On the phonological complexity scale in Korean diphthongs

There are seven monophthongs and ten diphthongs in Modern Korean as shown in (1), ignoring the parenthesized somewhat old pronunciations. Especially, as shown in (2), the <ɨy> (orthographically) represents another type of on-glide, i.e., [ɰi] phonetically, rather than a complex vowel sequence [ɨi] or an off-glide [ɨj], but it has often been described as the only off-glides resisting the massive change favoring on-glides in Korean (Kim 1968, Ahn & Iverson 2007, etc.).

Numerous studies have been made on the synchronic/diachronic formation of the *y*- and *w*-onglides from the earlier Middle Korean off-glide system (K.-M. Lee 1972, Ahn & Iverson 2007). Little, however, is known about whether there is a phonological complexity scale in Korean diphthongs. This study aims to provide functional Magnetic Resonance Imaging (fMRI) and external evidence for the existence of such a scale.

Previous fMRI studies on the processing of phonological complexity modulated complexity simply by varying syllable aggregation (Shuster & Lemieux, 2005; Soros et al., 2006); i.e., complexity was based on number of syllables rather than their internal organization. Thus, we design an experiment to distinguish differences in complexity based on phonological organization. Given that production of phonologically complex syllables requires more planning and more motor execution, our hypothesis is that producing phonologically complex syllables will generate activation in Broca’s area, a region known to be responsible for speech production planning, as well as motor-related areas to a greater extent than phonologically simple syllables. This is in line with the results of a recent fMRI study (Park 2008), which are consistent with Ahn and Iverson’s (2004) minimalized representation hypothesis about Korean’s three-way contrasts of lax vs. tense vs. aspirated plosives. According to the model, the aspirated series is marked by the privative feature [spread glottis], meaning that it is more phonologically complex than the other two. In support of the model, Park revealed that the perception of Korean word-initial plosives with more complex phonological representation, i.e., aspirated stops, elicited more activation in Broca’s area or the left Inferior Frontal Gyrus (IFG) than that of their tense and lax counterparts with simpler phonological configuration, as illustrated in (3).

To test the hypothesis of this study, we first conduct a ratings study outside the scanner with a separate subject group in which subjects rated linguistic complexity (w.r.t. ease of articulation) of bisyllabic pronounceable nonwords (i.e., “pseudowords”) on a scale of 1 (simple) to 5 (complex) (n=30). Based on the ratings, we divide the bisyllabic pseudoword stimuli into three different complexity conditions as shown in (4) (COMPLEX, Mid-COMPLEX, and SIMPLE). These are: COMPLEX stimuli *ɰi*, Mid-COMPLEX stimuli including some diphthongs such as *ye, yo, yu*, as well as *wi*, *we*, and SIMPLE stimuli all the monophthongs such as *i, a, o*, etc.

Stimuli are carefully controlled on a variety of nuisance variables such as number of phonemes and syllables, imageability, etc. We use a “slow event-related” design with clustered image acquisition. Clustered acquisition is implemented with a TR of 10 *s*, and the data from all slices are acquired in 2000 ms of this time interval. The hemodynamic response function is known to peak approximately 5 *s* after the presentation of an auditory stimulus (Hulvershorn et al., 2005). Based on this observation, the stimuli are presented 5 *s* prior to the midpoint of the data acquisition. To assess response accuracy, the subjects’ vocal responses are recorded via the microphone channel of the Silent Scan Audio System (Avotec, Stuart, FL, USA) and stored on a

PC. For the measurement of response latency, the latency between the onset of the cue and the onset of speech is calculated. For the statistical analysis of functional imaging data, three different one-sample *t* tests for the SIMPLE, Mid-COMPLEX and COMPLEX conditions with random subject effects are employed to locate activations unique to each complexity condition and common to all subjects. A voxel-level threshold with an uncorrected $p < 0.001$ ($T = 4.3$, $df = 9$) with cluster size>50 is used to detect statistical difference at the voxel-level. Clusters with a corrected threshold $p < 0.05$ at the cluster-level are considered to be significantly different in consideration of multiple comparison.

Ten male speakers of native Korean participate in this fMRI experiment, in order to control for sex differences in cerebral laterality of language processing (Clements et al., 2006) and in brain activation during cognitive tasks (Bell et al., 2006) that have been found in recent neuroimaging studies. The task is to read aloud pseudowords presented every 10 seconds following visual instructions (e.g., "say *tɨɰi*"). Visual instructions are delivered through a screen above a participant's head at a constant onset-to-onset interstimulus interval of 10 s with Eprime 1.1 (Psychology Software Tools, Pittsburgh, PA, USA). Participants are requested to produce the required overt response immediately after the end of the instruction. We use overt production paradigm because covert speech paradigms do not provide normal auditory and kinesthetic feedback mechanisms important in speech production (Jones et al., 2000).

The experiment consists of five sessions lasting approximately 9 minutes each. Each session is made up of 6 separate blocks of speech, 2 blocks of Broca's area localizer (60 s each), and 3 blocks of baseline (30 s). During the baseline, no visual instructions are given, and no responses are taken. All instructions and responses are made within the silent interval between the fMR image acquisitions. In the Broca's area localizing block, participants perform a verb generation task (e.g., a participant should complete an aurally-presented sentence *Na.neun hak.kyo-e* (나는 학교에 ____ "I __ to school")) according to the instructions on the screen, over the course of twelve 30-second blocks (6 per condition). The instructions (*Mun.jang.eul wan.seong.ha.se.yo* (문장을 완성하세요) "complete the sentence") will be presented every 10 seconds. Participants are trained before scanning using feedback to familiarize them with the procedure and the nature of the pseudowords. The overall experimental procedure is illustrated in (5).

According to behavioral results, the response latency of COMPLEX syllables is significantly higher than that of Mid-COMPLEX, which in turn shows significantly longer response latency than SIMPLE syllables. As for the fMRI results, both brain activities unique to the COMPLEX relative to Mid-COMPLEX and SIMPLE conditions and activities unique to the Mid-COMPLEX relative to SIMPLE condition are observed in Broca's area and sensory-motor areas. Moreover, we attempt to propose an internal complexity scale within the Mid-COMPLEX structures. To this end, we investigate the various compositionally distinctive structures, i.e., *y*- vs. *w*- onglide distinction, *y/w* + vowels with different height, *y/w* + vowels with different backness, etc. These experimental results lend support to our hypothesis, providing neural support for the existence of phonological complexity scale in Korean diphthongs.

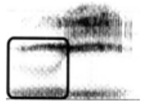
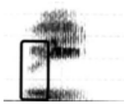
(1)

i	ɨ	u
e	ə	o
(ɛ)	a	

-	wi	-	-	yu	-
ye	we	yə	wə	yo	-
(yɛ)	(wɛ)	ya	wa		

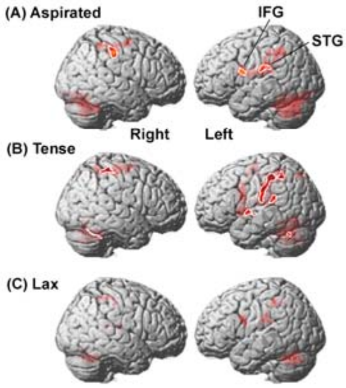
\+ ɰi <ɨy>

(2)


[ɨi] vs.

[ɰi]

(Shin 2000: 191)

(3)

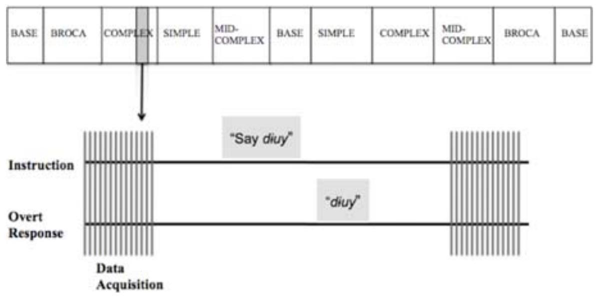


Surface rendering representation of the activation areas for the Aspirated, Tense, and Lax conditions. The activation maps of the Aspirated, Tense, and Lax conditions are displayed in (A)-(C), respectively, for $p < 0.005$ with cluster size > 100. Clusters with a threshold $p < 0.001$ and cluster size > 50 were displayed with white solid lines (Park 2008).

(4)

COMPLEX	Mid-COMPLEX	SIMPLE
ɰi	(A) wi, we, wə, wa (B) ye, yə, ya, yo, yu	i, a, u, ….

(5)



Ahn SC and Iverson GK. 2004. Dimensions in Korean laryngeal phonology. J. of East Asian Linguistics 13, 345-379.
Ahn SC and Iverson GK. 2007. Structural imbalances in the emergence of the Korean vowel system. In J. Salmons and S. Dubenion-Smith (eds.) *Historical Linguistics 2005*. Pp. 275-293. John Benjamins.
Bell EC, Willson MC, Wilman AH, Dave S, and Silverstone PH. 2006. Males and females differ in brain activation during cognitive tasks. *Neuroimage* 30, 529-538.
Clements AM, Rimrodt SL, Abel JR, Blankner JG, Mostofsky SH, Pekar JJ, Denckla MB, and Cutting LE. *2006*. Sex differences in cerebral laterality of language and visuospatial processing. *Brain and Language* 98, 150-158.
Hulvershorn J, Bloy L, Gualtieri EE, Redmann CP, Leigh JS, and Elliott MA. 2005. Temporal resolving power of spin echo and gradient echo fMRI at 3T with apparent diffusion coefficient compartmentalization. *Human Brain Mapping* 25, 247-258.
Jones JA and Munhall KG. 2000. Perceptual calibration of F0 production. *JASA* 108, 1246-1251.
Kim CW. 1978. The vowel system of Korean *Language* 44, 516-527.
Lee KM. 1972. *Guk.eo.eum.sa Yeon.gu* (History of Korean Phonology). Seoul: Tower Press.
Park HI. 2008. Brain activation patterns in perception of Korean laryngeal contrasts. PhD. Diss. U Wisconsin-Milwaukee.
Park HI, Lee EK, and Park HJ. 2009. Neural Correlates in the Processing of Phonological Complexity: fMRI study. Ms. Yonsei U
Shuster LI and Lemieux SK. 2005. An fMRI investigation of covertly and overtly produced mono- and multisyllabic words. *Brain and Language* 93, 20-31.
Shin JY. 2000. *Mal.so.li-ui I.hae* (Understanding Speech). Seoul: Hankook Publishing co.
Soros P, Sokoloff LG, Bose A, McIntosh AR, Graham SJ, and Stuss DT. 2006. Clustered functional MRI of overt speech production. *Neuroimage* 32, 376-387.