Linguistic Research 30(2), 293-312

Dynamic spectral patterns of American English front monophthong vowels produced by Korean-English bilingual speakers and Korean late learners of English*

Eunjin Oh
(Ewha Womans University)

Oh, Eunjin. 2013. Dynamic spectral patterns of American English front monophthong vowels produced by Korean-English bilingual speakers and Korean late learners of English. *Linguistic Research* 30(2), 293-312. This study investigated how Korean-English bilingual speakers and Korean late learners of English realize dynamic patterns of front monophthong vowels in American English. Twenty subjects read monosyllabic [CVC] words containing the vowels [i, ɪ, ɛ, æ] in a carrier sentence. F1, F2, and F3 frequency values were measured at 11 points for each speaker by dividing their vowel duration into ten equal segments. The results indicated that, compared to Korean late learners of English (Group NK), the Korean-English bilingual speakers (Group KE) exhibited F1, F2, and F3 transition shapes and overall formant values of the English front vowels that are more similar to those of native speakers of American English (Group NE). It was also found that the movement directions of the vowels [i] and [æ] varied widely across the three groups, those of the vowel [ɪ] were similar across the three groups, and those of the vowel [ɛ] were similar for Groups NE and KE, but not for Group NK. For the vowel [ɪ], the ranges of the formant movements overlapped for Groups NE and KE, but that of Group NK was quite distinct. For Groups NE and KE, the trajectories of the vowels [ɛ] and [æ] were not as distinct as the trajectories of [i] and [ɪ], though they did not overlap. The vowels [ɛ] and [æ] completely overlapped for Group NK. In general, Group KE exhibited more native-like spectral changes of the American English vowels than Group NK, indicating that there is an effect of language experience in learning the dynamic patterns of the vowels. **(Ewha Womans University)**

Keywords Dynamic spectral patterns, front monophthong vowels, American English, formants, F1, F2, F3, native speakers of English, Korean-English bilingual speakers, Korean late learners of English

* I am grateful to anonymous reviewers for their helpful comments and suggestions.

1. Introduction

Traditional studies on vowel qualities assumed that inherent qualities of monophthong vowels are determined mainly by the steady-state frequencies of their first three formants (e.g. Peterson and Barney 1952, Lehiste and Peterson 1961, Lindblom 1963). However, subsequent studies have demonstrated that the whole dynamic trajectory contains important information, not only for diphthongs but also for monophthongs, and that time-varying spectral information provides cues for vowel identification (e.g. Strange, Jenkins, and Johnson 1983, Nearey and Assmann 1986, Benguerel and McFadden 1989, Huang 1992, Hillenbrand and Gayvert 1993, Zahorian and Jagharghi 1993, Harrington and Cassidy 1994, Hillenbrand, Getty, Clark, and Wheeler 1995, Watson and Harrington 1999).

Broadly speaking, there are three arguments supporting the view that the dynamic pattern constitutes an important property of monophthong vowels. The first is that the dynamic information is necessary to improve vowel perception. For example, Strange, Jenkins, and Johnson (1983) demonstrated that the initial and final formant transitions in CVC syllables provide crucial cues for vowel identification in English. They tested three types of stimuli: (1) unmodified CVC stimuli as a control; (2) silent-center (SC) stimuli which replaced the center parts of vowels with silence, retaining only the initial and final transitions; and (3) variable-center (VC) stimuli, which discarded vowel transitions and retained only the vowel centers. They found that SC stimuli were perceived as accurately as unmodified ones, but that VC stimuli were identified poorly compared to both SC and unmodified stimuli (see also Benguerel and McFadden 1989). Likewise, Nearey and Assmann (1986) found that when two 30 ms spectral slices extracted at the 24% and 64% of vowel duration were presented in their natural order, vowel qualities were identified as accurately as when the full vowels were presented. If the first slice was repeated or the two slices were reversed in order, the identification score was reduced significantly. Similarly, Huang (1992) reported that when three spectral slices (25%, midpoint, 75%) were taken from a vowel, identification performance was significantly better than when only a single slice was taken from the vowel midpoint. In a similar vein, Hillenbrand and Gayvert (1993) explored how well listeners identify vowels based solely on their steady-state spectral cues. Using the values of F0 and the first three formant frequency values (F1, F2, F3) measured by Peterson and Barney (1952), they asked listeners to identify the synthesized stimuli of the steady-state vowel values. While the listeners of Peterson and Barney (1952) showed a 5.6 % error rate, those who listened to the static synthesized stimuli displayed a 27.3 % error rate. All these studies suggest that the entire range of dynamic properties provide cues that can improve the identification of vowels.[1]

The second argument supporting the importance of the dynamic pattern is related to the fact that the steady-state values of formant frequencies among vowels can overlap. This overlap can occur due to diachronic vowel changes, speaker variabilities, and/or phonetic context variabilities. For example, Hillenbrand, Getty, Clark, and Wheeler (1995) reported that the vowels [ɛ] and [æ] exhibited more overlap in the F1 versus F2 plot in their study than in Peterson and Barney (1952). However, listeners in Hillenbrand, Getty, Clark, and Wheeler (1995) still identified the two vowels quite well. The authors suggested that their subjects amplified spectral changes in the vowels (i.e. diphthongized the vowels more) compared to those of Peterson and Barney (1952), which had the effect of decreasing perceptual confusion which could have arisen from the overlap in the static values. Consequently, they suggested that American English vowels need to be characterized in terms of the entire trajectory, not just the static target values. Nearey and Assmann (1986) and Nearey (1989) also claimed that monophthong vowels in English have considerable formant movements or diphthongizations and are characterized by vowel-inherent dynamic properties. The third argument for considering the dynamic pattern arises from the fact that the relative time to target value varies across vowels in English. For example, tense vowels tend to delay the target timings compared to lax vowels (Watson and Harrington 1999).

The evidence supporting these three arguments come from studies examining a number of different aspects of vowel dynamics including: how dynamic information

[1] There are also studies in which the steady-state spectral information was found to be sufficient for monophthong perception. For example, Harrington and Cassidy (1994) reported that three spectral slices taken from the vowel midpoint and the two transitions in [CVd] words of Australian English helped diphthong classification, but they found that spectral information from the midpoint was sufficient for monophthongs. Watson and Harrington (1999) also claimed that it is necessary to model dynamic formant trajectories for diphthong discriminations, but that steady-state spectral values and durations are sufficient for identifying monophthong vowels. Putting these results together, it seems to be the case that there are experimental results in which vowels with consonant transitions are identified more accurately than vowels in isolation, but no studies in which isolated vowels are perceived more accurately than vowels with consonant transitions.

aids in monophthong perception; the overlap problem of static formant frequencies among vowels; the diphthongization of monophthong vowels in English; and the difference in the relative timings of vowel targets. What they all hold in common is the suggestion that the full range of spectral trajectories provide vowel cues well above and beyond those just provided by steady-state frequencies.

Studies of vowels produced by non-native speakers have mainly focused on differences between the steady-state values of vowel formants produced by native and non-native speakers (e.g. Kondo 2000, Guion 2003, Oh 2011). One study which investigated the dynamic patterns of vowels produced by non-native speakers is Yang (2010). Yang compared the spectral changes of English tense and lax vowels ([i] and [ɪ], [u] and [ʊ]) produced by native speakers of American English and Korean learners of English. The formant values were measured at seven time points. The results indicated that the Korean group did not produce tense-lax distinctions as clearly as the American group. Clearly, there is reason to think that focusing only on static differences is insufficient for the task of fully understanding how non-native speakers learn and produce English vowels.

Based on these motivations, the present study investigates how Korean-English bilingual speakers and Korean late learners of English realize dynamic patterns of front monophthong vowels in American English. It is explored whether language experience affects how dynamic patterns are learned by examining the spectral dynamics of vowels produced by simultaneous or early bilinguals of Korean and English, as well as late learners of English. The relative timings of vowel target values and the vowel durations will also be compared between native and non-native speakers of English.

2. Experimental methods

Twenty speakers participated in this production study. Eight were native speakers of American English (Group NE, age range 20-24, mean age 21.3), six were Korean-English bilingual speakers (Group KE, age range 19-21, mean age 19.8), and six were Korean learners of English (Group NK, age range 18-28, mean age 23.8). Several of the speakers overlap with those of Oh (2011). All participants were male.

All members of Groups NE and KE were undergraduate students at a university located in the American state of Oregon. All members of Group NE were native speakers of western American English, born and raised in Oregon or California. All the speakers in Group KE were born and raised in the U.S. except KE1 who was born in Seoul, Korea and raised in the U.S. from age 5. Speakers KE3, KE4, KE5, and KE6 were exposed to both Korean and English from birth, and speakers KE1 and KE2 were exposed to Korean from birth and English from age 5. At the time of recording, KE1, KE2, KE3, and KE4 used both English and Korean daily. KE5 and KE6 used English daily and Korean only when necessary. No member of Groups NE and KE could speak any other languages at more than an intermediate level. All the speakers of Group NK were undergraduate or graduate students at a university located in Seoul, Korea. All were native speakers of Seoul Korean, who were born and raised in Seoul. None had resided in an English-speaking country for more than one year.

For test materials, the monosyllabic words “beat [bit],” “bit [bɪt],” “bet [bɛt],” and “bat [bæt]” were used, containing a voiced bilabial stop in the initial position, a front vowel in the medial, and a voiceless alveolar stop in the final. The test words were read in the carrier sentence ‘Say “___” to me.’

Regarding the vowels examined, when one begins enunciating the English high front tense vowel [i], the tongue is already in the high and front position, but moves to a position that is even higher and further forward. Therefore, this vowel is sometimes “symbolized as a vowel plus a glide [iy] (Prator and Robinett 1985: 126),” and not considered a pure monophthong (e.g. Celce-Murcia, Brinton, and Goodwin 1996, Davenport and Hannahs 2005). When one articulates the English high front lax vowel [ɪ] after [i], the jaw relaxes and drops, and “the forced spreading of the lips disappears (Prator and Robinett 1985: 126).” For the articulation of the English mid front lax vowel [ɛ], the tongue body is between the mid and low positions in the mouth. For the English low front lax vowel [æ], the tongue body is lower than [ɛ] and “may drop a bit lower during articulation (Celce-Murcia, Brinton, and Goodwin 1996: 103).”

Recordings were made in sound-proof rooms in phonetics laboratories located in the universities attended by the speakers. A MARANTZ PMD670 recorder and a Shure BG5.1 microphone were used to record Groups NE and KE, while a TASCAM CC-222MKII CD recorder and a Shure KSM32 microphone were used for Group NK. The test sentences were presented on a computer screen one at a time at an interval of two seconds. The production time of each sentence was limited to two seconds so as to limit the vowel duration difference among speakers and between groups. The recordings were digitized at a sampling rate of 22,050 Hz, and stored as WAV files on a PC. The data in the WAV files were then analyzed using the speech analysis program PCquirerX.

Duration and formant frequency values (F1, F2, and F3) of the vowels were measured for each speaker. Vowel duration was measured on a waveform using a cursor from right after the release of the word-initial stop to the beginning point of the closure of the word-final stop. The measurements were confirmed using spectrograms. Formant frequencies were estimated by linear predictive coding (LPC), then confirmed also using spectrograms. To measure formant frequencies, the vowel duration was segmented into 10 equal parts. A measurement was taken at the ends and each cutpoint, resulting in 11 measurements, with the first occurring after the release of the initial stop and the last at the beginning of the closure of the final stop. All in all, a total of 2720 measurements were made ([3 formants * 11 measurement points * 4 vowels * 20 speakers] + [vowel duration * 4 vowels * 20 speakers]).

3. Experimental results

3.1 Dynamic spectral patterns of American English front monophthong vowels

Figure 1 presents the formant movements of the four American English front monophthong vowels produced by the three groups in the study. Eleven points are presented for each group, with each point representing the group average for the respective measurement point. The horizontal axis indicates the eleven measurement points and the vertical axis the formant frequencies. Table 1 presents the average formant values of the four front vowels produced by the three groups for the eleven measurement points.

For the vowel [i] ("beat"), the F1 and F2 transition shapes of Groups KE and NK look analogous to those of Group NE. The maximum and minimum F1 values are, respectively, 275 Hz (measurement point 1) and 234 Hz (11) for Group NE, 255

Hz (10) and 230 Hz (11) for Group KE, and 278 Hz (8) and 240 Hz (1) for Group NK. The differences between the maximum and minimum F1 values are 41 Hz for Group NE, 25 Hz for Group KE, and 38 Hz for Group NK. The average F1 values across the eleven measurement points are 260 Hz for Group NE, 242 Hz for Group KE, and 262 Hz for Group NK. The maximum and minimum F2 values are, respectively, 2284 Hz (measurement point 7) and 2111 Hz (1) for Group NE, 2241 Hz (7) and 2084 Hz (1) for Group KE, and 2292 Hz (6 and 8) and 2158 Hz (1) for Group NK. The differences between the maximum and minimum F2 values are 173 Hz for Group NE, 157 Hz for Group KE, and 134 Hz for Group NK. The average F2 values across the eleven measurement points are 2226 Hz for Group NE, 2192 Hz for Group KE, and 2248 Hz for Group NK. As for F3 values, while the transition shape of Group KE looks analogous to that of Group NE, Group NK's looks different from both. The maximum and minimum F3 values are, respectively, 3008 Hz (measurement point 6) and 2636 Hz (1) for Group NE, 2967 Hz (6) and 2665 Hz (1) for Group KE, and 3124 Hz (5) and 2684 Hz (1) for Group NK. The differences between the maximum and minimum F3 values are 372 Hz for Group NE, 302 Hz for Group KE, and 440 Hz for Group NK. The average F3 values across the eleven measurement points are 2881 Hz for Group NE, 2872 Hz for Group KE, and 2991 Hz for Group NK.

For the vowel [ɪ] ("bit"), the difference between Groups NE and NK looks larger than that between Groups NE and KE for all F1, F2, and F3 transition shapes. The maximum and minimum F1 values are, respectively, 433 Hz (measurement point 7) and 374 Hz (1) for Group NE, 436 Hz (6) and 335 Hz (11) for Group KE, and 327 Hz (5) and 285 Hz (1) for Group NK. The differences between the maximum and minimum F1 values are 59 Hz for Group NE, 101 Hz for Group KE, and 42 Hz for Group NK. The average F1 values across the eleven measurement points are 410 Hz for Group NE, 396 Hz for Group KE, and 313 Hz for Group NK. The maximum and minimum F2 values are, respectively, 1943 Hz (measurement point 4) and 1812 Hz (11) for Group NE, 1879 Hz (6) and 1804 Hz (10) for Group KE, and 2137 Hz (4) and 2012 Hz (10) for Group NK. The differences between the maximum and minimum F2 values are 131 Hz for Group NE, 75 Hz for Group KE, and 125 Hz for Group NK. The average F2 values across the eleven measurement points are 1905 Hz for Group NE, 1841 Hz for Group KE, and 2099 Hz for Group NK. The maximum and minimum F3 values are, respectively, 2713 Hz (measurement point 9) and 2562 Hz (1) for Group NE, 2713 Hz (9) and 2475 Hz (1) for Group KE, and 2850 Hz (5) and 2677 Hz (1) for Group NK. The differences between the maximum and minimum F3 values are 151 Hz for Group NE, 238 Hz for Group KE, and 173 Hz for Group NK. The average F3 values across the eleven measurement points are 2667 Hz for Group NE, 2631 Hz for Group KE, and 2781 Hz for Group NK.

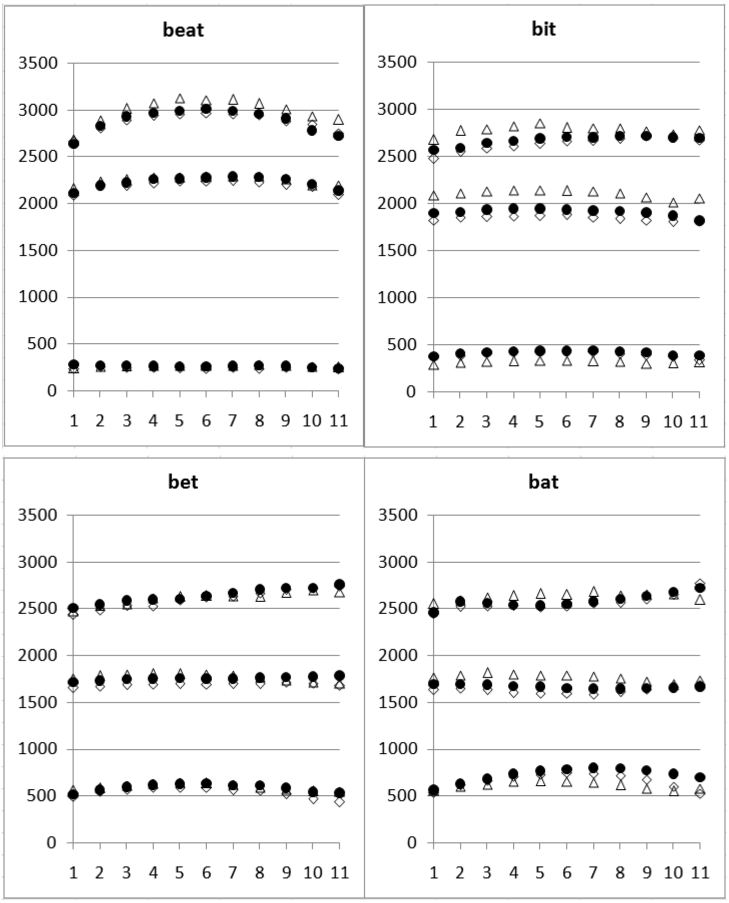

Figure 1. Transition shapes of F1, F2, and F3 (average values by measurement points) of American English front monophthongs produced by Group NE (●), Group KE (◇), and Group NK (△) (The horizontal axis indicates the measurement points and the vertical axis the formant frequencies (Hz).)

Table 1. Average formant values (Hz) of American English front monophthongs produced by the three groups at eleven measurement points

		Group NE			Group KE			Group NK		
		F1	F2	F3	F1	F2	F3	F1	F2	F3
[i]	1	275	2111	2636	233	2084	2665	240	2158	2684
	2	270	2189	2823	253	2207	2802	259	2237	2890
	3	265	2222	2926	243	2189	2887	263	2252	3012
	4	263	2258	2965	246	2214	2938	264	2277	3072
	5	258	2264	2990	240	2238	2958	265	2279	3124
	6	259	2273	3008	238	2239	2967	268	2292	3104
	7	260	2284	2988	243	2241	2954	269	2290	3111
	8	264	2282	2955	237	2225	2943	278	2292	3067
	9	261	2257	2908	248	2200	2877	267	2257	3008
	10	246	2206	2777	255	2179	2848	258	2197	2932
	11	234	2137	2719	230	2091	2748	252	2194	2896
[ɪ]	1	374	1895	2562	370	1821	2475	285	2086	2677
	2	405	1903	2584	381	1848	2549	311	2109	2772
	3	416	1929	2636	409	1857	2586	317	2126	2791
	4	423	1943	2662	421	1865	2604	323	2137	2822
	5	429	1940	2685	424	1866	2633	327	2135	2850
	6	431	1932	2699	436	1879	2659	326	2133	2805
	7	433	1921	2696	418	1854	2665	325	2129	2797
	8	427	1918	2710	406	1835	2685	316	2104	2796
	9	414	1899	2713	387	1821	2713	300	2065	2767
	10	382	1867	2698	374	1804	2697	301	2012	2732
	11	381	1812	2692	335	1806	2674	314	2054	2780
[ɛ]	1	512	1709	2508	491	1652	2436	566	1753	2471
	2	560	1730	2542	545	1671	2482	583	1783	2531
	3	592	1743	2586	566	1687	2531	612	1800	2551
	4	613	1751	2596	586	1688	2524	630	1804	2614
	5	626	1754	2601	588	1698	2587	644	1804	2627
	6	628	1750	2636	587	1691	2623	637	1800	2639
	7	611	1749	2668	562	1698	2630	619	1784	2634
	8	607	1759	2702	553	1698	2664	590	1766	2628
	9	581	1768	2718	519	1709	2702	567	1731	2669
	10	541	1773	2717	464	1705	2721	549	1715	2694
	11	529	1782	2756	433	1679	2741	547	1705	2679

[æ]	1	561	1689	2458	536	1631	2445	556	1764	2553
	2	627	1693	2574	614	1645	2520	600	1784	2590
	3	680	1686	2558	662	1633	2528	623	1814	2625
	4	733	1669	2534	704	1600	2531	654	1797	2638
	5	766	1664	2531	730	1594	2516	658	1789	2660
	6	782	1650	2548	752	1593	2527	650	1784	2649
	7	798	1642	2574	741	1582	2559	642	1774	2681
	8	788	1644	2603	714	1612	2564	613	1755	2647
	9	768	1649	2634	674	1636	2601	580	1725	2656
	10	734	1652	2677	600	1654	2650	551	1686	2655
	11	698	1663	2722	522	1674	2771	578	1732	2597

For the vowel [ɛ] ("bet"), the three groups showed analogous patterns for F1, F2, and F3 transition shapes. The maximum and minimum F1 values are, respectively, 628 Hz (measurement point 6) and 512 Hz (1) for Group NE, 588 Hz (5) and 433 Hz (11) for Group KE, and 644 Hz (5) and 547 Hz (11) for Group NK. The differences between the maximum and minimum F1 values are 116 Hz for Group NE, 155 Hz for Group KE, and 97 Hz for Group NK. The average F1 values across the eleven measurement points are 582 Hz for Group NE, 536 Hz for Group KE, and 595 Hz for Group NK. The maximum and minimum F2 values are, respectively, 1782 Hz (measurement point 11) and 1709 Hz (1) for Group NE, 1709 Hz (9) and 1652 Hz (1) for Group KE, and 1804 Hz (4 and 5) and 1705 Hz (11) for Group NK, respectively. The differences between the maximum and minimum F2 values are 73 Hz for Group NE, 57 Hz for Group KE, and 99 Hz for Group NK. The average F2 values across the eleven measurement points are 1752 Hz for Group NE, 1689 Hz for Group KE, and 1768 Hz for Group NK. The maximum and minimum F3 values are, respectively, 2756 Hz (measurement point 11) and 2508 Hz (1) for Group NE, 2741 Hz (11) and 2436 Hz (1) for Group KE, and 2694 Hz (10) and 2471 Hz (1) for Group NK. The differences between the maximum and minimum F3 values are 248 Hz for Group NE, 305 Hz for Group KE, and 223 Hz for Group NK. The average F3 values across the eleven measurement points are 2639 Hz for Group NE, 2604 Hz for Group KE, and 2612 Hz for Group NK.

For the vowel [æ] ("bat"), the difference between Groups NE and NK looks larger than the difference between Groups NE and KE for all F1, F2, and F3 transition shapes. The maximum and minimum F1 values are, respectively, 798 Hz (measurement point 7) and 561 Hz (1) for Group NE, 752 Hz (6) and 522 Hz (11) for Group KE, and 658 Hz (5) and 551 Hz (10) for Group NK. The differences between the maximum and minimum F1 values are 237 Hz for Group NE, 230 Hz for Group KE, and 107 Hz for Group NK. The average F1 values across the eleven measurement points are 721 Hz for Group NE, 659 Hz for Group KE, and 610 Hz for Group NK. The maximum and minimum F2 values are, respectively, 1693 Hz (measurement point 2) and 1642 Hz (7) for Group NE, 1674 Hz (11) and 1582 Hz (7) for Group KE, and 1814 Hz (3) and 1686 Hz (10) for Group NK. The differences between the maximum and minimum F2 values are 51 Hz for Group NE, 92 Hz for Group KE, and 128 Hz for Group NK. The average F2 values across the eleven measurement points are 1664 Hz for Group NE, 1623 Hz for Group KE, and 1764 Hz for Group NK. The maximum and minimum F3 values are, respectively, 2722 Hz (measurement point 11) and 2458 Hz (1) for Group NE, 2771 Hz (11) and 2445 Hz (1) for Group KE, and 2681 Hz (7) and 2553 Hz (1) for Group NK. The differences between the maximum and minimum F3 values are 264 Hz for Group NE, 326 Hz for Group KE, and 128 Hz for Group NK. The average F3 values across the eleven measurement points are 2583 Hz for Group NE, 2565 Hz for Group KE, and 2632 Hz for Group NK.

Table 2 presents the average F1, F2, and F3 values across the eleven measurement points for the three groups. The average formant values of Group KE were smaller than those of Group NE for all four vowels. All the average formant values of Group NK were larger than those of Group NE except the F1 values of [ɪ] and [æ], and the F3 value of [ɛ].

Table 2. Average formant (F1, F2, and F3) values across eleven measurement points for Groups NE, KE, and NK (Hz)

	F1				F2				F3			
	[i]	[ɪ]	[ɛ]	[æ]	[i]	[ɪ]	[ɛ]	[æ]	[i]	[ɪ]	[ɛ]	[æ]
NE	260	410	582	721	2226	1905	1752	1664	2881	2667	2639	2583
KE	242	396	536	659	2192	1841	1689	1623	2872	2631	2604	2565
NK	262	313	595	610	2248	2099	1768	1764	2991	2781	2612	2632

3.2 Dynamic movements of English front monophthong vowels on F1–F2 plots

Figures 2 and 3 present the F1 (vertical axis) and F2 (horizontal axis) plots of the dynamic movements of the American English front vowels produced by Groups NE, KE, and NK. Figure 2 presents results for the vowels [i] and [ɪ], while Figure 3 does so for [ɛ] and [æ]. The arrows indicate the starting points of the vowel trajectories (i.e. measurement point 1 for each vowel).

As seen in Figure 2, the ranges of the formant movements for the vowel [i] overlap for all three groups. However, the movement directions of the vowel are quite different across the groups. For Group NE, the vowel starts from a lower-retracted position, moves higher-fronter, then to a higher-retracted position (assuming that F1 values inversely relate to vowel height and F2 values positively relate to vowel frontness, e.g. Ladefoged 1993). For Group KE, the vowel starts from a higher-retracted position, moves lower-fronter, then to a lower-retracted position. For Group NK, the vowel starts from a higher-retracted position, moves lower-fronter, then to a higher-retracted position.

While the ranges of the formant movements for the vowel [ɪ] overlap for Groups NE and KE, the formant range for Group NK (indicated by the dashed circle) is considerably different from both. Group NK produced the vowel with smaller F1 and larger F2 values compared to Groups NE and KE. The trajectories of [i] and [ɪ] produced by Groups NE and KE are completely separated, ensuring sufficient discrimination between the vowels, but the trajectories of the two vowels produced by Group NK are close to each other, though not overlapping. On the other hand, the movement directions of the vowel are similar for all three groups. Across the groups, the vowel starts from a higher-fronter position, moves lower, then to a higher-retracted position.

Figure 3 shows that the ranges of the formant movements for the vowel [ɛ] overlap for Groups NE and NK, but the formant range for Group KE is a bit higher and more retracted. However, the movement directions of the vowel are similar for Groups NE and KE, but not NK. For Groups NE and KE, the vowel starts from a higher-retracted position, moves lower-fronter, then to a higher-further forward position. For Group NK, the vowel starts from a higher-retracted position, moves lower-fronter, then to a higher-retracted position.

The ranges of the formant movements for the vowel [æ] are separate for all three groups. Compared to Groups NE and KE, the F1 and F2 ranges for Group NK are smaller and larger, respectively (i.e. the vowel is higher and fronter). The F2 range is larger (i.e. the vowel is fronter) for Group NE than for Group KE. The movement directions of the vowel are similar between Groups NE and KE, but not NK. For Groups NE and KE, the vowel starts from a higher-fronter position, moves lower-retracted, then to a higher-fronter position. For Group NK, the vowel starts from a high-retracted position, moves lower-forward, then to a higher-retracted position.[2] Though they do not overlap, the trajectories of the vowels [ɛ] and [æ] are not as distinct as the trajectories of [i] and [ɪ] for Groups NE and KE (see section 1 and Hillenbrand, Getty, Clark, and Wheeler 1995). The trajectories of [ɛ] and [æ] completely overlap for Group NK, as shown in the dashed circle in Figure 3.

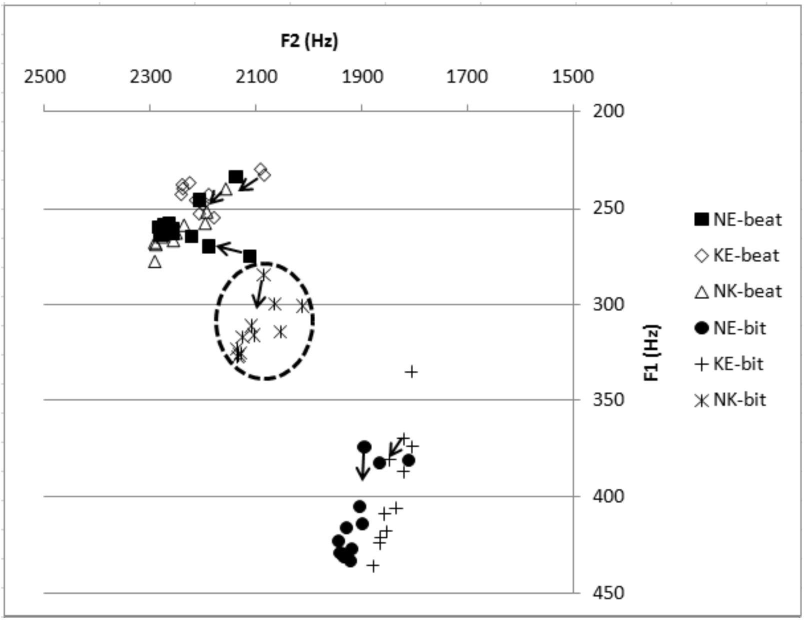

Figure 2. F1–F2 plots of dynamic movements of the English front vowels [i] and [ɪ] produced by Groups NE, KE, and NK (Arrows indicate the starting points of the vowel trajectories.)

2 The vowel raising which appeared in the latter half of the vowels [ɪ, ɛ, æ] for all three groups is presumed to occur in order to articulate the following alveolar stop.

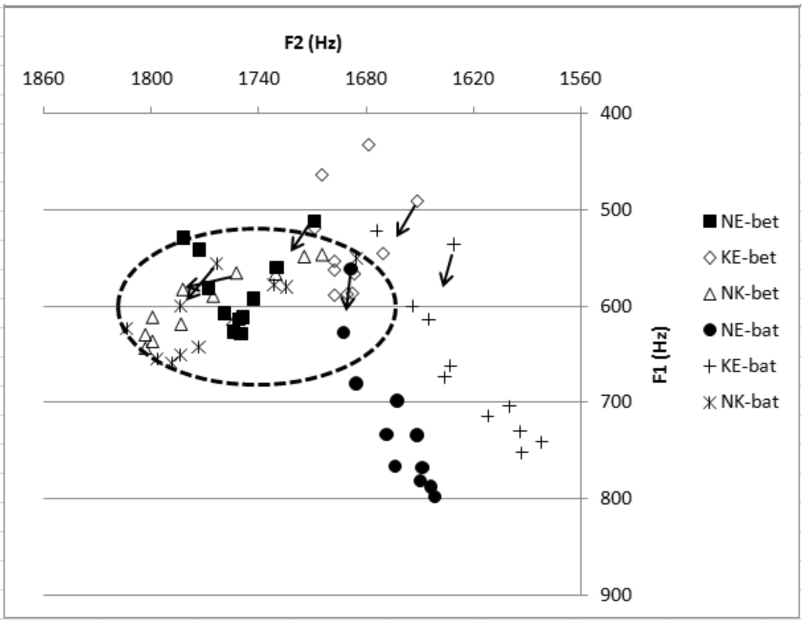

Figure 3. F1-F2 plots of dynamic movements of the English front vowels [ɛ] and [æ] produced by Groups NE, KE, and NK (Arrows indicate the starting points of the vowel trajectories.)

3.3 Relative timing of F2 targets for [i] and [ɪ]

The difference in the relative timing of the vowel targets for the tense and lax vowels was examined by finding the measurement points showing the maximum F2 values of the vowels [i] and [ɪ]. Since the locus F2 values of the preceding bilabial and following alveolar consonants are lower than the target F2 values of the high front vowels, the F2 trajectories of the vowels, in general, show inverse *u* shapes (e.g. Sussman, Hoemeke, and Ahmed 1993, Ladefoged 2003). Therefore, the maximum F2 values are expected to be the target F2 values of the vowels.

Table 3 shows that for all individuals in Group NE, the maximum F2 values for the vowel [i] occur at or after measurement point 6, and those for [ɪ] occur at or before point 6. As we would expect, the relative timing of the vowel target of the tense [i] tended to occur later than that for the lax [ɪ] (see section 1 and Watson and Harrington 1999). For Group KE, the maximum F2 value for [i] was exhibited at measurement point 2 by subject KE2, and at measurement point 5 by KE1 and KE4.

The maximum F2 value for [ɪ] was exhibited by KE6 at measurement point 8. For Group NK, the maximum F2 value for [i] was exhibited by subject NK5 at measurement point 2, and that for [ɪ] was observed from NK4 at point 11. In summary, there were individual speakers in Groups KE and NK who did not exhibit native-like patterns for the relative timing of the F2 targets for the vowels [i] and [ɪ].

Table 3. Measurement points showing the maximum F2 values of vowels [i] and [ɪ] by individual speaker

	Group NE								Group KE						Group NK					
	1	2	3	4	5	6	7	8	1	2	3	4	5	6	1	2	3	4	5	6
[i]	8	7	6	11	8	6	7	7	5	2	8	5	6	9	7	7	7/8	11	2	6
[ɪ]	5	4	5	4	6	3	3	1	6	6	3	1	4	8	4	6	4	11	4	4

3.4 Durations of English front monophthong vowels by speaker group

Figure 4 presents the average durations of the English front monophthong vowels produced by the three groups. Although the production time of a test sentence was restricted to two seconds (see section 2 for explanation), there appeared to be systematic differences in the durations of the vowels. All three groups produced vowel durations in the order [æ] (longest) > [i] > [ɛ] > [ɪ] (shortest). This pattern is explained by speech physiology: the lower vowels were longer than the higher vowels (i.e. [æ] was longer than [i], [ɪ], and [ɛ], and [ɛ] was longer than [ɪ]), and the tense vowel was longer than the lax vowels (i.e. [i] was longer than [ɛ] and [ɪ]). The vowel durations of Group KE tended to be slightly longer than those for Group NE. Vowel durations for Group NK were shorter than those for Group NE. One noticeable difference among the speaker groups was that the average duration of [æ] was considerably shorter for Group NK than for Group NE ($p = 0.056$, independent samples *t*-test for Group NE and Group NK).

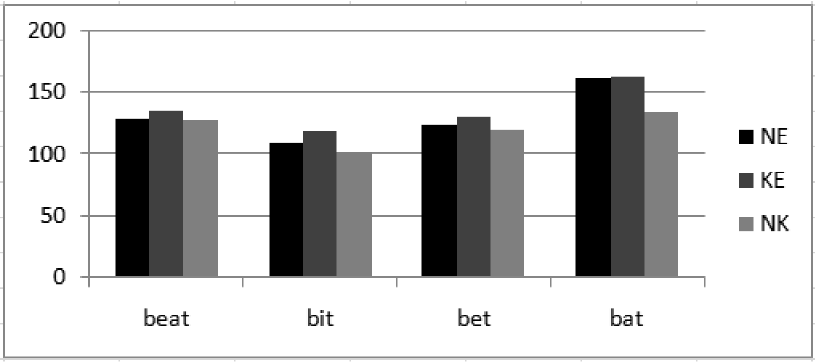

Figure 4. Average durations of the English front vowels produced by Groups NE, KE, and NK (ms)

4. Summary and discussion

Overall, the Korean-English bilingual participants in this study (Group KE) produced the front vowels more similar to those of the native speakers of American English (Group NE) than did the Korean late learners of English (Group NK). Aspects of vowel production that were examined were F1, F2, and F3 transition shapes (Figure 1), as well as overall formant values of the English front vowels. The movement direction of the vowel [i] was different across all three groups, that for the vowel [ɪ] was similar across the three, and those for [ɛ] and [æ] were similar for Groups NE and KE, but not for NK (Figure 2). For the vowel [ɪ], the ranges of the formant movements overlapped for Groups NE and KE, but the formant range for Group NK was quite different from those of the other two groups. The trajectories of [i] and [ɪ] produced by Groups NE and KE were completely separate, ensuring sufficient discrimination between the vowels. On the other hand, the trajectories of the vowels produced by Group NK were close to each other, possibly impeding discrimination between the vowels.[3] For Groups NE and KE, the distinction between the trajectories of [ɛ] and [æ] was not as clear as that between [i] and [ɪ], though they did not overlap (cf. Hillenbrand, Getty, Clark, and Wheeler 1995). In the case

3 According to Yang (2010) which studied the spectral changes of the English vowels [i, ɪ, u, ʊ] produced by native speakers of English and Korean late learners of English, the two groups were especially different in the F2 values for [i] and in all F1, F2, and F3 values of [ɪ] for its male subjects.

of Group NK, [ɛ] and [æ] actually overlapped (Figure 2).

As for the relative timing of the F2 targets for the vowels [i] and [ɪ], the relative timing of the vowel target tended to be later for the tense [i] than for the lax [ɪ] for Group NE (e.g. Watson and Harrington 1999). Within Groups KE and NK, there were speakers who showed maximum F2 values for [i] and [ɪ] that were respectively earlier and later than those of the speakers in Group NE. As for vowel durations, all three speaker groups exhibited longer durations for the lower vowels than the higher ones, and for the tense vowel than the lax ones. Group KE tended to produce vowel durations that were slightly longer than those for Group NE. The vowel durations for Group NK tended to be shorter than those for Group NE.

Summing up the results, Group KE exhibited more native-like dynamic patterns of the American English front monophthongs than did Group NK. This is seen in the transition shapes of formants, the overall formant ranges, and the movement directions of vowels. These results indicate that language experience affects the learning of dynamic patterns of non-native vowels. To some extent, these findings support extant theories of phonetics learning, for example, the exemplar-based model, which claims that phonetic representations of words are the sum of the language experience that a speaker accumulates through his or her life span (see Foulkes and Docherty 2006 and references mentioned there). However, Group KE did show differences from Group NE, for example, in the movement direction of the vowel [i] and the F2 target timing of the vowels [i] and [ɪ]. It remains to be explained why these phonetic aspects were particularly difficult for the bilingual speakers to acquire, whether due to perceptual difficulty or some other reason.

One direction for future research is to determine whether there is a transfer effect from the dynamic shapes of Korean vowels in the English vowels produced by the speakers in Group NK. Such a study may be conducted by examining the dynamic spectral patterns of Korean vowels. In this study, Group NK produced native-like static and dynamic patterns for the English vowel [i], but produced the English vowel [ɪ] quite differently from both the [i] and [ɪ] produced by Group NE (Figure 1). It needs to be examined whether the dynamic patterns of the Korean [i] are analogous to those of the English [i]. In the case of the English vowels [ɛ] and [æ], Group NK produced native-like static and dynamic patterns for the former, but not the latter. The vowel trajectory of [æ] completely overlapped with that of [ɛ] (Figure 2). It may be the case that due to the merging of the Korean vowels [ɛ] and [æ], the speakers in Group NK lost the ability to distinguish between the two vowels in the low front vowel region (e.g. Oh 2011). As a result, they may have merged the two English vowels into the Korean values of [ɛ]. It needs to be explored whether the merged value of the two English vowels produced by Group NK are analogous to the native Korean [ɛ]. By comparing the dynamic patterns of vowels between the languages, it may be possible to separate, to a certain extent, the language-universal aspects of vowel dynamics from the language-specific ones. In addition, it would be interesting to perform a perception study of the vowel dynamics produced by non-native speakers in order to see whether the perception of stimuli produced by non-native speakers also benefit from dynamic information and whether vowel perception scores differ in the stimuli produced by speakers with varying degrees of language experience.

References

Benguerel, A.-P., and T. U. McFadden. 1989. The effect of coarticulation on the role of transitions in vowel perception. *Phonetica* 46: 80-96.

Celce-Murcia, M., D. M. Brinton, and J. M. Goodwin. 1996. *Teaching Pronunciation: A Reference for Teachers of English to Speakers of Other Languages*. Cambridge University Press.

Davenport, M., and S. J. Hannahs. 2005. *Introducing Phonetics and Phonology* (second edition). Hodder Arnold.

Foulkes, P., and G. Docherty. 2006. The social life of phonetics and phonology. *Journal of Phonetics* 34: 409-438.

Guion, S. G. 2003. The vowel systems of Quichua-Spanish bilinguals: Age of acquisition effects on the mutual influence of the first and second language. *Phonetica* 60: 98-128.

Harrington, J., and S. Cassidy. 1994. Dynamic and target theories of vowel classification: Evidence from monophthongs and diphthongs in Australian English. *Language and Speech* 37: 357-373.

Hillenbrand, J., and R. T. Gayvert. 1993. Identification of steady-state vowels synthesized from the Peterson and Barney measurements. *Journal of the Acoustical Society of America* 94: 668-674.

Hillenbrand, J., L. A. Getty, M. J. Clark, and K. Wheeler. 1995. Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America* 97: 3099-3111.

Huang, C. B. 1992. Modelling human vowel identification using aspects of formant trajectory and context. In Y. Tohkura, E. Vatikiotis-Bateson, and Y. Sagisaka (eds.), *Seech Perception, Production and Linguistic Structure*, 43-61. Amsterdam: IOS Press.

Kondo, Y. 2000. Production of schwa by Japanese speakers of English: An acoustic study of shifts in coarticulatory strategies from L1 to L2. In M. B. Broe, and J. B. Pierrehumbert (eds.), *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, 29-39. New York: Cambridge University Press.

Ladefoged, P. 1993. *A Course in Phonetics* (3rd edition). Harcourt Brace.

Ladefoged, P. 2003. *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*. Blackwell Publishing.

Lehiste, I., and G. E. Peterson. 1961. Transitions, glides and diphthongs. *Journal of the Acoustical Society of America* 33: 268-277.

Lindblom, B. 1963. Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America* 35: 1773-1781.

Nearey, T. M. 1989. Static, dynamic, and relational properties in vowel perception. *Journal of the Acoustical Society of America* 85: 2088-2113.

Nearey, T. M., and P. F. Assmann. 1986. Modelling the role of inherent spectral change in vowel identification. *Journal of the Acoustical Society of America* 80: 1297-1308.

Oh, E. 2011. The mid and low vowel systems of Korean-English bilingual speakers. *Korean Journal of Linguistics* 36: 635-657.

Peterson, G. E., and H. L. Barney. 1952. Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24: 175-184.

Prator, C. H., and B. W. Robinett. 1985. *Manual of American English Pronunciation* (fourth edition). Harcourt Brace.

Strange, W., J. J. Jenkins, and T. L. Johnson. 1983. Dynamic specification of coarticulated vowels. *Journal of the Acoustical Society of America* 74: 695-705.

Sussman, H. M., K. A. Hoemeke, and F. S. Ahmed. 1993. A cross-linguistic investigation of locus equations as a phonetic descriptor for place of articulation. *Journal of the Acoustical Society of America* 94: 1256-1268.

Watson, C. I., and J. Harrington. 1999. Acoustic evidence for dynamic formant trajectories in Australian English vowels. *Journal of the Acoustical Society of America* 106: 458-468.

Yang, B. 2010. Formant trajectories of English high tense and lax vowels produced by Korean and American speakers. *Korean Journal of Linguistics* 35: 407-423.

Zahorian, S. A., and A. J. Jagharghi. 1993. Spectral-shape features versus formants as acoustic correlates for vowels. *Journal of the Acoustical Society of America* 94: 1966-1982.

Eunjin Oh
Department of English Language and Literature
Ewha Womans University
52 Ewhayeodae-gil, Seodaemun-gu
Seoul 120-750, Korea
Email: ejoh@ewha.ac.kr

Received: 2013. 05. 20
Revised: 2013. 08. 10
Accepted: 2013. 08. 10