Production of Aspiration in English Voiceless Stops in Terms of Place of Articulation*

Ji Yea Kim
(Seoul National University)

Kim, Ji Yea. 2012. Production of Aspiration in English Voiceless Stops in Terms of Place of Articulation. *SNU Working Papers in English Linguistics and Language X, XX-XX.* This paper aims to investigate whether place of articulation (POA) has an influence on the production of aspiration in English voiceless stops. A production experiment was conducted with native speakers of English and Korean. The participants were required to read a set of stimuli, which contained English voiceless stops in the word-initial position and in the position immediately following /s/. The VOT values in their production were measured. The results showed that in both of the groups, which were classified according to the subjects' native language, dorsal stops were the highest in VOT values. This is ascribed to the language-universal aerodynamic reason, which is related to the nature of the oral cavity. **(Seoul National University)**

Keywords: aspiration, English voiceless stops, place of articulation (POA), production, voice onset time (VOT)

1. Introduction

Much research has been conducted on the difficulties of producing L2 phonemes which are not observed in one's L1 phonetic system. B.-E. Cho (2004), for example, pointed out that it appears to be difficult for Korean learners of English to produce the English consonants /f/ and /v/ as native speakers of English do because these voiceless and voiced labiodental fricatives do not exist as phonemes in the Korean phonetic system.

---

* This paper is a shortened version of a part of the author's master's thesis. Portions of the first draft of the thesis were presented at the 2011 Annual Conference of the Applied Linguistics Association of Korea (ALAK), Seoul, Korea, November 12, 2011.

What makes L2 production more challenging is that phonemes, which are regarded to exist both in L1 and L2, are in some cases categorized differently based on which language they belong to. Keating's (1984) analysis on the voiced stops /b, d, g/ and the voiceless stops /p, t, k/ are intriguing in this sense. It was argued in her paper that, even though the two kinds of stops are phonologically contrastive (i.e., the contrast between voiced stops and voiceless stops) both in English and Polish, their phonetic implementation might be different. While phonologically voiceless stops in Polish are at all times phonetically realized as voiceless unaspirated stops, English voiceless stops have two options for their phonetic implementation: voiceless unaspirated stops and voiceless aspirated stops.

Consequently, what influences aspiration in English can be a crucial question. The answer to this question is closely related to the environment in which a voiceless stop appears within a word. On the basis of the relation between aspiration and context, the current study seeks out to examine whether place of articulation (henceforth, POA) also has an influence on the production of aspiration in English voiceless stops by native speakers of English and Korean. The degree of aspiration is represented as voice onset time (henceforth, VOT) in this research. The research questions are as follows:

Research questions

1. Does POA have a significant effect on the VOT values of English voiceless stops in the word-initial position and those in the position following /s/?
2. Is there any difference between native speakers of English and Korean? How are they different or similar?

In order to answer these research questions, one production experiment was conducted.

2. Previous studies

Aspiration has been defined as “a puff of air” which occurs during the period of the vocal folds being open before the voicing of the following vowel begins (Gussenhoven and Jacobs 2005, Ladefoged 2006). By mentioning the phonological feature [spread glottis], Vaux (2002) argued that voiceless obstruents are the target of aspiration.

Voiceless stops in English and Korean undergo the unique language-specific implementations. In English, the environment where a voiceless stop appears within a word decides whether the stop is aspirated or unaspirated. The following rule-based theoretic process encapsulates what was argued by most researchers as the environment of aspiration in English.

(1) A rule-based theoretic description of the environment of aspiration in English (based on Clark and Yallop 1995, Morton and Tatham 1980, Silverman 1998)[1]

$$\overset{\text{C}}{[+\text{stop}, -\text{voice}]} \rightarrow [+\text{aspiration}] \ / \ \#\ \underset{[+\text{stress}]}{(\text{V})}$$

The rule above indicates that, when placed in the word-initial position and followed by a (stressed) vowel (V), voiceless stop consonants (C) are realized as voiceless “aspirated” stops (e.g., $[p^h]$ in *pill*).[2] On the

---

1 It was argued in Clark and Yallop (1995) that voiceless stops surface with aspiration when they are in the word-initial position. Morton and Tatham’s (1980) description is in more detail, mentioning the presence of a vowel immediately following the voiceless stops. Besides, Silverman (1998) required that the vowel should bear a primary stress. Following Silverman (1998), the present study seeks to examine English aspirated stops that occur in the word-initial position and make the consonant cluster /s + p, t, k/.

2 As for the environment in which voiceless stops are aspirated in English, the syllable-initial position and the foot-initial position as well as the word-initial position should

contrary, voiceless stops are unaspirated when they immediately follow the fricative /s/ and make the consonant clusters /s + p, t, k/ (e.g. [p] in *spill*).[3] It is notable that the two different kinds of stops (i.e., aspirated stops and unaspirated stops) in English are not regarded as allophones but as different phonemes, resulting in a difference in meaning. Furthermore, unaspirated stops are subdivided into 2 kinds of phonemes: lenis (or lax) stops and fortis (or tense) stops. Thus, the Korean language is considered to be a three-category language (Lisker and Abramson 1964) consisting of 3 types of voiceless stops: aspirated stops /p^h, t^h, k^h/, unaspirated lax stops /p, t, k/, and unaspirated tense stops /p’, t’, k’/ (e.g., /p/ in [p^hul] ‘grass’, /p/ in [pul] ‘fire’, and /p/ in [p’ul][4] ‘horn’, respectively).

3. PRODUCTION EXPERIMENT

A production experiment was conducted according to the following procedure.

3.1 Methods

3.1.1 Subjects

A total of 25 subjects were recruited in this production experiment. Five of them were native speakers of English (3 male speakers and 2 female

---

be mentioned (C.-W. Kim 1970, Iverson and Salmons 1995, Vaux 2000). In the present study, however, only the voiceless stops /p, t, k/ in the word-initial position are examined as the ones bearing aspiration for the sake of convenience.

[3] However, Vaux (2002) pointed out that this simplistic /s/ theory fails to explain the case in which English voiceless stops following /s/ do surface with aspiration. In his examples *dis*[t^h]*aste* and *dis*[t^h]*end*, voiceless stops are aspirated even though they are preceded by the fricative /s/. It is attributed to the fact that /s/ and /t/ do not belong to the same syllable. In this sense, the present study restricts its scope only to English voiceless stops which follow the fricative /s/ and make a single syllable with the fricative.

[4] The diacritic “’” is used to represent tense stops. In some research including M.-R. Kim *et al*. (2002), tense stops are represented with the diacritic “*”.

speakers) and the rest were native speakers of Korean (12 male speakers and 8 female speakers). The subjects were either students or faculty members of Seoul National University in Korea.

The group of native English speakers (henceforth, NES) consisted of two American English speakers, one Canadian speaker, and two British English speakers[5]. They ranged from 23 to 48 in age (mean = 34.6, SD (Standard Deviation) = 9.07).

As for the 20 native Korean speakers, their ages ranged from 20 to 30 years (mean = 23, SD = 2.66). They were subdivided into two groups depending on their English proficiency levels, which were represented as TEPS[6] scores: Korean high-level proficiency speakers of English group (henceforth, KHS: 10 students whose TEPS scores were over 800 (mean = 864.4, SD = 44.29)) and Korean low-level proficiency speakers of English group (henceforth, KLS: 10 students whose TEPS scores were below 800 (mean = 581.2, SD = 123.52)).[7] None of the Korean participants had lived in any English speaking countries. All subjects were paid for their participation.

3.1.2 Stimuli

Thirty-six English real words which contain the voiceless stops /p, t, k/ in their word-initial position and the in the sequence /s + T[8]/ were selected as the stimuli of the production experiment. According to Table

---

5 The rationale behind the inclusion of British English speakers along with North American English speakers can be found in Docherty (1992). According to his account of the aspects of British English quoting Davidsen-Nielsen (1974), the environments of voiceless stops being aspirated and unaspirated in British English are not different from those in North American English. That is, voiceless stops are accompanied by aspiration when they are in the sequence /s$pl/ while they are unaspirated in the sequence /$spl/ (where the sign "$" represents a syllable boundary).

6 Test of English Proficiency developed by Seoul National University

7 Since the subgroup KLS includes native Korean speakers in the low English proficiency level and the ones in the intermediate English proficiency level, it would sound more logical if this group was named the non-high-proficiency group. However, the simpler one was used in this study for the sake of convenience.

8 Voiceless stops /p, t, k/ are represented as "T".

1, POA (3 kinds: labial, coronal, and dorsal stops) and the height of the following vowel (3 kinds: high, mid, and low vowels) as well as the target voiceless stops' position within a word (2 kinds: the word-initial position (i.e., #_V) and the position following /s/ (i.e., #s_V) were taken into consideration.

Table 1. The stimuli used for the production experiment

POA		Labial			Coronal			Dorsal		
V height		H	M	L	H	M	L	H	M	L
Context	#_V	pill	pun	par	tick	tuck	tar	kick	cunning	car
		pin	purr	pot	tin	tongue	taco	kin	curt	cocktail
	#s_V	spill	spun	spar	stick	stuck	star	skid	scunner	scar
		spin	spur	spot	sting	stung	stop	skin	skirt	scarf

Based on the 2 types of context, each stimulus in the #CV syllable (where C here represents the voiceless stops /p, t, k/) was paired with the other one in the #CCV syllable (where the first C here represents /s/ and the second /p, t, k/). They share the information of POA and vowel height. In this way, *pill* and *spill*, for example, made one pair, and *kick* and *skid* made another pair[9].

The stimuli were scrambled in a random order to prevent the participants from recognizing the objective of this research.

3.1.3 Procedure

To begin with, the subjects were required to fill out a questionnaire which was designed to ask them about their language backgrounds.

The next procedure was to read the 36 randomly ordered stimuli with 3 repetitions. The subjects were asked to read each of them in the carrier phrase "*I said* ____ *today.*" in a natural tone and at a normal

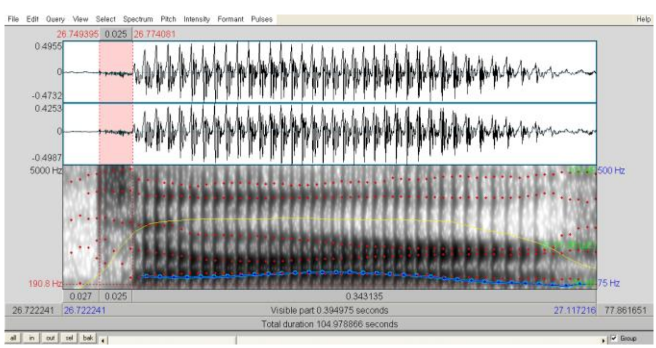
[9] Since the 2 kinds of context examined in this research were related only to the onset of a syllable, the difference in codas (e.g., /k/ in *kick* and /d/ in *skid*) was ignored.

speed. The objective of using a carrier phrase was to prevent the participants from making excessively unnatural productions. The 108 stimuli (36 words in a set of stimuli x 3 repetitions) that the subjects read were recorded by a Zoom H4 Handy digital recorder in a silent room.

3.1.4 Measurement

An acoustic analysis of the recordings was conducted with the aid of the computer program Praat[10]. H. Kim *et al*. (2004) and Chang (2007) served as the guidelines on measuring the VOT values of the voiceless stops in the recordings: VOT values were measured from the onset of the release burst of a stop to the point where the vibration of the glottis begins. An example which shows the way to measure the VOT value of /t/ in *star* is in Figure 1.

Figure 1. The waveform and spectrogram of the VOT of /t/ in *star* on Praat



The selected area shows the duration of the aspiration produced by NES 5 when he was reading /t/ in *star*. The value was then converted into

---

[10] *Praat: doing phonetics by computer* (http://www.fon.hum.uva.nl/praat/).

milliseconds (msec.). In the same way, the aspiration of the voiceless stops in the 2,700 words (36 words x 25 subjects x 3 repetitions) were measured in terms of their VOT values.

4. Results and discussion

The main purpose of this section is to examine the relation between VOT values and POA as one of various factors influencing the VOT values[11]. The means and standard deviations of the VOT values were classified according to the groups of the subjects (i.e., NES, KHS, and KLS), the contexts, and the POAs.

Table 2 shows NES' means and standard deviations of VOT values.

Table 2. The means and standard deviations of NES' VOT values categorized according to the context and POA

Context	#_V			#s_V		
POA	Labial	Coronal	Dorsal	Labial	Coronal	Dorsal
mean (msec.)	76.87	82.71	94.05	15.17	20.18	24.11
SD	21.48	52.11	86.56	16.53	9.01	10.03

The voiceless stops in the word-initial position (i.e., #_V) were higher than those in the position following /s/ (i.e., #s_V) when it comes to NES' mean VOT values. It is notable that the VOT values show a gradual increase from labial stops to coronal stops and dorsal stops in

---

11 Other factors in addition to POA do have an impact on VOT values. It is argued that 4 factors are worthy enough to mention: the native language of a speaker (i.e., whether the speaker is a native speaker of English or not), the English proficiency level of a learner, the context in which the English voiceless stops appear (i.e., either in the word-initial position and in the position immediately following /s/), and POA of voiceless stops. For further details, see J. Y. Kim (2012).

both of the environments: in the #_V position, /p/ (mean = 76.87, SD = 21.48) < /t/ (mean = 82.71, SD = 52.11) < /k/ (mean = 94.05, SD = 86.56), and in the #s_V position, /p/ (mean = 15.17, SD = 16.53) < /t/ (mean = 20.18, SD = 9.01) < /k/ (mean = 24.11, SD = 10.03).

Table 3 demonstrates KHS' means and standard deviations of VOT values.

Table 3. The means and standard deviations of KHS' VOT values categorized according to the context and POA

Context	#_V			#s_V		
POA	Labial	Coronal	Dorsal	Labial	Coronal	Dorsal
mean (msec.)	59.93	57.27	75.31	18.87	22.56	32.74
SD	20.34	20.89	20.44	15.41	9.43	11.29

In KHS' production, the VOT values of the word-initial voiceless stops were higher than the VOT values of the voiceless stops immediately following /s/. In the context #s_V, the values varied in order according to POA: /p/ (mean = 18.87, SD = 15.41) < /t/ (mean = 22.56, SD = 9.43) < /k/ (mean = 32.74, SD = 11.29). Unlike NES' distribution in Table 2, it seems that in the #_V position, coronal stops were lower than labial stops in VOT values: /t/ (mean = 57.27, SD = 20.89) < /p/ (mean = 59.93, SD = 20.34) < /k/ (mean = 75.31, SD = 20.44)

KLS' mean VOT values and standard deviations are shown in Table 4. KLS' distribution of VOT values was similar to NES and KHS' even though the difference between the value of labial stops (i.e., 56.47 msec.) and that of coronal stops (i.e., 56.36 msec.) was not large enough. VOT values of the stops in the #s_V position showed a gradual increase as the values in the same environment did in Tables 2 and 3: /p/ (mean = 31.46, SD = 16.07) < /t/ (mean = 38.16, SD = 17.83) < /k/ (mean = 48.74, SD = 16.56).

Table 4. The means and standard deviations of KLS' VOT values categorized according to the context and POA

Context	#_V			#s_V		
POA	Labial	Coronal	Dorsal	Labial	Coronal	Dorsal
mean (msec.)	56.47	56.36	75.62	31.46	38.16	48.74
SD	21.72	21.12	18.76	16.07	17.83	16.56

Based on the raw values discussed in Tables 2, 3, and 4, bar graphs are provided below.

Figure 2. The mean VOT values of NES, KHS, and KLS depending on POA in #_V

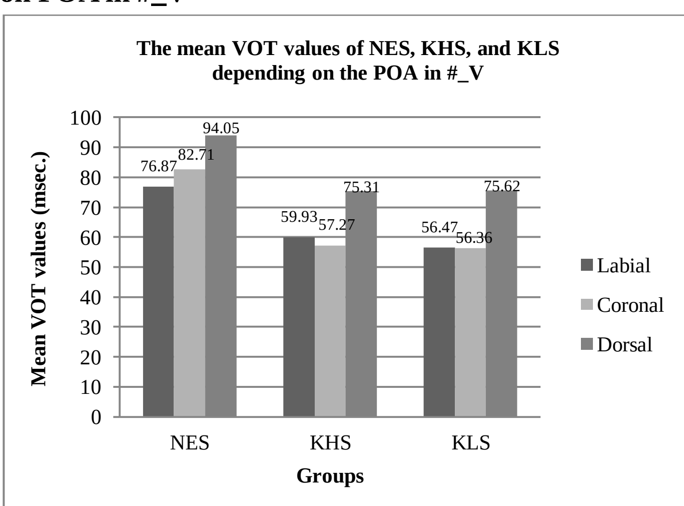
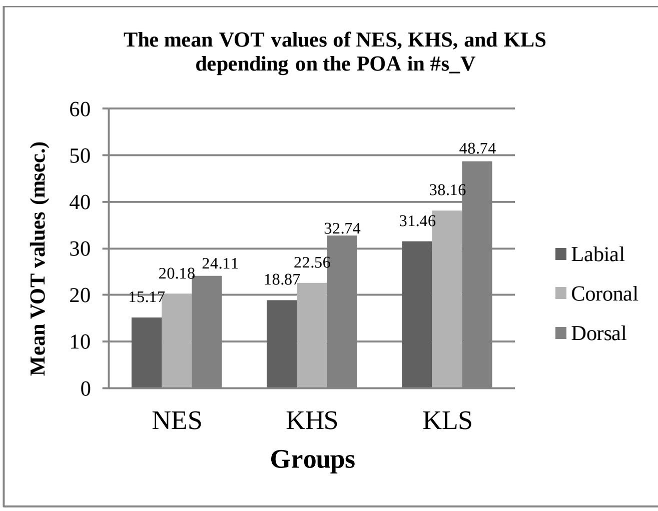


Figure 2 encapsulates the overall trends of NES, KHS, and KLS in terms of the mean VOT values only in the word-initial position. While the mean VOT values of labial stops and coronal stops did not have a big difference, the values of dorsal stops in all groups appeared to be significantly higher than the 2 other kinds of stops.

Figure 3 summarizes the distributions of the 3 groups' mean VOT values in the #s_V position.

Figure 3. The mean VOT values of NES, KHS, and KLS depending on POA in #s_V



The distribution of the VOT values' increasing from labial stops and coronal stops to dorsal stops was observed more clearly in the position following /s/. It is intriguing that the dorsal stop /k/ has the highest VOT values among the stops of 3 POAs.

In order to examine whether the tentative interpretation of the raw VOT values was convincing or not, a statistical analysis was also conducted by calculating linear models with the statistics package R[12]. The statistical results are show in Table 5.

[12] *The R Project for Statistical Computing* (http://www.r-project.org/).

Table 5. The linear model for the effects of the POA and its combination with the group and context on the VOT values

	df	Sum of Square	Mean square	*F* Value	*p* Value
POA	2	128230	64115	102.4949	< 2.2e-16 ***
POA: Group: Context	4	1198	300	0.4788	0.75133
Residuals	2682	1677701	626		

*** statistically significant ($p < .001$), **($p < .01$), *($p < .05$)

According to Table 5, the factor POA *per se* had a strong impact on the VOT values ($F(2,2682) = 102.4949$, $p < .001$), which means that the gradual increase in the VOT values of English voiceless stops in the order of labial, coronal, and dorsal stops is a statistically significant trend. The evidence for this conclusion can be found in H. Kim *et al*. (2004) and Yao (2009). They claimed that POA functions as a strong factor, which influences VOT.

The most interesting result of all was the highest VOT value of dorsal stops. According to Tsuchida *et al*. (2000), the length of aspiration in the velar stop /k/ is naturally longer than those in /p/ and /t/. T. Cho and Ladefoged's (1999) aerodynamic reason supports Tsuchida *et al*. (2000) and the results of this research. Because the oral cavity behind dorsal stops is the smallest, the air pressure in it must be the highest. Consequently, it takes the longest time until the vocal folds start to vibrate. Moreover, T. Cho *et al*.'s (2002) argued that there is no significant difference between labial stops and alveolar stops, which is substantially in agreement with the findings of the present study.

As for the research questions, it was found out in this paper that POA has a highly significant effect on the VOT values of English voiceless stops in both of the contexts (i.e., #_V and #s_V) regardless of a speaker's native language. The VOT values of both native speakers of

English and Korean were in a gradual increase, having dorsal stops with the highest VOT values.

5. Conclusion

The purpose of the present study was to investigate whether POA among various factors has a significant effect on aspiration in English voiceless stops, which is represented as VOT values. In the production experiment, native speakers of English and Korean were asked to read English real words containing English voiceless stops in the word-initial position and those in the position following /s/. The statistical analysis indicated that POA do have an impact on VOT values and supported the distribution of VOT values with a gradual increase according to POA: from labial and coronal stops to dorsal stops. Although the first 2 kinds of stops did not have a large difference in VOT values from each other, it turned out that dorsal stops were always in the longest in VOT. This is due to the aerodynamic reason, which is related to the structure of the oral cavity. In sum, it was discovered in this study that both native speakers of English and Korean showed the significant effect of POA on VOT values. It seems that the phenomenon —the different VOT values based on the different POAs—is language-universal because no difference was observed between the 2 groups: NES and NKS.

References

Boersma, P. and D. Weenink. (2011). Praat: doing phonetics by computer (Version 5.2.44) [Computer program].

Chang, C. B. (2007). *The status of voicing and aspiration as cues to Korean laryngeal contrast*. Paper presented at the 43rd Annual

Meeting of the Chicago Linguistics Society. Chicago, IL: University of Chicago, May 5.

Cho, B. E. (2004). Issues concerning Korean learners of English: English education in Korea and some common difficulties of Korean students. *The East Asian learner, 1(2).* Retrieved 19 March 2011 from: http://www.brookes.ac.uk/schools/education/eal/eal-1-2/vol1-no2 koreanlearnersofenglish.pdf

Cho, T. and P. Ladefoged. (1999). Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics, 27,* 207–229.

Cho, T., S. A. Jun, and P. Ladefoged. (2002). Acoustic and aerodynamic correlates of Korean stops and fricative. *Journal of Phonetics, 30,* 193-228.

Clark, J. and C. Yallop. (1995). *An introduction to phonetics and phonology*. Oxford: Blackwell.

Davidsen-Nielsen, N. (1974). Syllabification in English words with medial sp, st, sk. *Journal of Phonetics, 2,* 15–45.

Docherty, G. J. (1992). *The timing of voicing in British English obstruents*. Berlin: Walter de Gruyter.

Gentleman, R. and R. Ihaka. (2011). The R Project for Statistical Computing (Version 2.13.2) [Computer program].

Gussenhoven, C. and H. Jacobs. (2005). *Understanding phonology*. London: Hodder Arnold.

Iverson, G. and J. Salmons. (1995). Aspiration and laryngeal representation in Germanic. *Phonology, 12,* 369-396.

Keating, P. A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language, 60(2),* 286-319.

Kim, C. W. (1970). A theory of aspiration. *Phonetica, 21,* 107-116.

Kim, H., J. Cole, H. Choi, and M. Hasegawa-Johnson. (2004). The effect of accent on acoustic cues to stop voicing and place of articulation in radio news speech. In *Proceedings of Speech Prosody* (pp. 29-32). Nara: Japan.

Kim, J. Y. (2012). *Perception and production of aspiration in English voiceless stops: an experiment-based study of native speakers of English and Korean*. MA thesis, Seoul National University, Seoul, Korea, February 2012.

Kim, M. R., P. S. Beddor, and J. Horrocks. (2002). The contribution of consonantal and vocalic information to the perception of Korean initial stops. *Journal of Phonetics, 30,* 77-100.

Ladefoged, P. (2006). *A course in phonetic*s. Boston: Wadsworth.
Lisker, L. and A. S. Abramson. (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word, 20*, 384-422.
Morton, K. and M. Tatham. (1980). Devoicing, aspiration, and nasality: cases of universal misunderstanding? *Occasional Papers, Department of Language and Linguistics, University of Essex, 23*, 90-103.
Silverman, D. (1998). Alveolar stops in American English, and the nature of allophony. *NELS, 28*, 425–435.
Tsuchida, A., A. C. Cohn, and M. Kumada. (2000). Sonorant devoicing and the phonetic realization of [spread glottis] in English. *Working Papers of the Cornell Phonetics Laboratory, 13*, 167-181.
Vaux, B. (2002). *Aspiration in English*. Paper presented at the University of Wisconsin-Madison.
Vaux, B. and B. Samuels. (2005). Laryngeal markedness and aspiration. *Phonology, 22*, 395-436.
Yao, Y. (2009). Understanding VOT variation in spontaneous speech. In *Proceedings of 18th International Congress of Linguists (CIL XVIII)* (pp. 29–43). Seoul: Korea.

Ji Yea Kim
jiyeakim@snu.ac.kr