

말레이시아어와 인도네시아어 모음 공간의 지형도

박한상(홍익대), 박정숙(한국외대), 전태현(한국외대)

<차 례>

- | | |
|---------|---------|
| 1. 서론 | 3. 연구결과 |
| 2. 연구방법 | 4. 결론 |

<Abstract>

Geophysics of Vowel Space in Bahasa Malaysia and Bahasa Indonesia

Hansang Park, Jeongsook Park, Taihyun Chun

The present study investigates the vowels of Bahasa Malaysia and Bahasa Indonesia in terms of the first two formant frequencies and provides a three-dimensional formant chart of vowels by plotting F1, F2, and the frequency of <F1,F2> datapoints on 4 different scales: Hz, mel, bark, and the number of ERB. For this study, we recruited 30 male native speakers of Bahasa Malaysia and Bahasa Indonesia (15 each) which include 6 vowels (i, e, a, o, u, ə) in various contexts. The three-dimensional formant chart showed geophysics of vowel space, such that mountain peaks stand in particular locations with a higher frequency of occurrence of datapoints. The geophysics of vowel space may shed lights on the perceptual structure of vowel space. The results also showed that vowels in utterance-final positions have a significantly higher F1 and a significantly lower F2 than those in utterance-medial or utterance-initial positions, which means that vowels in utterance-final positions are more back and lower in vowel space than those in utterance-medial or utterance-initial positions.

* Keywords: Bahasa Malaysia, Bahasa Indonesia, Geophysics, Vowel space, Formant, F1, F2.

1. 서론

모음의 포먼트는 모음의 음향학적 특성을 연구하는 데 널리 사용되어 왔다. 그 중 F1은 혀의 높이와 관계가 있고 F2는 혀의 위치와 관계가 있다고 알려져 있다 [1]. 모음의 음향적 특성에 관한 연구는 여러 연구를 통해 이론적으로 설명되었고 경험적으로 증명되었다[1][2][3][4][5][6][7][8][9]. 모음의 특성 및 모음 공간에 관한 연구는 정규화(normalization)의 방법에 따라 두 가지로 나눌 수 있다. 하나는 성별과 연령 등의 화자 특성으로 인한 목표 지점의 차이, 환경과 맥락으로 인한 목표 지점의 차이, 모음의 길이나 종류 등 모음의 고유한 특성으로 인한 목표 지점의 차이 등을 설명하는 것이고 다른 하나는 청취 체계와 인지적 특성을 고려하여 모음 공간을 설명하는 것이다.

초기 연구에서는 정적인 안정 구간에서 측정된 모음의 음향적 특성을 설명하였으며 이른바 “Simple Target Model”로 알려져 있다[10][11][12][13]. Peterson and Barney(1952)는 미국 영어 모음의 F1과 F2 값을 측정하여 평면에 배치한 포먼트 차트를 모음 기술에 사용하였고 성인 남성, 성인 여성, 아동은 모음 공간에서 목표 모음의 위치가 다르다는 실험 결과를 제시하였다[3]. 목표 모음의 위치는 성별과 모음뿐만 아니라 “target undershoot”[14][15], 모음의 고유한 길이(intrinsic vowel duration), 긴장 모음(tense vowel)과 이완 모음(lax vowel)을 포함한 모음의 종류[16], 인접한 자음에서 모음으로 혹은 모음에서 인접한 자음으로의 포먼트 전이[17] 등에 따라서도 차이가 난다고 하였다.

Simple Target Model의 한계를 극복하기 위한 대안으로 제시된 것이 이른바 “Elaborated Target Model”이다[10][11][12][13]. Hz 단위로 측정된 모음의 포먼트 주파수를 mel, bark, K, ERB의 수 등 인지적 및 심리음향학적 단위로 변환하여 모음 공간의 인지적 특성을 연구하였다[6][7][18][19][20][21]. Strange(1989)는 두 모델 중 어떤 모델도 모음의 인지를 충분히 설명하지 못한다고 주장하면서 모음을 식별하는 데 시간적인 그리고 동적인 정보가 이용되는 “Dynamic Specification Model”을 제시하였다[10]. 그러나 Dynamic Specification Model을 포함해서 어떤 모델도 모음 공간의 변이를 모두 적절하게 설명하지는 못한다. 결국 모음 공간의 연구는 인접한 자음으로부터 혹은 인접한 자음으로의 동적인 전이 정보가 안정 구간에 있는 정보를 보완한다고 볼 수 있다.

기존의 연구들은 모음의 안정 구간에서 포먼트 값을 측정하든지 아니면 측정된 포먼트의 평균값을 이용하여 모음 공간을 구성함으로써 포먼트 값들의 발생 빈도에 관한 정보를 제공하지 못했다. 포먼트 값들의 발생 빈도에 관한 정보를 제시한다 하더라도 단순히 겹쳐 놓음으로써 모음의 범주적 구분을 방해하는 요소로 인식한 측면이 있었다. 본 연구에서는 기존의 연구 결과들을 포괄적으로 설명할 수 있고 빈도를 고려하여 모음 공간을 표현할 수 있는 하나의 대안으로 모음의

지형도를 제시하고자 한다. 모음의 목표 지점이 성별과 연령 등의 화자 특성으로 인한 차이, 환경과 맥락으로 인한 차이, 모음의 길이나 종류 등 모음의 고유한 특성으로 인한 차이 등에 따라 달라진다 하더라도 목표 모음은 빈도가 높은 지점에 나타나게 된다. 목표 모음은 모음의 인지 공간에서 빈번히 자극이 가해지는 지점으로 이해할 수 있을 것이다. 모음 공간에 목표 모음의 위치를 표시하는 것은 바둑판에 바둑돌을 놓는 것에 비유할 수 있고 모음 공간에 포먼트의 발생 빈도를 표시하는 것은 바둑판의 교차점에 칩을 쌓는 것에 비유할 수 있다.

본 논문에서는 말레이시아어와 인도네시아어를 대상으로 F1, F2를 측정하고 포먼트 차트의 단위 영역에 나타나는 포먼트 값의 빈도를 하나의 차원으로 설정하여 모음 공간을 3차원으로 제시한다. 즉 모음 공간을 평면으로만 보는 것이 아니라 모음의 사용 빈도를 고려하여 모음 공간을 입체적으로 볼 수 있는 지형도를 제시한다. 이 지형도에는 포먼트 값이 발생 빈도에 따라 산과 평원으로 나타날 것이며 산은 다시 여러 개의 산꼭대기와 능선 계곡 등으로 나타날 것으로 예상된다. 산꼭대기는 모음의 목표 지점에 해당하며 능선은 인접한 자음의 조음 위치에 관한 정보를 제공할 것으로 예상된다. 이러한 지형도는 모음의 안정 구간을 찾아 포먼트를 측정하거나 측정된 포먼트의 평균값을 구해서 모음 공간을 구성할 필요가 없다. 즉 인접한 자음의 동시조음적 특성이 나타나는 모음의 시작과 끝을 포함해서 모음 구간 내의 모든 지점에서 일정한 간격으로 포먼트 값을 측정하여 그 값들을 모두 모음 공간에 나타낼 수 있다는 장점이 있다.

말레이시아어 및 인도네시아어 모음의 음향학적 특성은 종합적이고 체계적인 방법으로 분석된 적이 없다. 본 논문은 말레이시아어 및 인도네시아어 모음의 음향학적 특성을 종합적이고 체계적으로 규명한다. 본 연구의 결과는 말레이시아어 및 인도네시아어 연구의 기본 자료로 사용될 수 있으며 말레이시아어와 인도네시아어를 전공하는 한국 학생들과 한국어를 배우고 있는 말레이시아 및 인도네시아 학생들의 음성 교육을 위한 기초 자료가 될 것이라는 점에서 의의가 있다. 본 논문은 F1과 F2의 모음 평면에서 발생빈도를 하나의 차원으로 설정하여 모음의 지형도를 살펴봄으로써 모음의 인지적 공간을 새롭게 조명해 볼 수 있는 계기를 제공한다. 또한 모음의 지형도는 모음의 시작과 끝을 포함한 전 구간에서 포먼트를 측정함으로써 인접한 자음의 발생유형, 조음 위치, 조음 방법 등에 관한 정보가 지형적 특성으로 설명될 수 있다는 점에서도 의의가 있다.

2. 연구 방법

말레이시아어는 말레이시아의 국어이며 인도네시아어는 인도네시아의 국어이다. 말레이시아어와 인도네시아어 모두 표준 말레이어의 한 변종이다. 두 언어 모

두 로마자를 이용하여 언어를 기록하고 있으며 음운 목록도 차이가 없다.

말레이시아어는 수도 쿠알라 룸푸르(Kuala Lumpur)가 소재하고 있는 슬랑오르(Selangor) 지방에서 출생하고 자란 화자들을 선정하여 녹음하였으며 인도네시아어는 수도 자카르타(Jakarta)가 있는 자바(Java 혹은 Jawa) 섬 출신 화자들을 선정하여 인도네시아어를 녹음하였다. 화자들은 2005년 9월 현재 서울 및 경기 소재의 대학에 유학하고 있는 학생들이었으며 본인과 부모의 출생지 및 거주 기간 등도 포함된 설문 조사를 실시하여 각각 15명의 화자를 선정하였다. 화자들에 관한 정보는 아래의 <표 1>에 제시되어 있다.

<표 1> 말레이시아어 및 인도네시아어 화자 정보

화자	생년	출생 도시	출생 주	국적
KHA	1983	Kajang	Selangor	Malaysia
FAD	1983	Kuala Lumpur	Kuala Lumpur	Malaysia
IZZ	1985	Kuala Lumpur	Selangor	Malaysia
NOO	1987	Kajang	Selangor	Malaysia
FAI	1985	Kuala Lumpur	Kuala Lumpur	Malaysia
FAK	1984	Petaling Jaya	Selangor	Malaysia
ADB	1987	Klang	Selangor	Malaysia
SHA	1984	Ampang	Selangor	Malaysia
HEL	1984	Kajang	Selangor	Malaysia
AFF	1984	Kuala Lumpur	Kuala Lumpur	Malaysia
NAS	1987	T Karang	Selangor	Malaysia
SHL	1984	Klang	Selangor	Malaysia
ADI	1984	Kuala Lumpur	Kuala Lumpur	Malaysia
WAZ	1983	Sungai Buloh	Selangor	Malaysia
HAN	1974	Kuala Lumpur	W. Persekutuan	Malaysia
HAR	1981	Bandung	Jawa	Indonesia
WAR	1978	Jakarta	Jawa	Indonesia
HAZ	1967	Jakarta	Jawa	Indonesia
AZI	1982	Jakarta	Jawa	Indonesia
SAP	1981	Jakarta	Jawa	Indonesia
IRF	1982	Jakarta	Jawa	Indonesia
DEN	1976	Jakarta	Jawa	Indonesia
SET	1975	Kediri	Jawa	Indonesia
YUH	1971	Maglang	Jawa	Indonesia
ARI	1981	Surabaya	Jawa	Indonesia
BEN	1979	Tangerang	Jawa	Indonesia
HER	1969	Yogyakarta	Jawa	Indonesia
JEL	1975	Yogyakarta	Jawa	Indonesia
UTA	1972	Surabaya	Jawa	Indonesia
ROS	1980	Banyuwangi	Jawa Timur	Indonesia

본 연구에서는 발화의 처음과 중간 그리고 마지막에서 실현되는 모음을 얻기 위하여 /CVda/와 /bVC/ 두 가지 형태의 단어를 사용하였다. /CVda/는 발화의 처음과 중간에 사용하였으며 /bVC/는 발화의 마지막에 사용하였다. C는 말레이어의 자음이며 V는 말레이어 단순 모음이다. 어두와 어말에서 실현되는 모든 자모의 조합을 포함시키기 위하여 C와 V에 자음과 모음을 번갈아 넣었다. 말레이시아어와 인도네시아어에서 어두에서는 25 개의 자음이 모두 나타나며 어말에서는 /ŋ, w, j, v, ʃ, q/를 제외한 19 개만이 나타난다. 본 연구에서 사용된 토큰 중 일부는 실제 단어이고 일부는 무의미 단어다. 본 연구에서 사용된 실험 문장과 말레이시아어 및 인도네시아어의 자음과 모음은 다음과 같다.

#CV: _____ yang saya kata. “내가 말하는 것은 “_____”다”

VCV: kata _____ itu. “그 ‘_____’를 말해”

VC#: coba kata _____ . “_____’를 말해 봐.”

C: /p, b, t, d, k, g, s, h, ʃ, ʈ, m, n, ŋ, ŋ, l, r, w, j, f, v, z, ʃ, x, ʁ, q/

V: /i, e, a, o, u, ə/

녹음된 실험 문장은 모두 로마자로 씌어졌으며 녹음된 문장의 수는 2개 언어 * 15 화자 * 414개 단어(발화의 시작과 중간: 25 자음 * 6 모음; 발화의 마지막: 19 자음 * 6 모음)로 총 12,420 문장이었다. 녹음은 한국외국어대학교 언어연구소의 녹음실에서 진행되었다. 화자의 음성은 AUDIO TECHNICA ATM 75 Headset을 이용하여 녹음하였으며 TEAC사의 TASCAM US 122 Audio Interface를 이용하여 녹음 즉시 컴퓨터 파일로 저장하였다. 녹음된 음성의 표본추출률을 44,100 Hz였으며 양자화 비트는 16이었다.

녹음된 화자의 모음은 Praat 4.4.13의 TextGrid를 이용하여 모음의 시작과 끝을 정하여 분절하였다. 모음의 시작과 끝은 파형과 역동영역(dynamic range)을 35 dB로 정한 스펙트로그램을 참조하여 정하였다. 모음의 시작점은 성대의 진동이 시작된 후 스펙트로그램에서 포먼트 값들(F1, F2, F3, F4)이 뚜렷이 보이기 시작하는 시점으로 하였고 모음의 끝은 파형에서 진폭이 현저히 줄어들고 스펙트로그램에서 포먼트 값들(F1, F2, F3, F4)이 모두 보이는 마지막 시점으로 하였다. 모음을 경계가 설정된 TextGrid 파일로부터 추출한 다음 Formants & LPC: To Formant(Burg)...를 이용하여 포먼트 개체로 변환하였다. 각 프레임은 20 ms의 Gaussian 윈도우를 사용하였고 10 ms 간격으로 계산하였다. 최대 포먼트 수는 5로 하였으며 최대 포먼트 값은 남자의 경우에 일반적으로 사용하는 5,000 Hz로 하였다. 모음의 길이에 따라 프레임의 수가 달리 나타났다. 포먼트 값이 나타난 총 프레임의 수는 85,675였다. 따라서 모음당 평균 프레임의 수는 약 7개였으며 (85,675/12,420=6.89) 모음의 평균 길이는 프레임 수에 10 ms를 더한 값, 즉 약 80

ms였다.

모음의 F1과 F2는 Hz 단위로 측정된 후 mel, bark, ERB의 수 등으로 변환하였다. 그 이유는 mel, bark, ERB의 수 등의 단위들이 주파수의 인지적 특성을 반영하고 있기 때문이며 말레이시아어와 인도네시아어 모음 공간의 인지적 특성을 볼 수 있기 때문이다. 주파수를 mel 척도로 변환하기 위하여 공식 (1)을 이용하였다 [6].

$$\left(\frac{1000}{\log 2}\right) * \log\left(\frac{f}{1000} + 1\right) \quad (1)$$

식 (1)에서 f 는 주파수를 Hz 단위로 나타낸 값이다. 주파수를 bark 척도로 변환하기 위하여 공식 (2)를 이용하였다[22].

$$7 * \ln\left(\frac{f}{650} + \sqrt{1 + \left(\frac{f}{650}\right)^2}\right) \quad (2)$$

식 (2)에서 f 는 주파수를 Hz 단위로 나타낸 값이다. ERB를 계산하기 위하여 공식 (3)을 이용하였다[21].

$$ERB = 24.7\left(4.37 \frac{f}{1000} + 1\right) \quad (3)$$

식 (3)에서 f 는 주파수를 Hz 단위로 나타낸 값이다. ERB는 주파수 특성을 나타내는 방법이기도 하나 모음 공간에서 표현할 때는 ERB의 수(the number of ERB), 즉 E로 나타내는 것이 일반적이다. ERB의 수를 계산하기 위하여 공식 (4)를 이용하였다[21].

$$E = 21.4 * \log_{10}\left(4.37 \frac{f}{1000} + 1\right) \quad (4)$$

식 (4)에서 f 는 주파수를 Hz 단위로 나타낸 값이다.

3. 연구 결과

말레이시아어와 인도네시아어 모음의 환경 및 모음별 포먼트 값이 측정단위별로 아래의 <표 2>와 <표 3>에 제시되어 있다.

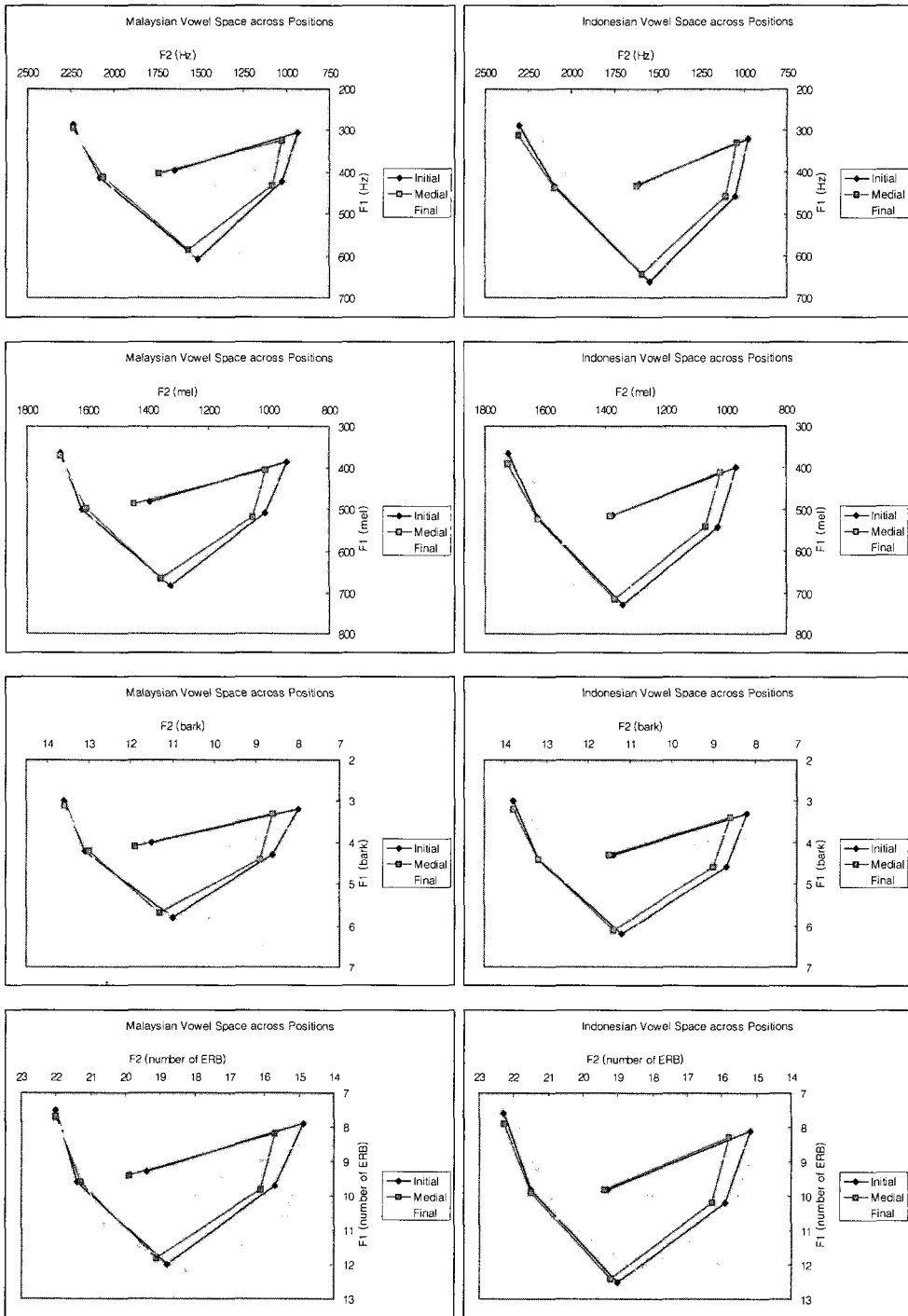
<표 2> 말레이시아어 모음의 환경별 포먼트 평균값.

Position	Vowel	Hz		mel		bark		number of ERB	
		F1	F2	F1	F2	F1	F2	F1	F2
Initial	i	286	2237	363	1690	3	13.6	7.5	22
	e	413	2082	498	1618	4.2	13.1	9.6	21.4
	a	607	1511	683	1324	5.8	11	12	18.8
	o	423	1027	509	1013	4.3	8.6	9.7	15.7
	u	306	933	385	940	3.2	8	7.9	14.9
	ə	395	1649	480	1396	4	11.5	9.3	19.4
Medial	i	294	2237	371	1689	3.1	13.6	7.7	22
	e	412	2060	497	1604	4.2	13	9.6	21.3
	a	586	1568	664	1357	5.7	11.3	11.8	19.1
	o	433	1084	518	1054	4.4	8.9	9.8	16.1
	u	323	1027	403	1009	3.3	8.6	8.2	15.7
	ə	402	1741	486	1446	4.1	11.9	9.4	19.9
Final	i	316	2204	396	1678	3.3	13.5	8	21.9
	e	526	1865	608	1515	5.2	12.4	11	20.5
	a	635	1367	708	1239	6	10.4	12.3	18
	o	488	978	572	979	4.8	8.3	10.6	15.4
	u	339	894	421	916	3.5	7.8	8.4	14.7
	ə	432	1391	517	1252	4.4	10.5	9.8	18.1
Total	i	295	2230	373	1687	3.1	13.6	7.7	22
	e	439	2026	523	1590	4.4	12.9	9.9	21.2
	a	607	1494	683	1313	5.8	10.9	12	18.7
	o	441	1032	526	1017	4.4	8.6	9.9	15.8
	u	318	947	397	952	3.3	8.1	8.1	15.1
	ə	405	1609	490	1373	4.1	11.4	9.4	19.2

<표 3> 인도네시아어 모음의 환경별 포먼트 평균값.

Position	Vowel	Hz		mel		bark		number of ERB	
		F1	F2	F1	F2	F1	F2	F1	F2
Initial	i	289	2302	366	1720	3	13.8	7.6	22.3
	e	434	2096	519	1626	4.4	13.2	9.8	21.5
	a	661	1548	729	1345	6.2	11.2	12.5	19
	o	458	1051	543	1029	4.6	8.7	10.2	15.9
	u	319	976	398	966	3.3	8.2	8.1	15.2
	ə	429	1605	514	1375	4.3	11.4	9.8	19.3
Medial	i	311	2307	390	1723	3.2	13.8	7.9	22.3
	e	438	2097	522	1626	4.4	13.2	9.9	21.5
	a	643	1590	714	1369	6.1	11.4	12.4	19.2
	o	457	1106	542	1068	4.6	9	10.2	16.3
	u	330	1043	411	1020	3.4	8.6	8.3	15.8
	ə	432	1623	516	1385	4.3	11.5	9.8	19.4
Final	i	316	2169	396	1662	3.3	13.4	8.1	21.8
	e	493	1900	577	1534	4.9	12.5	10.6	20.7
	a	627	1434	701	1280	6	10.7	12.2	18.4
	o	487	1037	571	1022	4.8	8.7	10.6	15.8
	u	338	936	419	946	3.5	8	8.4	15
	ə	448	1390	533	1253	4.5	10.5	10	18.1
Total	i	300	2280	378	1711	3.1	13.7	7.8	22.2
	e	448	2054	532	1606	4.5	13	10	21.3
	a	648	1535	718	1338	6.1	11.1	12.4	18.9
	o	464	1063	549	1038	4.6	8.8	10.3	16
	u	325	985	405	976	3.4	8.3	8.2	15.3
	ə	432	1577	517	1359	4.4	11.3	9.8	19.1

각 환경별로 모음 공간에서 차지하는 각 모음의 위치를 살펴보고 모음 공간을 비교할 수 있는 모음 공간이 <그림 1>에 제시되어 있다. <그림 1>에서 좌측에 있는 그림들은 말레이시아어의 모음 공간을 나타내고 오른쪽에 있는 그림들은 인도네시아어의 모음 공간을 나타낸다. 위에서부터 아래로 는 Hz, mel, bark, ERB의 수의 순서로 단위를 달리하여 나타낸 그림들이다. 각각의 그림에서는 발화 위치, 즉 발화의 시작(Initial), 발화의 중간(Medial), 발화의 끝(Final)별로 모음 공간이 그려져 있다.



<그림 1> 말레이시아어와 인도네시아어의 환경별 모음 공간

<그림 1>에 나타나 있듯이 Hz, mel, bark, ERB의 수 등 측정 단위별로는 두드

러진 차이를 발견할 수 없으나 환경별로는 주목할 만한 차이를 발견할 수 있다. 인도네시아어의 /a/를 제외한 모든 모음에 대하여 F1의 평균값은 발화의 시작이나 중간에서보다 발화의 끝에서 더 높았고 F2의 평균값은 발화의 시작이나 중간에서보다 발화의 끝에서 더 낮았다. 즉 발화의 끝에서 목표 모음의 저모음화와 후설모음화가 동시에 관찰되었다. 인도네시아어의 /a/는 F1의 평균값이 다른 두 환경보다 발화의 끝에서 더 낮았으며 F2의 평균값은 다른 모음과 마찬가지로 다른 환경에서보다 발화의 끝에서 더 낮았다. 즉 인도네시아어의 /a/는 후설모음화만 관찰되었다. 두 언어 모두에서 발화의 시작과 발화의 중간 사이에는 F1과 F2의 평균값에 두드러진 차이가 없었다.

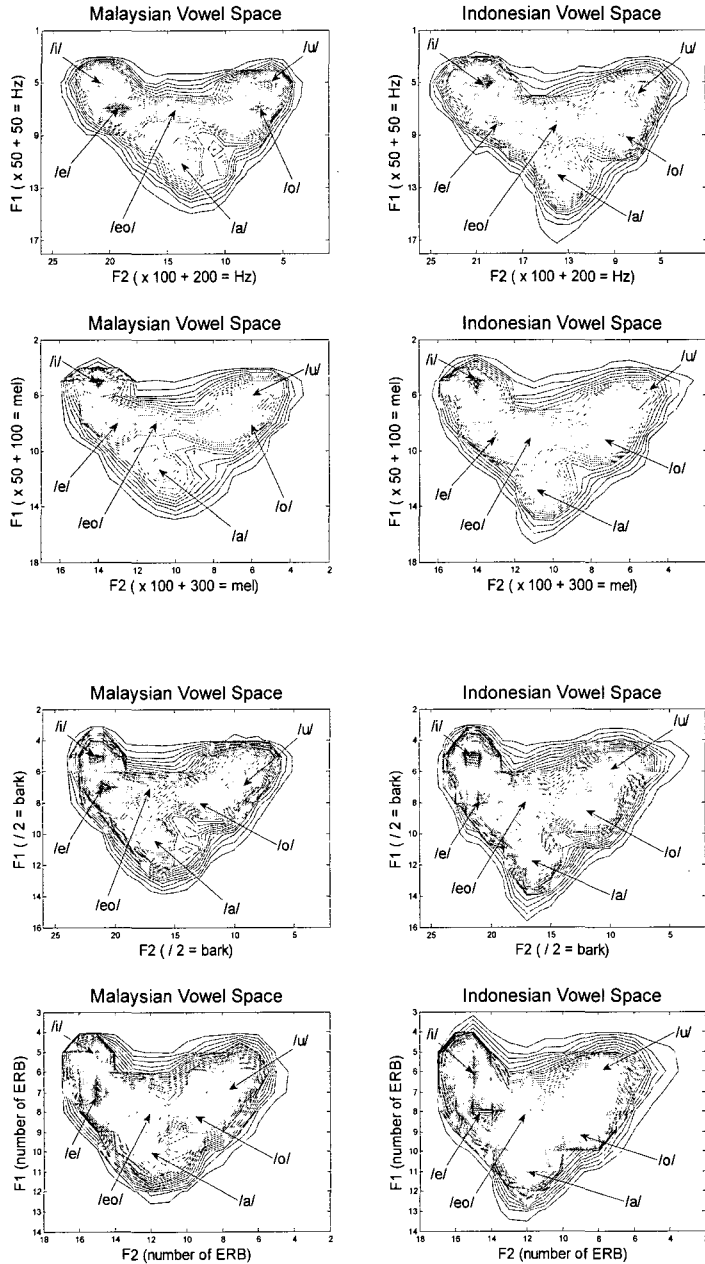
말레이시아어와 인도네시아어에서 발화의 위치에 따른 포먼트 값의 차이가 유의미한지 알아보기 위하여 F1과 F2 각각을 종속 변수로 하고 환경과 모음을 독립 변수로 하여 일원량이원분산분석을 실시하였다. 먼저 말레이시아어를 살펴보면, F1에 대하여 모음 사이에 유의미한 차이가 있었다[F(5, 40427)=27127.506, $p < 0.0001$]. 사후 분석 결과 /e/와 /o/는 유의미한 차이가 없었고 기타 모음들 사이에는 서로 유의미한 차이가 있었다. 환경 사이에도 유의미한 차이가 있었다[F(2, 40427)=3294.413, $p < 0.0001$]. 사후 분석 결과 환경 사이에 서로 유의미한 차이가 있었다. 그리고 모음과 환경 사이에 유의미한 교호작용이 있었다[F(10, 40427)=314.023, $p < 0.0001$]. F2에 대해서는 모음 사이에 유의미한 차이가 있었다[F(5, 40427)=28102.721, $p < 0.0001$]. 사후 분석 결과 모음 사이에 서로 유의미한 차이가 있었다. 환경 사이에도 유의미한 차이가 있었다[F(2, 40427)=1422.448, $p < 0.0001$]. 사후 분석 결과 환경 사이에 서로 유의미한 차이가 있었다. 그리고 모음과 환경 사이에 유의미한 교호작용이 있었다[F(10, 40427)=127.945, $p < 0.0001$].

다음으로 인도네시아어에 대해서 살펴보면, F1에 대해서는 모음 사이에 유의미한 차이가 있었다[F(5, 45197)=22816.168, $p < 0.0001$]. 사후 분석 결과 모음 사이에 서로 유의미한 차이가 있었다. 환경 사이에도 유의미한 차이가 있었다[F(2, 45197)=303.570, $p < 0.0001$]. 사후 분석 결과 환경 사이에 서로 유의미한 차이가 있었다. 그리고 모음과 환경 사이에 유의미한 교호작용이 있었다[F(10, 45197)=157.184, $p < 0.0001$]. F2에 대해서는 모음 사이에 유의미한 차이가 있었다[F(5, 45197)=33241.998, $p < 0.0001$]. 사후 분석 결과 모음 사이에 서로 유의미한 차이가 있었다. 환경 사이에도 유의미한 차이가 있었다[F(2, 45197)=1264.598, $p < 0.0001$]. 사후 분석 결과 환경 사이에 서로 유의미한 차이가 있었다. 그리고 모음과 환경 사이에 유의미한 교호작용이 있었다[F(10, 45197)=77.412, $p < 0.0001$].

저모음화와 후설모음화가 나타난 것은 발화 환경에서 사용한 토큰의 차이, 특히 모음에 인접한 자음의 영향으로 인해 발화의 끝에서 F1의 평균값이 상대적으로 높고 F2의 평균값이 상대적으로 낮기 때문으로 보인다. 발화의 시작과 중간에서 사용한 토큰은 CVda로서 동일하지만 발화의 끝에서 사용한 토큰은 bVC로서

발화의 시작이나 중간에서 사용한 형태와 다르다. CVda 형태에서는 C에 25개의 자음 모두가 나타나는 반면 발화의 끝에서 사용한 bVC 형태에서는 C에 /n, w, j, v, ʃ, q/ 등 6개의 자음이 나타나지 않고 19개의 자음만 나타난다. 이들 6개의 자음이 후행 모음으로 이어질 때 나타나는 포먼트 전이를 살펴보면 다른 자음으로부터 시작되는 포먼트 전이에 비해 F1의 값이 상대적으로 낮다. 반면에 F2는 /w, v/를 제외하고는 모두 상대적으로 높다. 이러한 자음으로부터 시작되는 포먼트 전이로 인해 F1의 평균값이 낮아지고 F2의 평균값이 높아져 발화의 시작과 중간에서 상대적으로 고모음화하고 전설모음화한 반면 발화의 끝에서는 상대적으로 저모음화하고 후설모음화했는지도 모른다. CVda와 bCV에서 나타나는 /d/와 /b/의 차이도 발화의 끝에서 나타나는 저모음화와 후설모음화에 기여했는지 모른다. bVC에서 처음에 나타나는 자음 /b/는 F1이 낮고 F2의 로커스가 높은 반면에 CVda의 끝에서 나타나는 /d/는 F1은 /b/와 차이가 없지만 F2의 로커스는 /b/에 비해 상대적으로 높다. 이로 인해 포먼트 전이 과정에서 나타나는 F2의 포먼트 평균값이 상대적으로 더 높아져 발화의 시작과 중간에서 상대적 고모음화와 전설모음화가 나타나고 발화의 끝에서는 상대적 저모음화와 후설모음화가 나타났는지도 모른다. 한편 발화의 끝에서 사용된 실험 문장이 F1과 F2의 값이 더 높게 실현되도록 하는 운율적 특징을 가지고 있는지도 모른다. 그러나 발화의 시작과 중간에서 나타난 포먼트 값들이 사용한 실험 문장이 다름에도 불구하고 별 차이를 보이지 않는다는 점을 고려할 때 큰 영향을 끼쳤다고 보기는 어려운 것으로 보인다.

말레이시아어와 인도네시아어 모음 전체를 대상으로 Hz, mel, bark, ERB의 수 등의 단위로 나타낸 <F1, F2> 값의 발생 빈도를 고려한 등고선 그림이 <그림 2>에 제시되어 있다. <그림 2>에서 X 축은 F2를 나타내고 Y 축은 F1을 나타낸다. <그림 2>에서 좌측의 그림은 말레이시아어의 모음 공간을 나타내고 우측의 그림은 인도네시아어의 모음 공간을 나타낸다. 각각의 그림에서 수평 축은 F2를 나타내고 수직 축은 F1을 나타낸다. 위에서부터 아래로 Hz, mel, bark, ERB의 수로 나타낸 등고선 그림이 제시되어 있다. <그림 2>에서 Hz, mel, bark, ERB의 수 등의 척도로 나타낸 그림 모두에서 일정한 간격으로 구간을 나누어 값을 재조정하였다. Hz와 mel 단위의 경우 F1은 50 Hz와 50 mel 단위로 구간의 대표값을 이용하였으며 F2는 100 Hz와 100 mel 단위로 구간의 대표값을 이용하였다. 예를 들면, F1의 경우 225 Hz부터 275 Hz까지를 250 Hz로 나타내었으며 F2의 경우 350 Hz부터 450 Hz까지를 400 Hz로 조정하여 나타내었다. bark의 경우 F1과 F2 값을 2로 나누어 반올림하여 정수로 나타내었다. 예를 들면 3.5 bark를 2로 나누면 1.75이고 이것을 다시 반올림하면 2가 된다. ERB의 수의 경우 F1과 F2 값을 반올림하여 정수로 나타내었다. 이러한 방법으로 F1과 F2 값을 구간별로 조정한 다음 모두 85,675 프레임의 F2와 F1 값을 X 축과 Y 축에 각각 배치한 다음 <F1,F2>의 빈도를 MATLAB을 이용하여 64단계의 등고선 척도로 나타내었다.



<그림 2> <F1,F2> 값의 발생 빈도를 고려한 말레이시아어와 인도네시아어 모음 공간에 대한 등고선 그림. Hz, mel, bark, ERB의 수 등 각각 다른 단위를 사용하였다. /e/는 /ɛ/를 나타낸다.

<그림 2>에서 등고선이 밝은 색으로 표시된 부분은 모음 포먼트 값의 빈도가

높음을 나타낸다. 이 부분을 모음 공간에서 나타나는 모음의 목표 지점으로 볼 수 있다. 모음의 목표 지점은 텍스트 화살표로 나타나 있고 /eo/는 /ə/를 나타낸다. 말레이시아어와 인도네시아어 모두에서 /i, e, o, u/는 목표 지점이 밝은 색으로 뚜렷이 나타나지만 /a, ə/는 상대적으로 밝은 색이 뚜렷하지 못하다. 인도네시아어 모음 /a/의 목표 지점이 말레이시아어 모음 /a/의 목표 지점보다 아래로 더 치우쳐 있어 인도네시아어의 모음 공간이 말레이시아어의 모음 공간보다 다소 넓게 나타난다. Hz 척도로 나타낸 그림에서는 /i/와 /e/ 사이의 거리가 좁지만 mel 척도로 나타낸 그림에서는 /i/와 /e/ 사이의 거리가 다소 넓게 나타난다. 이것은 /i/와 /e/의 차이가 음향적 공간에서보다 인지 공간에서 더 쉽게 구별될 수 있음을 의미한다.

4. 결론

본 연구에서는 말레이시아어와 인도네시아어를 대상으로 모음의 F1과 F2 값의 환경별 차이를 살펴보고 <F1,F2> 값의 빈도를 하나의 차원으로 설정하여 모음 공간을 3차원으로 구성하였다. 그 결과 발화의 끝에서 F1의 평균값은 다른 환경에서보다 더 높았고 F2의 평균값은 다른 환경에서보다 더 낮았다. 모음의 <F1,F2> 값의 빈도를 고려한 등고선 그림에서 모음의 목표 지점이 뚜렷이 나타났으며 mel 척도를 사용한 결과 인접한 모음간의 차이가 다소 더 크다는 것을 확인하였다. 본 연구를 통해 인도네시아어의 모음 공간이 말레이시아어의 모음 공간보다 다소 넓다는 점도 확인할 수 있었다.

본 연구의 시도는 음성 생성과 음성 인지의 연구 그리고 음성 교육에 시사하는 바가 있다. 음성 생성의 연구 측면에서는 Simple Target Model이나 Elaborated Target Model 등에서 소규모 자료를 이용하여 조음상의 차이를 설명함으로써 정규화(Normalization)의 문제를 다루었지만 본 논문에서는 상대적으로 대규모 자료를 이용하여 빈도라는 측면에서 살펴봄으로써 정규화의 문제를 바라보는 새로운 시각을 얻었다. 즉 대규모의 자료를 통하여 빈도가 높은 곳에 목표 모음이 나타난다는 것이다. 음성 인지의 연구 측면에서 살펴보면 오랜 세월을 통한 언어 습득 과정에서 다양한 언어 입력이 있음에도 불구하고 모음의 목표 지점이 형성되는 것은 모음 공간의 특정 위치에 높은 빈도가 형성되고 그 지점이 바로 인지 공간에서 목표 지점이 됨으로써 인지상의 혼란이 사라진다고 볼 수 있다. 그리고 빈도가 높을수록 목표 지점의 고착화가 일어나 웬만한 양의 새로운 인지 자극이 없으면 기존의 목표 지점과 다른 새로운 목표 지점을 모음 공간에 구축하기 힘들다고 볼 수 있다. 이런 시각이 음성 교육에 시사하는 바는 분명하다. 어떤 화자가 새로운 방언이나 외국어를 습득할 때 기존에 노출되어 있던 방언이나 모국어에 모음의 목표 지점이 뚜렷이 형성되어 있으면 빈도가 높은 새로운 목표 지점이 생기기 어

렵다. 비록 새로운 목표 지점이 생긴다 하더라도 꾸준히 빈도를 유지하지 않으면 오래지 않아 그 새로운 목표 지점이 기존의 목표 지점에 파묻히는 결과를 초래한다. 따라서 언어 습득에서 가장 중요한 것은 화자가 배우고자 하는 방언이나 언어에 꾸준히 노출됨으로써 새로운 목표 지점을 형성하도록 노력해야 하며 새롭게 목표 지점이 형성되면 빈도를 일정한 수준으로 유지하여 기존의 목표지점에 파묻히지 않도록 노력해야 한다.

본 연구에서 언어별로 전체 모음의 지형도는 흥미로운 특성을 보여주지만 동시조음의 특성이나 환경과 맥락에 따른 지형도는 제시하지 못하였다. 동시조음의 특성이나 환경과 맥락에 따른 의미 있는 지형도를 보기 위해서는 각 모음별로 정밀한 데이터 처리가 필요한 것으로 생각된다. 이런 특성을 반영한 세분화된 지형도는 차후의 과제로 남겨 둔다.

참 고 문 헌

- [1] K. N. Stevens, *Acoustic Phonetics*, Cambridge, MIT Press, 1999.
- [2] T. Chiba and M. Kajiyama, *The vowel: Its Nature and Structure*, Tokyo, Tokyo- Kaiseikan, 1941.
- [3] G. Peterson and H. Barney, "Control methods used in a study of vowels", *Journal of the Acoustical Society of America*, Vol. 24, pp. 175-184, 1952.
- [4] G. Fant, *Acoustic Theory of Speech Production*, The Hague, Mouton, 1960.
- [5] K. N. Stevens, "The quantal nature of speech: Evidence from articulatory-acoustic data", In P. B. Denes and E. E. David Jr. (eds) *Human communication: A unified view*, New York, McGraw Hill, 51-66, 1972.
- [6] G. Fant, *Speech Sounds and Features*, Cambridge, MIT Press, 1973.
- [7] A. K. Syrdal and H. S. Gopal, "A perceptual model of vowel recognition based on the auditory representation of American English vowels", *Journal of the Acoustical Society of America*, Vol. 79, pp. 1086-1100, 1986.
- [8] K. N. Stevens, "On the quantal nature of speech", *Journal of Phonetics*, Vol. 17, pp. 3-46, 1989.
- [9] J. Hillenbrand, L. A. Getty, M. J. Clark, and K. Wheeler, "Acoustic characteristics of American English vowels", *Journal of the Acoustical Society of America*, Vol. 97, pp. 3099-3111, 1995.
- [10] W. Strange, "Evolving theories of vowel perception", *Journal of the Acoustical Society of America*, Vol. 85, pp. 2081-2087, 1989.
- [11] T. M. Neary, "Static, dynamic, and relational properties in vowel perception", *Journal of the Acoustical Society of America*, Vol. 85, pp. 2088-2113, 1989.
- [12] J. D. Miller, "Auditory-perceptual interpretation of the vowel", *Journal of the Acoustical Society of America*, Vol. 85, pp. 2114-2134, 1989.

- [13] W. Strange, "Dynamic specification of coarticulated vowels spoken in sentence context", *Journal of the Acoustical Society of America*, Vol. 85, pp. 2135-2153, 1989.
- [14] B. E. F. Lindblom, "Spectrographic study of vowel reduction", *Journal of the Acoustical Society of America*, Vol. 35, pp. 1773-1781, 1963.
- [15] S.-J. Moon and B. E. F. Lindblom, "Interaction between duration, context, and speaking style in English stressed vowels", *Journal of the Acoustical Society of America*, Vol. 96, pp. 40-55, 1994.
- [16] G. E. Peterson and I. Lehiste, "Duration of syllable nuclei in English", *Journal of the Acoustical Society of America*, Vol. 32, pp. 693-703, 1960.
- [17] I. Lehiste and G. E. Peterson, "Transitions, glides, and diphthongs", *Journal of the Acoustical Society of America*, Vol. 33, pp. 268-277, 1961.
- [18] E. Zwicker and E. Terhardt, "Analytical expressions for critical-band rate and critical bandwidth as a function of frequency", *Journal of the Acoustical Society of America*, Vol. 68, pp. 1523-1525, 1980.
- [19] W. Koenig, "A new scale for acoustic measurements", *Bell Laboratories Record*, Vol. 27, pp. 299-301, 1949.
- [20] B. R. Glasberg and B. C. J. Moore, "Derivation of auditory filter shapes from notched-noise data", *Journal of Speech and Hearing Research*, Vol. 47, pp. 107-138, 1990.
- [21] B. C. J. Moore, "Aspects of Auditory Processing Related to Speech Perception", In W. Hardcastle and J. Laver (eds) *The Handbook of Phonetic Sciences*, pp. 539-565, 1995.
- [22] P. Boersma and D. Weenink, *Praat: Doing Phonetics by Computer* (Version 4.4.19) [Computer program]. Retrieved April 28, 2006, from <http://www.praat.org/>, 2006.

접수일자: 2006년 5월 15일

게재결정: 2006년 6월 22일

▶ 박한상(Hansang Park) : 교신저자

주소: 121-791 서울시 마포구 상수동 72-1 홍익대학교

소속: 홍익대학교 영어교육과

전화: 02) 320-1867

E-mail: phans@hongik.ac.kr

▶ 박정숙(Jeongsook Park)

주소: 130-791 서울특별시 동대문구 이문동 270번지 한국외국어대학교

소속: 한국외국어대학교 동양어대학 말레이-인도네시아어과

전화: 02) 2173-3076

E-mail: parkhami@hufs.ac.kr