Vowel Formant Trajectory Patterns for Shared Vowels of American English and Korean

Chung, Hyunju¹) · Kong, Eun Jong,²) · Weismer, Gary³)

ABSTRACT

The purpose of this study was to explore the cross-linguistic difference in the spectral movement pattern of American English and Korean vowels. Eight American vowels $/\alpha/$, /e/, /i/, /i/, /i/, $/\alpha/$, $/\alpha/$, $/\alpha/$, $/\alpha/$, /e/, /i/, /i/, $/\alpha/$, $/\alpha/$, $/\alpha/$, $/\alpha/$, $/\alpha/$, /e/, /i/, $/\alpha/$, $/\alpha/$, /

Keywords: vowels, Korean, American English, spectral movement

1. Introduction

The phonological categories of vowels have typically been differentiated from each other by their formant frequency values. One of the most common ways to describe vowels is to measure the first two formant frequencies (F1 and F2) at the vowel midpoint. This measurement approach is one of the most widely used ones in studies of vowel acoustics (e.g., Hillenbrand *et al.*, 1995; Peterson & Barney, 1952). The assumption behind this approach is that the vowel midpoint is the place that has minimal shift in formant values, and is believed to be a "target" point that speakers try to reach when producing vowels (simple target theory, Strange, 1989) which thus best represents the acoustic characteristics of vowels. The result of the classic study of Peterson & Barney (1952), however, showed a pattern that raised

Received: November 3, 2010 Revision Received: December 8, 2010 Accepted: December 10, 2010 questions about this previously held assumption. When vowels of American English were plotted in the F1-F2 acoustic space, substantial overlap in F1-F2 values was observed. More interestingly, despite this extensive overlap in F1-F2 values, however, listeners could successfully identify vowels 94% of the time. It has been suggested, therefore, that although target formant frequency values define vowel characteristics well and play an important role in vowel identification, they are not the only cues for vowel identification and listeners might attend to other cues. Spectral change patterns in vowels, which characterize vowels with formant values not just at a single time point, but across multiple time points, therefore, have received attention as another potentially important set of cues for characterizing vowels and their identification (Nearey & Assmann, 1986; Strange, 1989; Hillenbrand, Getty, Clark, & Wheeler, 1995).

From the 1980s, studies showed that vowel formant trajectories or vowel dynamic information is important in vowel perception. Nearey and Assmann (1986) and Strange (1989) both showed that higher error rates were found in vowel identification tasks when vowel dynamic information was absent from the signal as compared to when this information was given. Nearey and Assmann (1986) showed that identification error rates increased

¹⁾ University of Wisconsin-Madison, hchung23@wisc.edu, corresponding author

²⁾ University of Wisconsin-Madison, ekong@wisc.edu

³⁾ University of Wisconsin-Madison, gweismer@wisc.edu

greatly when listeners were given stimuli (a sequence of nucleus-10ms silence- nucleus) without an offglide in the vowel. In another study, Strange (1989) showed that listeners could identify vowels 85% of the time even without the presence of the vowel nucleus in the signal. Both of these results suggest that information in the vowel onglide and offglide are important cues for vowel identification. Along with evidence for the role of vowel dynamic information in vowel perception, several studies showed acoustic evidence of vowel inherent changes. Hillenbrand et al., (1995) showed patterns of vowel spectral change in American English vowels and Nearey and Assmann (1986) showed the same patterns in Canadian English. Both studies showed that /e/ and /o/ had the greatest spectral change while /a/, /i/, and /u/ remained relatively stable. Also, in both studies the directions of spectral movement for /e/ and /o/ tended towards the spectral "target" for /i/ and /u/, respectively. Likewise, recent studies on vowels have shown the importance of vowel dynamic information in characterizing vowels both in terms of perception and production.

To date, no studies have characterized the spectral movement pattern of Korean vowels although there are studies of static vowel characteristics with F1 and F2 values measured at the vowel midpoint (Yang, 1992; Yang, 1996). Yang (2010) studied vowel spectral movement patterns of vowels produced by Korean speakers. However, this study focused on examining the difference between English and Korean speakers' production of English tense versus lax vowels (where he found evidence of cross-linguistic differences in the F1 and F2 trajectory patterns between vowels of two languages), rather than on formant trajectory patterns of Korean vowels produced by native speakers of Korean. In this study, we examine vowel spectral movement patterns of five Korean vowels /a/, /e/, /i/, /o/, and /u/ produced by native adult speakers of Korean, and compare these patterns with those of both tense and lax versions of English vowels produced by native adult speakers of American English.

2. Methods

The data used in this study come from a larger study, the *Paidologos* project (Edwards & Beckman, 2008; Edwards & Beckman, 2009), which examined the effects of phoneme and phoneme sequence frequency on phonological acquisition of word-initial lingual stops across languages.

2.1 Participants

The participants included 6 adult speakers (3 males and 3 females) of American English and 10 adult speakers (5 males and 5 females) of Korean. All of their age ranged from 18 to 30 years. All participants were monolingual native speakers of American English and Korean, respectively. Dialectal influences on vowel production were controlled by recruiting subjects from only one region of each country; recruitment was done by a native speaker of the same dialect region. English-speaking subjects were recruited in Columbus, Ohio and Korean-speaking subjects from Seoul, Korea. All participants passed a hearing screening and reported no history of speech, language, or hearing problems.

2.2 Procedures

					5		
	English	ı			Korea	n	
	/s/	/∫/		/s/	/∫/	/s'/	/ʃ'/
/α/	15	17	/a/	17	-	17	-
/e/	12	-	/e/	15	-	16	-
/ɛ/	6	18	/i/	-	17	-	17
/i/	-	6	/0/	16	-	9	-
/I/	15	11	/u/	17	-	6	-
/0/	18	17					
/u/	-	5					
/ʊ/	16	12					

Table 1. Word initial consonant environment and number of vowels used for the analysis

Speech samples were collected using a word repetition task. The stimuli were familiar real words. Participants were asked to repeat each word after viewing a color picture and hearing an auditory recording of word to be produced. Productions were digitally recorded. The stimuli were words that began with obstruent-vowel sequences. Only tokens beginning with fricatives were used in this study to remove any possible effect of different consonantal environments on vowel formant patterns. Therefore, the word initial consonants were alveolar /s/ and post-alveolar $/\int/$ for English and denti-alveolar /s/ and /s'/ for Korean. We also included post-alveolar fricative $/\int/$ and $/\int'/$, which allophonically appear before /i/ vowels in Korean. Target vowels were /a/, /e/, (ϵ) , (i), (I), (o), (u), and (v) for English and (a), (e), (i), (o) and (u)for Korean. In Korean, the two fricatives /s/ and /s'/ share the same place of articulation, but they are different in terms of larvngeal contrasts (Cheon & Anderson, 2008). The summary of word-initial consonants preceding the target vowels for each language is provided in <Table 1>.

2.3 Acoustic analysis

The first two formant frequency values (F1 and F2) were extracted from the vowel onset to the vowel offset, with a step size of 6ms, using the LPC solution in *Praat* (version 5.0.29) (Boersma & Weenink, 2006). The size of the analysis window was 25ms and the dynamic range was set at 30dB. Vowel onset was defined as the time at which the first clear glottal pulse was observed in both the waveform and the spectrographic display and vowel offset was defined as the final glottal pulse extending at least through F1 and F2.

2.4 Statistical analysis

The temporal scale of each vowel was transformed to have seven different proportional time points by binning formant values extracted over time. The bin sizes differed from token to token due to varying durations and a fixed size of analysis window. We treated the second and fifth time point among seven discrete values as the onset and the midpoints of the vowel and used F1 and F2 values between those times to depict trajectory plots and the formant movement analysis. We ignored the formant values at the first time point to avoid potential mistrackings caused by the preceding consonant and ignored the formant values at the latter half of the duration to control the effect of coda on vowels in the analysis. Therefore, the spectral movement patterns only from vowel 'onset' to the 'midpoint' were the focus of the analysis.

The formant trajectories were explained across time using a mixed effects regression (Singer & Willett, 2003; Mirman, Aslin & Magunuson, 2007; Barr, 2008). The dependent variable was formant frequencies and the independent variable was time. The regression model used up to three orders of an orthogonal polynomial to describe the trajectory shapes in detail. Each order of polynomial terms explains different components of the shape in an independent manner. The first order of orthogonal polynomial describes the linear slope of shape. A negative coefficient indicates a decreasing linear trend over time, and a positive coefficient an increasing trend. The second and the third orders describe the symmetric and asymmetric curve component of the trajectory, respectively. We used all three orders to capture variations of the formant frequency trajectories, but we will focus on discussing the linear slopes. Two separate regression models for each vowel type were made for F1 and F2 trajectories.

3. Results

3.1 Vowel spectral movement

<Figure 1> shows the spectral movement patterns of eight English vowels, / α /, /e/, / ϵ /, /i/, / π /, / α /, / α /, and / σ /, in the vowel acoustic spaces of F2 as a function of F1. The arrow indicates the direction of movement connecting the formant pattern measured at the onset to the midpoint of vowels produced by adult speakers of American English. The top panels are for / α /, /i/, and /u/, the middle panels are for / ϵ /, / π /, and / σ /, and bottom panels are for /e/ and /o/.



Figure 1. Spectral movement patterns of eight American English vowels ("A"=/α/, "e"=/e/, "E"=/ε/, "i"=/i/, "T"=/I/, "O"=/ο/, "u"=/u/, and "U"=/σ/). Dotted lines represent trajectories of /α/, /e/, /ε/, /i/, /I/, and /ο/, and solid lines represent those of /u/ and /σ/. The averaged formant trajectories of each vowel were indicated by thick letters and lines.

The results of the trajectory movement patterns for American English vowels are in line with the findings of previous studies by Hillenbrand *et al.*, (1995) and Nearey and Assmann (1986). In both male and female productions, a good deal of spectral movement was observed in nearly all vowels. As described in a vowel space, the directions of formant trajectories were consistent

across vowel types, where the formant frequencies departed from the relative center of the vowel space and ended at the relative periphery of the vowel space. The three corner vowels /i/, / α /, and /u/ moved toward the periphery of the vowel acoustic space, while the non-corner vowel /e/ moved toward /i/ over time. The direction of spectral movement patterns of the vowel /o/ was also very consistent, moving toward the periphery of the vowel space. <Table 2> shows the average F1 and F2 values of onset and midpoint of eight English and five Korean vowels.

Table 2. Average F1 and F2 values (in Hz) of American English and Korean vowel onsets and midpoints

		F1			F2			
Engli	English							
	ma	ıle	female		male		female	
	onset	mid	onset	mid	onset	mid	onset	mid
/a/	621	734	762	859	1385	1278	1636	1595
/e/	546	461	609	471	1798	2102	2278	2469
/ɛ/	575	625	675	700	1728	1653	2023	1959
/i/	374	333	393	374	2156	2177	2635	2770
/1/	363	440	446	504	1997	1791	2367	2292
/0/	537	557	676	630	1426	1239	1663	1249
/u/	408	417	475	511	1863	1398	2195	1353
/ʊ/	391	397	467	462	1789	1600	2032	1706
Korean								
	ma	ıle	fem	ale	male		female	
	onset	mid	onset	mid	onset	mid	onset	mid
/a/	713	743	883	912	1600	1573	1764	1645
/e/	513	537	599	619	2000	2119	2288	2373
/i/	352	340	442	401	2245	2140	2485	2408
/0/	437	447	490	526	1109	1026	1147	1132
/u/	430	499	444	532	1757	1495	1979	1664

<Figure 2> shows the spectral movement pattern of five Korean vowels, /a/, /e/, /i/, /o/, and /u/. The top panels are for /a/, /i/, and /u/ and the bottom panels are for /e/ and /o/. Contrary to the patterns found in English vowels, there was no noticeable difference of formant frequencies between vowel onset and midpoint in Korean vowels. Furthermore, there was no consistent direction of movement for the three corner vowels /i/, /a/, and /u/ or in the non-corner vowel /o/. Relatively consistent direction of movements was found for /e/. For this vowel, the trajectories moved from the centralized position toward the periphery of the vowel acoustic space, although the pattern was less straightforward than that observed in English vowels. In general, these results suggest that for Korean vowels, a vowel space formed by F1 and F2 values measured at the vowel onset will

show no significant difference from that formed by F1 and F2 values measured at the vowel midpoint.



Figure 2. Spectral movement patterns of five Korean vowels, /a/, /e/, /i/, /o/, and /u/. Dotted lines represent trajectories of each of /a/, /i/, /e/, and /o/, and solid lines represent those of /u/. The averaged formant trajectories of each vowel were indicated by thick letters and lines.

Overall, a cross-linguistic difference was observed in the vowel spectral movement patterns of American English versus Korean vowels produced by adult speakers of each language. While English vowels had consistent directions of formant frequency movement from the vowel onset to the midpoint, moving from the centralized position of the vowel space toward the periphery side with greater degree of movement, Korean vowels showed relatively less degree of movement with no consistent direction.

3.2 Mixed effects regression models

The changes of each formant frequency (F1 and F2) over time were modeled using mixed effects regression models. Sixteen individual models were made for F1 and F2 trajectories for each of eight American English vowels and ten models for F1 and F2 of five Korean vowels. <Figure 3> shows F1 trajectories (the *y*-axis) of vowels as a function of time (the *x*-axis) predicted by regression models. The left panels of <Figure 3> are for male speakers and the right panels are for female speakers. The top panels are for American English vowels and the bottom panels are for Korean vowels.



Figure 3. F1 trajectories of vowels as a function of time as estimated from the mixed effects regression.

For F1 patterns of American English vowels, /I/, /u/, / σ / and /e/ showed a great degree of overlap in their static measured F1 values. Even the spectral movement pattern showed great overlap for /I/, /u/, and / σ /. The vowel /e/, however, showed a pattern of spectral change that was distinct from those of other high vowels, with a larger decrease in movement. The coefficient of the mixed effects model confirmed this observation. As shown in <Table 2>, the coefficients for both /i/ and /e/ were negative and the coefficient for /e/ was considerably larger than those of other vowels (except / α /), indicating relatively great magnitude of F1 movements. All other vowels had positive coefficients and had smaller absolute values of coefficients. This pattern was found in vowel productions of both males and females.

In contrast, Korean vowels did not display a clear increasing or decreasing trend in their F1 trajectories. That is, there was hardly any change in the movement patterns for all five vowels. Compared to English, static measured F1 values of Korean vowels were well-differentiated from each other, except /o/ and /u/. Even the spectral movement patterns of these two vowels showed no large change over time. These patterns were similar for both males and females. The coefficients showed that the degree of movement was greater for /u/ as compared to /o/ vowels for males, but vice versa for females. In general, the coefficient of the mixed effect regression model were much smaller for Korean vowels as compared to English vowels, indicating that the magnitude of change in F1 values from vowel onset to midpoint was smaller for Korean values relative to American English vowels.



Figure 4. F2 trajectories of vowels as a function of time as estimated from the mixed effects regression.

Table 3. Coefficients of the linear slope (the first order term) from the mixed effects regression models.

		English			Korean		
		male	female		male	female	
F1	/a/	80.49	76.74	/a/	13.84	6.50	
	/e/	-60.45	-106.97	/e/	20.72	5.86	
	/ɛ/	43.24	31.63	/i/	-2.92	6.32	
	/i/	-26.36	-17.45	/0/	7.99	27.90	
	/1/	44.66	29.40	/u/	22.91	-0.29	
	/0/	3.88	-26.74				
	/u/	5.21	30.50				
	/ʊ/	6.70	-5.11				
F2	/a/	-81.44	-55.74	/a/	3.52	-86.47	
	/e/	225.21	178.32	/e/	86.52	55.45	
	/ɛ/	-38.15	-51.37	/i/	-92.72	-4.58	
	/i/	34.80	95.71	/0/	-86.24	-26.11	
	/1/	-105.76	-11.76	/u/	-174.20	-165.39	
	/0/	-145.00	-308.70				
	/u/	-359.98	-645.92				
	/ʊ/	-137.12	-213.65				

<Figure 4> shows the predicted F2 trajectories (the y-axis) of vowels as a function of time (the x- axis). Similar to the findings in F1 trajectories, relatively greater changes in F2 values over time were observed in English vowels than in Korean vowels. In general, back vowels (/ α /, /o/, /u/, and / σ /) in American English had negative direction of movements. In particular, /u/ produced by English-speaking adults showed the greatest movement among all other vowels. The vowel /o/ also followed the same negative

direction of movement. For front vowels (/i/, /ɪ/, and /e/), there was less overlap in F2 values at vowel midpoint as compared to the patterns found in F1. The vowel /e/ showed the greatest degree of positive direction of movement.

For Korean vowels, front vowels (/i/ and /e/) were clearly distinguished from non-front vowels (/a/, /u/, and /o/) in terms of static measured F2 values. However, the spectral movement of F2 over time was relatively stable across vowels. In other words, static measured F2 values alone were enough to differentiate the different vowel categories in Korean. While /u/ and /o/ showed overlapping static measured F1 values, static measured F2 values of these two vowels were well-differentiated from each other. The specific direction and the degree of spectral movement of each vowel indicated by coefficients of linear slope estimated from the mixed effects regression models can be found in <Table 3>.

4. Discussion

This study explored cross-linguistic differences in spectral movement patterns of five Korean and eight American English vowels in word-initial CV (fricative-vowel) contexts by measuring F1 and F2 values from the vowel onset to the midpoint. The findings were that American English vowels had greater magnitude of spectral movements and had consistent direction of movements across vowels in the same phonemic categories. In contrast, Korean vowels were well-differentiated by static F1 or F2 values, and did not show consistent patterns of spectral change. The directions of spectral movements were almost random among vowels of the same phonemic categories, especially for the three corner vowels /i/, /a/, and /u/. For /e/ and /o/, however, there was some consistency in terms of the direction of movement, each moving toward the periphery of the vowel space.

This cross-linguistic difference in vowel spectral movement patterns has important implications for theories of vowel perception and speech sound development in young children. First, the results of this study showed that American English vowel pairs that are not distinguishable by static measured F1 and F2 values showed distinctive spectral change patterns. As suggested by recent findings on the importance of vowel dynamic information for vowel identification (i.e., Strange, 1989), this study provides additional evidence that spectral movement patterns play an important role for the identification of vowels, and especially so when the static F1 and F2 values do not differentiate the vowels. In other words, American English vowels are better characterized by a combination of spectral movement patterns of F1 and F2 and static measures of F1 and F2 values. On the other hand, the fact that no such movement patterns were found in Korean vowels suggests that spectral change information will not improve the separability of vowel categories in Korean. Rather, it is mainly static measures of F1 and F2 values that are sufficient for vowel identification in Korean.

Secondly, the fact that vowels of American English are defined not only by static measures of F1 and F2, but also by trajectories implies that during the vowel acquisition process, young children need to learn not only how to reach the target F1 and F2 values in the acoustic space, but also the direction and magnitude of the gestures. In other words, the fine phonetic detail that produces language-specific characteristics of vowel articulation includes gesture plus target for English-speaking children, but perhaps only target for Korean-speaking children. An example of this is found in the comparison of English and Korean /u/. The relatively flat F2 trajectory in Korean suggests that Korean speakers are close to the low F2 target typically associated with /u/ in any language (i.e., Hillenbrand et al., 1995); on the other hand, American speakers start at a higher F2 frequency and show a decreasing pattern throughout the trajectory. Another way to say this is that the English vowel /u/ seems less coarticulated with the preceding consonant as compared to Korean /u/. If American children show this "less coarticulated" version of /u/, it suggests they are very sensitive to language-specific, phonetic details of their vowel system. Taken together, this result suggests that vowel acquisition of American English could be more difficult, and thus mastered later than Korean vowels. Whether or not this claim is true for /u/ specifically or applies in general to the acquisition of English versus Korean vowels is a matter for further research.

Some limitations of this study should be noted. First, the durational difference between Korean and English vowels could influence the magnitude of spectral movements in vowels. In general, Korean vowels are shorter in duration than vowels produced by English-speaking adults (Yang, 1996). This could explain why American English vowels have a greater degree of movement as compared to Korean vowels. Secondly, although this study tried to remove the effect of consonant environment on vowel formant values by confining consonant environment to fricatives, the trajectory patterns shown in the results could be attributed to coarticulation between word-initial fricatives and vowels. That is, the cross-linguistic differences in the spectral movement patterns of American English vowels and Korean vowels could be due to the language-specific differences in the degree of coarticulation between fricatives and vowels.

In order to understand the precise nature of vowel inherent changes of both American English and Korean vowels, additional studies examining formant trajectory patterns of vowels following consonants of different place or manner of articulation would be necessary. Moreover, studies examining the spectral movement patterns of vowels produced by children of native American English and Korean are needed to understand whether differences in the magnitude and direction of spectral movement patterns affect the process of vowel acquisition.

Acknowledgments

We thank Dr. Jan Edwards for her comments on the previous version of this paper.

References

- Barr, D.J. (2007). "Analyzing 'visual world' eyetracking data using multilevel logistic regression", *Journal of Memory and Language*, Vol. 59, pp. 457-474
- Boersma, P. & Weenink, D. (2008). Praat: doing phonetics by computer [Computer program]. Retrieved Aug, 2008, from http://www.praat.org/.
- Cheon, S. Y., & Anderson, V. B. (2008). "Acoustic and perceptual similarities between English and Korean sibilants: implications for second language acquisition", *Korean Linguistics*, Vol. 14, pp. 41-64.
- Edwards, J. & Beckman, M. E. (2009). "Methodology questions in studying consonant acquisition", *Clinical Linguistics and Phonetics*, Vol. 22, pp. 937-956.
- Edwards, J. & Beckman, M. E. (2008). "Some cross-linguistic evidence for modulation of implicational universals by language-specific frequency effects in the acquisition of consonant phonemes", *Language Learning and Development*, Vol. 4, pp. 122-156.
- Hillenbrand, J. M., Getty, L. A., Clark, M. J. & Wheeler, K. (1995). "Acoustic characteristics of American English vowels", *Journal* of the Acoustical Society of America. Vol. 97, pp. 3099-3111.
- Mirman, D, Dixon, J. A. & Magnuson, J. S. (2008). "Statistical and computational models of the visual world paradigm: Growth curves and individual differences", *Journal of Memory and Language*, Vol. 59, pp. 475-494.

- Nearey, T. M. & Assmann, P. F. (1986). "Modeling the role of inherent spectral change in vowel identification", *Journal of the Acoustical Society of America*, Vol. 80, pp. 1297-1308.
- Peterson, G. & Barney, H. L. (1952). "Control methods used in a study of the vowels", *Journal of the Acoustical Society of America*, Vol. 24, pp. 175-184.
- Singer, J. D. & Willett, J. B. (2003). Applied longitudinal analysis: Modeling change and event occurrence. New York: Oxford University Press.
- Strange, W. (1989). "Dynamic specification of coarticulated vowels spoken in sentence context", *Journal of the Acoustical Society of America*, Vol. 85, pp. 2135-2153.
- Yang, B. (1992). "An acoustical study of Korean monophthongs produced by male and female speakers", *Journal of the Acoustical Society of America*, Vol. 91, pp. 2280-2283.
- Yang, B. (1996). "A comparative study of American English and Korean vowels produced by male and female speakers," *Journal* of *Phonetics*, Vol. 24, pp. 245-261.
- Yang, B. (2010). "Formant trajectories of English high tense and lax vowel produced by Korean and American speakers", *Korean Journal of Linguistics*, Vol. 35, pp. 407-421.

Chung, Hyunju: Corresponding author Dept. of Communicative Disorders, Waisman Center, University of Wisconsin-Madison
1500 Highland Ave.
Madison, WI 53705, U.S.A. Tel: +1-608-215-6975
Email: hchung23@wisc.edu
Research Interests: Vowel acoustics, Phonological acquisition, Speech production and perception, Speech motor control
2006-present: Doctoral Candidate and Project Assistant

• Kong, Eun Jong

Waisman Center,
University of Wisconsin-Madison
1500 Highland Ave.
Madison, WI 53705, U.S.A.
Tel: +1-614-493-7370
Email: ekong@wisc.edu
Research Interest: Acoustic phonetics, cross-linguistic phonological development
2009~present: Post-doctoral research associate
Ph.D., Dept. of Linguistics, The Ohio State University, 2009

• Weismer, Gary

Dept. of Communicative Disorders, University of Wisconsin-Madison Goodnight Hall 1975 Willow Dr. Madison, WI 53706, U.S.A. Tel: +1-608-262-6472 Email: gweismer@wisc.edu Research Interest: normal speech production, speech motor disorder Professor, University of Wisconsin-Madison

Ph.D., University of Wisconsin-Madison