

화자 인식을 위한 모음의 포먼트 연구*

안병섭(고려대), 신지영(고려대), 강선미(서경대)

<차 례>

- | | |
|----------------------|------------------------|
| 1. 서 론 | 3. 결과 및 논의 |
| 2. 연구 방법 | 3.1 모방 발화 |
| 2.1. 모방 발화 | 3.1.1. 모음의 포먼트 특징 |
| 2.1.1. 피험자 | 3.1.2. 포먼트 차이값 |
| 2.1.2. 실험 문장 및 녹음 과정 | 3.1.3. 효용성 검토: F-ratio |
| 2.1.3. 자료 분석 | 3.2. 위장 발화 |
| 2.2. 위장 발화 | 3.2.1. 모음의 포먼트 특징 |
| 2.2.1. 피험자 | 3.2.2. 포먼트 차이값 |
| 2.2.2. 실험 문장 및 녹음 과정 | 3.2.3. 효용성 검토: F-ratio |
| 2.2.3. 자료 분석 | 4. 결 론 |

<Abstract>

A Study on Formants of Vowels for Speaker Recognition

Byoung-seob Ahn, Jiyoung Shin, Sunmee Kang

The aim of this paper is to analyze vowels in voice imitation and disguised voice, and to find the invariable phonetic features of the speaker. In this paper we examined the formants of monophthongs /a, u, i, o, ʊ, ε, ʌ/. The results of the present are as follows : ① Speakers change their vocal tract features. ② Vowels /a, ε, i/ appear to be proper for speaker recognition since they show invariable acoustic feature during voice modulation. ③ F1 does not change easily compared to higher formants. ④ F3-F2 appears to be constituent for a speaker identification in vowel /a/ and /ε/, and F4-F2 in vowel /i/. ⑤ Resulting of F-ratio, differences of each formants were more useful than individual formant of a vowel to speaker recognition.

Keywords: Voice imitation, Disguised voice, formant, speaker recognition, individuality

* 본 논문은 한국과학재단 목적기초연구(R01-1999-000-00229-0) 지원으로 수행되었습니다.

1. 서 론

사람의 목소리는 언어적(linguistic) 정보뿐만 아니라 감정이나 태도와 같은 준언어적(paralinguistic) 정보 및 성별이나 나이 등과 같은 발화자의 개인성(individuality)을 담고 있다. 따라서 청자는 화자의 목소리를 듣고 발화 의미 및 화자의 감정이나 심리 상태뿐만 아니라 그 화자가 누구인지도 알아낼 수 있다. 이러한 말소리의 특징은 음성을 시각적으로 분석할 수 있게 됨으로서 화자 인식(speaker recognition)을 가능하게 하는 음향적 단서가 무엇인지를 밝혀 내려는 연구로 이어졌다.

화자 인식에 이용되는 개인성(individuality) 정보에는 성도 길이나 성대의 특성과 같이 발성 기관의 선천적인 개인차에 의한 것과 방언적 특성과 같은 후천적인 발성 습관에 기인한 것이 있다. 전자에 대한 연구는 주로 포만트와 기본주파수를 대상으로 진행되어 왔으며, 후자는 지역 방언이나 개인 방언의 특성을 이루는 언어적 특징들을 대상으로 하였다.[1][2][3][4][5] 그런데 후천적인 발성 습관은 개인의 노력 여하에 따라 화자가 쉽게 변화시킬 수 있을 뿐만 아니라 그 특징을 기계적으로 처리하기에는 여러 면에서 어려운 점이 있다. 그래서 상대적으로 고정적인 값을 보이는 발성 기관의 선천적인 개인차를 이용한 화자 인식 연구가 주로 진행되어 왔다. 특히 모음의 각 포만트를 이용해 화자를 인식하려는 연구가 많았는데, 그 결과 F1과 F2는 모음의 종류를 구분 짓는 중요한 자질이며, F3 이상의 고차 포만트는 모음의 종류보다는 상대적으로 화자의 개인성과 관련한 것으로 밝혀지기도 했다.[6][7] 그러나 고차 포만트일수록 발화에 따라서 혹은 심지어 단일 발화내에서도 변화의 폭이 커서 고차 포만트 정보를 화자 인식에 이용하는 데 어려움이 있다.[8]

따라서 본 연구는 모음 포만트를 이용한 화자 인식의 이러한 한계를 극복하기 위하여 다양한 발화에서도 적용될 수 있는 화자 특징을 찾는 데 목적을 두었다. 이를 위해서 본 연구에서는 변조된 음성(modulated voice)을 분석 대상으로 하였으며, 변조된 음성을 본인 발화와 비교하는 연구를 진행하였다. 본 연구는 음성 변조의 방법으로 특정 대상인의 음성을 모방하는 모방 발화(voice imitation)와 자신의 정체체를 숨기기 위해 목소리를 바꾸는 위장 발화(disguised voice)를 사용하였다. 모방 발화와 위장 발화는 음성 변조라는 측면에서는 동일하지만 서로의 발화 전략은 다르다. 모방 발화는 특정 대상인의 음성을 목표로 삼기 때문에 모방 발화자는 목표 음성에 접근하기 위하여 발성 유형뿐만 아니라 성도의 변형까지 구사한다. 반면에 위장 발화자는 주로 발성 유형을 변화시키는 발화 전략을 구사한다.

이 논문에서는 모음의 포만트 값과 포만트 간의 차이값을 계산하여 각 측정값들의 화자 내 변이(intraspeaker variability)와 화자 간 변이(interspeaker variability)를 비교한 후 포만트 간의 차이값이 화자 특징이 될 수 있음을 밝힐 것이다. 아울러 그 결과를 바탕으로 화자 인식을 위한 비밀 단어 설정에 유리한 모음이 무엇인지

를 제안하고자 한다.

모음 포먼트에 대한 이러한 연구는 인간의 음성적 특징을 좀더 잘 이해하는 과정이 될 것이며, 한 걸음 더 나아가서는 음성공학 분야나 범음성학(forensic phonetics) 분야의 발전에 기여할 수 있는 정보를 제공할 수 있을 것으로 기대된다.

2. 연구 방법

2.1. 모방 발화

2.1.1. 피험자

모방 발화 연구를 위한 피험자들은 단일 화자의 음성을 모방하는 집단과 복수 화자의 음성을 모방하는 집단으로 구분하여 선정하였다.

단일 목표 음성의 모방을 위해서는 피험자를 공개 모집하고 지원한 사람들 가운데 음성 모방 능력이 보통 이상의 수준으로 판단되는 네 명을 선정하였다. 이 가운데 강한 기식성으로 인해 모음의 포먼트가 잘 관찰되지 않는 한 명을 제외하였다.

복수 목표 음성에 대한 모방은 두 명의 피험자를 선정하였다. 복수 목표 음성에 대한 모방은 특출한 모방 능력이 필요하므로, 성대 모사를 위한 인터넷 동호회에서 회장으로 활동하고 있는 피험자와 MBC 예능 프로그램에서 성대 모사를 위주로 활동하고 있는 개그맨을 섭외하였다.

모방 발화 녹음에 참여한 피험자 정보는 다음의 <표 1>과 같다.

<표 1> 모방 발화의 피험자 정보

피험자	나이	신장(cm)	성장지	성별	모방 방법
화자 A	19	174	서울	남	복수 모방
화자 B	25	177	서울	남	복수 모방
화자 C	26	176	서울	남	단일 모방
화자 D	23	178	서울	남	단일 모방
화자 E	24	176	서울	남	단일 모방

2.1.2. 실험 문장 및 녹음 과정

모방 발화를 위해서는 모방의 대상이 되는 목표 발화가 필요한데, 목표 발화의 음성 자료로는 청와대 인터넷 홈페이지에 구축된 김대중 전 대통령의 '2003년 신년사'를 이용하였다. 김대중 전 대통령의 음성적 특징을 많은 사람들이 기억하고 있어서, 다른 사람의 음성보다 김대중 전 대통령의 음성을 피험자들이 공통적으로 모방하기가 용이할 것으로 판단했기 때문이다. 22개 문장으로 구성된 목표 음성 자료를 텍스트와 함께 녹음 일주일 전에 피험자들에게 전달하여 피험자들이 충분히 연습할 시간을 주었다. 단일 모방 피험자들은 22개 문장과 실험을 위해 선정한 30개 단어를 본인 발화와 모방 발화로 2회 반복하여 녹음하였다.¹⁾ 그리고 복수 모방 피험자들은 본인 발화와 피험자들이 자신 있게 구사할 수 있는 대상자들의 음성을 모방하여 각각 2회 반복하여 녹음하였다. 피험자들이 모방한 모방 대상자는 다음의 <표 3>과 같다.

<표 3> 피험자들의 모방 대상자

피험자	모방 대상자
화자 A	김대중, 권영길, 노무현, 멀더(극중 성우 음성), 앙드레김, 이회창, 정몽준, 전도연, 엄앵란
화자 B	김대중, 김제동, 박준형, 신문선
화자 C	김대중
화자 D	김대중
화자 E	김대중

2.1.3. 자료 분석

피험자들의 음성 자료는 Cool Edit Pro를 이용하여 16000Hz와 16bit로 디지털화하였으며, wavesurfer 1.5.2로 분석하였다. 포먼트 측정은 단어 및 문장 내의 동일한 환경에 위치한 각 모음의 안정 구간(F2 기준)의 정중앙을 측정하였다. 모음의 포먼트 값들은 wavesurfer의 formant plot 기능을 이용하였으며, 각 측정 지점의 FFT값과 대조하였다.

본 연구에서는 모음 /a, ʌ, ε, o, u, ʊ, i/를 분석 대상으로 하였으나, 예비 실험 결과 모음 /ʌ, o, u, ʊ/는 화자 및 발성 유형에 따라 포먼트의 일관된 측정이 어려

1) 모방 발화와 위장 발화에 사용된 단어와 문장 목록은 논문 끝에 부록으로 덧붙였다.

왔다. 이 가운데 모음 /u, u/는 내재적 길이가 비고모음에 비해 상대적으로 짧아 발성 환경이나 발성 종류에 따라 F2 이상에서 변이가 컸다. 게다가 기식성이 강한 화자의 발화에서는 F3과 F4의 측정이 어려웠다. 비고모음 /o, ɔ/에서는 화자 모두 두 모음의 위치가 고모음 쪽으로 편향되어 있어(특히 /o/) 앞에서 언급한 고모음의 특징이 대부분 관찰되었다.

이에 비해 모음 /a, ε, i/는 안정 구간이 뚜렷하게 나타나며, 분절음 환경이나 발성 유형에 거의 영향을 받지 않아 고차 포만트의 측정이 안정적이었다. 이 세 모음 가운데 /a/가 분절음 환경이나 발성 종류에 관계없이 포만트 측정에 가장 안정적이었다. 고모음인 /i/는, /u, u/와 달리 짧은 내재적 길이에도 불구하고 안정 구간을 보이고 있었고 고차 포만트가 비교적 잘 관찰되었다. 그러므로 본 연구에서는 측정의 안정성이 보장되고 발성 환경의 영향이 적은 이 세 모음을 중심으로 연구를 진행하였다.²⁾

2.2. 위장 발화

2.2.1. 피험자

위장 발화 연구를 위한 자료를 얻기 위하여 남성 열 명의 발화를 녹음하였다. 이 가운데 본인 발화나 위장 발화 시 강한 기식성으로 인해 포만트 측정이 어려운 두 명의 자료는 분석의 대상에서 제외하였다. 위장 발화의 피험자 정보는 다음의 <표 2>와 같다.

<표 2> 위장 발화의 피험자 정보

피험자	나이	신장(cm)	성장지	성별
화자 1	25	182	서울	남
화자 2	36	184	서울	남
화자 3	27	167	서울	남
화자 4	31	170	서울	남
화자 5	29	170	서울	남
화자 6	26	170	서울	남
화자 7	24	178	인천	남
화자 8	24	177	경기	남

2) 우리말 단모음 /a, ε, i, o, u/를 대상으로 벡터 양자화를 이용해 실험한 [9]의 결과에서도 모음 /i/, /ε/, /a/에서 높은 인식률을 보였다.

2.2.2. 실험 문장 및 녹음 과정

위장 발화를 위해서는 단모음 7개가 모두 고려된 11개 단어와 이것을 틀문장 '나는 []라고 말한다'에 넣은 11개 문장을 정상 발화, 위장 발화 유형별로 2회 반복하여 녹음하였다. 위장 방법으로는 '가성, 저음, 짜내기'의 세 가지를 피험자에게 요구하였다.

2.2.3. 자료 분석

모방 발화의 방법과 동일한 방법으로 녹음된 음성 자료를 디지털화하였다. 또한 모음의 포먼트도 모방 발화와 동일한 방법으로 측정하였으며, 측정 역시 모방 발화와 동일하게 모음 /a, ε, i/를 대상으로 하였다.

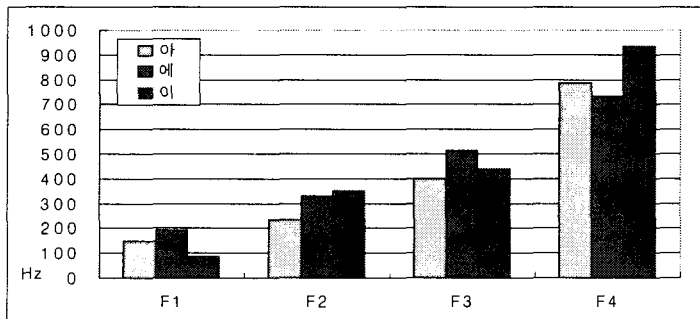
위장 발화의 피험자 가운데는 실험자가 요구한 위장 방법을 성공적으로 수행하지 못한 경우가 있었다. 예를 들면 화자의 2의 경우 '짜내기' 위장을 하지 못하고 가성으로 발화하였는데, 이런 경우의 음성 자료는 분석에서 제외하였다. 또한 화자가 적극적으로 위장을 하는 과정에서 포먼트 측정이 불가능할만큼 강한 기식을 사용한 경우가 있었다. 이런 예는 주로 '가성'을 사용한 위장 발화 자료인데, 이런 경우의 음성 자료도 분석에서 제외하였다.

3. 결과 및 논의

3.1 모방 발화

3.1.1. 모음과 포먼트 특징

모방 발화의 모음 포먼트를 측정한 결과 피험자 모두의 경우에서 모음과 발화의 종류에 관계없이 F1은 비교적 일정한 값을 보인 반면 F2, F3, F4는 큰 변화 범위를 보였다.



<그림 1> 모방 발화의 포먼트 변화 범위

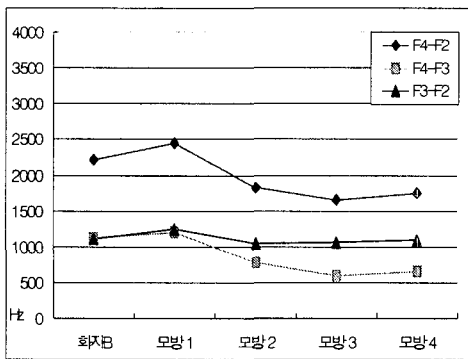
<그림 1>은 피험자들의 모음별 포만트 값의 평균을 나타낸 것인데, 이 결과는 [6][7][10]의 실험과도 일치한다. 다만 이 결과는 모방 발화를 대상으로 하였기 때문에 정상 발화만을 대상으로 연구한 [6]의 결과보다 각 포만트의 변화 범위가 크게 나왔다. 화자별로 살펴 볼 때도 이러한 경향은 지켜져 고차 포만트로 갈수록 그 변화 범위가 더 크게 나타났다.

3.1.2 포만트 간의 차이값

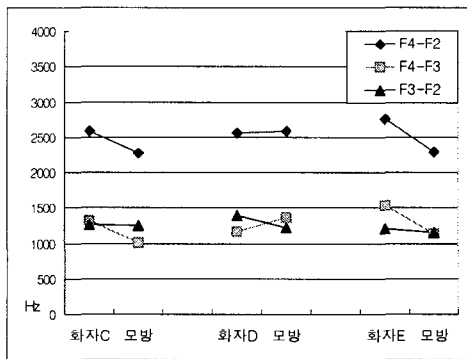
포만트는 성도의 특정 공명강의 형태를 반영한 음향적 값들이다.[11] 성도의 길이가 어느 정도 일정하다고 보면 모음에 따라서 만들어지는 공명강의 크기는, 일정한 성도의 공간을 각 공명강들이 분할하는 것으로 볼 수 있다. 따라서 모음마다 공명강들이 그 크기를 조절하는 데는 특정 공명강끼리 상관성이 있을 것으로 보았다.

특정 공명강들의 상관성을 알아보기 위해 본 연구에서는 각 포만트 간의 차이값을 계산하였다. 포만트 간의 차이값을 계산할 때는 다른 포만트 값에 비해 상대적으로 변이가 작은 F1 값을 제외하였다. F1 값은 다른 포만트에 비해 일정한 값을 보이므로 포만트 값의 변화 정도가 큰 다른 포만트와 상관성을 가지기 어려울 것으로 판단했기 때문이다. 따라서 포만트 간의 차이는 F3-F2, F4-F2, F4-F3만을 대상으로 하였다.

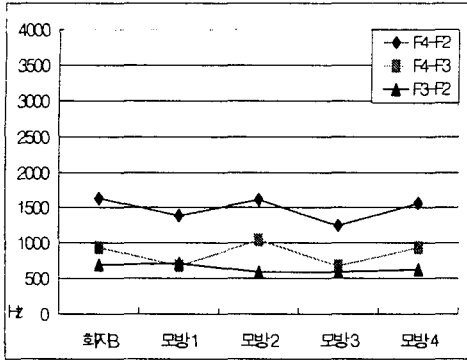
분석 결과 모음 /a, ε/에서는 F3-F2 값이, 모음 /i/에서는 F4-F2 값이 상대적일 일정한 것으로 나타났다.



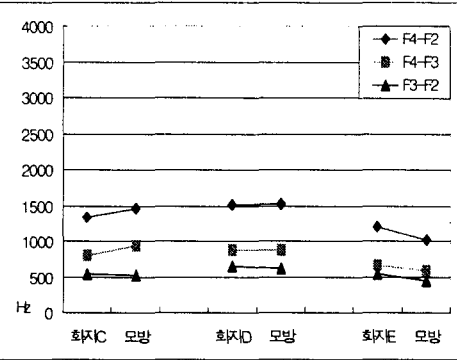
<그림 2> 화자 B의 /a/ 포만트 차이값



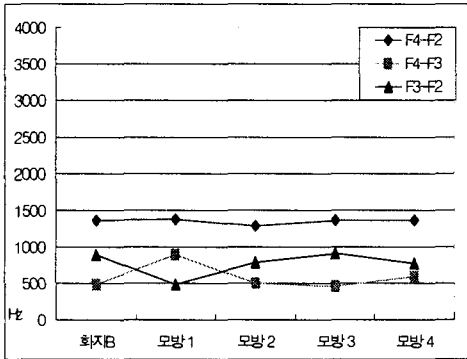
<그림 3> 화자 C~E의 /a/ 포만트 차이값



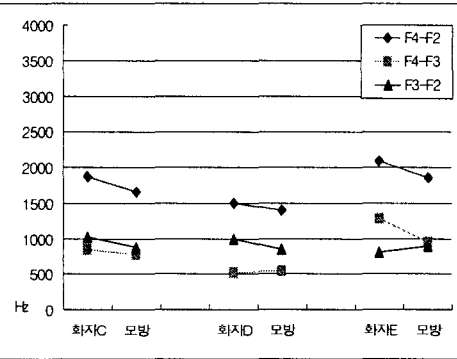
<그림 4> 화자 B의 /ɛ/ 포먼트 차이값



<그림 5> 화자 C~E의 /ɛ/ 포먼트 차이값



<그림 6> 화자 B의 /i/ 포먼트 차이값



<그림 7> 화자 C~E의 /i/ 포먼트 차이값

<그림 3>의 화자 D는 /a/ 모음의 경우에서 F3-F2 값보다 F4-F2 값이 더 일정한 것으로 나타났고, <그림 7>의 화자 C의 모음 /i/에서도 F3-F2와 F4-F3이 F4-F2보다 오히려 더욱 고정된 값을 보여주기도 하였다. 그러나 이러한 실현들은 화자 개별적이어서 화자 전체에서 관찰되지는 않는다. 결국 화자 간 공통된 특징은 모음 /a/와 /ɛ/에서는 F3-F2 값이, 모음 /i/에서는 F4-F2가 상대적으로 안정된 값들을 가진다는 것이다.

3.1.3. 효용성 검토 : F-ratio

화자 인식에 이용되는 특징들은 무엇보다도 다음과 같은 조건을 갖추고 있어야 한다. ① 화자 간 정보를 효과적으로 나타낼 수 있어야 한다. ② 측정이 쉬워야 한다. ③ 시간에 대해 안정된 값을 가져야 한다. ④ 음성신호에서 자연적이고 반복적으로 관찰되어야 한다. ⑤ 발성 환경이 바뀌어도 영향이 적어야 한다. ⑥ 쉽게 모방할 수 없어야 한다.[12] 이러한 조건들을 많이 충족할수록 좋은 요소라고

할 수 있다.

모음의 포맷트는 앞에서 언급한 결과처럼 고차 포맷트로 갈수록 변이 정도가 커서 화자 인식의 한 자질로 이용하기에는 무리가 있다. 더구나 일정한 값을 보여 주었던 F1은 화자 간 변별성이 크지 않았기 때문에 화자 인식의 자질(feature)로 이용하기에는 위험성이 크다. 화자 간 변별성이 크지 않다는 점은 승인 오류(false acceptance)를 범할 가능성이 크기 때문이다.[12]³⁾ 따라서 단순히 모음의 포맷트를 이용하여 화자 인식을 하기에는 다음과 같은 문제점이 지적될 수 있다.

- ① F1이 화자 내 변이가 가장 작으나 화자 간 변이가 크지 않아서 신체 구조가 비슷하거나 자매 간, 형제간의 경우에서 화자 간 변별이 더욱 어렵다.
- ② 고차 포맷트가 개인성을 보여주는 화자 특징이라는 기존의 연구에도 불구하고 화자 내 변이가 크기 때문에 화자 인식에 이용하는 것은 무리가 있다.
- ③ 고차 포맷트일수록 발성 유형에 따라, 특히 기식성의 강한 정도에 따라 포맷트 측정에 어려운 경우가 있다.

따라서 포맷트 이용의 문제점들을 보완해 줄 새로운 자질이 필요한데, 필자들은 본 연구에서 다루었던 포맷트 차이값들을 화자 인식에 한 자질로 추가하기를 제안한다. 그렇게 하기 위해서는 먼저 포맷트 차이값이 화자 인식에서 과연 효율적인 자질이 될 수 있는지를 검토해야 하는데 본 연구에서는 F-ratio 값을 이용하고자 한다.

F-ratio는 화자 인식 시스템에서 특징 변인의 효용성을 측정하는 데 사용되는 척도이다.[12][14] 원래 F-ratio는 좋은 자질을 선택하기 위한 목적보다 나쁜 특징을 가려내는 목적으로 많이 사용되는데, 본 연구에서는 이 척도의 기능을 역으로 이용해 포맷트 값과 포맷트 차이값 가운데 어떤 것이 화자 인식의 효용성 측면에서 우위를 보이는가를 검토하고자 한다. F-ratio 값은 다음과 같이 정의된다.

$$F = \frac{\text{전체 화자들의 평균의 변이값 (variance of speaker means)}}{\text{화자 각각의 변이의 평균 (mean intraspeaker variance)}}$$

결국 F-ratio는 화자 내 편차로 화자간의 편차를 나눈 것과 동일한 효과인데, F-ratio 값이 클수록 상대적으로 효용성이 큰 자질이라 할 수 있다.[12] 그것은 화자 내 편차는 작을수록, 화자 간 편차는 클수록 화자 인식에서 유리하기 때문이다. 모방 발화의 각 포맷트 값과 포맷트 차이값에 대한 F-ratio를 계산한 결과는 다음과 같다.

3) 화자 인식의 오류에는 거부 오류(false rejection)와 승인 오류(false acceptance)가 있는데, 승인 오류가 거부 오류보다 더 큰 문제를 야기 시킨다.

<표 4> 모방 발화의 포먼트 값과 포먼트 차이값에 대한 F-ratio

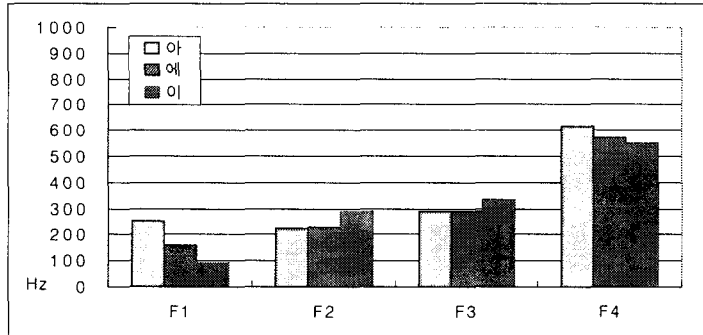
	/a/	/ɛ/	/i/
F 1	0.79	0.66	1.29
F 2	0.78	0.83	0.47
F 3	0.27	0.35	0.85
F 4	0.54	0.93	0.82
F4 - F2	1.31	1.82	1.80
F4 - F3	1.05	1.60	1.58
F3 - F2	1.12	2.10	1.21

<표 4>를 보면 세 모음 모두에서 포먼트 차이값이 포먼트 값보다 효용성에서 우위를 점하고 있다. F3-F2가 일정한 값을 보여 화자 특징적이었음을 나타냈던 모음 /a/에서는 F-ratio 측정 결과 F4-F2가 F3-F2보다 상대적으로 효용성이 뛰어난 것으로 밝혀졌다. 모음 /ɛ/와 /i/에서는 포먼트 차이값에서 화자별로 가장 일정한 값을 보여주었던 F3-F2(/ɛ/), F4-F2(/i/)가 가장 높은 효용성을 보여주었다. 특이한 것은 포먼트 값 가운데 모음 /i/의 F1에 대한 F-ratio 값이 포먼트 차이값의 F-ratio 결과만큼이나 높은 점수로 나타났다는 것인데, 이것은 모음 /i/의 F1 값이 7개 변인 가운데 가장 작은 화자 간 변이를 가졌음에도 불구하고 화자 내 변이가 상대적으로 더욱 작았기 때문이다.

3.2. 위장 발화

3.2.1. 모음의 포먼트 특징

위장 발화에서도 모음별, 화자별 포먼트 값들의 변화 양상이 모방 발화와 같은 경향으로 나타났다. 다음은 위장 발화 피험자들의 모음별 포먼트 값의 평균의 평균을 나타낸 것이다.

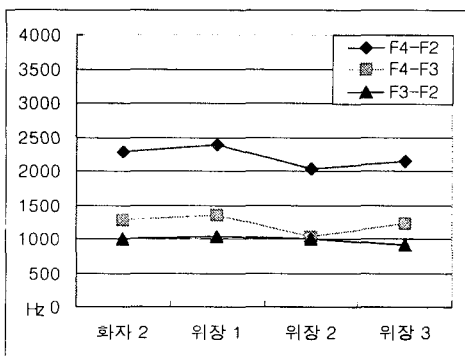


<그림 8> 위장 발화의 포맷트 변화 범위

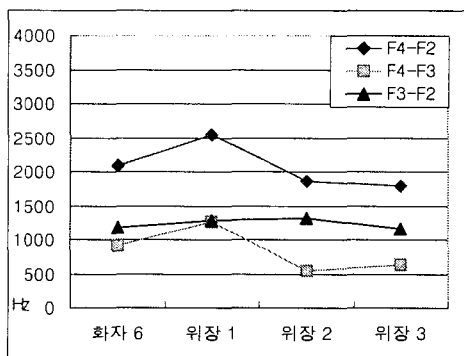
위장 발화의 포맷트 변화 범위의 전체적인 경향성은 모방 발화와 비슷하지만 변이 정도가 모방 발화보다 전반적으로 크지 않은 것으로 나타났다. 그것은 모방 발화와 위장 발화의 발화 전략이 다르기 때문인 것으로 판단되는데, 위장 발화는 성도 변형보다는 발성 유형을 주로 변화시킴으로써 자신의 목소리를 숨기는 음성 변조 방법이라면 모방 발화는 모방 목표에 따라 발성 유형의 변화와 함께 성도의 크기를 적극적으로 변화시키는 음성 변조 방법이기 때문이다. 특히 모음간에 동일하게 F3과 F4의 변화가 모방발화보다 현저하게 작게 나타났는데, 이 점에 미루어 성도를 적극적으로 변화시키면 고차 포맷트 값의 변이가 커진다는 사실을 확인할 수 있었다.

3.2.2 포맷트 간의 차이값

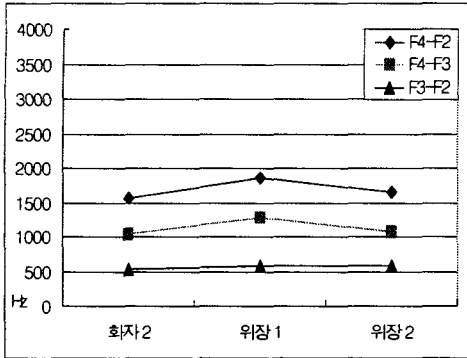
모방 발화와 마찬가지로 위장 발화에서도 모음 /a/, /ε/에서는 F3-F2 값이, 모음 /i/에서는 F4-F2 값이 일정한 것으로 나타났다.



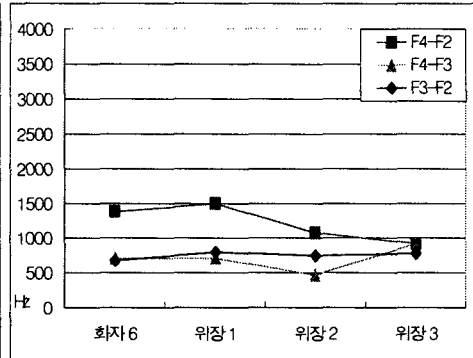
<그림 9> 화자 2의 /a/ 포맷트 차이값



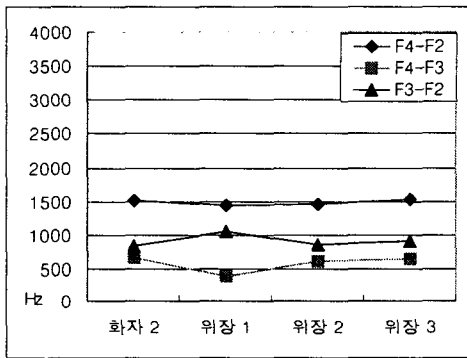
<그림 10> 화자 6의 /a/ 포맷트 차이값



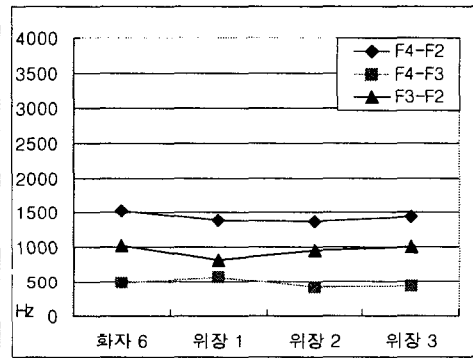
<그림 11> 화자 2의 /ɛ/ 포먼트 차이값



<그림 12> 화자 6의 /ɛ/ 포먼트 차이값



<그림 13> 화자 2의 /i/ 포먼트 차이값



<그림 14> 화자 6의 /i/ 포먼트 차이값

위장 발화에서도 모방 발화와 마찬가지로 화자에 따라서는 모음 /a/나 /ɛ/에서 F4-F2, F4-F3 값들이 모음 /i/에서는 F3-F2값이 F4-F2과 같이 일정하게 나타나기도 하지만 전반적인 경향은 아니며 화자 개별적이라고 할 수 있다. 결국 모음별로 일관된 값들을 나타내는 특징은 모방 발화와 마찬가지로 모음 /a/, /ɛ/에서는 F3-F2 값이, 모음 /i/에서는 F4-F2 값이다.

3.2.3. 효용성 검토 : F-ratio

위장 발화에서 7개 측정값들에 대한 F-ratio를 계산하여 비교한 결과, 위장 발화에서도 모방 발화와 같이 모음의 포먼트 값보다는 포먼트 차이값이 더욱 효용성이 있는 것으로 나타났다. <표 5>는 그 측정 결과이다.

<표 5> 위장 발화의 포맷 값과 포맷 차이값에 대한 F-ratio

	/a/	/ɛ/	/i/
F 1	0.47	0.66	0.95
F 2	0.98	0.64	1.06
F 3	0.35	0.67	1.24
F 4	0.90	0.92	1.27
F4 - F2	1.39	1.09	1.46
F4 - F3	1.14	1.04	0.97
F3 - F2	1.34	1.68	0.55

모음별로도 모방 발화와 같은 결과가 나왔는데, /a/에서는 F4-F3가, /ɛ/에서는 F3-F2가, /i/에서는 F4-F2가 모음별 가장 높은 결과로 나타났다. 모음 /a/에서 F4-F3가 가장 높은 값으로 나타났지만 F3-F2 값과 근소한 차이(0.05)밖에 나지 않으므로 모음 /a/의 경우 F3-F2 값도 효용성 측면에서 유용하다 하겠다. 그런데 한 가지 눈여겨볼 것은 점은 /i/ 모음의 각 포맷의 F-ratio 값이 모방 발화나 위장 발화에서의 다른 모음인 /a/나 /ɛ/와 큰 차이로 높은 값을 나타내고 있다는 점이다. 비슷한 경우로 모음의 포맷을 대상으로 한 화자 인식 연구인 [9]의 결과에서도 모음 /i/에서 가장 높은 인식률을 보여 주었는데, /i/ 모음의 각 포맷들이 다른 모음보다 화자 내에서 작은 변이를 가지기 때문인 것으로 풀이된다.

4. 결 론

본 연구에서는 본인 발화와 변조된 발화(모방 발화, 위장 발화)를 대상으로 하여 포맷 값과 포맷 차이값을 비교하고, 두 특징에 대한 효용성을 검토하였다.

우선 측정 대상인 7개 모음 중에서 /a/, /ɛ/, /i/가 /o, u, ʊ, ɪ/보다 발화 종류나 화자에 상관없이 측정에 안정적이어서 화자 인식에 적합한 모음으로 보았다.

모음의 포맷들은 화자 내 변이가 커서 화자 특징적이지 못하다는 결론을 얻었으며, 화자 내 변이가 가장 작았던 F1 조차도 화자 간 변이가 작아 화자 특징적이라고 할 수 없었다. 특히 기존 연구에서 화자 개인성과 유관한 것으로 알려진 고차 포맷트는 큰 화자 내 변이로 인해 화자 인식에 고차 포맷트를 이용하기는 어려울 것으로 보았다.

다음으로 비교했던 것은 포맷 간의 차이값이다. 그 결과 모음별로 일정한 값

을 나타내는 변인이 있었는데, 모음 /a/와 /ε/에서는 F3-F2가, 모음 /i/에서는 F4-F2가 화자 간에서 일정한 값들로 나타났다.

마지막으로 포먼트와 포먼트 차이값에 대한 효용성을 검증해 보았다. 이 변인들에 대한 상대적인 효용성 검증을 위해 화자 인식에서 이용되는 F-ratio 측정을 하였는데, 각 포먼트 값보다는 포먼트 차이값이 상대적으로 높은 값을 보였다. 포먼트 차이값 중에서도 화자별 일정한 값을 보여주었던 모음 /ε/의 F3-F2(/ε/)가, 모음 /i/의 F4-F2가 높은 점수를 나타내어 역시 효용성 측면에서도 다른 변인보다 우월했음을 알 수 있었다. 반면에 모음 /a/에서는 포먼트 차이값 결과와 달리 F4-F2가 가장 뛰어난 효용성을 가진 것으로 나타났지만 F3-F2의 F-ratio 값에서 근소한 차이밖에 나지 않으므로 여전히 효용성이 있는 것으로 보인다. 그러므로 모음의 포먼트를 이용한 화자 인식에서 단순히 모음의 포먼트값을 이용하는 것보다는 포먼트의 차이값을 병용하는 것이 효과적이라고 할 수 있다.

향후 연구에서는 포먼트 간 차이값들이 과연 성공적인 화자 변별력을 가지는가에 대해서 정상 발화의 대단위 음성 자료를 대상으로 한 다양한 방법의 검증이 이루어져야 할 것이다.

참고문헌

- [1] 차일환·조병모·윤대희, “성문에 의한 화자 인식에 관한 연구” 논문집(연세대 산업기술연구소) 21-2, pp. 313~316, 1989.
- [2] E. Wetterholm, “The Significance of Phonetics in Voice imitation, In Proc” *Australian International Conference on Speech Science and Technology*, Canberra, pp. 342~347, 2000.
- [3] J. Laver, *Principles of Phonetics*, Cambridge: Cambridge University Press, 1994.
- [4] 김영일·차일환, “한국어 화자 인식에 관한 연구” 논문집 17-1,(연세대학교 산업기술연구소), 1985.
- [5] 박미영·신지영·강선미, “위장·모방 발화에서 나타나는 기본주파수 변화 양상에 관한 연구” *한국어학(한국어학회)* 24, pp. 65~98, 2004.
- [6] 유영화·박종철·김윤희, “Sound Spectrograph에 의한 우리말 단모음 분석에 관한 연구” *국립과학수사연구소연보* 17, pp. 222~254, 1985.
- [7] Do-Hong Ko, “Voice Similarities between Sisters” *Speech Sciences* 8-3, pp. 43~50, 2001.
- [8] 안병섭·신지영·강선미, “모방 발화의 모음 포먼트 연구” *대한음성학회 2004년 봄학술대회 발표집*, 2004.
- [9] 김민수 외, “모음의 특성을 고려한 화자인식 시스템 구현” *자연과학논문집* 19(한양대학교 자연과학연구소), pp. 113~117, 2000.
- [10] A. Eriksson and P. Wretling, “How Flexible is the Human Voice?-A Case study of Mimicry” *In Proc. Eurospeech* Vol 2, pp. 1043~1046, 1997
- [11] 신지영, *말소리의 이해*. 한국문화사, pp. 175~189, 2000.

- [12] 김진현 · 홍수기 · 장성길, “자동화자식별법에 관한 연구” *국립과학수사연구원연보* 28, pp. 478 ~ 488, 1996.
- [13] F. Nolm, “Speaker Recognition and Forensic Phonetics” *The Handbook of Phonetic Sciences*, Blackwell Publishers Ltd, pp. 744 ~ 767, 1997.
- [14] D.O'Shaughnessy, “Speaker recognition” *IEEP ASSP magazine*, pp. 4 ~ 17. 1986.

접수일자 : 2004년 8월 22일

게재결정 : 2004년 9월 13일

▶ 안병섭(Byoung-seob Ahn)

주소: 136-701 서울시 성북구 안암동 5가 1번지 고려대학교 민족문화연구원 A 107호

소속: 고려대학교 민족문화연구원 음성언어정보연구실

전화: 02) 3290-2505

FAX: 02) 926-8385

E-mail: abs001@empal.com

▶ 신지영(Jiyoung Shin)

주소: 136-701 서울시 성북구 안암동 5가 1번지 고려대학교 국어국문학과

소속: 고려대학교 문과대학 국어국문학과

전화: 02) 3290-1973

FAX: 02) 926-8385

E-mail: shinjy@korea.ac.kr

▶ 강선미(Sunmee Kang)

주소: 136-704 서울특별시 성북구 정릉동 16-1 서경대학교 컴퓨터과학과

소속: 서경대학교 컴퓨터과학과

전화: 02) 940-7291

FAX: 02) 940-7291

E-mail: smkang@skuniv.ac.kr

[부 록]

□ 모방 발화의 문장 목록

1. 존경하고 사랑하는 국민 여러분!
2. 희망의 2003년, 새아침이 밝았습니다.
3. 국민 여러분 모두에게 더 큰 행복과 희망의 한해가 되기를 중심으로 기원합니다.
4. 전 세계에 600만 해외 동포와 2200만 북한 동포들에게도 따뜻한 새해 인사를 전합니다.
5. 돌이켜 보면, 지난 한해는 엄청난 변화 속에 우리 대한민국 위상이 크게 도약하는 뜻깊은 해였습니다.
6. 월드컵 4강 신화와 남북이 하나 된 부산 아시안 게임이 있었습니다.
7. 역사상 가장 공명한 지방 선거와 대통령 선거도 실시했습니다.
8. 세계가 부러워하는 경제 발전도 있었습니다.
9. 감동의 한해였습니다.
10. 국민 여러분이 해냈습니다.
11. 깊은 감사의 말씀을 드리는 바입니다.
12. 국민 여러분! 올해는 현 국민의 참여 속에 선출된 노무현 대통령 당선자의 새 정부가 출범하는 해입니다.
13. 우리 모두 힘을 모아 새 정부를 지원해야겠습니다.
14. 그리하여 현재 제기되고 있는 북한 핵 문제를 해결하여 한반도 평화를 정착 시켜야겠습니다.
15. 경제와 민생을 더 한층 공고히 발전시킵시다.
16. 우리 함께 손을 잡고 나아갑시다.
17. 우리 후손들에게 세계 일류 국가, 자랑스런 대한민국을 물려줍시다.
18. IMF 외환위기 극복과 월드컵 4강 신화의 저력과 자신감을 가지고 힘차게 견인합시다.
19. 저는 남은 임기 동안 국정을 안정적으로 마무리하여 다음 정부가 보다 나은 여건에서 출발할 수 있도록 최선을 다하겠습니다.
20. 다시 한번 여러분께서 보내주신 그간의 협력에 대해 깊이 감사 드리며 계속되는 성원을 바라마지 않습니다.
21. 국민 여러분 가정에 만복이 깃드시기를 기원합니다.
22. 감사합니다.

□ 모방 발화의 단어 목록

열매, 노릇노릇, 마무리, 국민, 한남동, 산나물, 무당벌레, 아동문제, 동네노인, 노란 나비, 감사합니다, 여보세요, 고맙습니다, 안녕하세요, 반갑습니다, 기원합니다, 가정에, 국민 여러분, 한반도, 올해는, 역사상, 사랑하는, 선출된, 국정을, 월드컵, 다시 한번, 모두에게, 견인합니다, 2003년.

□ 위장 발화의 단어 목록

단 어 : 아가, 주거, 옥이, 개그, 어가, 기구, 수고, 우기, 애기, 가게, 이기
 틀문장 : 나는 []라고 말한다.