

한글 단모음의 포만트 분석과 성도내의 공명효과에 관한 연구

A Study on the Formant Analysis of Korean Monophthongs and their Resonance Effect in Vocal Tract

*신 현 재(Shin, H. J.)

**윤 석 왕(Yeun, S. W.)

요 약

한글 단모음의 음향특성을 기본진동수와 배진동수를 고려하여 포만트 분석하였으며, 성도내의 공명현상과 포만트진동수와의 상관관계를 고찰하였다.

성악을 전공하는 남성으로 하여금 한글 단모음 12개를 5개의 기본진동수에 맞추어 3초동안 발음하게 하여 FFT스펙트럼 분석기를 통해 진동수 스펙트럼을 얻었다.

포만트 분석에 의해 제 1 포만트는 인두강, 그리고 제 2 포만트는 구강의 공명효과에 의함을 밝혔고, 원순화가 일어나므로서 제 2 포만트 진동수가 낮아짐을 발견하였다. 제 1 포만트와 제 2 포만트 진동수로는 “어”의 [ə]와 [ʌ], “이”의 [a]와 [i], 그리고 “에”와 “애”의 음향학적 차이를 뚜렷이 구분짓기는 어려웠다.

ABSTRACT

Twelve Korean monophthongs were studied by formant analysis, fundamental frequencies and their harmonics were considered as the parameters of analysis. The analyzed data were twelve Korean monophthongs which were pronounced with the five fundamental frequencies by the five male vocal musicians. The study shows that the first and the second formants are characterized by the resonance of the cavities of pharynx and mouth, respectively. The lip rounding effect decreases the second formant frequency. The phonemes of [a]/[i], [e]/[ɛ] and [ə]/[ʌ] were not distinguished well in this formant analysis.

*,** 성균관대학교 물리학과 음향학 연구실

I. 서론

한글은 한정된 수의 음소로 구성되어진 표음문자이므로, 각 음소가 규칙적으로 배열되어 단어나 문장을 만든다. 김영열⁽¹⁾ 등은 한글을 자음의 음소와 모음의 음소로 분리하면 작은 데이터 양으로 처리 시간을 단축시켜 음성을 인식하고 합성할 수 있음을 밝혔다.

국어학자들은 조음방법이나 조음기관의 위치를 고려하는 발생학적 관점에서 한글의 모든 음소를 분류하고 있다. 조음 방법이나 조음기관의 위치는 항상 음성의 음향특성에 영향을 미치므로, 서로 다른 방식으로 발생한 음성은 다른 음향특성을 가지게 된다. 그러므로, 발생학적 관점에서 연구된 내용은 음향학적 관점에서 연구된 내용과 긴밀한 연관성을 가지며, 실질적인 예로서 성도내의 공명특성을 반영하는 포만트(formant)를 들 수 있다.

포만트는 모음을 특징짓는 중요한 요소로 인지되어왔고 김영일⁽¹⁾ 등은 음소를 이용한 한글 단음 인식 실험에서 포만트를 이용하여 97.6%로 모음을 인식하였다. 모음이 발생될 때는 혀파로부터 매초 40~200cm³의 공기가 성대를 진동시키며 성대의 진동은 기본진동수를 결정한다.⁽²⁾ 사람의 말이 기계소리와는 달리 부드럽게 들리는 것은, 사람의 말속에는 기본진동수의 변화가 있기 때문이다. 따라서 음성의 인식 및 합성에 있어서 기본진동수의 변화가 고려되어야만 한다.

본 연구에서는, 진동수 스펙트럼에서 기본진동수와 배진동수를 고려한 새로운 포만트 진동수 측정법을 소개하였다. 그리고, 한글에서 발음되는 12개 단모음의 제 1 포만트 진동수와 제 2 포만트 진동수를 측정하였으며, 성도내의 공명장에서 공명효과와 포만트 진동수와의 상관관계를 밝혔다. 또한, 한글 단

모음의 발생학적 연구 내용과 음향학적 실험결과를 비교하여, 한글 단모음에 대한 음향학적 분류를 시도 하였다.

II. 실험 및 분석

(1) 자료수집 및 분석방법

성악을 전공하는 남자(20~24세) 7명으로 하여금 한글에서 다르게 발음될 수 있는 12개의 단모음(표 1)⁽³⁾을 5개의 기본진동수에서 3초 동안 지속 발음하게 하였다. 소음이 없는 자료를 얻기 위해 방음 처리가된 녹음실에서 녹음(KUDELSKI:Nagra4.2)하였고, 화자와 마이크(Beyer Dynamic:M88N) 사이의 거리는 1m를 유지하였다. 정확한 기본진동수에서 발음하도록 하기 위해 모든 화자로 하여금 미리 녹음된 피아노음을 헤드폰을 통해 들으면서 발음하게 하였다. 녹음된 자료에서 기본진동수의 정확성과 화자의 음성 상태를 검토하여, 7명의 화자중 5명을 선택하였으며, 한개 모음당 분석한 자료의 수는 25개로, 12개 모음에 대해 총300개의 자료를 분석하였다.

토통 남자의 기본진동수 영역이 100~150Hz인 점을 고려하여⁽⁴⁾ 피아노 건반의 G(98Hz), A(110Hz), C(130Hz), D(147Hz), 그리고 E(165Hz) 음을 택하여 발음하게 하였다.⁽⁵⁾ Nobaki Hiraoka⁽⁶⁾ 등이 사용한 방법으로 측정된 본 연구의 기본진동수들을 [표 2]에서 보는 바와같이 모든 화자들이 정확히 각 단모음들을 발음 하였음을 확인하였다.

녹음된 자료는 발화후 800ms에서 안정된 모음구간을 보였고, 이를 분해능이 12.5Hz인 0~5Hz의 진동수영역으로 스펙트럼 분석기(B&K:Type 2033)를 이용하여 FFT(Fast Fourier Transform) 하였다.⁽⁷⁾ 이때, 샘플링 주파수는 12.8KHz였고, 80ms(1024 point)에 Hanning window를 취하여 실질적인 분석 구간이 30ms가 되게하였다⁽⁸⁾

(2) 포만트 진동수의 측정 및 분석

모음의 진동수 스펙트럼은 [그림 1]과 같이 기본진동수와 배진동수들로 이루어진다. 포만트 진동수를 측정하는 방법은 여러가지로 연구되어지고 있으며,⁽⁹⁾ 본 연구에서는 포만트 진동수 영역에서 배진동수를 고려하여 성도내에서 효율적으로 공명하는 진동수를 잘 나타내는 포만트 진동수 측정법에 대해 고찰 하겠다.

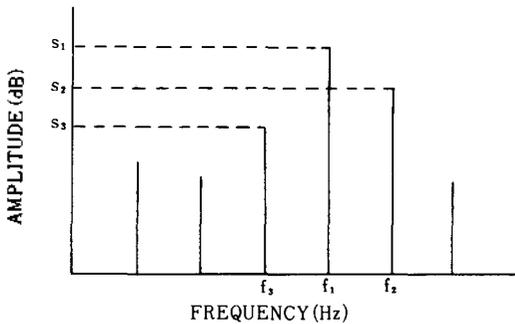


그림 1 모음의 진동수 스펙트럼 (제 1 포만트 영역).

FMDH (Frequency of the Most Prominent Harmonic)는 [그림 1]과 같은 스펙트럼에서 가장큰 진폭을 가진 f_1 을 포만트 진동수로 정의하는 방법이다.⁽¹⁰⁾ 두개의 가장 큰 진폭의 배진동수를 택하여 중가평균을 취하는 FMPH 2 (weighted mean of the two most prominent harmonics) 방법과 세개의 가장 큰 진폭의 배진동수를 택하여 중가평균을 취하는 FMPH 3 (weighted mean of the three most prominent harmonics) 방법은 [그림 1]에서 f_1, f_2 또는 f_1, f_2, f_3 를 측정하여 다음 관계식으로 포만트 진동수를 계산한다.⁽¹¹⁾

$$F_m^N = \frac{\sum_{i=1}^N f_i S_i}{\sum_{i=1}^N S_i} \quad (1)$$

여기서, f_i 는 기본진동수의 배진동수들이고 S_i 는 f_i

의 6dB/oct pre-emphasized amplitude value이다. S_i 는 [그림 1]에 보이바와 같이 S_i 의 극대값이 S_1

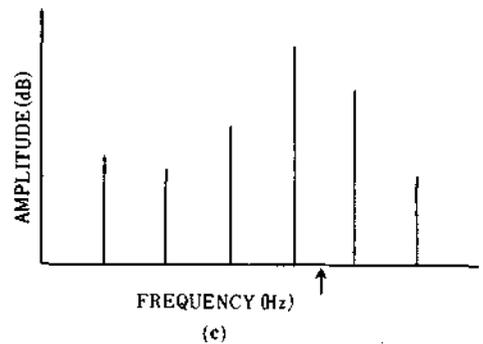
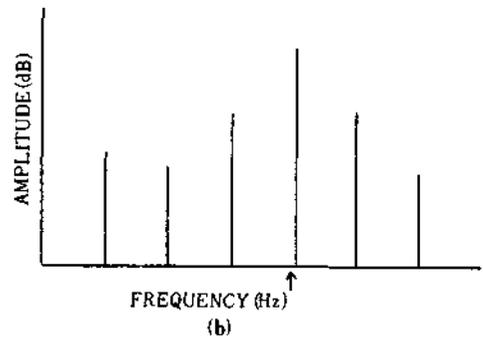
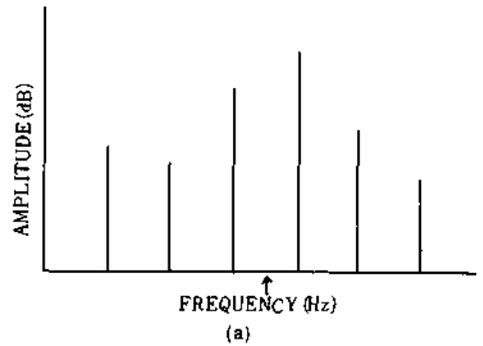


그림 2 포만트 진동수 (↑)와 배진동수와의 관계. 포만트 진동수가 극대진폭의 배진동수 (a)보다 작은 경우, (b)와 같은 경우, (c)보다 큰 경우.

일때 $S_{i-1} > S_i$ 의 관계로 가지도록 i 를 결정한다. 이 때 FMPH2는 $N=2$ 일때 포먼트 진동수 F_{2m} 를 계산하고, FMPH3는 $N=3$ 일때 포먼트 진동수 F_{3m} 를 계산한다.

[그림 2]는 성도내에서 효율적으로 공명될 수 있는 진동수와 배진동수들 사이의 관계를 나타내고 있다. FMPH 2와 FMPH 3로 측정된 포먼트 진동수는 성도내에서 효율적으로 공명할 수 있는 진동수와 배진동수들 중의 하나가 일치하지 않는 경우에도 어느정도 성도의 효율인 공명 진동수에 근접한 값임을 보인다.

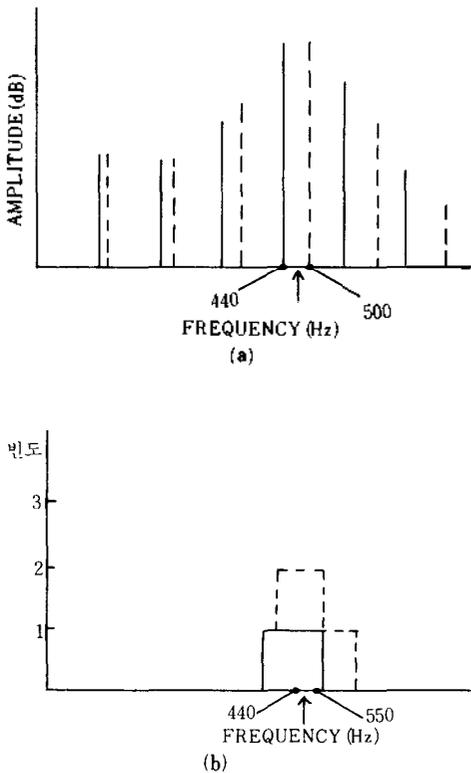


그림 3 기본진동수와 배진동수를 고려한 포먼트 진동수(↑) 측정법.
 (a) 기본진동수가 110Hz(실선), 125Hz(점선) 일때 모음의 진동수 스펙트럼,
 (b) 기본진동수의 폭을 갖는 rectangle 함수이용.

본 연구에서는 5개의 기본진동수에서 FMPH 방법으로 측정된 포먼트 주파수를 다음과 같은 방법으로 통계처리 하여 성도의 효율적인 공명진동수에 근접하도록 하였다. [그림3-(a)]와 같이 성도의 효율적인 공명진동수는 화살표 위치의 470Hz이고, 실선은 기본진동수 F_0 가 110Hz일때 진동수 스펙트럼이며, 점선은 F_0 가 125Hz일때 진동수 스펙트럼이라고 가정하자. FMPH방법으로 각 기본진동수에 대하여 제 1포먼트 진동수를 440Hz와 500Hz로 측정할 수 있다. 이를 [그림3-(b)]와 같이 가로축은 진동수의 축으로 정의된 좌표에 기본진동수의 폭을 갖는 rectangle 함수로 중첩되는 부분은 세로축의 빈도로 표시하였다.

본 연구에서는 한개 모음당[표 2]와 같은 5개의 기본진동수에 대해 5명이 발음한 것을 각 모음에서

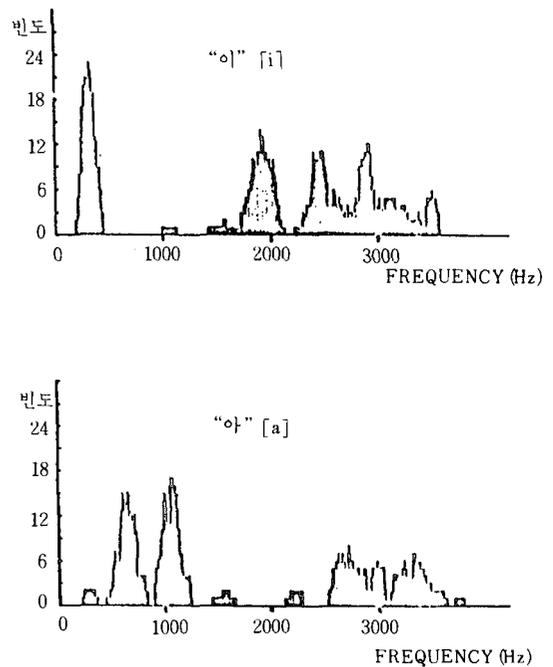


그림 4 "이"와 "아" 모음의 포먼트 진동수 분포.

생기는 3개혹은 4개의 포만트에 대해 전술한 방법으로 통계처리 하였다. “이”와 “아” 모음의 경우에는 (그림 4)와 같은 결과를 얻었다. [그림 4]에 나타난 극대값 들은 낮은 진동수의 것부터 제1포만트 제2포만트 진동수분포에 해당한다.

음성 합성 및 인식에 있어서 제 1 포만트와 제2포만트가 특별히 중요한 요소라는 점을 고려하여, 제 1 포만트와 제 2 포만트 진동수를 각각의 표준편차에 함께 [그림 5]에 제시하였다.

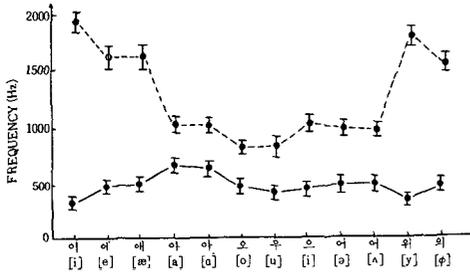


그림 5 제 1 포만트 진동수(실선)와 제 2 포만트 진동수(점선)의 평균과 표준편차.

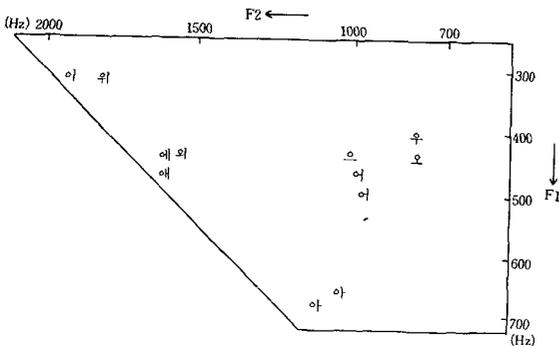


그림 6 F1-F2도표(●; F1, F2진동수).

Ⅲ. 결과 및 고찰

(1) 포만트와 공명강의 관계

각 단모음의 포만트 진동수를 세로축은 제 1 포만트

트 진동수 F_1 , 가로축은 제 2 포만트 진동수 F_2 로 정의된 [그림 6]과 같은 F_1-F_2 도표에 표시하면, 음성학자들이 경험에 의해 얻은 각모음 발음시 혀의 위치를 표시하는 [그림 7]의 모음삼각도와 잘 일치하고 있음을 알 수 있다.¹²⁾ 이는 포만트가 성도내의 공명특성을 반영하고 있기 때문이며, 각 모음 발음시 성도내의 공명강들의 공명현상을 연구 하므로써 설명되어질 수 있다.

성도내에는 [그림 8]에 보인 4개의 중요한 공명강이 있다.¹⁴⁾ 인두강은 혀뿌리와 인두벽 사이에 형성되는 공명강으로서 혀뿌리의 움직임이 인두강의 크

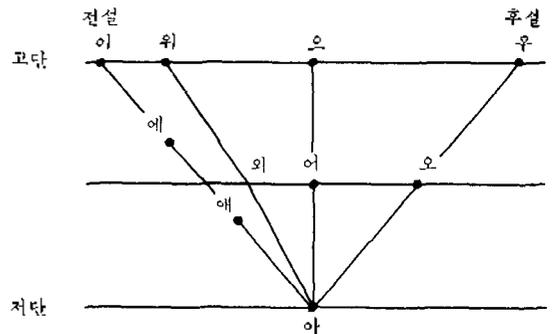


그림 7 모음삼각도.

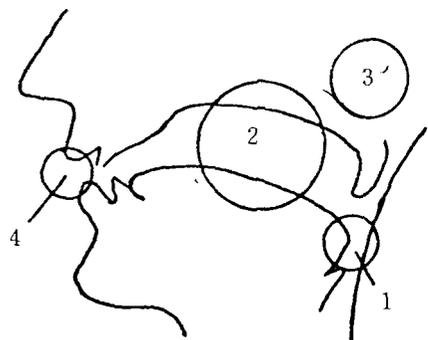


그림 8 성도내의 4개 중요 공명강.

1-인두강, 2-구강, 3-비강, 4-입술의 공명강.

기를 변하게 한다. 구강은 그바닥을 형성하는 혀로 대부분이 가득차 있고, 혀의 운동에 의해서 구강의 모양과 용적은 다양하게 변할 수 있다. 비강은 모양과 용적이 거의 변하지 않으므로 비강의 공명효과는 거의 동일하며 특히, 한글 단모음중에서 비강의 공명은 없는 것으로 알려져 있다. 입술 공명장은 원순모음인 경우에 입술을 내밀어 둥글게 하므로서 만들어지고, 구강의 공명효과를 변화 시킨다.

조음기관들 중 혀는 가장 능동적으로 움직여서 여러가지 음소를 발음하는데 사용되는 모든 공명효과를 얻을 수 있다.

[그림 6]의 F_1-F_2 도표가 혀의 위치를 나타낸 [그림 7]의 모음삼각도와 일치한다는 것은 제 1포먼트 진동수와 제 2포먼트 진동수가 혀의 위치에 따라 변하는 공명장에서 공명되어지는 진동수라는 것을 암시한다. 혀의 위치에 따라 용적이 바뀌는 공명장은 인두강과 구강이며 [그림 9]는 세계의 모음 발음시에 혀의 위치와 인두강과 구강의 용적을 보여 주고 있다.⁴⁾ “아”, [a]의 경우에 혀는 입안에서 거의 평탄해서 휴식상태에 가깝고, “이”, [i]의 경우에 혀의 위치는 고단절설이다. [a]가 가[i] 일때 보다 인두강은 작고, 구강은 크다. 여기서, 큰공명장에서는 낮은 소리가 공명되고, 작은 공명장에서는 높은 소리가 공명되는 물리음향학적 현상을 고려하면, 공명장의 용적 변화와 F_1-F_2 도표에서 진동수 변화는 쉽

게 연관 되어진다. 즉, [a] 일때 보다 [i] 일때 제 1포먼트 진동수가 낮게 나타나고 제 2포먼트 진동수가 높게 나타나는 이유는 인두강은 커졌고 구강은 작아졌기 때문이다. 결과적으로, 제 1 포먼트는 인두강의 공명효과를 반영하고 제 2포먼트는 구강의 공명효과를 반영한다.

한글 단모음 중 입술공명장에서 공명을 일으키는 원순모음은 “우”와 “오”이다. 또한, “위”와 “외”발화시 혀의 위치는 각각 “이”와 “에”발화시 혀의 위치와 동일하고, 다만 원순화(이→위, 에→외)가 되어 입술의 공명을 일으킨다. F_1-F_2 도표에서 “우”와 “오”의 경우에는 다른 모음들 보다 낮은 제 2포먼트 진동수를 가지고 있으며, “위”와 “외”의 경우는 “이”와 “에”보다 각각 낮은 제 2포먼트 진동수를 가지나, 제 1포먼트 진동수는 거의 같다. 따라서, 입술 공명장은 제 2포먼트 진동수에 변화를 주고있음을 알 수 있다. 제 2포먼트의 공명이 구강에서 형성된다는 앞의 결과와 종합하면, 원순모음의 경우 입술의 공명장이 구강에 합쳐져 좀 더 큰 공명장을 만들어 줌으로서, 제 2포먼트 진동수의 감소를 초래한다고 할 수 있다.

(2) 한글 단모음의 음향학적 고찰

국어학자들은 [표 1]에 주어진 12개의 단모음중 “어” 발음의 [ə]와 [ʌ]를 묶어 [ə]로 표기하고, “아”

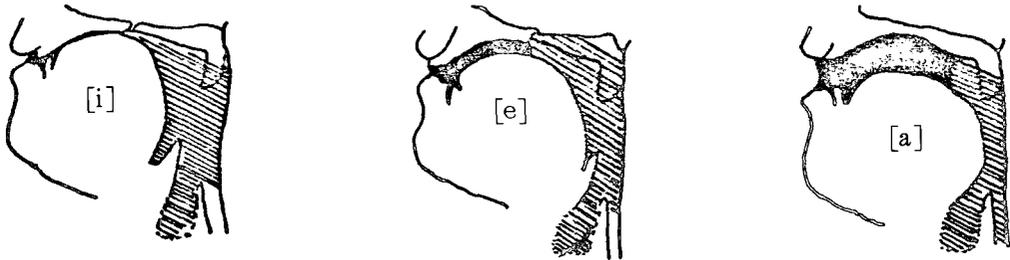


그림 9 혀의 위치와 구강 및 인두강의 용적 관계.

이, [i]	아, [a]	우, [u]	어, [ʌ]
에, [e]	야, [a]	으, [i]	위, [y]
애, [ɛ]	오, [o]	어, [e]	외, [ɸ]

표 1 한글 단모음 12개
[]안은 I.P.A(International Phonetic Alphabet)
표기

기본진동수(F ₀)	99.0	110.6	132.2	147.9	165.4
표준편차	1.16	0.87	1.45	1.44	1.73

표 2 발음된 한글 단모음의 평균과 표준편차
한개 기본진동수당 처리된 자료의 수;
60개(5명×12모음)

발음의 [a]와 [ʌ]를 묶어 [a]로 표기한다. 한글에서 이 음소들의 차이를 [표 3]에 나타내었다.³⁾

먼저, [ə]와 [ʌ]에 대해 고찰하면, 한글에서 “어”는 연결되어지는 음소에 따라 [ə]나 [ʌ]중의 하나로 발음되어 지며, 이들은 서로 배타적인 결합변이음이다. 제 1포먼트 진동수에 있어서는 [ə]가 [ʌ]보다 13Hz 낮은 값을 가지고, 제 2포먼트 진동수에서는 17Hz 높은 값을 가진다. (그림 6)의 F₁-F₂ 도표에서 이들 두 모음은, 비슷한 혀의 위치에서 발음되고 있음을 알 수 있다. 그러나, 두 모음이 단어 속에서 장음과 단음에 대응하여 구분되어 지므로, 단어 속에서 두모음의 지속시간을 측정하여야만 음향학적 차이를 얻을 수 있으리라 사료된다.

“아”에 대한 국어학자들의 견해는 [표 3]에서 보는 바와 같이 구분하는 사람도 있으나, 일반적으로 변이음을 가지지 않는다고 보는 경향도 있다. [a]는 [a]보다 제 1포먼트 진동수에서 30Hz 높은 값을 가지고, 제 2포먼트 진동수에 있어서는 54Hz 높은 값을 가진다. F₁-F₂도표에서 이 두개의 모음은 거의 동일한 혀의 위치에서 조음되고 있음을 보인다.

이상의 모음 이외에도 “에”와 “애”는 현 세대에서 구분없이 사용되고 있다는 것이 국어학자들에 의해

한글표기	I.P.A	보 기	혀 의 위 치	비 고
어	ə	없다	“에”, “오”와 같은 높이의 중설	장음
	ʌ	없다	“오”보다 조금 아래의 후설	단음
아	a	없는다	저단 전설	·
	ɑ	안는다	저단 후설	·

표 3 국어학자에 의한 한글단모음 “어”, “아”의 발음상 분류.

기술 되어 진다.³⁾ F₁-F₂도표에서 이 두개의 모음은 (그림 7)의 모음삼각도의 경향과 일치하지 않는 유일한 것들이다. 비록 “에”와 “애”로 표기는 다르지만 음향학적으로 큰 차이를 보이지 않고 있다.

본 연구에서는 모든 발화자의 출신도가 경상도이기 때문에 이상의 결과중 특히, “으”나 “외”와 “애”에서 지역 특색이 나타났으리라 생각된다. 표준말을 사용하는 사람을 대상으로 단어나 문장속에서 이상과 같은 단모음을 분석 하므로써, 한글에서 구분 발음되는 단모음의 표준적인 정보를 얻을 수 있으리라 사료된다.

IV. 결 론

한글 단모음의 포먼트 분석과 성도내의 공명효과에 관한 연구결과 다음 두가지 결론을 얻었다. 첫째, 한글 단모음이 발음될 경우 제 1포먼트는 인두강, 제 2포먼트는 구강의 공명효과에 의함을 밝혔다. 둘째, 원순모음 발음시, 입술의 공명강을 구강에 더하여져 더욱 큰 공명강을 만들므로 제 2포먼트 진동수가 감소함을 확인하였다.

한글 단모음의 음향학적 고찰에서는 첫째, “어”발음의 [ə]와 [ʌ], “아”발음의 [a]와 [ɑ] 그리고 “에”와 “애”는 제 1포먼트 진동수와 제 2포먼트 진동수에 있어서, 본 연구의 음향학적 분석으로는 뚜렷이 구분할 수 없었다. 둘째, “어”발음의 [ə]와 [ʌ]는 단어나 문장중에서 지속시간을 고려한 연구가 수행

되어야 음향학적 구분이 가능할 것이다.

참 고 문 헌

1. 김영일, 조문재, 차일환, "음소를 이용한 한국어의 인식", 한국음향학회지, 3⁽²⁾: 35, (1984)
2. E. MEYER, E.G. NEUMANN, *Physical and Applied Acoustics*, (Academic Press, New York, 1972), p.257.
3. 허용, 국어 음운학, (샘문화사, 1985)
4. B. Malmberg, *La phonetique*, 음성학 오원교 역 (新雅社, 1984)
5. A.D. Pierce, *Acoustics*, (McGraw-Hill Book Co., New York, 1981), p.57.
6. N. Hiraoka, Y. Kitazoe and H. Ueta, "Harmonic-intensity analysis of normal and hoarse voices," *J. Acoust. Soc. Am.* 76(6), 1648, (1984).
7. J.T. Broch, *Mechanical Vibration and Shock Measurements*, (Bruel & Kjaer, 1984), p.160.
8. N. Thrane, *Application of a long Memory FFT Analyzer in Speech Analysis*, Bruel & Kjaer application notes, 066-81.
9. 정의진, 이형준, 김순협, "선형예측 위상 스펙트럼을 이용한 한국어 숫자음의 포만트 추출에 관한 연구", 음성통신 및 처기기술 WORKSHOP 논문집, 18, (1986).
10. P.F. Assmann and T.M. Nearey, "Perception of front vowels; The role of harmonics in the first formant region²⁾", *J. Acoust. Soc. Am.* 81(2), 520, (1987).
11. R. Carlson, G. Fant, and B. Granstrom, "Two-formant models, pitch and vowel perception", in *Auditory Analysis and Perception of Speech*, G. Fant and M.A.A. Tatham, Ed., (Academic Press, London, 1975).
12. P. Ladefoged, *A Course in Phonetics*, 2nd Ed., (Harcourt Brace Jovanovich, New York, 1982).