

한국어 이중모음의 음향학적 연구
An Acoustical Study of Korean Diphthongs

양 병 곤(동의대 영어영문학과)

<ABSTRACT>

The goals of the present study were (1) to collect and analyze sets of fundamental frequency (F0) and formant frequency (F1, F2, F3) data of Korean diphthongs from ten linguistically homogeneous speakers of Korean males, and (2) to make a comparative study of Korean monophthongs and diphthongs.

Various definitions, kinds, and previous studies of diphthongs were examined in the introduction. Procedures for screening subjects to form a linguistically homogeneous group, time point selection and formant determination were explained in the following section.

The principal findings were as follows:

1. Much variation was observed in the ongliding part of diphthongs.
2. F2 values of [j] group descended while those of [w] group ascended.
3. The average duration of diphthongs were about 110 msec, and there was not much variation between speakers and diphthongs.
4. In a comparative study of monophthongs and diphthongs, F1 and F2 values of the same offgliding part at the third time point almost converged.
5. The gliding of diphthongs was very short beginning from the h-noise.

Perceptual studies using speech synthesis are desirable to find major parameters for diphthongs. The results of the present study will be useful in the area of automated speech recognition and computer synthesis of speech.

I. 서론

오늘날 첨단기기의 눈부신 발전과 더불어 음성분석 및 합성분야에서 영어 사용권을 비롯한 각 선진국에서의 언어 연구가 상당한 진척을 보이고 있으나, 한국어 이중모음에 대한 자료는 지금껏 표준자료로 이용할 만큼 충분한 연구대상과 여러 가지 변수를 고려한 정밀한 실험 방법을 통해 이뤄 놓은 것이 부족하며, 특히 기존 연구에서는 방언과 같은 언어학적 요소와 나이 성별 등과 같은 비언어적 변수를 고려치 않아 음성인식 또는 합성 분야에 보편적으로 이용할 수 있는 기초 자료로 쓰기에는 부적합하다. 또한, 이중모음은 모음의 시작부터 끝까지 계속적으로 변하기 때문에 측정에 상당한 어려움이 있다. 그 결과, 조직적인 측정방법을 통해 연구해놓은 기존 자료가 별로 없다. 이중모음에 대한 연구는 기존 단모음의 연구와 함께 보다 간편하고 정확한 음성 인식이나 합성 장치 개발에 필수적이라 할 수 있는 중요한 기초 연구 과제이다.

본 연구는 비언어적 요소를 최대한 제거한 피험자의 발음을 통해 발생된 한국어 이중모음의 형성음을 음향학적으로 연구 조사하여 순수 언어학적인 형성음의 기초자료를 제공하고 이미 완성한 한국어 단모음에 대한 음향학적인 연구와 비교 검토하여 그 관계를 정립함으로써 음성합성 및 자동 음성 인식 장치 개발에 기여함을 목적으로 한다.

허파에서 나온 기류가 성대를 울려서 성도를 통해 걸러 나올 때, 도중에 막히거나 좁혀짐으로써 소음을 내는 자음과는 달리 막힘없이 조음되는 음을 모음이라한다. 모음 중 일정한 조음에 의해 음색이 변하지 않는 것을 단모음(monophthong, 또는 simple vowel)이라하고 모음이 두 개 또는 세 개 결합한 음을 각각 이중모음(diphthong), 삼중모음(triphthong)이라 하며 이것들을 복모음(compound vowel)이라 부른다.

이중모음에 대한 여러가지 정의를 살펴보면, 먼저, 영어학사전(조성식편, 1990)에는 이중모음이란 두 개의 모음을 말하는데 하나의 서서히 변화하는 모음, 즉 전이모음(gliding vowel)으로 보거나 한 개의 완전한 모음(full vowel)과 한 개의 전이음(glide)이 결합되어 있다고 보거나 또는 두 개의 안정된 모음 사이에 빠르게 변하는 전이음으로 되어있다고 본다. 이렇게 학자들마다 다른 정의를 내리게 되는 이유는

조음적, 지각적, 음향적인 기준 또는 이들의 조합에 의해 이중모음을 해석하기 때문이다.

Kenyon (1966:208) 은 이중모음이란 두 개의 소리라는 뜻이지만, 엄격하게 말하면, 한 음절내에서 한 개의 계속 전이하는 모음 ("one continuous gliding vowel sound within one syllable")으로 되어있다고 했으며, Gerhard (1954:21) 는 한 모음의 위치에서 다른 모음의 위치로 전이하는 단음절 모음 ("monosyllabic vowel glide commencing at the position for one vowel and moving toward another")을 말한다고 정의했다. 전상범(1988:225)에 따르면 이중모음이란 "한 음절 안에서 계속적으로 미끄러져 가는 하나의 음"으로 정의하고 [au]의 예를 들어 [a]라는 음에서 시작하여 [u]음을 지향하여 미끄러지듯이 발음하되 [u]음의 위치에 도착하지는 않는다는 것이다. Ladefoged (1982:281) 는 이중모음을 "a vowel in which there is a change in quality during a single syllable."으로 정의하고, 한 음절내에서의 자질적 변화가 있는 모음이라고 했다. 만약에 모음의 음성학적 표기와 조음학적인 발성 몸짓 (vocalic gesture)의 목표가 동일하다면 이중모음은 두 개의 다른 목표모음을 가질 것이다. 이렇게 보면 동일한 두 개의 목표모음을 갖는 장모음과 이중모음을 구분하기가 어렵다. 따라서 우리는 이중모음이란 반드시 두 개의 "다른" 목표모음을 갖는 것으로 규정해야 한다 (Ladefoged 와 Maddieson, 1990:120). 지금까지의 이중모음에 대한 정의에서 공통된 사실은, 이중모음이란 두 개의 변별적인 요소가 있고 그 사이에 빨리 변하는 전이부가 있다는 것이다.

그러면 이러한 이중모음의 정의에 따라 어떠한 종류가 있는지 알아보자. 이중모음은 앞에서 정의한 대로 두 개의 다른 목표음을 가지고 그 사이의 최단 거리를 전이하는 부분으로 되어있는 순정이중모음과 이중모음화된 모음으로 구분할 수 있으며, 조음기관인 혀의 이동 방향과 이동 거리, 음향적 성분소 및 올림도 등과 같은 기준에 따라 여러가지로 나뉘어 진다. 먼저, 혀의 이동 방향에 따라 구분한 것을 살펴보면, Trager와 Smith (1966)의 이중모음에서는 두번째 요소가 반모음 /y/인 이중모음, 즉 /iy, ey, ay, oy/ 는 전향이중모음 (fronting diphthong) 또 /w/의 방향으로 변하는 /uw, ow, aw/는 후향이중모음 (retracting diphthong) 이라한다. 또 /h/의 방향으로 변하는 /ih, eh, ah, əh, əh, uh/ 또는 /r/이 따라오는 /ir, er, ar, ər, ɔr, ur/은 내향이중모음 (centering diphthong)이라한다. 혀의 이동이 모음 사변형을

가로질러 가는 것을 대각이중모음 (diagonal diphthong)이라고 하며 언제나 상향으로 수직이동하는 것은 수직이중모음 (vertical diphthong)이라 부른다. 이들 외에도 모음사변형의 외곽에서 중앙지역으로 이동하는 중앙집중이중모음 (centering diphthong) 등이 있다 (Sloat 등 1978:17-18). 혀의 이동 거리에 따라서는, [ai, ɔi]처럼 혀의 이동거리가 비교적 긴 것을 전이중모음 (full diphthong)이라 부르고 [ei, ou]처럼 이동하는 거리가 짧은 것을 반이중모음 (half diphthong)이라고 부르기도 한다. 한편, 올림도에 따라 분류하면 첫음의 올림도가 뒷부분보다 두드러지게 들리면 하강 이중모음 (falling diphthong), 그 반대인 경우는 상승이중모음 (rising diphthong)이라고 부른다. 이 때 두드러진 부분이 음절의 중핵 (syllable nucleus 또는 syllabic)이 되고 약화된 부분은 음절의 부음(nonsyllabic)이 된다 (김재민 등 1985:355). 영어의 이중모음은 하강 이중모음에 해당한다. 이중모음의 또 하나의 범주로 /w, j/에 모음이 후속되는 예 (we, you)는 후속모음의 올림도가 크기 때문에 상승이중모음에 해당한다. 우리말 /왜,애/ 등이 여기에 해당한다. 다만 이 전이음을 자음으로 규정할 때는 순수한 이중모음이라 할 수 없다 (김재민 등, 1985:356). Lehiste 와 Peterson (1961)은 영어에서 두 가지 복합음절핵 (complex syllable nuclei)을 가진 이중모음을 더 세분하여, 음소적 이중모음(phonemic diphthong)과 비음소적 이중모음(non-phonemic diphthong)으로 구분했다. 음소적 이중모음은 세 가지 음향 성분소로 되어있다. 즉, 형성음 변화가 적은 시작부분과 종결부분의 안정된 상태 부분과 그 사이에서 형성음 값이 변하는 전이 부분이 있다. 이 때 시작부분과 종결부분의 두 개의 목표모음이 존재한다. 한편 한 개의 목표모음을 갖는 비음소적 이중모음은 음절핵 전역에 걸쳐서 형성음 값이 서서히 변하는 것이다. 이와 비슷한 것으로써, /ai, au, oi/와 같은 순정이중모음외에도 영어에는 두 개의 형성음 목표를 갖는 /ey, ow, uw/ 등이 있는데 이는 이중모음화된 모음 또는 종결전이부를 가진 모음 (vowels with offglides)으로 분류되기도 한다. 이들은 순정 이중모음과 지속 시간에 있어서 차이가 나므로 "gliding vowels(전이모음)"이라고 불렀다. 구체적으로 보면, Gay (1968)에서 지적했듯이, 이중모음의 특징이 되는 긴 초기 안정상태가 없이 단지 전이부분만 존재한다.

순이중모음은 음향적으로 볼 때, 시작부분의 안정된 부분과 전이 부분 그리고 비교적 짧은 종결부분의 안정된 부분으로 되어있다 (Holbrook 와 Fairbanks, 1962:

Lehiste 와 Peterson, 1961). 다시 말해서, 이중모음은 두 개의 목표모음 (vocalic target)으로 되어 있고 이중모음의 지속시간의 대부분이 전이음에 할애되어 있다. 그러나 빠른 속도의 발화에서는 이중모음의 두번째 안정된 부분이 제거되는 경향이 있다 (Gay, 1968). Gay(1968)는 빠른 발화에서의 이중모음의 내부 전이부분은 느린 발화에서 보다 짧다는 것을 보였다. 그러나 그 전이 부분이 같은 비율로 변했고, 이중모음의 개시 주파수 (onset frequency)도 같았다. 그리고 이중모음의 마지막 부분의 고유한 형성음 값에 도달하지 못하고 발음되는 "undershoot" (Lindblom, 1964)과 같은 현상으로 끝난다.

이중모음에 대한 연구자료들은 주로 단모음의 범주를 이용하여 시작과 종결부분으로 구분한 뒤 이 점들 사이의 지름길인 전이부를 연구했다. 이 때 전이부의 형성음 값의 변화, 특히, 혀의 움직임에 반영하는 제 2형성음 전이를 연구하거나, 전이부의 지속시간에 대한 발성과 지각에 대한 연구가 많다. 예를 들어, 형성음 변화의 지속시간을 바꾸었을 때 판단이 어떻게 달라지는가에 대한 연구로는 O'Connor 등 (1957)과 Liberman 등 (1956)이 있다. 이들은 전이부분의 지속시간이 서로 다른 소리를 구분하는 중요한 단서로 작용한다는 점을 발견했다. O'Connor 등 (1957)에서는 초성인 /w, r, l, j/에 대한 음향적 단서를 찾는 연구에서, /j, w/에 적절한 형성음을 가진 안정된 부분이 완전 모음이라는 인상을 갖지 않게 하려면 40 msec를 넘을 수 없음을 밝혔다. 또한 적절한 형성음 전이기간은 약 100 msec이며, 그보다 길면, 초기 자음에 대한 부분이라기보다는 음색이 바뀌는 모음, 즉 이중모음의 인상을 주게됨을 밝혔다. Liberman 등 (1956)은 폐쇄음-전이음-이중모음의 연속을 전이부의 지속시간의 변화에 따라 지각되는 부분을 조사했다. /wε/는 그 변이하는 시간이 150에서 200 msec일 때 /uε/로 지각되기 시작했다. 이와 비슷한 /j/ - /i/ 연속점에서도 자음 + 모음과 이중모음의 판단 경계는 200 - 230 msec 근처임이 나왔다. 이러한 결과는 형성음 전이의 시간이 이중모음 식별의 중요한 단서임을 나타낸다. 이런 연구들은 다양한 전이음 길이가 이중모음의 구별에 매우 중요한 역할을 함을 보인다. Jha (1985)는 인도아리아어인 마이틸리어의 /ai/와 /au/라는 음운적으로 구분되는 이중모음이 처음부터 끝까지 이중모음인가에 대한 스펙트로그램의 조사를 실시했다. 그 결과, 제 2형성음의 변화 속도와 개시목표음 위치가 이중모음의 중요한 단서이며, 종결 목표치는 지속시간에 따라 변하여, 느린속도로 말할 때는 도달하고, 보통 또는

빠른 속도일 때는 목표지점에 도달하지 못하고 끊어져버린다. 따라서 이중모음이란 느린 속도에서는 두 개의 단모음의 연결이라고 할 수 있으나 보통 또는 빠른 속도일 때는 두 개의 목표모음이 분명히 구분되지 않는음을 알 수 있다.

한국어음소는 자음 19, 반모음 2, 단모음 10, 중모음 12가 된다. 반모음은 /j, w/ 이고, 단모음은 /이 위 으 우 에 외 어 오 애 아 어/ 중모음은 두가지로 구분했다. 먼저 [반모음 + 단모음]의 종류로서 모음점 (성절음)이 뒤에 있으므로 상승적 이중모음 (rising diphthong)이라고 불렀다. 반모음은 중-저의 네 평순모음에 공통적으로 연결되어 도표 1과 같이 나타난다 (허웅, 1983:204).

도표 1. 한국어 반모음과 평순모음의 결합 유형

전		후		전		후	
중	je	ja	중	we	wa		
저	jɛ	ja	저	wɛ	wa		

A. 반모음 [j]와 결합

/야, 여(비강세), 여(강세), 요, 유, 예, 애/

B. 반모음 [w]와 결합

/위, 웨, 왜, 와, 워(비강세), 워(강세)/

/위/는 이중모음으로 발음되기도 한다. 상승적 중모음에 있어서 /j/의 자질은 모음 /이/와 유사한 위치이므로 두 개음을 결합하면 변별이 되기 힘들다. 또한 /w/의 변별적 자질은 입술이 원순 --> 평순으로의 이동이 불가능하여 변별적인 자질이 실현되지 못한다 (허웅, 1983:204). 이로써, /요, 유, 위/의 세 이중모음이 더 나타난다. [단모음 + 반모음]의 예로는 하강적 이중모음으로써 /의/가 있다. 이 하강적이 중모음은 다른 중모음과는 고립되어 불안정하므로 상승적 중모음과 비슷하게 상승적으로 발음하기도하나 안정성이 없어서 앞으로 소멸될 것이라 예측된다 (허

용, 1983:205). 한국어의 [아 + 우]는 영어의 [au] 와 달리 두 개의 요소를 같은 올림도로 발음하므로 영어의 이중모음과는 다르다 (원경식, 1990: 82).

II. 실험방법

연구대상으로는 텍사스 오스틴 주립대학에서의 한국어 단모음의 연구를 위한 녹음에 참가한 한국인 남녀 유학생 중 서울 출신의 건강하고, 정상적인 청력을 가지고, 언어적으로 동질적인 그룹으로 선별된 남자 10명으로 했다. 이들은 두가지 선별과정에서 뽑혔다 (Yang, 1990, 1992). 먼저 설문지에 응답한 내용에서 방언구사자를 제외했다. 설문지에는 피험자의 방언, 주로 생활한 곳, 나이, 키, 부모의 방언, 발화 및 청취 장애 여부 등을 기록하게 되어 있다. 이어서 남자 및 여자그룹 중에서 남녀 각각 다섯 명씩을 임의로 선정하여 이들에게 모든 남자 피험자가 발성한 단어 중에서 주변음에 해당하는 [하다, 해다, 후다, 히다]로 된 네 가지 단어를 편집하여 한꺼번에 들려준 뒤, 본인의 발음과 다르게 방언으로 들리는 단어 옆에 임의의 표시를 하게 했다. 이 표시를 모두 합하여 보다 적은 수를 얻은 남녀 피험자 각각 두 명을 선정한 뒤 이들의 판정을 기준으로 하여 발음한 단어의 34% 이상이 방언으로 표시된 녹음 대상자는 제외시켰다. 단모음의 판정으로 이중모음도 같은 결과를 가져올 것이라 가정하고 이중모음에 대한 판정을 따로 시행하지는 않았다.

발음 표본으로는 열 한 개의 이중 모음을 [ㅎ_다]의 환경에서 다섯 번씩 겹치지 않게 임의의 순서로 배열했다. 이런 환경을 이용한 이유는 [ㅎ]를 발음할 때의 조음기관의 모습은 모음에 가장 가까운 것으로써 따라오는 모음에 가장 적은 영향을 주기 때문에 선택했고, 종결어미를 [__다]로 선정한 이유는 한국어의 동사의 기본형으로 자연스러운 단어를 만들고 동시에 마지막으로 혀의 움직임은 화자마다 같게 하여 동일한 환경을 조성하기 위한 것이었다.

피험자들이 방음 장치된 녹음실에서 정상 속도로 자연스럽게 발음한 자료 중 변화가 많은 시작과 끝 부분을 제외한 세번씩의 발음을 수집했다. 녹음된 소리는 The University of Texas at Austin (UT)의 음성학 실험실의 VAX 컴퓨터에 수록되어 있는 KLSPEC 소프트웨어 (MIT대학의 D.H.Klatt이 쓰고 UT의 Jerry Lane이 수정했음.)를 사용하여 이중모음에 대한 형성음 및 기본 주파수를 측정하여 개인의 평균치를

환산했다. 표본음성은 YAMAHA Model KX-800U 카세트를 이용하여 틀고, 4 KHz 저주파 여파기를 거쳐서 10 kHz 표본속도로 계수화시켰다. 스펙트로그램은 매 1 msec마다 6.4 msec 해밍 윈도우를 이용하여 256 표본 이산 푸리에 변환 분석(DFT)을 이용했다. 이 때 이중모음의 형성음은 계속하여 변하기 때문에 스펙트로그램에 나타난 모음의 시작과 끝 부분사이의 총 지속시간을 측정하여 이를 네 등분한 각 시간점을 포착하여 각 지점에서의 형성음 값과 기본주파수를 수집했다. 이 때 스펙트로그램과 함께 진폭자료도 함께 이용하여 규칙적인 진폭변화 부분을 모음으로 선정했다.

이렇게 하여 총 5,280개 (10명 X 3번 발성 X 11 모음 X 4개의 시간점 X 4개 측정치 = 5,280)를 수집했다. 이 측정값에서 세 번 발성한 자료의 평균을 내어 개인별 자료로 최종 선정했고, 열 명의 측정값의 평균을 내어 그래프로 그려서 검토했다. 이들의 자료 처리는 Macintosh의 Excel 3.0으로했고, 통계분석은 StatView 512+를 이용하여 도표화했다. 마지막으로 이렇게 수집한 자료와 기존 한국어 단모음 자료와의 관계를 탐구했다.

연구에서는 각 모음의 스펙트로그램을 형성음 값의 최초의 참고 및 최종 결정의 기준으로 사용했다. 먼저, 스펙트로그램상의 각 형성음 중심에 자를 대어 연필로 선을 그은 뒤 좌우에 인쇄된 눈금으로 부터 추정값을 정한 뒤, 컴퓨터가 계산한 두 개의 DFT배음 스펙트럼 (평균 포락선과 선형예측계수에 의한 포락선)의 서너 개의 추정값과 비교했다. 대부분의 경우에 이 수치는 일치되었다. 또한, 잘못 측정하는 경우가 생기지 않도록 동일 피험자가 세 번 발성한 단어의 스펙트로그램을 동시에 비교했고, 500 Hz이상 형성음 값의 차이가 났을 때, 이것은 분명히 다른 목표 모음으로 발성한 것으로 간주하고 평균치 합산에서 제외했다. 또한 소음으로 인한 단절된 변화를 보이는 수치도 제외했다. 처음부터 끝까지 이중모음의 형성음값 결정은 단모음 자료와 비교하여 대체적인 형성음 주파수값을 예상하고 수집했다. 또한 제 1형성음의 값이 기본주파수값과 일치하거나 스펙트로그램상의 추정값과 다를 때는 스펙트로그램상의 추정값을 채택했다. 제 3형성음이 예상되는 위치에 두 개의 진폭 정점이 보일 경우는 두 정점의 주파수 값을 나누어 평균값을 취했다. 특히 네 개의 시간점에서 수집한 형성음 주파수 변화를 연속으로 프린트하여 연속선상에서 극단적으로 벗어난 수치는 앞 뒤의 형성음값에서 추정하여 선정했다.

기본 주파수값은 자기상관계수법에 의해 KLSPEC 소프트웨어로 추정한 값을 수집했

다. 이 때 소음으로 인하여 최초와 최후의 시간점 (a, d)에서 발생한 잘못된 기본주파수 값은 각 개인별 기본주파수 평균값이나 다른 두 개의 동일한 음성표본에서의 중간점의 두 시간점에서의 값과 비교하여 정정하거나 또는 삭제했다. 중간점의 두 시간점에서의 값이 없을 때는 정상적인 변화를 추정하여 값을 지정했다. 기타 형성음 및 기본주파수 값의 결정은 음향음성학적 상식을 이용하여 정했다.

III. 실험결과 분석

3.1 한국어 이중모음의 기본주파수와 형성음 값

이중모음을 음향학적으로 분석한 결과를 모음별로 살펴보면 도표 3.1과 같다. 이 도표에는 피험자 10명이 발성한 이중모음의 각 측정 시간점에 대한 기본주파수(F0) 및 형성음 값 (F1, F2, F3)의 평균과 각각에 대한 표준편차값(sd)을 수록하고 있다. 소숫점 아래 수치는 모두 반올림했다.

도표 3.1을 살펴보면, 각 개인별 수치를 더하여 평균을 냈기 때문에 개인별 특성이 뚜렷이 나타나지 않는 단점이 있기는 하지만, 대체적인 이중모음의 윤곽을 알 수 있다. 먼저, 각 모음의 기본주파수와 형성음값의 표준편차를 살펴보면 집단내의 변화를 포착할 수 있다. 기본주파수에서는 평균 26 Hz의 변화가 있으며 그 범위는 22에서 31 Hz사이로 상당히 안정되어 있음을 알 수 있다. 기본주파수값은 측정지점에 따라 달라지긴하지만, 주된 경향은 시작부분에서 종결부분에 갈수록 급격히 하강하고 있으며, [왜, 웨, 위] 등에서는 다소 처음 세지점에서 안정되어 있다가 종결부분에서 하강하는 패턴을 보이고 있다.

도표 3.1 이중모음의 10명 피험자의 개인별 기본주파수(F0) 및 형성음 값

(F1, F2, F3)과 그에 대한 표준편차. Tmp 는 측정한 시간점을 나타내며 a는 모음 시작부, d는 모음 종결부분에 해당한다. 총평균은 모든 모음에 대한 평균값을 말한다.

모음	Tmp	F0	sdF0	F1	sdF1	F2	sdF2	F3	sdF3
의	a	178	23	300	40	2079	329	2869	213
	b	176	27	343	38	2110	233	2879	228
	c	174	28	341	34	2155	207	2863	211
	d	166	30	329	48	2090	193	2831	147
와	a	169	22	732	112	1137	108	2492	176
	b	163	24	731	64	1222	84	2488	157
	c	162	25	701	48	1306	81	2519	161
	d	159	24	546	101	1367	95	2584	160
왜	a	171	23	496	122	1664	167	2423	104
	b	169	25	535	71	1769	59	2491	88
	c	168	26	532	66	1809	127	2595	112
	d	162	26	432	78	1704	137	2638	140
웨	a	168	24	374	59	1607	179	2371	114
	b	169	25	463	45	1819	100	2430	112
	c	167	26	490	62	1839	101	2564	116
	d	160	27	417	88	1716	115	2629	140
워	a	167	23	518	73	923	51	2585	189
	b	167	25	539	64	1006	66	2577	198
	c	166	26	559	69	1105	63	2565	172
	d	160	28	474	86	1293	133	2673	129
야	a	168	23	619	115	1689	216	2562	219
	b	164	25	726	75	1479	124	2476	188
	c	161	26	734	71	1400	79	2459	180
	d	157	25	578	65	1403	94	2561	150
애	a	171	25	506	69	1983	177	2683	152
	b	169	26	547	57	1910	161	2693	117
	c	166	26	564	47	1830	159	2700	89
	d	160	27	461	84	1714	127	2684	133
모음	Tmp	F0	sdF0	F1	sdF1	F2	sdF2	F3	sdF3
예	a	170	28	396	114	2083	188	2746	169
	b	168	31	455	81	2015	173	2699	156
	c	165	31	475	70	1915	119	2681	128

요	d	158	29	408	86	1765	117	2697	151
	a	174	25	377	56	1322	237	2419	146
	b	169	26	440	37	1068	61	2477	168
	c	166	26	441	45	1084	86	2542	208
유	d	162	26	413	60	1282	141	2634	215
	a	176	24	304	42	1661	209	2323	176
	b	172	27	356	43	1458	116	2327	193
	c	169	29	362	47	1401	135	2386	188
여	d	162	31	349	54	1495	172	2514	185
	a	169	22	532	73	1382	176	2597	118
	b	167	25	577	57	1205	52	2561	102
	c	165	26	560	60	1211	75	2580	82
	d	161	27	483	90	1307	138	2670	154

총평균		167	26	489	67	1563	136	2585	155

제 1형성음에서는 표준편차의 평균이 67 Hz이고 그 범위는 34 에서 122 Hz에 걸쳐 있다. 제 2형성음에서는 평균 136 Hz의 표준편차를 보였고, 그 범위는 51 Hz에서 329 Hz로 나타났는데 특히 [의]의 발음에 있어서 상당한 변이를 보임으로서 허웅(1985)에서 지적한대로 [의]의 발음이 안정되어 있지않음을 증명해준다. 가장 적은 변이를 보인 [워]에서는 다음에 이어지는 [__다]에 가까이 갈수록 편차가 심한데 이는 전이음 [w]를 발음하기위해 혀를 뒤로 당겼다가 다시 경구개 부위로 [__다]를 발음하는 과정에서 피험자마다 다른 속도와 방식으로 조음하기 때문인 것으로 여겨진다. 제 3형성음에서는 평균 155 Hz의 변화를 보이고 있으며 그 범위는 83에서 228 Hz로 나타나고 있다. 여기서도 [의]에서 가장 많은 변이를 나타나고 있다. 전체적으로 보았을 때, 모든 이중모음들이 상당히 안정되게 발음되었음을 알 수 있으나, 특기할 사항은, 모음 시작부분인 a시점에서 가장 많은 변화가 나타났는데 이는 바로 [w, j]의 전이음 발생 방식에 피험자마다 서로 차이가 있기 때문인 것으로 여겨진다. 이러한 사실은 스펙트로그램을 이용하여 각 형성음값 등을 수집할 때 개인별로 독특한 발생방식에서 관찰할 수 있었다. 예를 들어, 스펙트로그램에 나타난 개인별 특징을 살펴보면, s1과 s6은 측정지점 d에서 불연속적으로 제 3형성음 값이 끊어졌다가 상승했고, 제 2형성음에서는 [ㅎ] 소음이 [이]의 위치에서 출발하여 일직선으로 비

스듬히 내려오는 특징이 있었다. 이러한 제 2형성음의 특징은 s3에서도 2 kHz전후해서 출발하여 나타난다. 또한, s4, s5, s6, s7, s8 등은 [의]의 발음을 단모음처럼 발음해서 제 2형성음의 변화가 거의 없었다. s5와 s7의 경우에는 [훼]와 [훼]를 비슷하게 발음했다. s7의 경우에는 [애]를 단모음으로 발음했다. 일반적으로 이중모음에 대한 음향학적 연구시 적절하고 규칙적인 측정 시간점을 선택하는 것은 상당히 중요한 일이라 할 수 있다.

각 모음을 [w]와 [j]로 시작하는 그룹으로 구분했을 때는 어떤 결과를 보일까? 여기에 대한 자료는 도표 3.2에 나타나 있다.

도표 3.2. 각 모음을 [w]와 [j]로 시작하는 그룹으로 구분하여 평균한 값

A [w] 그룹 (와, 왜, 웨, 워)

시간점	F0	sdF0	F1	sdF1	F2	sdF2	F3	sdF3
a	169	23	530	92	1333	126	2468	146
b	167	25	567	61	1454	77	2497	139
c	166	26	571	61	1515	93	2561	140
d	160	26	467	88	1520	120	2631	142
평균	165	25	534	76	1455	104	2539	142

B. [j] 그룹 (야 애, 예, 요, 유, 여)

시간점	F0	sdF0	F1	sdF1	F2	sdF2	F3	sdF3
a	171	25	456	78	1687	201	2555	163
b	168	27	517	58	1523	115	2539	154
c	165	27	523	57	1474	109	2558	146
d	160	28	449	73	1494	132	2627	165
평균	166	27	486	67	1544	139	2570	157

C. A, B그룹을 합친 이중모음의 시간점별 기본주파수 및 형성음 평균값

시간점	F0	sdF0	F1	sdF1	F2	sdF2	F3	sdF3
a	170	24	493	85	1510	163	2511	155
b	168	26	542	60	1488	96	2518	146
c	166	27	547	59	1494	101	2559	143
d	160	27	458	81	1507	126	2629	153
총평균	166	26	510	71	1500	121	2554	149

도표 3.2에서 살펴보면, 시간이 지남에 따라 기본주파수의 변화는 [w]그룹에서는 9 Hz의 범위에서 내려갔고 [j]그룹에서는 11 Hz로서 거의 비슷한 변화 유형을 보이고 있다. 제 1형성음에서는 [w]그룹에서 상승하다가 d시점에서 100 Hz하강했으며 [j]그룹도 비슷한 양상을 보였다. 그러나 제 2형성음에서는 두 그룹간에 뚜렷한 차이가 보인다. 즉, [w]그룹에서는 200 Hz정도 상승한 반면 [j]그룹에서는 170 Hz 하강했다. 이것은 전이음의 주파수 위치가 [w]는 낮고 [j]는 높기때문에 발생한 것으로 여겨진다. 특히 [w]그룹에서는 시작부분과 종결부분의 변이가 많았다. 제3형성음에서는 두 그룹 모두 종결부분에서 주파수가 상승했고 시작부에서 세번 째 시점까지는 [w]그룹에서는 100 Hz 상승했고, [j]그룹에서는 큰 변화가 별로 없었다. 도표 가운데 두 그룹을 합친 이중모음의 시간점별 기본주파수 및 형성음 평균값을 살펴보면 기본 주파수에서는 모음 시작부분에서 종결부분에 까지 10 Hz 감소했고 주파수범위는 제 1형성음이 510 Hz 제 2형성음이 1500 Hz 제 3형성음이 2554 Hz 로써 균일관의 형태인 모음 [어]의 형태가 되므로써 이는 바로 모든 모음의 평균이 조음상 제일 발음하기 쉽고 운동을 요하지 않는 중모음의 위치로 환원함을 알려준다. 따라서 음성 합성에 있어서도 기본 형인 [어] 모음을 기준으로하여 설정하여 여러 가지 모음으로 변환 또는 수정하는 것이 빠른 방법이 될 수 있을 것이다. 표준편차 값에 있어서, 제 1, 2, 3 형성음으로 갈수록 높아지고 있는데 이는 바로 형성음 값이 높아지므로써 변별하는 범위가 넓어져서 지각상 예민도가 떨어지고 있음을 보인다. 제 3형성음의 편차값의 평균은 제 1형성음의 값보다 약 두 배가 된다.

지금까지 각 시간점에서 변화하는 경향을 살펴보았는데 이번에는 각 모음별로 그 변화폭을 알아보기로 한다. 도표 3.1에서 볼 수 있듯이, 첫번 째 시간점에서 출발하여 모음 종결부분인 네번 째 시간점은 다음에 오는 [__다] 발음을 예상하고 조음했기 때문에 다시 첫번 째 시간점의 값으로 돌아가는 경향이 보이므로 충분한 변화를 포착하기 어렵다. 따라서 네번 째 시간점은 여기서 제외하고 첫번 째와 세번 째 사이의 주파수의 변화만 (a-d)을 관찰했다. 다음 도표 3.3을 살펴보자.

도표 3.3 첫번 째 시간점과 세번 째 측정시간점 사이의 주파수 변화폭.

모음	F0	F1	F2	F3
의	4	-41	-76	6
와	7	31	-169	-27
왜	3	-36	-145	-172
웨	1	-116	-232	-193
워	1	-41	-182	20
『 평균	3	-41	-182	-93』* [w]그룹 평균
야	7	-115	289	103
애	5	-58	153	-17
예	5	-79	168	65
요	8	-64	238	-123
유	7	-58	260	-63
여	4	-28	171	17
『 평균	6	-67	213	-3』** [j]그룹 평균
『 총평균	5	-55	39	-37』*** 모든 이중모음 평균

도표 3.3에서 살펴보면, 기본주파수에서는 평균 5 Hz의 감소를 보였고, 제 1형성음에서는 약 55 Hz 상승했고 제 2형성음에서는 [w]로 시작하는 그룹에서는 평균 182 Hz 상승했다. 이중 가장 많은 변화를 보인 모음은 [웨]로서 232 Hz의 변화를 보인 반면 가장 적은 변화를 보인 모음은 [왜]였다. 이것은 단모음의 위치에서 볼 때 이동거리가 짧을 수록 적은 변화를 했다고 여겨진다. [j]로 시작하는 그룹에서는 평균 213 Hz의 변화를 보였는데, 이중 가장 많은 변화를 보인 모음은 [야]로서 289 Hz의 변화를 보인 반면 가장 적은 변화를 보인 모음은 [애]였다. 여기서 조음위치상 [j]와 상당히 떨어져 있는, [야, 요, 유]에서는 230 Hz이상의 변화가 있는 점이 주목된다. 제 3형성음에서는 [w]그룹에서는 평균 93 Hz의 상승이 있었고 [j]그룹에서는

3 Hz의 상승이 있었다. [웨]와 [왜]에서 급작스런 상승이 있었다.

지금까지 살펴본 바에 의하면, 제 2형성음에서 가장 많은 변화가 있었고, 다음으로는 제 1형성음, 이어서 제 3형성음 순서였다. 기본주파수에서는 큰 변화가 없었다. 따라서, 다음으로는 제 1,2 형성음에 초점을 두고 분석해 본다. 도표 3.1 가운데 제1,2 형성음을 그림 3.1처럼 나타내 보았다. 그림에서 x축은 제 1형성음을 나타내고 y축은 제 2형성음을 나타낸다. 단위는 Hz이다. 이중모음은 해당위치에 가까운 곳에 표시했다. 각 시점을 표시하기위해 사각형을 이용했다. 모음 시작부를 빈사각형으로 시작하여 단계별로 채워간 뒤, 모음 종결부분은 사각형이 완전히 까맣게 채워지도록 표시했다. 선은 각 시간점으로의 이동을 나타낸다.

그림 3.1 각 시간점에서 살펴본 한국어 이중모음의 형성음 이동.

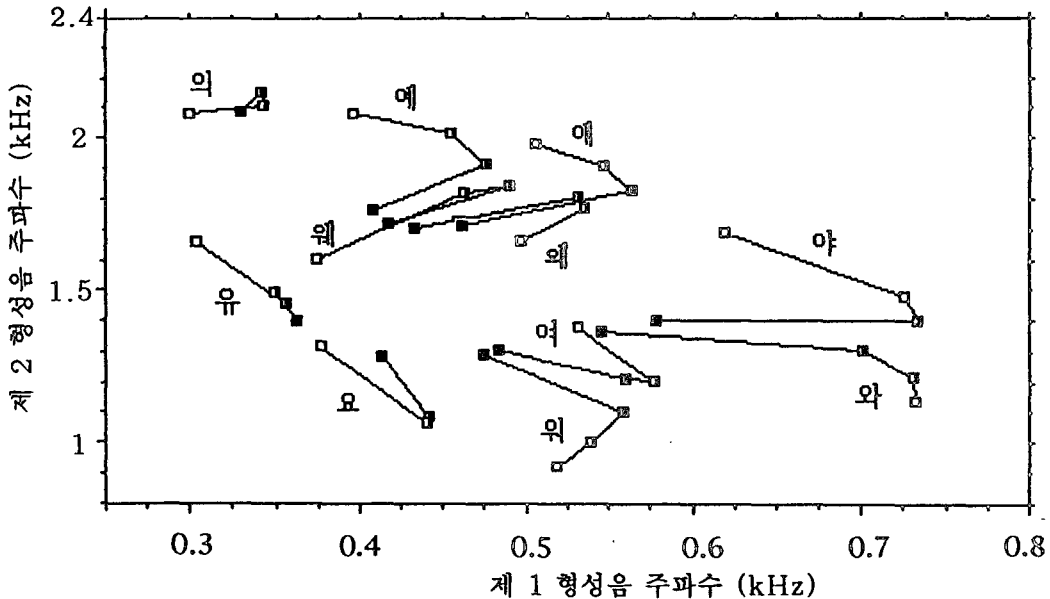


그림 3.1은 각 시간점에 대한 변화를 나타내고 있는데, 제 1형성음에서는, [와]를 제외하고는 모두 상승했다가 종결부분에서 하강했다. 제 1형성음은 턱의 움직임에 반영하는 것임을 생각해 볼 때, 당연한 결과라고 여겨진다. 왜냐하면, 각 이중모음을 발음하기위해 턱을 내려 발음했다가 마지막 [다]를 발음하기위해 턱을 올렸기 때문일 것이다. 혀의 움직임을 주로 반영하는 제 2형성음에 있어서는 전설모음보는 후설모음에 있어서 보다 많은 변화가 있었다. 각 그룹별로 살펴보면, [w]로 시작하는 그룹은 [와]를 제외하고 모음 시작부분에서 세번째 부분 까지 제 1형성음과 제 2형성음 주파수가 모두 점점 상승하는 경향이 있고, [j]로 시작하는 그룹은 제 1형성음은 [이]쪽을 향한 낮은 값에서 모두 상승하고있고, 제 2형성음은 [이]의 높은 형성음 값에서 출발하여 하강하는 경향을 보인다. 이는 바로 성도의 움직임과 일치하는 결과라 여겨진다. 특히, 다음에 이어지는 환경인 [__다]의 경구개 파열음을 조음하기위해 종결부분이 모음 공간의 중앙으로 집중하는 현상을 볼 수 있다. [의]의 경우에는 제 1형성음이나 제 2형성음의 이동거리가 별로 없어서 거의 [이]발음에 가까운 모습이다. 이것은 스펙트로그램에서 모음 시작부분 이전의 [ㅎ__]소음에서 상당한 전이부분이 나타나 있었는데, 본 연구에 있어서는 이 부분에 대한 측정값이 없어서 볼 수가 없다.

3.2 이중모음과 단모음의 비교

여기서는 이중모음의 형성음 값과 사전 연구된 단모음의 값을 비교 검토하기로 한다. 단모음에 대한 형성음 및 기본주파수 값은 Yang(1990)의 논문을 기준으로 설정했다. 단모음 실험에 참여한 각각의 피험자는 이중모음의 연구에도 참여했기 때문에 피험자의 번호는 동일함을 밝혀둔다. 각 개인별로 본 단모음의 형성음 값은 Yang(1990:139-143)에 실려 있다. 이중모음의 자료가운데 네번째 측정시간점은 그림 3.1에서 살펴본 것처럼 다음에 오는 [__다]의 영향을 상당히 받아서 가운데로 집중하는 현상이 나타났으므로 여기서는 제외하고 또한 변화가 많은 제 1,2형성음에 대하여 주로 분석해 보기로 한다. 단모음의 자료는 수집한 시점에 있어서 이중모음과 약간 다르게 되어있기 때문에 (Yang, 1990), 서로 일치가 되는 B와 D지점을 굵은 선으로 연결하고 이중모음은 모음개시부에서 세번째 시간점까지를 가는 선으로 연결

하여 표기하기로 한다. 그림 3.2에서도 그림 3.1과 같은 요령으로 나타내었다.

그림 3.2 이중모음과 단모음의 비교 분석.

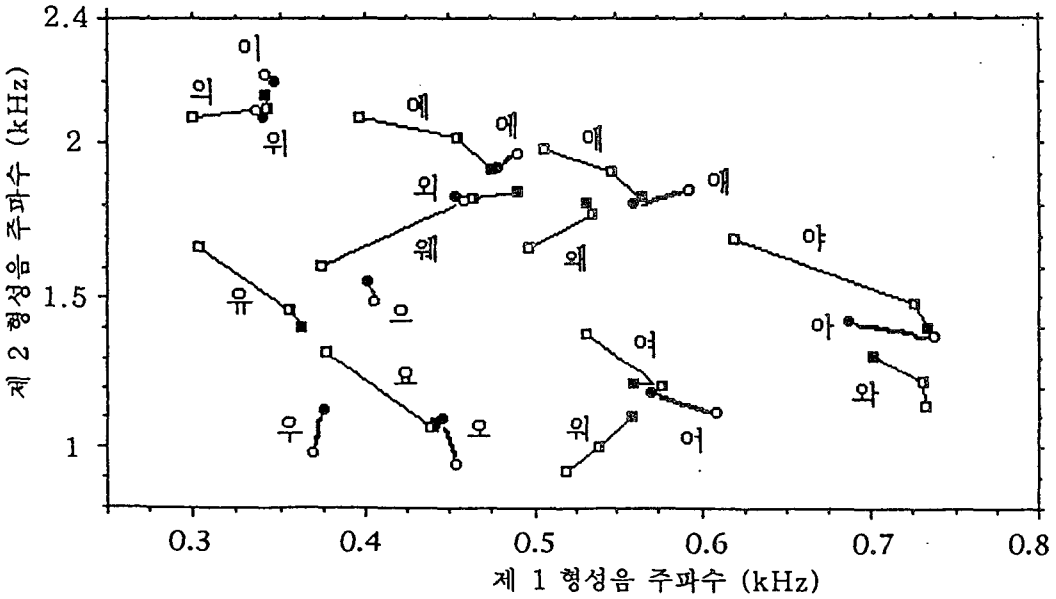


그림 3.2에서 살펴보면, 단모음과 비교해 볼 때 이중모음은 제 1,2형성음 모두에서 상당한 변화를 보이고 있다. 특히, 세번째 시간점은 단모음의 위치와 거의 일치하는 경향을 보이고 있다. 예를 들어, [에]는 [에]로 [애]는 [애]로 정확히 일치하며, [웨]나 [왜] 도 각각의 종결부분이 단모음의 위치에 근접한다. [여, 위]도 [어]의 위치에 돌아가며, [요]는 [오]와 겹쳐진다. 각 모음 그룹별로 살펴보면, [j]그룹의 경우는 단모음 [이]를 향하고 있지만, 상당히 떨어진 위치에서 시작하고 있고, [w]그룹의 경우는 모두 [우]의 위치를 향해 있음을 알 수 있다. 구체적으로 그 궤적을 살펴보면 [야]인 경우는 단모음 [애]의 위치에 가깝게 시작하여 [아]로 향했고, [웨]의 경우는 단모음 [외]의 위치에서 시작했다. 그 변이폭은 이중모음에서는 뚜렷

이 나타나고 있는데 비해 단모음에서는 아주 짧은 거리를 가로질러 가고 있음을 알 수 있다. 모음 [의]의 경우에는 단모음이나 [위]의 모양과 비슷한 위치에서 이뤄짐으로써 모음 [으]와 [이]의 결합이라기 보다는 영어의 강세모음 [이]에서 약세모음 [이]로의 이동과 같은 음향적 이동을 보인다. [외]나 [위]의 이중모음적 성격규명은 단모음의 측정 지점에 있어서 첫 전이부분이 빠져있기 때문에 여기서는 분명히 구분할 수는 없으나, 대체적으로 봐서는 단모음의 변화에 해당한다고 볼 수 있다. 앞으로 이 두 모음에 대해서는 소음부분 [ㅎ] 부터 시작하여 변화를 측정 분석함으로써 결정적인 판단을 내릴 수 있을 것이다.

3.3 이중모음의 지속시간과 진폭

여기서는 이중모음의 총 지속시간에 대한 개인별 자료와 평균값, 그리고 진폭의 변화를 알아보기로 한다. 도표 3.4에서 각 개인별 및 각 이중모음별로 구분하여본 지속시간의 평균값을 낮은 값부터 정렬하여 나타내 보았다.

도표 3.4 각 이중모음별로 구분하여 본 지속시간의 평균값 및 표준편차.

이중모음	평균	표준편차
여	99	32
유	100	33
와	101	36
의	101	37
위	110	34
왜	113	32
요	114	37
야	117	32
예	119	26
애	119	31
웨	122	33
총평균	110	33

이중모음의 지속시간을 모음별로 평균하여 본 결과 110 msec이고 표준편차의 평균은 33 msec이었다. 그 범위는 모음 [여]의 99 msec에서 모음 [예]의 122 msec로써 23 msec의 범위에서 발생했다. 또한 [w]그룹을 평균하여 보면 112 msec 이고 [j]그룹의 평균은 111 msec로써 두 그룹간에 뚜렷이 구분하여 유형화할 차이는 없는 것으로 여겨진다. 이러한 지속시간은 Liberman 등(1956)의 영어 지각실험에서 200-- 230 msec 근처에 [자음+모음]의 경계가 나타난 것에 비추어볼 때 매우 짧은 시간이고 따라서 이중모음이라기 보다는 [자음 +모음]으로 보는 것이 타당할 것이다. 하지만 한국인들의 지각체계가 다를수도 있으므로, 이것은 앞으로 지각실험을 통해 밝혀야할 과제이다. 이어서 각 피험자별로 평균 및 표준편차를 살펴보면 도표 3.5와 같다.

도표 3.5 피험자별 지속시간의 평균 및 표준편차

피험자	평균	표준편차
s9	78	12
s3	80	12
s4	82	14
s2	94	17
s8	99	16
s7	109	11
s10	117	20
s6	122	13
s1	152	16
s5	171	10
총평균	110	14

피험자별로 살펴본 결과는 그 범위는 s9의 78 msec 에서 s5의 171 msec로 약 100 msec에 걸쳐 있으나, 표준편차는 평균 14 msec로써 대체로 개인별로는 안정되게 발음했음을 알 수 있다. 덧붙여, 보다 나은 양질의 음성합성시 진폭의 변화도 중요한 변수가 될 수 있는데 각 피험자들의 진폭을 조사해 본 결과, 피험자마다 개인별 특성이 다르고, 측정값의 기복이 심하여 평균했을 때 어떤 중요한 의미를 발견하기가

어려웠다. 따라서 여기에서는 진폭의 변화를 피험자 선발시 동료들의 평가로 방언적 요소가 가장 적게 지적된 s10에 한정하여 구체적으로 살펴보면 아래 도표 3.6 과 같다.

도표 3.6 s10이 발생한 이중모음의 각 측정 시간점에 있어서 각 형성음 값에 대한 진폭값 총평균

시간점	aF1	aF2	aF3
a	36	32	28
b	39	37	33
c	39	38	33
d	38	35	32
평균	38	36	32

도표 3.6에서 살펴보면 제 1, 2, 3형성음 각각 모음 시작부분에서 상승했다가 종결 부분에서 하강하는 모습을 보인다. 그 변화폭은 3에서 6 dB로 크지않다. 시간점에 관계없이 평균한 값으로 보면 진폭값은 제 1형성음이 가장 높고 다음으로 제 2, 3형성음 순으로 내려간다. 이를 각 그룹별로 분리하여 살펴보면 도표 3.7과 같다.

도표 3.7 s10이 발생한 이중모음의 각 측정 시간점에 있어서 각 형성음값에 대한 진폭값의 그룹별 총평균

A. [w] 그룹				B. [j] 그룹		
시간점	aF1	aF2	aF3	aF1	aF2	aF3
a	38	33	25	35	31	30
b	43	39	35	38	36	32
c	43	40	36	38	36	31
d	40	38	34	38	34	30

두 그룹 사이에 뚜렷이 구분될 만큼 큰 차이는 없으나 [w]그룹보다는 [j]그룹에서 진폭 값이 대체로 낮았다. 이는 [이]발음에 가까운 [j]그룹에서 보다 좁은 모양의 조음때문이라 여겨진다. 또한 [w]그룹에서의 시간점에 따른 진폭변화 범위가 다소 [j]그룹보다는 넓었다.

IV. 결론

본 논문에서는 서울출신의 언어적으로 동질적인 그룹을 이루는 10명의 피험자를 선정하여 이들이 발음한 열 한개의 이중모음을 음향학적으로 분석하고 단모음과의 비교도 했다. 지금까지 살펴본 한국어 이중모음의 음향학적 연구 결과를 정리해보면 다음과 같다.

1. 이중모음의 기본 주파수와 형성음의 피험자간의 변화를 표준편차값을 통해 분석해 본 결과, 모음 시작점에서 상당한 변화가 있었다.
2. 모든 모음의 형성음 평균값을 내 본 결과 균일관인 중모음 [어]에 가까운 값이 나와서, 앞으로 이중모음 합성의 기본형으로 설정하여 여러 가지 모음으로 변환하는데 이용할 수 있다.
3. 각 모음의 형성음과 기본주파수의 변화를 분석해 본 결과, 제 2형성음에서 가장 많은 변화가 있었고, 그 다음으로 제 1, 3형성음 순서로 나타났다. 기본주파수에서는 별다른 큰 변화가 없었다. 특히, 제 2형성음에 있어서, [j]로 시작하는 그룹은 하강하고, [w]로 시작하는 그룹은 상승하는 서로 대조된 모습을 보였다.
4. 각 모음별로 네 개의 시작점에서의 변화를 추적해 본 결과, 턱의 움직임은 주로 반영하는 제 1형성음에서는 [와]를 제외하고는 모두 상승했다가 종결부분에서 하강했다. 이에 반해 혀의 움직임을 주로 반영하는 제 2형성음에서는 전설모음 보다는 후설모음에서 보다 많은 변화를 보였다.
5. [__다]의 경구개 파열음을 발음하기위해 종결부분은 모음 공간의 중앙으로 집중했다. [의]의 경우에는 제 2형성음의 이동거리가 별로 없어서 거의 [이] 발음에 가까운 형성음 분포를 보였다.
6. 총지속시간에 대해서는 평균 110 msec로써 발생했고, 모음간에는 표준편차가 33 msec 화자간에는 14 msec로써 대체로 피험자나 모음별로 안정되게 발음했음을 알

수 있었다.

7. 이중모음과 단모음의 비교에서는 단모음은 이중모음에 비해 뚜렷한 변화를 보였고, 특히 세번 째 측정시간점은 단모음과 이중모음이 거의 일치했다. 이는 영어의 이중모음과는 반대로, 첫번 째의 전이음은 약하고 짧게 발음되고, 두번 째 목표모음이 강하고 정확히 발음되었음을 보여준다.

8. 이중모음은 단모음보다는 제 1,2형성음에서 상당한 변화를 보여서 턱과 혀의 움직임이 다소 활발하다고 볼 수 있다. 또한, 이중모음의 이동 경로를 살펴보면, [w]그룹에서는 제 1,2형성음이 모두 상승한다면, [j]그룹에서는 제 1형성음은 상승하나 제 2형성음은 주로 하강한다. 이중모음의 종결부분은 항상 단모음의 모음점으로 돌아간다. 이중모음의 이동 경로는 대체로 짧아서 반이중모음에 해당한다고 볼 수 있으며, 전이음의 시작은 [ㅎ] 소음에서 부터 일어난다.

본 연구에서 나타난 문제점이나 제안사항으로는 이중모음의 총지속시간을 네 등분하여 측정시간점을 정한 관계로 스펙트로그램에서 나타난 [ㅎ] 소음에서 나타나는 전이부분이 측정되지 않았다. 따라서 일부 이중모음의 발음에 있어서 소음부분에 대한 분석도 포함시킨다면 보다 정확한 결과가 나오리라 여겨진다. 또한 실험실에서 측정된 자료보다 현실적인 발음 환경에서 발생된 자료의 검토가 앞으로 이뤄져야 할 것이다. 마지막으로 본 실험은 주로 발생한 모음의 분석에 치중했으나, 음성합성을 통하여 지각적인 실험을 통해 중요한 변수를 찾는 것도 앞으로의 중요한 연구과제가 될 것이다.

본 연구의 결과는 한국어 이중모음에 대한 순수 언어학적인 기초자료로 이용될 것이며, 또한 단모음과 이중모음간의 상관 관계를 밝힘으로써 음성합성 및 음성 자동 인식장치 등의 개발이나 한국어 교육에 활용되리라 기대된다. 특히, 이중모음에 대한 복잡한 연구 방법 때문에 등한시해 왔던 이 분야의 연구에 조직적인 분석 방법을 제시함으로써 앞으로 더 많은 연구가 가능하게 되리라 여겨진다.

<참고 문헌>

- 김승곤. (1987). *음성학*. 서울: 정음사.
- 김재민, 이계순, 한영희, 전상범, 신상순. (1985). *영어음성학*. 서울: 신아사.
- 고도홍, 구희산, 김기호, 양병곤. (1994 인쇄 예정). *음향음성학 개론*. 서울: 한신문화사.
- 양병곤. (1992). *음성학 입문*. 부산: 진영문화사.
- 양병곤. (1993). *한국어 이중모음의 음향음성학적 연구*. 부산: 진영문화사.
- 원경식. (1990). *영어음성학*. 음운론. 현대영어를 중심으로. 서울: 탐출판사.
- 조성식 편. (1990). *영어학 사전*. 서울: 신아사.
- 전상범. (1988). *영어음성학*. 서울: 을유문화사.
- 허응. (1983). *국어음운학*. 서울: 정음사.
- Bond, Z.S. (1978). "The effects of varying glide durations on diphthong identification." *Language and Speech* Vol. 21 No. 13, pp. 253-63.
- Bond, Z.S. (1982). "Experiments with synthetic diphthongs." *Journal of Phonetics* 10, pp. 259-64.
- Gay, T. (1968). "Effect of speaking rate on diphthong formant movements." *Journal of Acoustical Society of America (JASA)* 44(6), pp. 1570-73.
- Gay, T. (1970). "A perceptual study of American English diphthongs." *Language and Speech* Vol. 13 (2)., pp. 65-88.
- Gerber, S.E. (1972). "Perception of segmented diphthongs." In *Proceedings of the Seventh International Congress of Phonetic Sciences*. (A. Rigault and R. Charbonneau, eds), pp. 480-92. The Hague: Mouton and Co.
- Gerhard, R.H. (1954). *General American Pronunciation*. Tokyo: Meirindo.
- Hertz, S.R. (1991). "Streams, phones and transitions: toward a new phonological and phonetic model of formant timing." *Journal of Phonetics* 19(1), pp. 91-110.
- Holbrook, A. & Fairbanks, G. (1962). "Diphthong formants and their movements." *Journal of Speech and Hearing Research* 5(1), pp. 38-58.
- Jha, S.K. (1985). "Acoustic analysis on the Maithili diphthongs." *Journal of Phonetics* 13, pp. 107-115.

- Kenyon, J.S. & Knott, T.A. (1953). *A Pronouncing Dictionary of American English*. Springfield: Merriam.
- Kenyon, J.S. (1966). *American Pronunciation*. Ann Arbor: Wahr.
- Klatt, D.H. (1973). "Durational characteristics of prestressed word initial consonant clusters in English." *Quarterly Progress Reports*. (Research laboratory of Electronics, MIT, Cambridge, Mass.). QPR 108, pp.253-260.
- Ladefoged, P and Maddieson, I.(1990). "Vowels of the world's languages." *Journal of Phonetics* 18, pp.93-122.
- Ladefoged, P. (1982). *A Course in Phonetics*. New York: Harcourt Brace Jovanovich, Inc.
- Lehiste, I and Peterson, G.E. (1961). "Transitions, glides, and diphthongs." *JASA* 33(3). pp.268-77.
- Lehiste, I. (1964). "Acoustical characteristics of selected English consonants." Bloomington: Indiana Univ. *Research Center in Anthropology, Folklore and Linguistics* P. No. 34.
- Lieberman, A.M., Delattre, P.C., Gerstman, L.J. and Cooper, F.S. (1956). "Tempo of frequency change as a cue for distinguishing classes of speech sounds." *Journal of Experimental Psychology* 52, pp.127-37.
- Lindblom, B. (1964). "Articulatory activity in vowels." *Quarterly Progress Status Report Speech Transmission Laboratory*. Royal Institute of Technology. Stockholm. No.2, pp.1-5.
- O'Connor, J., Gerstman, L.J., Liberman, A.M., Delattre, P. and Cooper, F.S. (1957). "Acoustic cues for the perception of initial /w, j, r, l/ in English. *Word* 13, p.24-43.
- Pickett, J.M. (1987). *The Sounds of Speech Communication*. Austin, Tx:Pro-Ed.
- Sloat, C., Taylor, S.H., and Hoard, J.E. (1978). *Introduction to Phonology*. Englewood Cliffs, N.J.: Prentice-Hall, Inc.
- Trager, G.L. & Smith, H.L. (1966). *An Outline of English Structure*. Washington: American Council of LEarned Societies.
- Ward, L.C. (1972). *The Phonetics of English*. Cambridge: Heffer.
- Yang, B.G. (1990). *Development of Vowel Normalization Procedures: English and Korean*. Seoul:Hanshin.
- Yang, B.G. (1992). "An acoustical study of Korean monophthongs." *JASA* 91(4), pp.2280-83.