Topics in Korean-Japanese Phonetics: Or a layman's view on Korean phonetics

Kikuo Maekawa (National Language Research Institute)

1. Prosodic phrasing in Japanese and Korean

Traditional studies of intonation has been concerned mostly with the classification of the sentence-final local renditions; i.e. falling, rising, flat etc. In Maekawa (1991a,b), I tried to show that the intonational difference between WH and Yes-No questions in Tokyo Japanese could not be captured in a satisfactory way by this kind of analysis. These studies showed that it was the *prosodic phrasing* of an utterance that played a crucial role in the perceptual discrimination of the two question types.

The same conclusion was obtained by the analyses of *Kumamoto* Japanese: so-called *accentless dialect*. An accentless dialect is a dialect which has no pitch accent at the level of lexicon, and is therefore typologically very similar to Standard Korean. Maekawa (1990, 1994b) showed that the speakers of Kumamoto Japanese were very sensitive to the existence, as well as its precise location, of a prosodic boundary.

Jun and Oh (1994) analyzed almost the same problem in Korean (Seoul dialect). They share the same conclusion that the prosodic phrasing is primarily important for the intonational discrimination between WH and Yes-No questions. Besides, Jun and Oh reported considerable amount of individual difference in the choice of sentence-final boundary tone with respect to the WH question.

Independently from Jun and Oh's study, Min (1994) reported a precise Korean analogue of Maekawa's (1991a) experiment. Although Min did not analyze his data in term of prosodic phrasing, it is perfectly possible to interpret his data as showing the importance of prosodic phrasing for the distinction between the WH and Yes-No questions in Seoul Korean.

To recapitulate, the two languages show remarkable similarity regarding the intonational characteristics of WH and Yes-No questions. It is a quite interesting issue in Korean-Japanese contrastive phonetics, as well as in the teaching of the two languages as foreign language, to know how far this similarity goes. Perhaps, we would overestimate the similarity if we pay too much attention to the case of WH and Yes-No questions. My experience as a learner of Korean —a very limited one, unfortunately— tells me that there are several cases where Korean prosodic phrasing shows marked difference from the Japanese phrasing.

For example, in Japanese a verb and the following auxiliary verb are said to make a single accentual phrase like,

/ tabe'ru / + / kudasa'i / = / ta'betekudasai /　　(An apostrophe denotes an accent.)

[Note: The above analysis involves deaccenting of the auxiliary verb, which turns out to be untrue by a fine phonetic analysis (Maekawa 1994a). Here, I adopt a simplified view for the sake of simplicity.]

In Korean, this is not necessarily the case. The following example is from a textbook of Korean and its accompanying recorded tape (Umeda and Kim, 1989, p.74).

[hanbon soge̲h̲e̲ dʒusejo]

The underscore stands for a local rise in pitch followed by a sharp fall, which usually marks the existence of a prosodic boundary (accentual phrase boundary, according to Jun, 1993). In this case, the verb and the auxiliary verb have their own independent accentual phrases. The following case is from the same textbook (*ibid.* p.68).

[atʃʰimenin ujuna ho̲n̲tʃʰarɯl maʃidʒo]

Here the tonal index is not linked to the end of the prosodic phrase, rather it is on the first syllable of the following phrase.

2. Perception of Japanese long vowels by Korean learners of Japanese

The second topic is concerned with the well known difficulty shared by virtually all Korean learners of Japanese: the production and perception of so-called special morae in Japanese. Very recently I started an experimental study of crosslinguistic comparison of the perception of Japanese quantity by foreign learners in collaboration with Yasuhiko Sukegawa of Tohoku University, and this is the first preliminary report from the project.

The experiment was concerned with the distinction of three Japanese words: /obasan/ (aunt), /oobasan/ (a Japanese family name), and /oba'asan/ (grand mother). The opposition between /obasan/ and /oobasan/ consists in the quantity difference of the word-initial syllable, while the opposition between /obasan/ and /oba'asan/ consists in the quantity difference in word-medial position as well as the absence vs. existence of an accent.

All stimuli used in the experiment were synthesized by mean of the LPC synthesis from one and the same token of /obasan/ uttered by a male Japanese subject. Three different sets of stimuli were synthesized from the original utterance. What follows is the description of the stimuli used in the experiment.

Set-O was concerned with the distinction between /obasan/ and /oobasan/. This set included the following seven synthetic stimuli.

o-0 : Synthetic version of the original /obasan/　　(Typical /obasan/.)

o-1 : The duration of /o/ 20 ms longer than o-0.

o-2 : The duration of /o/ 40 ms longer than o-0.

o-3 : The duration of /o/ 60 ms longer than o-0.

o-4 : The duration of /o/ 80 ms longer than o-0.

o-5 : The duration of /o/ 100 ms longer than o-0.

o-6 : The duration of /o/ 120 ms longer than o-0.　　(Typical /oobasan/.)

Set-A was concerned with the distinction between /obasan/ and /oba'asan/. This set included the following seven synthetic stimuli.

a-0 : Synthetic version of the original /obasan/ with accent.

a-1 : The duration of /a/ of syllable /ba/ 20 ms longer than a-0.

a-2 : The duration of /a/ 40 ms longer than a-0.

a-3 : The duration of /a/ 60 ms longer than a-0.

a-4 : The duration of /a/ 80 ms longer than a-0.

a-5 : The duration of /a/ 100 ms longer than a-0.

a-6 : The duration of /a/ 120 ms longer than a-0.　　(Typical /oba'asan/.)

In the synthesis of set-A, the accentual phrase of /oba'asan/ was considered to be a "degenerate accentual phrase" (Maekawa, 1994a). Degenerate accentual phrase is, roughly speaking, an accentual phrase which lacks the initial rise of pitch that usually marks the beginning of an accentual phrase in Tokyo Japanese. It was because of the career sentence used in the experiment (See below). Set-A', the last set, was identical to the set-A except that the members of this set did not have accent. The members of this set will be refereed to as a'-0 and so on.

Each of the stimuli in Set-O and Set-A was presented ten times in random order to experimental subjects. The task was a three-way forced distinction among /obasan/, /oobasan/ and /oba'a san/. Each stimulus was presented in a career sentence of /da'rega X ni at'tano?/ (Who met X?), where X is the designated location of a stimuli. Since the stimuli was located immediately after a WH word in a WH question sentence, the initial rise of the accentual phrase corresponding to X was not usually realized under the effect of focus on the WH word (See Maekawa, 1991a,b; 1994a).

The experiment was conducted both in Tohoku University in Sendai and The National Language Research Institute in Tokyo. The subjects of the experiment involved seven Koreans. They were graduate students majoring in various research fields including linguistics and Japanese. Preliminary result of this experiment revealed two interesting facts. First, Korean subjects were generally *more* sensitive to the phonetic difference of the vowel duration both in word-initial and word-medial positions. According to the judgment given by two Japanese subjects, the perception

threshold (the 50% probability point) between word-initial short /o/ and long /oo/ was in the vicinity of o-2, while the threshold between word-medial /a/ and /a'a/ was in-between a-2 and a-3. The thresholds of Korean subjects located much earlier on the time axis. As for Set-O, the threshold was mostly on stimuli o-1. But some others' threshold was much earlier and was in-between o-0 and o-1. As for Set-A, the threshold was in-between a-1 and a-2.

Does this mean that Korean are more sensitive to the durational difference than Japanese are? Definitely not. The fact is that Koreans listened to the stimuli phonetically, while Japanese listened to them phonologically. In other words, Japanese subjects did not judge a stimulus to have a long vowel unless the duration of the target vowel was long enough to be counted as two morae long, while Koreans judged any stimuli having longer-than-usual vowels to be long.

Another finding, and perhaps the most interesting one, was concerned with the individual difference among Korean subjects. There were three Korean subjects who responded to the stimuli of Set-A in a peculiar way. Their response showed marked tendency of perceiving the younger stimulus of Set-A (i.e. a-0, a-1, a-2, a-3 and a-4) as /oobasan/. One of them even judged the stimulus a-0 and a-1 to be /oobasan/ nine and six times respectively over the ten presentations. Surprisingly, all of these subjects were from the cities of *Chollanamdo*. If this is really the interference caused by their native dialect, it seems that, in Chonnam, the distribution of a (word-initial) long vowel is closely correlated with the pitch fall somewhere in the word.

In order to clarify this point, we have conducted the second perception experiment, in which the Set-A' was used in place of the Set-A. For now, only two Chollanamdo speakers have participated in this experiment. But the results were quite striking ones. When the accent was deprived of the stimuli of Set-A, no Chonnam subjects showed the tendency to perceive the stimuli of Set-A' as /oobasan/. This result confirms strongly the above mentioned view on Chonnam correlation between a long vowel and a pitch fall.

Unfortunately, there seems to be no description regarding this correlation in Chonnam. Jun (1993), the only source of information regarding the tonal structure of Chonnam dialect as far as I know, tells nothing about this. But it is very interesting to note that Jun reported that in Chonnam it was always the initial consonant of an accentual phrase that determined the overall tonal structure of the phrase. If a consonant can determine the tonal structure of a phrase, a long vowel would be able to do it as well, though it is only a logical inference at the time being.

3. Note on 'Closed syllable' in Japanese

The last topic of this talk is concerned with the phonetic realization of syllable. As is well known, many linguists suppose that (Tokyo) Japanese has a prosodic constituency of *syllable* as distinct from *mora*, at least at the level of phonology. In this analysis, Japanese has a series of closed syllables like /sak/ in /sakka/ (writer) and /saN/ in /saNka/ (participation). These seemingly "closed" syllables are counted as two morae long. Although it seems to be true that the notion of

syllable is indispensable for the adequate description of the accentual system of Tokyo and other dialects, the nature of the "syllable" in this language has never been investigated from a point of view of speech production and perception. Here, I would like to show one important, but unknown, characteristics of the "syllable" in Japanese.

In Maekawa (1984), I compared the syllable structure of *Akita* Japanese and Tokyo Japanese from a point of view of segmental duration. Akita Japanese is known to be a syllable dialect. A syllable dialect is a language that lacks the prosodic constituency of mora. In this sense, it could be stipulated that the syllable structure of Akita is more similar to that of the languages like Korean than Standard Japanese. More specifically, it was expected that syllable final consonant —so-called *coda*— of Akita was more closely connected to the preceding vowel —the *core* of a syllable— than that of Tokyo Japanese; because Akita does not have mora, which gives an independent status to the syllable final consonant in Tokyo, the consonant has no other choice but to be affiliated to the preceding vowel. For example, /sakka/ should have the syllabic structure of /sak/ + /ka/, and , accordingly the vowel duration of /a/ in /sak/ should be different from that of /a/ in /saka/, which is analyzed as /sa/ + /ka/.

However, what I found was completely counterintuitive. It was not the preceding vowel but the following vowel that was affected by the gemination of the word-medial consonant. The duration of the word-medial consonant was doubled by the effect of gemination. And the duration of the preceding vowel showed no change. But the duration of the following vowel was reduced to some extent as if it compensated the durational increment caused by the gemination. As the result, the duration of the whole word does not reflect the increment caused by the gemination. This is why the auditory impression of Akita pronunciation is often described as "choked" by the speakers of Tokyo and other mora dialects. Exactly the same conclusion was drawn by Takada (1985) in his study of *Aomori* dialect, another syllable dialect.

These findings might be considered to be irrelevant for the understanding of Tokyo, or Standard, Japanese, which is the target language of the teaching of Japanese. But it is not so. The data presented in Maekawa (1984) and Takada (1985) show clearly that it is also true for Tokyo Japanese that consonant gemination and the resulting change in syllable structure does not affect the duration of the preceding vowel. The fact is, the duration of the preceding vowel is often longer in so-called closed syllables than in open syllables.

The shortening of the vowel duration in closed syllable (CSVS:Closed Syllable Vowel Shortening) is a very common phonetic cue to syllabification among very many numbers of world's languages (Maddisson, 1984). In this sense, Japanese can be said to be a very peculiar language regarding the phonetics of syllable. The "syllable" in Japanese is not a syllable in the common sense of the word, and is very dissimilar to the Korean syllable, because the latter shows clear CSVS.

Very recently, I had a chance to conduct an interesting, but very anecdotal, perception experiment regarding the syllabification in both languages. In this experiment, the phonemic

sequence of /kana/ was used as speech material. The duration of /n/ and its preceding /a/ were 60 and 90 ms respectively. In one set of stimuli, the duration of /n/ was lengthened by about 50 ms step until it reached the duration of typical realization of geminated /kanna/ (about 250 ms). This set involved 5 stimuli. Another set of stimuli was prepared by reducing the duration of /a/ in /na/ to about 60 ms. As the result we had 10 stimuli as a whole. All these manipulation was done by cutting and pasting the original waveform. These stimuli were presented in random order to two subjects, one Korean and one Japanese. Subjects were asked to judge if the word they heard was /kana/ or /kanna/.

The result showed clearly that vowel shortening could have different effects on both languages. In Korean, the shortened vowel favored the perception of /kanna/, while in Japanese it favored the perception of /kana/. I still don't know if this result can be generalized over to other subjects in both languages. But the result was exactly what was predicted by the fact stated above. The different natures of closed syllable in Korean and Japanese should be recognized by those who are involved in the education of Japanese as a foreign language, anyway.

REFERENCES

Jun, Sun-Ah (1993) *The phonetics and phonology of Korean prosody*, Ph.D. diss., The Ohio State University, Columbus.

Jun, Sun-Ah and Mira Oh (1994) "A prosodic analysis of three sentence types with "WH" words in Korean," in *Proc. International Conference on Spoken Language Processing*, pp.323-326, Yokohama.

Maddieson, Ian (1984) "Phonetic cues to syllabification," *UCLA Working paper in Phonetics*, 59, pp.85-101.

Maekawa, Kikuo (1984) "Akita hoogen sokuonno jizokujikan," (The duration of geminates in Akita), *Hoogen Kenkyuu Nenpoo*, 27, pp. 231-247, Izumi Shoten, Kyoto.

———— (1990) "Muakusento hoogenno intoneeshon: Shiron," (Intonation of accentless dialects; A preliminary), *Onseigengo*, 4,

———— (1991a) "Perception of intonational characteristics of WH and NON-WH questions in Tokyo Japanese," in *Proc. 12th International Congress of Phonetic Sciences*, Aix-en-Provence, France.

———— (1991b) "Tookyoo hoogen gimonbunno intoneeshon," (Intonation of Tokyo Japanese questions), in *Proc. of the 1991 Annual Convention of the Phonetic Society of Japan*, pp. 42-27.

———— (1994a) "Is there dephrasing of the accentual phrase in Japanese?" *OSU Working Paper in Linguistics*, pp. 146-165, The Ohio State University, Columbus.

———— (1994b) "Intonational structure of Kumamoto Japanese: A perceptual validation," in *Proc. International Conference on Spoken Language Processing*, pp. 119-122, Yokohama.

Min, Guan-Jun (1994) "Kankokugo gimonbun intoneeshonno onkyooteki bunsekito gouseionseiniyoru chikakujikken," (An acoustic analysis and perception experiment of intonation for WHJ and YN questions in Standard Korean), The Bulletin of Phonetic Society of Japan, 205, pp. 29-33.

Takada, Masaharu (1985) "Sokuonno chooonjoono tokuchooni tsuite," (On some articulatroy characteristics of the mora obstruent), *Occasional papers*, 6, pp.17-40, The National Language Research Institute, Tokyo.

Umeda, Hiroyuki and D. Kim (1989) *Standard Hanguru Kooza*, 1, Taishuukan, Tokyo.