

복합음과 대학생이 발음한 모음 포먼트 측정*

Formant Measurements of Complex Waves and Vowels Produced by Students

양 병 곤**

Byunggon Yang

ABSTRACT

Formant measurements are one of the most important factors to objectively test cross-linguistic differences among vowels produced by speakers of any given languages. However, many speech analysis softwares present erroneous estimates and some researchers use them without any verification procedures. The purposes of this paper are to examine formant measurements of complex waves which were synthesized from the average formant values of five Korean vowels using three default methods in *Praat* and to verify the measured values of the five vowels produced by 20 students using one of the methods. Variances along the time axis are discussed after determining absolute difference sum from the 1/3 vowel duration point. Results show that there were smaller measurement errors by the *burg* method. Also, greater errors were observed in the *sl* or *lpc* methods mostly caused by the inappropriate formant settings. Formant measurement deviations were greater in those vowels produced by the female students than those of the male students, which were mostly attributed to the settings for the vowels /o, u/. Formant settings can best be corrected by changing the number of formants to the number of visible dark bands on the spectrogram. Those results suggest that researchers should check the validity of the estimates from the speech analysis software. Further studies are recommended on the perception test of the original sound with the synthesized sound by the estimated formant values.

Keywords: speech analysis, acoustic measurements, vowel formants

1. 머리말

최근에 많은 연구자들이 음성분석 소프트웨어를 이용하여 모음의 포먼트값을 측정하여 언어 간 비교나 사전 사후 학습효과 등을 객관적으로 제시해왔다. 일부 연구자들은 분석과정에 대한 별다른 검토 없이, 기존의 연구논문에서 제시한 결과와 큰 값 차이를 보이는 잘못 측정된 결과를 논문 초고에 기록하여 제출하는 경우가 종종 있었다. 이에 본 논문에서는 포먼트값을 바르게

* 이 논문은 부산대학교 자유과제 학술연구비(2년)에 의하여 연구되었음.

** 부산대학교 영어교육학과 교수

구할 수 있는 방법을 찾아보기 위해, 합성음과 대학생이 발음한 모음의 포먼트값을 구하는 과정에서 발생할 수 있는 문제점을 살펴보고자 한다. 이러한 연구는 보다 정확한 포먼트 측정에 대한 자료를 제공함으로써 앞으로 국어를 비롯한 외국어 모음 연구에 타당성을 부여하여 의미 있는 결과를 도출하는데 기여할 수 있을 것으로 기대된다.

포먼트값은 성도의 공명 특성을 음향적인 수치로 구한 값을 말한다(Fant, 1960). 포먼트는 스펙트로그램에서 에너지가 밀집된 검은 띠의 중심점을 말하는데, 좀 더 자세히 측정하기 위해서는 스펙트럼을 구하여 에너지 정점을 찾기도 한다. 모음스펙트럼에서 포먼트가 차지하는 강도는 보통 대역(bandwidth) 값으로 나타내는데, 이 값이 클수록 포먼트가 넓게 퍼져 있게 되어 에너지 정점의 진폭값은 낮게 되고 포먼트값으로 추정될 확률이 낮아지게 된다. 실제 측정된 모음의 포먼트는 안면의 쪽 들어간 입안에서 방출된 음성을 얼굴표면이 반사시켜 확대시키는 방사특성에 의해 옥타브당 6 dB정도로 고주파 영역의 에너지가 증폭되지만, 고주파로 갈수록 에너지가 급격히 낮아지기 때문에 음성분석소프트웨어에서도 고주파대역증폭(preemphasis)을 시켜서 음성분석 소프트웨어에서 고주파대역의 에너지는 실제보다 훨씬 더 높게 표시된다. 일반적으로 모음의 포먼트는 여러 개가 나타나지만 가장 낮은 것부터 3~4 개를 사용하는데, 제1 포먼트(F1)는 턱을 내리거나 올려서 입안의 공간인 구강을 넓히면 높아지고 그 반대로 좁히면 낮아진다(Pickett, 1987). 제2 포먼트(F2)는 주로 혀의 움직임을 나타내는데, 혀가 앞쪽으로 밀려갔을 때는 높아지고 뒤쪽으로 밀려갔을 때는 낮아지는 경향이 있다. 제3 포먼트(F3)는 입술을 둥글게 하여 성도의 길이가 늘어나거나 혀를 말아 올려 /r/을 발음할 때 낮아진다.

포먼트값은 모음 연구의 매우 중요한 부분을 차지한다. Kent와 Read(2002)에서는 가장 전통적인 연구인 Peterson과 Barney(1952)를 비롯하여 Hillenbrand 외 (1995), Yang(1996)을 포함하여 그 당시까지 발표된 8 개의 연구에서 밝혀진 영어 모음의 포먼트 측정값을 제시하고 이 값들의 평균과 표준편차를 표로 나타내었다(강석한, 2007:132-133, 표 4-1과 4-2 참고). 이렇게 측정된 값들의 표준편차를 보면, 남성의 F1은 모음 /a/에서 54 Hz로 가장 크고 다른 모음에서는 17 Hz에서 37 Hz에 이르지만, F2에서는 가장 변화가 큰 모음이 /u/로 241 Hz정도로 매우 높고, 인접한 후설모음 /ɔ, o, ɔ/ 등에서는 130 Hz로 /æ/를 제외한 전설모음에 비해 두 배 이상의 편차를 보이고 있다. 이에 반해 F3에서는 /æ/에서 102 Hz로 최고값을 보이고 다른 모음에서는 대체로 44 Hz에서 80 Hz 범위에 걸쳐 있다. 여성화자의 F1의 편차는 모음 /æ, ɔ/가 각각 115 Hz, 114 Hz를 보이고, 나머지는 14 Hz에서 80 Hz까지 퍼져 있으나 대체로 편차가 적은 편이다. F2에서는 모음 /u/가 무려 295 Hz의 편차를 보이고, 인접한 후설모음 /ɔ, o, ɔ/는 각각 204 Hz, 183 Hz, 150 Hz로 다른 모음에 비해 상대적으로 많은 편차가 있다. F3의 표준편차값은 모음 /a/에서 141 Hz로 최고값을 보이고 나머지 모음의 편차값은 39 Hz에서 82 Hz로 분포되어 있다. 이런 경향의 주된 원인은 모음 /o, u/와 같이 F1과 F2가 서로 근접하는 모음에서 두 개의 포먼트를 한 개로 잘못 추정하거나, hard라는 목표어에서 모음 /a/를 측정할 위치가 연구마다 달랐기 때문일 수 있다. 보통 /r/발음을 미국영어에서는 혀 앞부분을 입천장을 향해 말아 올려 발음함으로써 F3값이 아래로 움푹 내려오는 모양을 보인다(Edward, 2003:207). 일부 방언을 사용하는 화자에서는 영국영어의 방식으로 이런 혀의 움직임을 보이지 않았을 수도 있다. 덧붙여, 모음 /ɔ/는 목표어를 hawed로 주었는데, 이 발음을 할 때 입술을 둥글게 한 경우와 그렇지 않은 경우에는 포먼트값의 차이

를 가져 왔을 것으로 생각된다.

일반적으로 포먼트값은 발음한 사람이 목표음을 바르게 발음했는가? 발화음의 환경은 어떠한가? 발화음의 어느 시간 지점에서 측정했는가? 어떤 방식으로 분석했는가? 일관성 있는 기준을 적용했는가? 등에 따라 같은 음성파일에 대한 측정값이라도 달라질 수 있다. 만약 피험자가 다른 목표 모음으로 발음했다면, 분석된 포먼트값의 결과에 근거한 일반화는 문제가 될 수 있다. 이런 오류를 줄이기 위해서는 녹음된 음성을 여러 명이 들어서 옳은 목표 모음인지 판단한 다음 분석을 하면 된다. 두 번째로 발화음의 환경은 공동조음 현상으로 가장 빠른 이동 거리로 조음이 되므로 앞뒤에 오는 자음의 영향이 모음 발화에 영향을 주게 된다. Lindblom(1963)은 모음 /u/만 발음할 때는 F2가 F1에 아주 가까이 접근하였으나 /dud/와 같이 자음 d 사이에 넣어 발음했을 때 F2값은 F1에서 많이 떨어져 발음되었는데 이를 목표음 미달(undershoot)의 예로 제시했다. 이렇게 녹음 환경에 따라 포먼트 측정값이 달라질 수 있고 다른 환경에서 측정한 값들은 서로 비교하는데 어려움이 있다. 덧붙여, 발화음의 측정 지점이 어디인가에 따라 포먼트값이 많이 달라진다. 또한 사람의 발성기관은 발성 중 끊임없이 움직이기 때문에 측정위치에 따라 값들이 달라질 수 있다. Yang(1996)은 발화속도에 따라 측정지점이 이동하고 있으므로 모음의 전체 지속시간을 기준으로 일정한 비율지점에서 측정하면 화자마다 어느 정도 일관성이 있음을 보고했다. 많은 기존의 연구들이 모음의 시작부나 끝부분보다는 안정구간을 측정위치로 많이 사용하는데 Strange(1987)는 모음 안정구간보다는 시작부분과 끝부분에 더 중요한 정보가 들어있음을 지적했다. 실제 안정구간을 묵음으로 하고 시작부분과 끝부분만 들려준 결과 많은 사람들이 정확히 모음을 지각했음을 밝힌 바 있다.

2. 합성음에 대한 포먼트 측정

2.1 합성음 제작 방법

분석소프트웨어가 포먼트값을 바르게 측정하는지 판단하기 위해서는 사람의 목소리보다는 기계로 만든 합성음을 이용하는 것이 편리하다. 합성음은 프라트의 기본 메뉴 가운데 Create sound from formula를 이용하여 1 초 분량의 44.1 kHz로 된 표본속도의 음성을 다음 공식을 입력하여 만들었다.

$$1/3*\sin(2*\pi*500*x)+1/5*\sin(2*\pi*1500*x)+1/7*\sin(2*\pi*2500*x)+1/9*\sin(2*\pi*3500*x)+1/11*\sin(2*\pi*4500*x)$$

위의 공식은 기본적으로 500 Hz, 1500 Hz, 2500 Hz, 3500 Hz, 4500 Hz으로 된 정현파를 하나의 소리로 만든 복합음으로 5 개의 포먼트를 가진 모음과 비슷한 구조를 보여준다. 합성음을 사람의 음성과 비슷하게 만들기 위해 고주파로 갈수록 진폭값이 낮아지도록 각 항의 분수값을 점점 낮추었다. 이 합성음의 확대한 음성파형과 스펙트로그램은 <그림 1>과 같다. <그림 2>는 합성음의 스펙트럼을 프라트로 분석하여 5000 Hz까지 확대하여 보여준다.

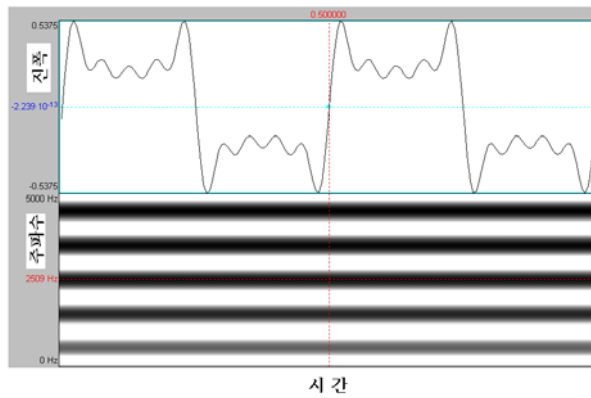


그림 1. 500 Hz의 홀수배음으로 된 5개의 정현파로 만든 합성음의 음성파형과 스펙트로그램

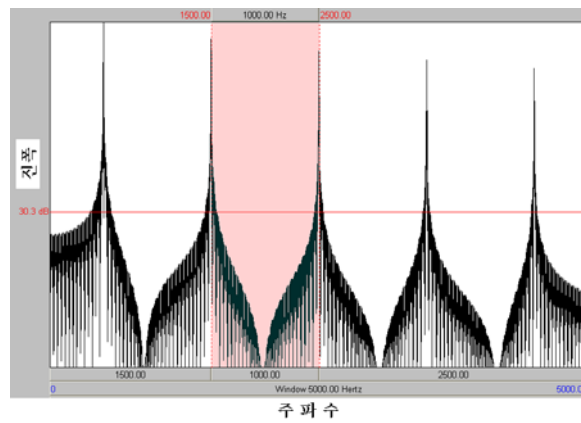


그림 2. 500 Hz의 홀수배음으로 된 5개의 정현파로 만든 합성음의 스펙트럼

2.2 실험 자료

실험 자료는 앞 절에서 제시한 합성음 제작 방법을 이용하여 10 명의 남자가 발음한 국어모음 “이, 에, 우, 오, 아”의 평균 포먼트값(Yang, 1996:251, Table 3 참고)을 가지는 소리를 만들었다. 연구자가 합성음을 직접 들어봤을 때 해당모음을 듣는 것 같은 청각적인 인상을 받았다. SenSyn, SpeechStation, SFS 등과 같은 기존의 포먼트 합성기를 이용할 때 수많은 변수값을 제대로 지정해야하고, 서로 인접한 포먼트값의 상호 충돌로 인한 잡음이 많이 생겨서 음성분석기의 성능을 점검하는데 이용하기가 불편한 점을 감안하면 이러한 합성음은 음성분석시스템 점검을 하는데 편리할 것으로 생각된다.

2.3 포먼트 추정 및 논의

프라트의 포먼트 추정방식에는 청각적으로 변환하는 것을 제외하고는 To Formant(burg), To Formant(sl)...과 To LPC(autocorrelation)...의 세 가지 방식으로 나눌 수 있다. To Formant(burg)의 처리과정은 먼저 최대 포먼트값의 두 배에 해당하는 표본추출속도로 변경한 뒤 고주파대역 강조를 적용하고, 가우시안창을 적용하여 LPC 계수를 Burg방식으로 계산한다(Press 외, 1992). 이 알고리즘은 50 Hz이하에 해당하는 포먼트값은 제거 한다. To Formant(sl)...메뉴는 Split-Levinson의 처리방식으로 되는데 지정한 주파수 범위에 강제로 일정한 개수의 포먼트를 찾아내도록 처리해준다(Willems, 1986). To LPC(autocorrelation)...는 자기상관방식으로 처리한 뒤 선형예측계수를 지정한 시간간격마다 구하여 포먼트에 해당하는 정점값을 차례로 나타내주며, 실험자는 선형예측계수를 구하기 위한 차수를 지정해야한다. 차수는 보통 포먼트 갯수의 두 배가 되도록 지정한다. 차수가 높을수록 보다 더 많은 포먼트값을 보여주기 때문에 한 개의 포먼트가 나타날 위치에 두 개 이상을 구해주게 된다(양병곤, 2003). 차수가 낮으면 모음 “우, 오”와 같이 두 개의 포먼트값이 가까이 접근한 경우에 하나의 포먼트로 처리하는 오류가 발생하기도 한다. 이 논문에서는 기본형인 16차를 분석 설정으로 하였다. 모든 포먼트값은 1초의 음성구간의 중간점인 0.5초 지점에서 구했다.

<표 1>은 원래 합성음에 사용한 포먼트 주파수와 합성음의 포먼트를 재분석한 burg방식, sl방식, 16차의 선형예측계수로 분석한 lpc방식으로 구한 포먼트값을 보여주고 있다. 원래의 합성음에 사용한 포먼트값과 구해진 값의 절대차이값을 모두 더했다. 전체적인 차이합계는 27505 Hz에 달한다. burg 방식의 차이합을 모두 더하면 107 Hz이고, sl 방식의 차이는 128 Hz이며, lpc방식에서 27270 Hz가 나와서 세 가지 방식 중에서는 lpc방식에서 차이가 가장 큼을 알 수 있다. 물론 lpc방식의 차수를 다르게 지정하거나, sl방식에서 포먼트 갯수를 원래 예측값보다 더 작거나 큰 수를 지정하면 차이합계는 훨씬 더 많아질 것으로 예상할 수 있다. 포먼트값에서는 F1에서 F3으로 갈수록 차이합계가 증가하고 있는데 이 값의 주된 원인은 lpc방식이 주도하고 있다. 실제 burg나 sl방식에서는 그 반대로 F1에서는 75~85 Hz차이를 보이고, F2에서는 13~44 Hz이나 F3은 두 가지 방식 모두 9 Hz의 차이밖에 나타나지 않아 반대 경향을 보여주고 있다. 모음별로는 전설모음 “이, 에”가 5554 Hz의 차이합계를 보인 반면 후설모음 “우, 오”에서는 그것의 세 배인 16786 Hz를 보여 후설모음에서 lpc 방식이 전체적인 변화량을 보임을 알 수 있다. 이는 lpc분석 방식에서 두 개의 포먼트가 가까이 있는 경우에는 두 개의 포먼트를 한 개로 처리함으로써 편차가 많이 나게 된 것으로 추정된다. 이렇게 포먼트 분석값은 분석방식과 파라미터 설정방법에 따라 항상 달라질 수 있음을 알 수 있다. 이 연구에서는 프라트의 burg방식을 이용하여 대학생들이 발음한 모음 목록에서 어느 측정지점에서 포먼트값의 변화가 가장 많은지를 살펴보고, 스펙트럼 정보에서 본 포먼트 추정값의 정확성을 확인하며 실제 포먼트 위치를 찾으려면 어떻게 해야 할 지 조사해보고자 한다.

표 1. 합성음의 포먼트값과 프라트의 기본 포먼트 분석방법(*burg*, *sl*, *lpc*)을 이용한 포먼트값과 절대차이합계

(단위: Hz)

모음	분석방법	F1	F2	F3	차이합계
이	합성음	341	2219	3047	0
	<i>burg</i>	342	2218	3044	5
	<i>sl</i>	342	2218	3044	5
	<i>lpc</i>	1915	3754	5473	5535
이 차이합계		1576	1537	2432	5545
에	합성음	591	1849	2597	0
	<i>burg</i>	591	1849	2599	2
	<i>sl</i>	591	1849	2599	2
	<i>lpc</i>	592	1850	2600	5
에 차이합계		1	1	7	9
우	합성음	369	981	2565	0
	<i>burg</i>	300	974	2566	77
	<i>sl</i>	413	999	2566	63
	<i>lpc</i>	1980	4631	7063	9759
우 차이합계		1724	3675	4500	9899
오	합성음	453	945	2674	0
	<i>burg</i>	466	944	2675	15
	<i>sl</i>	481	965	2675	49
	<i>lpc</i>	1823	3658	5414	6823
오 차이합계		1411	2734	2742	6887
아	합성음	738	1372	2573	0
	<i>burg</i>	740	1376	2575	8
	<i>sl</i>	740	1377	2575	9
	<i>lpc</i>	1120	3319	5392	5148
아 차이합계		386	1956	2823	5165
이에 차이합계		1577	1538	2439	5554
우오 차이합계		3135	6409	7242	16786
전체합계		5098	9903	12504	27505

3. 대학생 발화 모음에 대한 포먼트 측정

2.1 피험자와 음성수집

이 연구에 참여한 피험자는 대학교에 재학 중인 남학생 10 명과 여학생 10 명이다. 이들은 전용 연구실에서 삼성 DB-P73컴퓨터에 Sennheiser(PC151) 헤드셋 마이크를 연결하여 자신의 이름을 말하고 모음 “아, 에, 우, 오, 이, 으”를 보통 속도로 또렷하게 발음한 것을 GoldWave(v.

5.25)를 이용해 44.1 kHz의 표본속도로 저장했다. 마지막 발음 “으”는 분석에 포함하지 않고 다만 앞에 오는 모음 “우”와 동일하게 발음한 피험자를 실험에서 제외하는데 사용했다.

2.2 자료 분석 방법

자료 분석은 피험자가 발음한 모음 다섯 개(“아, 에, 우, 오, 이”)를 To Formant(burg)의 기본 설정 값으로 주어진 5500 Hz까지 5 개의 포먼트를 찾는 방식을 이용해 분석했다. 서론에서도 살펴본바와 같이 추정지점에 따라 포먼트값이 달라지기 때문에, 이 논문에서는 Yang(1996)에서 모음의 전체지속시간의 1/3지점에 해당하는 값을 구했던 점을 응용하여 균일한 간격의 7 개 지점마다 포먼트값 3 개(F1, F2, F3) 씩을 구했다. 구체적으로는 다음의 프라트 스크립트를 이용하여 피험자의 음성을 불러온 다음 모음 구간을 마우스로 선택하고, 선택된 구간 시작에서 30 ms, 끝에서 40 ms를 제외한 모음의 전체지속시간을 6 등분한 간격으로 시작점을 포함하여 7 개 지점의 포먼트값을 구했다. 20 명의 화자마다 다섯 개의 모음을 발음했고, 추정지점 7 개에서 각각 3 개의 포먼트값을 구하여 모두 2100 개의 자료를 구했다.

```
clearinfo
print Subject_1_이 'newline$'
name$=selected$("Sound")
Edit
  editor Sound 'name$'
  pause 분석할 모음구간을 선택하세요...
  cursorstart=Get start of selection
  cursorend=Get end of selection
  Select... 'cursorstart'-0.03 'cursorend'+0.03
  Extract sound selection (time from 0)
  Close
endeditor
Rename... vowel
dur=Get duration
start=0.06
end='dur'-0.07
voweldur='end'-'start'
step='voweldur'/6
To Formant (burg)... 0 5 5500 0.025 50
for i from 0 to 6
  timer='start'+i*'step'
  f1=Get mean... 1 timer timer+0.02 Hertz
  f2=Get mean... 2 timer timer+0.02 Hertz
  f3=Get mean... 3 timer timer+0.02 Hertz
  p='i'+1
  print 'p''tab$''f1:0''tab$''f2:0''tab$''f3:0''newline$'
endfor
fappendinfo D:\result\formant.txt
select Sound vowel
plus Formant vowel
Remove
select Sound 'name$'
```

선택 구간의 시작과 끝 부분을 피한 이유는 사전연구로 검토해 본 결과 25 ms로 된 포먼트 분석창에 묵음구간이 포함되어 측정값의 편차가 너무 많았기 때문이다. 덧붙여 각 측정 지점에 따른 변화를 줄이기 위해 20 ms 구간을 선택하여 평균 포먼트값을 구했다. 7 개 지점에서 측정된 포먼트값은 마이크로소프트 엑셀을 이용하여 3 번째 측정지점 값을 중심으로 절대차이합계를 구하여 남녀, 모음, 측정위치 등에 따른 변화 정도를 먼저 분석하였다. 3 번째 지점은 모음 구간의 지속시간을 3 등분했을 때 앞의 1/3 부분 해당한다. 참고로 프라트의 To Pitch(ac)...로 구한 남학생 10 명의 평균 피치값은 134.5 Hz이었고, 여학생의 평균 피치값은 226.8 Hz이었다. 남학생의 피치값의 범위는 111 Hz~153 Hz이었고, 여학생의 범위는 201 Hz~254 Hz이었다. 남녀 각각의 피치의 표준편차는 15.2 Hz와 15.9 Hz로 거의 비슷한 집단을 이루고 있었다.

2.3 분석결과와 논의

<표 2>는 성별, 모음별, 포먼트별 모음 포먼트값의 절대차이합계를 보여준다.

표 2. 남(m)녀(f), 모음별, 포먼트별 모음 포먼트값의 절대차이합계의 분포(단위: Hz)

모음	F1m	F2m	F3m	합계(m)	F1f	F2f	F3f	합계(f)
이	690	2420	2983	6093	932	8122	8620	17674
에	1151	2354	2580	6085	1644	8101	12372	22117
오	2092	14434	5720	22246	1020	2050	5409	8479
우	3176	13984	9257	26417	2116	7983	5169	15268
아	2490	3506	3894	9890	2766	2911	12944	18621
합계	9599	36698	24434	70731	8478	29167	44514	82159

<표 2>에서 보면 남학생 10 명에서 측정된 절대차이합계는 70731 Hz인데 비해 여성화자 10 명의 절대차이합계는 82159 Hz로서 남성보다 많다. 이는 여성 포먼트값이 대체로 높기 때문에 3 번째 지점값을 중심으로 절대차이합계가 높아졌기 때문으로 추정된다. 모음별로는 남성에서는 모음 “우”가 가장 높은 값을 차지하고 있고 모음 “에”가 가장 적은 차이합계를 보였고, 여성에서는 오히려 모음 “오”에서 변화량이 적고, 모음 “에”에서 가장 높은 절대차이합계를 보였다. 이는 포먼트 갯수가 5 개로 지정됨으로써 여학생이 발음한 모음 “오”를 분석하는 데는 적절하지만, 남녀 학생의 모음 “우”에서는 더 많은 갯수로 처리해야 할 것으로 여겨진다. 포먼트 별로 보았을 때, 남성에서는 모음 “오, 우”에서 두 개의 포먼트를 하나로 측정하는 오류가 많아서 F3을 F2로 측정하여 차이합계가 높아졌으며, 여성에서는 F1의 8478 Hz에 비해 F2는 3 배 이상, F3는 5 배 이상의 절대차이합계를 보였다. 이는 여성 화자의 F3나 F4가 매우 낮은 강도의 포먼트대역을 가지고 있어서 프라트에서 F4를 F3 값으로 잘못 추정한 경우가 많았기 때문으로 여겨진다.

표 3. 3번째 측정위치를 기준으로 본 모음 포먼트 추정지점별 절대차이합계의 분포 (단위: Hz)

측정위치	F1m	F2m	F3m	합계(m)	F1f	F2f	F3f	합계(f)
1	1736	7375	4610	13721	1299	5037	5403	11739
2	1503	6977	3215	11695	1084	6080	6972	14136
3	0	0	0	0	0	0	0	0
4	938	4741	4375	10054	700	4192	7879	12771
5	1374	3682	2923	7979	1045	4073	6450	11568
6	1731	5401	4170	11302	1649	4714	8633	14996
7	2317	8522	5141	15980	2701	5071	9177	16949

<표 3>은 포먼트 추정지점별로 구한 절대차이합계를 보여준다. 3 번째 측정지점 값을 기준으로 남학생들에서는 5 번째 측정지점에서 절대차이합계가 7979 Hz로 낮고, 7 번째 모음이 끝나는 지점 가까이에서 가장 높은 절대차이합계를 보였다. 여학생에서는 5 번째 측정지점에서 가장 낮은 절대차이합계를 보였고, 남학생들과 같이 7 번째 가장자리의 측정값의 절대차이합계가 가장 높았다. 이런 차이는 주로 어디서 발생하는 것일까? 이 문제의 해답을 찾기 위해 개별 남녀학생들의 분석 자료를 살펴보았다.

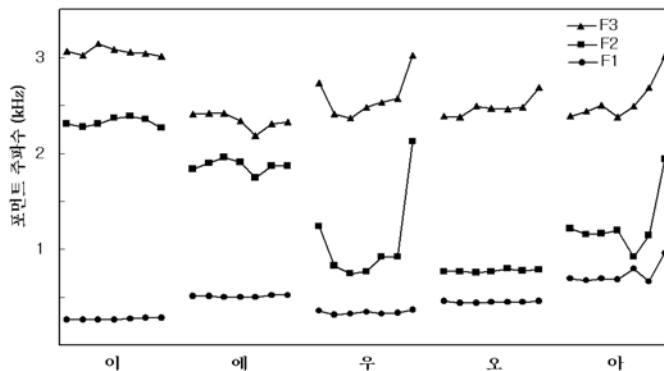


그림 3. 한 남학생이 발음한 모음의 포먼트 분석 결과

<그림 3>은 9079 Hz의 절대차이합계를 보인 남학생이 발음한 모음의 포먼트 분석 결과를 보여준다. 이 학생은 모음 “이”에서는 495 Hz의 차이를 보였고, “에”에서는 1180 Hz, 모음 “우”에서는 4009 Hz, 모음 “오”에서는 646 Hz, 모음 “아”에서는 2749 Hz의 절대차이합계를 보였다. 그림에서 보듯이 모음 “우”에서는 처음과 끝부분이 매우 변화가 심한 것을 볼 수 있고, 끝부분의 F2는 실제 F3값임을 알 수 있다. “아”에서도 음성과형에 작은 진폭을 보였던 끝의 세 개 지점에서 변화가 많음을 알 수 있다. 실제 포먼트의 변화는 성도의 모양이 바뀌는 것을 반영하는데, 그림의 “우, 아”와 같이 혀의 모양을 반영하는 F2가 갑작스레 변하기는 어렵기 때문에 측정오류로 판단해야 할 것이다.

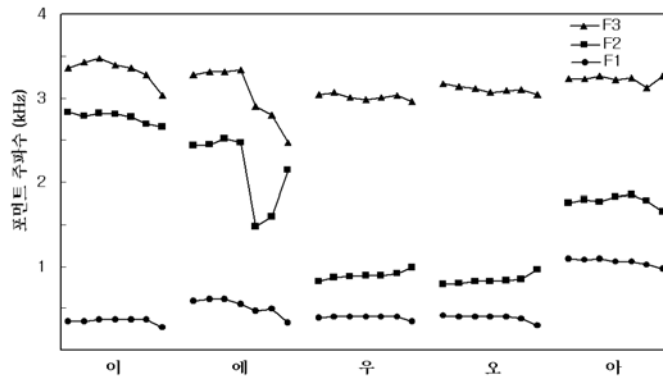


그림 4. 한 여학생이 발음한 모음의 포먼트 분석 결과

<그림 4>는 10069 Hz의 절대차이합계를 보인 여학생이 발음한 모음의 포먼트 분석 결과를 보여준다. 이 학생은 모음 “이”에서 1677 Hz의 차이를 보였고, “에”에서는 6361 Hz, 모음 “우”에서는 524 Hz, 모음 “오”에서는 632 Hz, 모음 “아”에서는 875 Hz의 절대차이합계를 보였다. <그림 4>에서 모음 “우, 오”는 포먼트값이 매우 안정적으로 측정되었음을 알 수 있고, 모음 “에”의 F2 값이 측정지점 5 번과 6 번에서 갑자기 떨어지는 모양을 보이고 있는데 이는 측정상의 에러라고 보아야 할 것이다. 이렇게 연속적으로 변하는 포먼트 궤적을 보면 갑작스런 변화지점에서 구한 포먼트값이 맞는지 확인하는데 도움이 될 것이다.

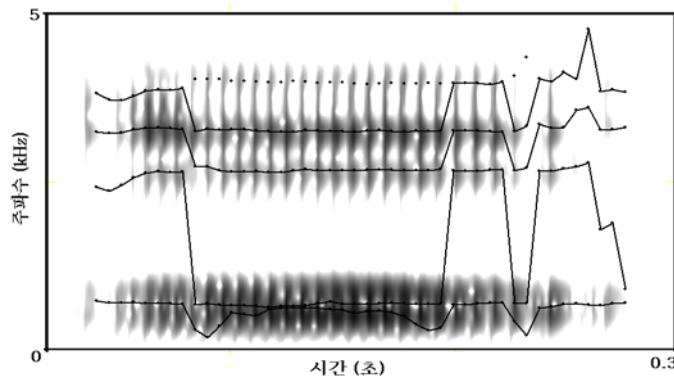


그림 5. 한 남학생이 발음한 모음 “오”의 스펙트로그램과 포먼트 궤적. 포먼트 갯수를 5개로 지정했음.

<그림 5>는 모음 “오”에서 가장 많은 오류를 보인 남학생의 스펙트로그램과 포먼트 궤적을 보여 준다. burg분석 설정에서 포먼트 갯수를 5개로 했기 때문에 F1과 F2 모두가 매우 가까이 접근하는 모양을 보여주고 있다. <그림 6>은 이 학생의 음성에 대한 분석 설정에서 포먼트 갯수를 6개로 바꾸어 포먼트 궤적을 추적한 그림이다. 스펙트로그램에 나타난 검은 포먼트 띠 중심부로 정확하게 F1과 F2가 분리되어 표시되었다.

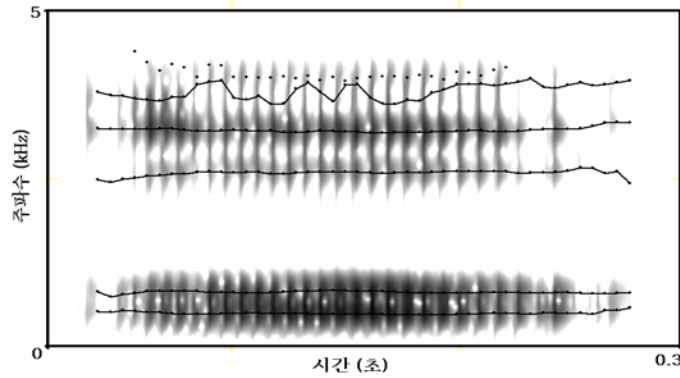


그림 6. 한 남학생이 발음한 모음 “오”의 스펙트로그램과 포먼트 궤적. 포먼트 갯수를 6개로 지정했음.

<그림 7>은 모음 “이”의 F2에서 5363 Hz로 가장 많은 오류를 보인 여학생의 스펙트로그램과 포먼트 궤적을 보여준다. 이런 지점은 비록 원래의 음성에 대해 고주파증폭을 하긴 했지만, F5부분의 스펙트럼에너지가 너무 낮고, 그림에 표시된 지점의 강도값이 상대적으로 높기 때문에 F2값으로 선택한 것으로 생각된다. 그래서 F2값이 스펙트로그램 상 포먼트 궤적이 전혀 없는데도 불구하고 오르내리는 모양을 보여준다.

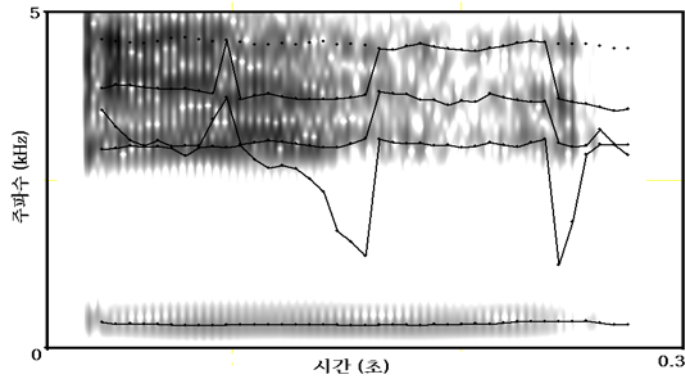


그림 7. 한 여학생이 발음한 모음 “이”의 스펙트로그램과 포먼트 궤적. 포먼트 갯수를 5개로 지정했음.

이 여학생의 포먼트 분석 설정에서 포먼트 갯수를 4개로 바꾸면 <그림 8>과 같이 F2값과 F3값을 제대로 구해준다. 결국 포먼트 갯수는 스펙트로그램의 검은 띠로 나타난 포먼트 띠의 갯수를 화자마다 확인하고 적절하게 조절해야만 바른 측정값을 구할 수 있다.

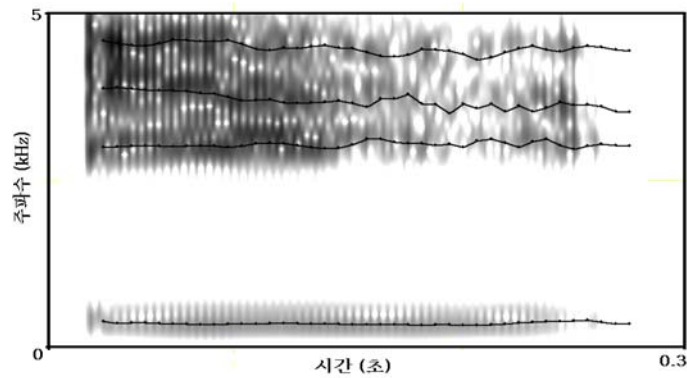


그림 8. 한 여학생이 발음한 모음 “이”의 스펙트로그램과 포먼트 궤적. 포먼트 갯수를 4개로 지정했음.

지금까지 두 남녀 학생의 포먼트 분석에서 나타난 오류의 예와 같이, 포먼트의 적절한 갯수를 판단하기 위해서는 <그림 5>와 같이 두 개의 포먼트가 너무 가까이 근접하거나 측정값들이 변화량이 심한 경우 또는 <그림 7>과 같이 스펙트로그램 상 포먼트가 없는 곳에 나타난 경우들을 찾아 분석 설정을 바꾸거나, 항상 스펙트로그램을 바탕에 놓고 측정값을 확인할 필요가 있다. 앞으로 프라트로 모음을 측정했을 때 개별 화자들에 대한 분석 오류의 비율이 얼마나 많을 지에 대한 연구가 필요하고, 포먼트 갯수 설정을 적절히 하기 위해 화자별 피치값이나 포먼트값의 범위 등 음향적 특징을 이용하는 방법에 대한 연구가 필요할 것으로 생각된다.

4. 맺음말

이 연구에서는 포먼트값 측정상의 문제점을 검토해보기 위해 프라트의 합성음 생성기로 모음과 유사하게 만든 합성음을 프라트의 세 가지 포먼트 분석방식을 적용하여 측정하고 원래의 값과의 절대적인 차이값을 구하여 어떤 문제점이 있는지 살펴보고, 20 명의 대학생 남녀학생들이 발음한 모음의 포먼트값을 burg방식으로 분석하여, 전체지속시간의 1/3 지점을 기준으로 모음구간 앞, 뒤의 7 개 측정지점에서 구한 포먼트값과의 절대차이합을 계산해 변화량을 측정하였고, 조음 동작으로 생각할 수 없는 포먼트 측정상의 오류를 수정하기 위해서는 어떻게 해야 할 지 살펴보았다. 그 결과는 다음과 같다.

첫째, 합성음의 포먼트 분석에서 프라트의 burg 방식에서 가장 적은 측정오류가 나왔다. sl방식이나 l pc방식에서는 인접한 포먼트를 가지는 모음이나 포먼트 갯수에 따른 오류가 크게 나타났는데, 프라트에서 제공하는 이 두 가지 방식을 이용하여 포먼트 분석을 할 때는 주의가 필요함을 알 수 있었다.

둘째, 20 명이 발음한 다섯 개 모음의 절대차이합계를 구해본 결과 포먼트값이 대체로 높게 나타나는 여성에게서 높은 변화량을 보였다. 모음 가운데서도 남학생과 여학생 모두의 모음 “우,

오”에서 높은 변화량을 보였고, 주된 원인으로는 두 개의 포먼트를 하나로 추정하는 오류나, F4를 F3으로 추정한 경우였다.

셋째, 남녀화자마다 적절한 포먼트 갯수를 지정해야만 정확한 포먼트를 구할 수 있음을 알 수 있었다. 개인별 자료를 살펴본 결과 남학생이 발음한 모음 “오”에서는 포먼트 갯수를 더 많이 지정했을 때 F1과 F2를 바르게 측정해 주었으며, 여학생이 발음한 모음 “이”에서는 오히려 포먼트 갯수를 더 적게 지정했을 때 F2의 에러를 줄일 수 있었다. 포먼트 갯수는 개인별로 스펙트로그램에서 볼 수 있는 검은 띠의 중심선으로 제대로 추정되었는지, 또 주위의 값에 비해 갑작스런 포먼트 변화가 일어나지 않았는지를 통해 확인할 필요가 있다.

이러한 결과들은 모음 포먼트 측정을 할 때 음성분석소프트웨어에서 구해진 값이 정확한지 반드시 검토해야 한다는 결론을 내릴 수 있다. 가능하면 잘못된 측정값에 근거를 두고 일반화를 하는 오류를 피해야 할 것이다. 앞으로 이렇게 측정된 포먼트값을 합성하여 원음과 어떤 부분에서 청각적으로 차이가 나는지 또 적절한 포먼트 갯수를 화자의 음향적 특징에 따라 지정하는 방안을 연구해볼 필요가 있다.

참 고 문 헌

- 강석한 역. 2007. *음향음성분석론*. 서울: Thomson Learning Korea, 박학사.
- 양병곤. 2003. *프라트(Praat)를 이용한 음성분석의 이론과 실제*. 부산: 만수출판사.
- Edward, H. 2003. (3rd ed.) *Applied phonetics: The sounds of American English*. New York: Delmar Learning.
- Fulop, S. 2007. “What’s wrong with these formants?.” *J. Acoust. Soc. Am.* 121, 3099.
- Kent, R. & Read, C. 2002. (2nd ed.) *Acoustic analysis of speech*. San Diego, CA: Singular Publishing Group.
- Lindblom, B. 1963. “Spectrographic study of vowel reduction.” *J. Acoust. Soc. Am.* 35, 1773-1781.
- Olive, J.P., Greenwood, A. & Coleman, J. 1993. *Acoustics of American English speech: A dynamic approach*. New York: Springer-Verlag.
- Peterson, G. & Barney, H. 1952. “Control methods used in a study of vowels.” *J. Acoust. Soc. Am.* 24, 175-184.
- Picket, J. 1987. *The sounds of speech communication: a primer of acoustic phonetics and speech perception*. Austin, Texas: pro-ed.
- Press, W., Teukolsky, S., Vetterling, W. & Flannery, B. 1992. *Numerical recipes in C: the art of scientific computing*. Cambridge, MA: Cambridge University Press.
- Stevens, K. 1998. *Acoustic phonetics*. Cambridge, MA: The MIT Press.
- Willems, L. 1986. “Robust formant analysis.” *IPO report 529*, 1-25. Eindhoven. Institute for Perception Research.
- Yang, B. 2002. “An acoustical study of English word stress produced by Americans and Koreans.” *Speech Sciences* 9(1), 77-88.

접수일자: 2008. 7. 24

수정일자: 2008. 8. 26

게재결정: 2008. 9. 10

▲ 양병곤

부산광역시 금정구 장전동 산 30 (우: 609-735)

부산대학교 사범대 영어교육과 교수

H.P.: 010-9618-7636

E-mail: bgyang@pusan.ac.kr

Website: <http://fonetiks.info/bgyang>