Acoustic Measurements of Korean Approximants

Mi-Ran Cho Kim and Andrew J. Lotto
University of Georgia *Boys Town National Research Hospital*

1. Introduction

Approximants are speech sounds that are produced with a markedly narrowed vocal tract but not narrowed enough to result in a turbulent airstream (Ladefoged, 2001; Kent & Read, 2002). Korean has three approximant phonemes, /l, w, j/. The phoneme /l/ is often described as a voiced alveolar lateral liquid. Although the Korean /l/ can occur in several different phonetic environments, it has only two distinct allophones, the flap [ɾ] and [l]. In word-initial position, /l/ is deleted when it is followed by [i] or [j] or changed to [n] elsewhere (Sohn, 1999). Word-initial phoneme /l/ is produced as the flap [ɾ] only in loan words (e.g., "radio", "Rome") or in a few Korean last names. The flap [ɾ] is produced with a very brief tongue movement touching the alveolar ridge. In inter-vocalic position, /l/ is also produced as the flap [ɾ]. In word-final position, /l/ is produced as the so-called light [l] which is made with the tongue tip (and/or blade) touching the alveolar ridge while lowering the sides of tongue to allow the airstream to escape.

The approximants /w/ and /j/ are also called glides. In Korean, the glides occur mostly in prevocalic position. The glide articulation involves a relatively slow movement of the vocal tract from an initial target towards the following vowel. The glide [w], which occurs before the vowels [i, e, ɛ, ə, ɑ], has a similar initial vocal tract configuration to the vowel [u]. The glide [j], which occurs before the vowels [e, ɛ, ə, ɑ, u, o], has a similar initial vocal tract configuration to the vowel [i].

In contrast to Korean, American English has four approximant phonemes, /l, r, w, j/. Phonetically, English /l/ has two allophones. The lateral /l/ in word-initial position is a light [l] which is very similar to the Korean [l] in word-final position. In contrast, the word-final /l/ is always produced as a dark [l] which is also called a velarized [ɬ], as in 'feel, field' (Ladefoged, 2001). The glides /w/ and /j/ also occur mostly in prevocalic position. Even though the flap [ɾ] is included in English phonetic inventory, it is a phonetic realization of both /t/ and /d/ when it occurs between a stressed vowel and an unstressed vowel.

The present study investigates the acoustic characteristics of Korean approximants [l, ɾ, w, j]. Even though both Korean and English contain all four approximant sounds in their phonetic inventory, the difference of the phonemic realization of each sound and the different phonetic context of each sound in the two languages may cause difficulties for second language learners (Ingram & Park,

1998). For Korean speakers, for example, it is very difficult to produce the light lateral [l] in word-initial position. It is also difficult for Korean speakers to produce [w] before [ʊ], as in 'wolf, wood'. For English speakers, it is difficult to produce the flap [ɾ] word-initially and the light lateral [l] word-finally. The present study provides more precise acoustic phonetic descriptions of the Korean approximants in different phonetic contexts. For the glides, we limit our attention to the syllables [wi] and [ju] in order to make the findings compatible to the previous measures, which are mostly from English. The close comparisons of Korean and English are expected to provide (1) better understanding of the language learners' difficulty and (2) more efficient teaching techniques in language teaching situation. Recent detailed acoustic comparisons of English liquids with Japanese phonemes have yielded insights into the task of the second language learner (Lotto, Sato & Diehl, 2004). This project is part of an effort to extend this methodology to the teaching of Korean and English

2. Acoustic Characteristics of Approximants

Phonetically, approximants are less sonorous than vowels but much more sonorous than other consonants. Like vowels, lateral [l] shows well defined formant-like resonances. Because of this, the acoustic characteristics of the lateral are often characterized in terms of the first two or three formants at the point of maximal vocal-tract narrowing of the articulation (Nolan, 1983; Lehiste, 1964). Also, a bifurcation of the vocal tract for /l/ produces antiformants. That is, the lateral /l/ has most of its acoustic energy below 5kHz, whereas there is not much (or any) energy above 5 kHz. The flap [ɾ] is acoustically characterized by a short duration with very brief closure and release periods.

The acoustic characteristics of the glides are also well represented by the first two or three formants. Both glides present characteristics intermediate between stops and vowels. For [wi], the formant transition pattern is very similar to the stop consonant [bi] and the vowel [ui]. For [ju], the formant transition pattern is very similar to the stop consonant [du] and the vowel [iu]. The main difference between these sounds is the duration of the transition. The transition duration of [bi] and [du] is the shortest, while the transition duration of [ui] and [iu] is the longest. The transition duration of [wi] and [ju] is intermediate (Kent & Read, 2002). Because the glides are produced with a more extreme constriction than the vowels, [w] has lower F1 and F2 frequencies than [u]. Likewise, the glide [j] has a lower F1 frequency and a higher F2 frequency than [i]. In general, [w, j] have lower F1 frequencies than the lateral [l], and [j] has a much higher F2 frequency than [w].

3. Experiment

3.1 Subjects

Twelve native speakers of Korean, six males and six females, participated in this experiment. They are all speakers of Seoul dialect, which is considered standard Korean. The subjects ranged from 22 to 34 years in age. Eleven participants were graduate students at the University of Georgia. One student was enrolled in the American Language Program. The subjects had lived in the U.S. for 3 months to 4 years at the time of recording. All subjects reported having normal speaking and hearing ability.

Table 1. List of Nouns

Segment	Noun
[l]	thal 'mask', tal 'moon', khal 'knife', phal 'arm', pal 'foot'
[ɾ]	salam 'person', palam 'wind', salang 'love', calang 'boast', talak 'attic'
[w]	wiki 'crisis', wipay 'violation', wisayng 'hygiene', wico 'forgery', wihem 'danger'
[j]	yucen 'heredity or oil field', yuhak 'studying abroad', yukyo 'Confucianism', yusan 'inheritance', yumwul 'relic'

3.2 Stimuli

For data collection, a reading list containing five tokens for each approximant was prepared. The approximants were embedded in slightly different sentences to make the segment more distinct from the neighboring segments and to make sentences more meaningful. For the flap [ɾ], nouns with CV.CVC syllable structure that contained /l/ in the onset position of the second syllable were embedded in the carrier sentence, "i (this) + kes (thing) + i (nominative particle) NOUN ipnita (is)" (in the Yale romanization system), which means "This is a noun" in English. The vowels were fixed as [a] in both syllables. For the light lateral [l], the mono-syllabic nouns with CVC structure that contained /l/ in the coda position were embedded in the carrier sentence, "i (this) + kes (thing) + un (topic-contrast particle) NOUN kathayo (looks like)", which means "(As for this,) this looks like a noun". In order to provide a clear boundary for the lateral, a stop consonant containing a complete closure was selected as the following segment. The topic-contrast particle –"un" was selected to make the sentence more natural. For [w, j], the two syllable nouns

starting with [wi] and [ju] were embedded in the carrier sentence, "i (this) + kes (thing) + un (topic-contrast particle) NOUN ipnita (is)", which means "(As for this,)This is a noun". Again the topic-contrast particle – "un" was selected in order to provide a clear boundary for the onset of each glide. The nouns used for the experiment are listed in Table 1. They are transcribed in the Yale romanization system.

3.3 Procedure

The recordings were made using a noise-canceling Shure SM48-LC microphone and a Sony Digital Audio Tape recorder with a 44-kHz sampling rate. The distance of the microphone from the mouth was approximately 4 inches. The recordings were made in a quiet chapel on the University of Georgia campus. Each subject was given a reading list and instructed to read the sentences as naturally as possible. The subject was also asked to keep speech rate constant. All of the sentences were read three times.

3.4 Acoustic Analysis

Upon completion of recordings, all signals were digitized at a sampling rate of 22 kHz. In total, 720 tokens (4 approximants x 5 tokens x 3 repetitions x 12 subjects) were acoustically analyzed from the spectrographic and waveform display of the TF32 program implemented in CspeechSP (Milenkovic & Read, 2000). The formant frequency measures were made from the Fourier transform spectrum. For the flap [ɾ], the duration of the segment was measured from the spectrographic display of the sentence. For the other approximants, since the acoustic change from/to vowels are often not distinct, the duration of the segment is not reported. For the light lateral [l], the first two formant frequencies were measured in the center of the segment that showed steady state formants. For [w], the first two formant frequencies were measured around the time of the F2 minimum. For [j], the first two formant frequencies were measured around the time of F2 maximum.

3.5 Results

3.5.1 Flap [ɾ]

Table 2 presents the mean duration with standard deviations of each subject. For male speakers, the average duration of the flap was 21.19 ms versus 27.06 ms for female speakers. This gender difference was statistically significant (t = 3.07, p<0.01). Previous acoustic analyses of Korean stop production demonstrated similar gender differences for temporal characteristics of speech (voice onset time and

closure duration; Kim, 1994). In general, Korean female speakers appear to produce slower articulations than males. This may be an attempt to enhance intelligibility in the face of a higher voice pitch. These durations are also quite a bit shorter than the flap durations measured by Sato, Lotto and Diehl (2003) for Japanese syllable-initial flaps (48.5 and 31.4 ms for males and females, respectively). The Korean flap is a particularly fast articulation, especially in contrast to the glides.

Table 2. Mean and standard deviation (in parentheses) of duration of flap (in ms.)

Subject	Mean (SD)	Subject	Mean (SD)
Male #1	18.0 (3.5)	Female #1	30.0 (8.8)
Male #2	21.4 (3.0)	Female #2	29.3 (7.1)
Male #3	19.9 (3.6)	Female #3	24.4 (5.2)
Male #4	26.4 (4.9)	Female #4	29.8 (4.8)
Male #5	21.0 (4.0)	Female #5	20.7 (1.9)
Male #6	20.4 (4.7)	Female #6	28.2 (5.1)

3.5.2 Lateral [l]

Table 3 presents the mean formant frequencies with standard deviations of the first two formants for the lateral [l] of each subject. The average male F1 and F2 values were 499.19 and 1644.41 Hz, respectively. The corresponding female means were 595.32 and 2092.48 Hz. (Higher female formants are expected given their typically smaller vocal tract.) The F2 values are quite a bit higher than typical values of [l] in American English for syllable-final or syllable-initial position with similar vowel context, which tend to be in the range of 980-1350 Hz (Sato *et al.*, 2003, see also, Espy-Wilson, 1992; Nolan, 1983; Lehiste, 1964). In fact, these F2 values are closer to those for English syllable-final [r] (1900-2250 Hz, Sato *et al.*, 2003).

Table 3. Mean and standard deviation (in parentheses) of F1 & F2 for lateral [l]

Subject	Mean (SD)		Subject	Mean (SD)	
	F1	F2		F1	F2
Male #1	474 (62)	1491 (59)	Female #1	587 (112)	2067 (149)
Male #2	462 (27)	1615 (65)	Female #2	616 (37)	2111 (60)
Male #3	567 (56)	1663 (49)	Female #3	536 (94)	2020 (114)
Male #4	450 (60)	1702 (50)	Female #4	539 (79)	2390 (123)
Male #5	541 (43)	1698 (55)	Female #5	627 (105)	1880 (164)
Male #6	500 (53)	1699 (56)	Female #6	677 (52)	2109 (70)

3.5.3 Glide [w]

Table 4 presents the mean formant frequencies with standard deviations of the first two formants for [w]. The average male F1 and F2 values were 265.41 and 1575.24 Hz, respectively. The corresponding female means were 371.32 and 1825.72 Hz. These F2 values are also quite a bit higher than averages reported for English productions of [w] (mean = 848 Hz; Espy-Wilson, 1992).

Table 4. Mean and standard deviation (in parentheses) of F1 & F2 for [w]

Subject	Mean (SD)		Subject	Mean (SD)	
	F1	F2		F1	F2
Male #1	249 (16)	1846 (113)	Female #1	423 (17)	1963 (208)
Male #2	231 (20)	1281 (116)	Female #2	429 (32)	2203 (205)
Male #3	297 (11)	1168 (251)	Female #3	364 (12)	1806 (186)
Male #4	292 (25)	1655 (108)	Female #4	366 (17)	1829 (369)
Male #5	258 (11)	1894 (114)	Female #5	309 (12)	1871 (125)
Male #6	266 (35)	1608 (378)	Female #6	336 (8)	1261 (184)

3.5.4 Glide [j]

Table 5 presents the mean formant frequencies with standard deviations of the first two formants for [j]. The average male F1 and F2 values were 265.72 and 2096.21 Hz, respectively. The corresponding female means were 367.80 and 2741.19 Hz. Note that the F1 values are nearly identical to those obtained for [w]. It is clear that these two glides are readily distinguished by the difference in F2 values. The F2 values of [l] are intermediate between the two glides.

Table 5. Mean and standard deviation (in parentheses) of F1 & F2 for [j]

Subject	Mean (SD)		Subject	Mean (SD)	
	F1	F2		F1	F2
Male #1	251 (14)	2000 (85)	Female #1	425 (19)	2785 (56)
Male #2	216 (11)	2039 (76)	Female #2	403 (23)	2496 (95)
Male #3	292 (10)	2417 (334)	Female #3	367 (34)	2627 (83)
Male #4	291 (30)	2056 (190)	Female #4	352 (30)	2743 (182)
Male #5	270 (14)	2029 (283)	Female #5	314 (16)	2968 (225)
Male #6	275 (43)	2099 (111)	Female #6	346 (22)	2842 (51)

4. Discussion

The purpose of this project was to measure some of the distinctive acoustic characteristics of the Korean approximants. The overall goal is to predict some of the difficulties that English speakers may have learning the Korean contrasts (and vice versa) and to provide norms for successful Korean productions. Several researchers have proposed that one can predict second language perception-production difficulties by examining the overlap between acoustic distributions for phonetic categories in the native and second language (e.g., Flege, 1995; Lotto *et al.* 2004). For example, the fact that the Korean word-final [l] has F2 values that overlap with English word-final [r] suggests that there may be some category interference for learners trying to acquire one of these allophones. The frequency of the third formant is likely to help disambiguate [l] from [r], but it is still likely that the mismatch on English and Korean [l] productions will lead to some difficulties.

Another potential difficulty for an English-Korean second language learner is the distinction between [w] and [j]. While these glides are distinguishable by F2 in both languages, the range of F2 values for [w] is much higher for Korean speakers. There are clearly some female productions of [w] that would be in the range of the English [j] (mean F2 = 2142 Hz, prevocalic, Espy-Wilson, 1992). The extent of the overlap between these category distributions is unknown at this point. Future work will examine the relationship between English and Korean glide distributions more extensively.

Lastly, it is interesting to note that [w] does not occur before high back vowels in Korean phonotactics. These vowels are distinguished by low F2 values. The higher F2 values of Korean [w] proceeding these vowels would lead to a pattern of decreasing F2 formant transitions; a pattern typical of [j]. This may be one reason why Korean speakers have problems with English words like "wolf" and "wood".

References

Espy-Wilson, C. Y. (1992). Acoustic measures for linguistic features distinguishing the semivowels /w j r l/ in American English. *Journal of the Acoustical Society of America, 92*, 736-757.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 233-273). Baltimore: York.

Kent, R. & Read, C. W. (2002). *The acoustic analyses of speech* (2nd ed.). San Diego: Singular Publishing.

Kim, M.-R. Cho (1994). *Acoustic characteristics of Korean stops and perception of English stop consonants.* Doctoral dissertation, University of Wisconsin, Madison.

Ingram, J. C. L. & Park, S-G. (1998). Language, context, and speaker effects in the identification and discrimination of English /r/ and /l/ by Japanese and Korean listeners. *Journal of the Acoustical Society of America, 103*, 1161-1174.

Ladefoged, P. (2001). *Vowels and Consonants: An introduction to the sounds of languages*. Oxford, UK: Blackwell.

Ladefoged, P. & Maddieson, I. (1996). *The sounds of world's languages*. Oxford, UK: Blackwell.

Lehiste, I. (1964). Acoustical characteristics of selected English consonants. *International Journal of American Linguistics, 30, No. 3*, 181-223.

Lotto, A. J., Sato, M., & Diehl, R. L. (2004). Mapping the task for the second language learner: The case of Japanese acquisition of /r/ and /l/. Presentation at *From Sound to Sense: 50+ Years of Discoveries in Speech Communication*, Boston, MA, June 2004.

Milenkovic, P. H. & Read, C. (2000). CSpeechSP [Computer software]. Madison, WI:Author.

Nolan, F. (1983). *The phonetic bases of speaker recognition.* Cambridge, UK: Cambridge University Press.

Sato, M., Lotto, A.J., & Diehl, R.L (2003). Patterns of acoustic variance in native and non-native phonemes: The case of Japanese production of /r/ and /l/. Presentation at the 146th Annual Meeting of the Acoustical Society of America, Austin, TX, November 12, 2003.

Sohn, H. M. (1999). *The Korean Language.* Cambridge, UK: Cambridge University