

# 위장발화의 단모음 포먼트 연구

노석은\* 박미경\* 조민하\* 신지영\*\* 강선미\*\*\*  
\* 고려대학교 민족문화연구원 음성언어정보연구실  
\*\*고려대학교 국어국문학과  
\*\*\*서경대학교 컴퓨터학과

## A Study on the Vowel Formants in Disguised Speech

Seokeun Noh\*, Mikyong Park\*, Minha Cho\*, Jiyoung Shin\*\*\* and Sunmee Kang\*\*\*\*

\* Spoken Language Information Lab, Institute of Korean Culture, Korea Univ.

\*\* Department of Linguistics and Literature, Korea Univ.

\*\*\* Department of Computer Science, Seokyeong Univ.

senoh@korea.ac.kr, miky@korea.ac.kr, freejj@hanmail.net,  
shinjj@korea.ac.kr, smkang@skuniv.ac.kr

### Abstract

The aim of this paper is to analyze the acoustic features for disguised voice. In this paper we examined the features such as pitch range, vowel formants(F1, F2, F3, F4) So the result of the analysis is as follows. : ① Pitch range and average of pitch value is very important cue for speaker verification. ② F3-F2 is also important cue for speaker verification ③ /a/ is more verified than other vowels.

### I. 서론

사람의 음성의 고유한 특징을 나타내주는 단서들로 기본주파수, 모음의 공명주파수(formant), 음의 길이, 음의 강도와 같은 것들이 있다고 알려져 왔다. 음의 높낮이를 나타내는 기본주파수는 성대의 길이, 두께 등 성대구조의 개인차, 성대를 지지하고 있는 후두연골의 크기와 형상, 성대근의 형상이나 동작 방법, 개인의 호기량 등에 따라 달라지기 때문에 화자의 특징을 잘 나타내 주는 특성으로 간주되어 왔다. 기본 주파수가 변화하는 모습은 억양을 보여주기 때문에, 화자의 발화습관에 대해 알 수 있다. 또 모음의 공명 주파수

역시 성도의 특징을 반영하기 때문에 모음의 종류를 구별 가능하게 하는 한편 그 절대값들은 사람에 따라 달라지는 모습을 보인다. 그뿐 아니라 발화 습관에 따라 좀더 명료한 발음을 구사하는 사람과 그렇지 못한 사람의 경우 역시 모음의 공명 주파수를 통해 알 수 있다. 음의 길이는 전체적인 발화 속도와 관련 깊다. 음의 강도 역시 개인차가 존재한다는 것을 경험적으로도 충분히 알 수 있을 것이다

이러한 단서들은 자음보다 모음에서 더 쉽게 그 특징을 드러낸다 따라서 본고는 특히 단모음을 중심으로 하여, 한 화자의 목소리를 특징짓는 데 주요한 역할을 하는 단모음의 음향적 특징은 어떤 것이 있는지 살펴보고, 이러한 특징들이 화자가 목소리를 변화시킬 때, 또 화자가 달라지는 경우에 어떻게 변화하는지 살펴볼 것이다.

### II. 연구 방법

본 연구에서는 보통발화와 위장발화를 비교하기 위해 서울·경기 출신 2, 30대 남녀 총 20명에게 목록을 제시하고,<sup>1)</sup> 자신의 정상 발화, 자유 위장발화, 최대한 음

본 논문은 한국과학재단 목적기초연구  
(R01-1999-000-00229-0) 지원으로 수행되었습니다  
1 남 10, 여 10명이다

의 높이를 높여서 발화, 최대한 낮게 발화, 짜내기(creaky) 소리로 발화, 속삭이는(whispered) 소리로 발화, 코를 막고 발화하는 방법을 사용하여 지시에 따라 2회씩 낭독하게 하였다.

대상 모음의 발화는 국어의 대표적 단모음인 /아/, /어/, /에/, /오/, /우/, /으/, /이/를<sup>2)</sup> 모음과 모음 사이에 적당한 휴지를 두어 단독발화하게 하였다.

녹음은 고려대학교 민족문화연구원 음성언어정보연구실 내에 있는 녹음실에서 하였고 오스트리아 AKG사의 C420B-Lock 마이크와 TASCAM DA-20 MK II DAT를 사용하여 녹음하였다. 16,000Hz, 16bit로 디지털화했으며 Centre for Speech Technology의 Wavesurfer 1.5.7을 이용해 녹음 자료를 분석하였다.

자료의 분석은 기본주파수의 측정과 7개 단모음의 포먼트 측정을 중심으로 하였다.

### III. 분석 결과

#### 1. 위장 수단의 선택

화자들이 최대한 자신의 목소리와 다르게 내기 위해 사용한 위장 수단으로 가성, 음높이 내림, 짜내기, 숨소리, 콧소리 등이 있었다.

위장 수단으로 가성을 사용한 경우가 9명으로 전체 20명 중에 절반에 가까운 수를 차지하였다. 단순히 음높이만 올린 것이 아니라, 가성에 약한 짜내기를 섞거나 숨소리(breathy voice)로 나타나는 경우도 있었다. 여성 화자들의 경우 가성을 주된 위장 수단으로 사용한 사람은 4명이었지만 음높이를 내리는 위장 수단을 사용한 경우를 제외하고, 다른 위장 수단(비음, 짜내기)을 사용하더라도 모두 평균 기본주파수가 올라가는 것을 알 수 있었다.

음높이를 내리는 것을 주요 위장 수단으로 사용한 경우는 모두 4명으로 이 중에서 3명이 여성이었다. 반면 숨소리를 위장 수단으로 사용한 것은 남성만 2명이었다. 콧소리로 발화한 경우는 남녀 2명씩 4명이었는데, 이 중 2명은 코막음을 하고 콧소리가 난 경우이다.

#### 2. 기본주파수

실험에서 문장이 아닌 단모음의 단독 발화를 요구했기 때문에 화자 내에서 기본주파수의 변화가 크지 않고 단조로운 모습을 보였다. 따라서 기본 주파수(F0) 곡선의 윤곽보다는 평균 기본주파수를 중심으로 살펴

보았다. 속삭이는 소리의 경우, 기본주파수를 측정할 수 없었기 때문에 제외하게 되었다.

정상 발화에서 남성 화자의 평균 기본 주파수는 99Hz ~ 129Hz, 여성 화자의 평균 기본 주파수는 164Hz ~ 214Hz 사이에서 나타났다.

지시에 따라 위장 수단을 변경하여 발화하였을 때, 코막음 소리에서는 정상 발화와 평균 기본 주파수의 차이가 거의 없었다. 소리를 최대한 낮추어서 발화하였을 때 여성화자의 경우 평균 기본 주파수는 150Hz ~ 180Hz, 남성 화자의 경우는 86Hz ~ 106Hz 정도로 낮게 나타났다. 자신의 낼 수 있는 최대한의 낮은 소리를 내라고 했을 때 여성 화자의 평균 기본 주파수라 하더라도 정상 발화의 남성 평균 기본 주파수의 범위보다도 높게 나타나는 것을 알 수 있다.

가성의 경우 남성 화자는 191Hz ~ 393Hz로 평균 기본 주파수가 화자에 따라 폭넓게 나타났다. 여성 화자의 경우 268Hz ~ 417Hz까지 나타났다. 다른 위장 수단에서 남성 화자의 평균 기본 주파수 범위와 여성 화자의 기본 주파수 범위가 겹치는 부분이 거의 없는 반면 가성의 경우 분명 여성의 기본 주파수 영역이 남성에 비해 높기는 하지만 화자에 따라서는 남성이 여성과 유사한 정도로 평균 기본 주파수를 올릴 수 있다는 것을 알 수 있다.

<그림 1>은 위장 수단이 달라질 때 여성 화자별 평균 기본 주파수의 변화를 보인 것이다. 정상 발화의 기본 주파수가 화자별로 변화하는 것에 비해 가성이나 짜내기의 평균 기본 주파수는 화자에 따라 훨씬 큰 폭으로 변화하고 있음을 알 수 있다. 저음의 경우는 정상 발화의 변화에 비하여서도 화자 별 차이가 크지 않았다. 이러한 경향은 남성 화자의 경우 역시 마찬가지였다.

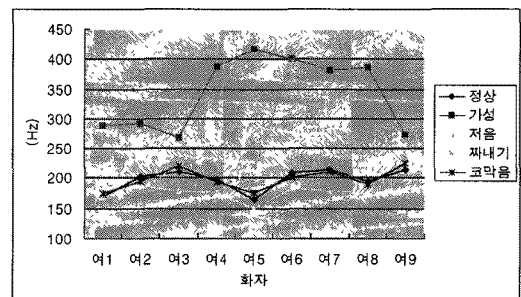


그림 1 위장 수단에 따른 여성 화자의 평균 기본 주파수

<그림2>와 <그림3>은 여성 화자와 남성 화자의 정상 발화의 기본주파수와 자유 위장 발화의 기본주파수의 차를 구한 것이다. 이때 화자의 배열은 가장 기본 주파수가 낮은 화자가 왼쪽에, 가장 기본주파수가 높

2 현대 서울 방언 화자들이 /에/와 /애/의 구별을 하지 못하고 /외/, /위/는 이중모음으로 발음한다는 7모음 체계를 따랐다 /에/의 경우 '에'의 표기로 제시되었다

은 화자는 오른쪽에 위치하도록 하였다. 이를 보면 상대적으로 평균 기본 주파수가 낮은 화자들이 자유 위장 수단으로 기본주파수를 올리는 경우가 더 많다는 것을 알 수 있다. 이는 여성의 경우에 기본주파수를 내려 발화한 화자가 3명이고, 남성이 더 가성을 많이 사용하여 위장하였다는 결과와 서로 통하는 바가 있다.

모든 화자가 그런 것은 아니었기 때문에 속단하기는 어렵지만 화자들은 자신의 기본주파수를 의식하고 있고, 이것이 가장 자신의 목소리와 다르게 발화하라고 하였을 때, 자신의 정상적인 기본주파수와 반대되는 방향으로 나가려는 경향이 있다는 것은 볼 수 있다. 이 부분에 대하여서는 차후 좀더 많은 화자들을 확보하여 좀더 연구해 볼 수 있을 것이다.

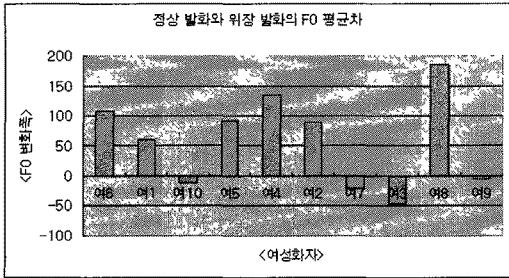


그림 2 여성 화자의 정상발화와 위장발화의 평균 기본주파수 차

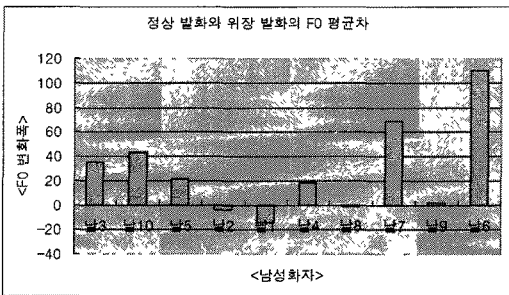


그림 3 남성 화자의 정상발화와 위장발화의 평균 기본주파수 차

### 3. 포먼트의 변화

본고에서는 단독발화한 단모음의 스펙트로그램에서 보이는 공명주파수의 안정적인 구간을 찾아 그 가운데를 측정하여 그 값을 구했다. 공명 주파수는 F4까지를 측정대상으로 하였다.

<그림 4>는 정상 발화된 /아/의 모음 포먼트들이 남성 화자들에 따라 어떤 차이가 나는지 보여준다. F1의 경우 최저 648Hz에서 최고 854Hz까지 차이가 난

다. F2의 경우 996Hz에서 1293Hz까지의 분포를 보인다. F3의 경우는 2530Hz부터 2880Hz까지, F4의 경우는 3380Hz부터 3900Hz까지 나타난다.

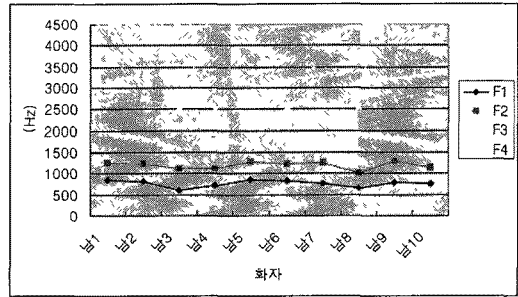


그림 4 정상 발화한 /아/의 모음 포먼트의 남성 화자별 변화

<그림 5>는 정상 발화된 /아/의 모음에서 F2-F1, F3-F2, F4-F3의 차를 보여준 것이다 이를 보면 F2-F1과 F4-F3에 비하여 F3-F2의 경우 화자별 차이가 더 크게 나타나는 것을 알 수 있다. 모음의 음가를 보여주는 F1, F2에 비하여 F3와 F2의 차이는 화자를 식별하는 데 이용될 수 있다는 중요한 단서가 될 수 있다는 증거가 될 수 있다. 이러한 경향은 위장 수단 이 변화했을 때도 크게 달라지지 않았다.

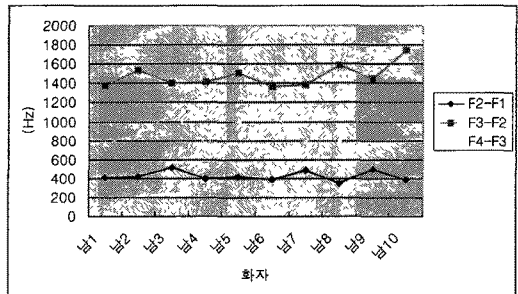


그림 5 남성화자의 정상 발화한 /아/의 모음 포먼트 차

모음별로 살펴볼 때, /아/의 경우 화자별 변이가 크게 나타나지만 /오/, /우/, /으/, /이/의 경우 포먼트의 화자별 변이가 그다지 크게 나타나지 않는다 또 후설 고모음인 경우 소리의 강도가 매우 약하고, 포먼트 역시 F1과 F2가 종종 통합되어 나타나는 등 측정에 어려움이 많이 발생한다. (그림 6)

이런 점들을 종합하여 볼 때 /아/모음이 화자 식별의 단서로 유용할 것이다. /아/의 경우 모음의 내재적 길이도 가장 길뿐 아니라, 대체로 모든 포먼트 값들이 안정적으로 보이고, 저차 포먼트의 화자별 차이도 다른 모음에 비하여 분명하게 나타난다.

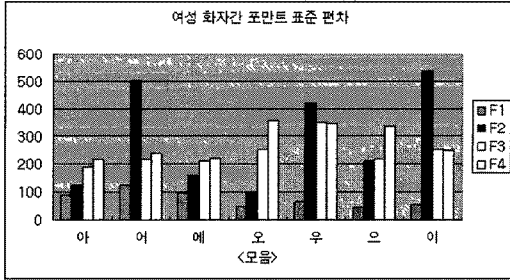


그림 6 여성화자의 모음별 포먼트 표준편차

#### IV. 결론

이상에서 국어 단모음의 기본 주파수와 공명 주파수가 발화 수단에 따라 어떻게 변화하는지 살펴보았다. 기본 주파수가 내려가면 F1과 F2도 내려가는 모습을 보이고, 올라가면 F1, F2도 따라서 올라가는 모습을 보이나 그 정도는 모음별로 차이가 크다.

/오/, /우/와 같은 모음은 화자와 발화 수단에 거의 영향을 받지 않고 비슷한 공명주파수 값을 갖는다. 따라서 이들 모음은 화자를 식별하기에 적절치 않을 것으로 보인다. 이와는 반대로 /아/는 화자와 발화 수단에 모두 큰 영향을 받는다. 본고의 실험은 극단적인 발성 상태에서 이루어졌으므로 실제 발화에서 남성 화자의 기본 주파수가 본고의 높은 소리까지 올라가는 경우는 흔치 않을 것으로 보인다. 그렇다면 일반 화자들을 서로 구분하는 데는 /아/ 모음의 저차 포먼트 역할을 살펴보는 것은 유용한 정보가 될 것이다.

모음의 F3-F2값을 보는 것도 화자 식별에 단서가 될 수 있다. 다른 포먼트값의 차이에 비해 F3-F2의 차는 화자별로 차이가 매우 크게 나타난다.

기본 주파수와 공명 주파수의 관계에 대해 생각해보면 기본 주파수가 낮아졌을 경우는 실험 대상이 된 남성 화자의 경우에 포먼트가 약간 변하긴 하지만 그 변화폭이 그다지 크지 않다. 그러나 가성으로 발화하거나 속삭이는 소리가 될 경우 그 포먼트 값이 상당히 큰 폭으로 변함을 알 수 있었다. 이는 가성이나 속삭이는 소리가 낮은 음성이나 코막은 음성에 비해 더 위장효과가 뛰어나다는 송민창(2003)의 연구결과와 일치하는 경향을 보인다.

기본 주파수와 공명 주파수 값은 대략적으로 화자 식별의 한 단서는 될 수 있지만 그것만으로 절대적으로 판단하기에는 부족한 점이 많다 추후 좀더 세밀한

연구를 통해 보충하기로 하겠다.

#### 참고문헌

- [1] 김진현, 홍수기, 장성길, “자동화자식별법에 대한 연구,” 국립과학수사연구소연보 제28권, pp.478-488, 1996.
- [2] 차일환, 유영화, 박종철, 김윤희, “음성개인식별에 관한 연구,” 국립과학수사연구소연보 제16권, pp.53-77, 1984.
- [3] Sun-Ah Jun, “K-ToBI (Korean ToBI) labelling conventions: Version 3”, *UCLA Working Papers in Phonetics* 99, pp.149-173, 2000.
- [4] 신지영, *말소리의 이해*, 한국문화사, 2000.
- [5] J. Laver, *Principles of Phonetics*, Cambridge University press, 1994.
- [6] W. Hardcastle & J. Laver, *The Handbook of Phonetic Sciences*, Blackwell Publishers Ltd, 1997.
- [7] Hansang Park, “Temporal and spectral Characteristics of Korean Phonation”, Ph.D dissertation, The University of Texas, 2002.
- [8] 홍수기, 김진현, 장성길, “음성의 Pitch 빈도분포의 개인성에 관한 연구”, *국립과학수사연구소연보* 제 27권, pp.366~371, 1995.