. Liberman & I. A. Sag, "Prosodic form and discourse function," *Papers of the tenth regional meeting of the CLS, Chicago*, pp. 416-427, 1974.
- [13] K. Maekawa, "Perception of Intonational Characteristics of wh and non-wh questions in Tokyo Japanese," *Proceedings of the XIIIth International Congress of Phonetic Sciences*, Vol 4: 202-205. 1991.
- [14] M. Oh & H.-Y. Lee, "Teaching Intonation of Korean as a foreign language," Paper presented at IAKL, Seoul, Korea. 1993.