

모방발화에 대한 음향음성학적 연구(2)

-운율 특징을 중심으로-

박미영* 박지혜* 신지영** 강선미***

*고려대학교 민족문화연구원 음성언어정보연구실

고려대학교 국어국문학과, *서경대학교 컴퓨터학과

An Acoustic Phonetic Study about Voice Imitation(2)

-Focusing on Prosody Feature -

Park, Miyoung* Park, Jihye* Shin, Jiyoung** Kang, Sunmee***

*Spoken Language Information Lab, Institute of Korean Culture, Korea University

**Department of Korean Language and Literature, Korea University

***Department of Computer Science, Seokyeong University

E-mail : pikel@unitel.co.kr, jh1905@orgio.net, shinjy@korea.ac.kr, smkang@skuniv.ac.kr

Abstract

The purpose of this paper is to research voice imitation. Voice imitation changes various phonetic feature. Also, in our experimental results, voice imitation has preferential prosody difference. For imitating voice, imitators change their fundamental frequency bandwidths for the most part. Imitative speakers change their high fundamental frequencies effectively while they maintain their low fundamental frequencies. Also, excellent group is distinctly superior to common group for imitating prosodic patterns. That is, the f0 bandwidth's change and the prosodic patterns are significant in imitating voice. But the low f0 is maintain by all speakers.

I. 서론

본 연구¹⁾는 모방발화에 대한 음향음성학적 연구로 음성모방시 실현되는 운율적인 특징을 살펴보는 것을 연

구목적으로 한다. 모방발화에 대한 본 연구는 본인발화가 아닌 특정 대상의 음성을 모방하는 과정에서 특정 모방 대상을 목표로 하였을 때, 일반 화자의 어떠한 음성적 특징이 변하고 변하지 않는가를 포착하는데에서 시작된다.

특히 본 연구는 각각의 강세구와 억양구에 대하여 K-ToBI 전사를 실시하여 모방자 본인의 발화에서 모방발화시, 특정 모방 대상의 발화에 대하여 음성학적으로 의미 있게 변화시켰는가, 각각의 억양 유형 모방이 유사한가를 살펴본다. 또한 음향음성학적인 분석으로는 모방발화와 본인발화, 대상발화별로 다양하게 실현된 강세구 및 억양구의 음높낮이값을 비교하는 데에 의의가 있다. 음높낮이값에서는 강세구 끝성조 La, Ha와 억양구의 경계성조 L%, H%와 더불어, 강세구 내의 위치 영향의 유무를 고려하여 억양구에 후행하는 강세구 초 성조 L, H(강세구 초 분절음이 기식음이 아닌 경우와 기식음인 경우) 등의 기본주파수를 분석한다. 또한 음성모방시 모방화자들이 본인발화와 모방발화에서 공통적으로 변하거나 변하지 않는 음향적 특징을 밝혀보고자 한다.

음성모방에 대한 기존의 연구가 부족한 시점에서, 본 연구는 화자 식별(identification) 및 화자 확인(verification) 기술에서 화자 사칭 문제를 고려할 수 있는 음성학적 근거를 제공할 수 있을 것이다.

1) 본 논문은 한국과학재단 목적기초연구(R01-1999-000-0022 9-0) 지원으로 수행되었습니다.

II. 실험 및 연구방법

본 실험은 모방대상을 '김대중 전 대통령'으로 선정²⁾하였다. 음성 녹음은 동영상 자료를 활용하였으며, 담화 및 연설문 중 일부를 녹취하여 이 가운데 화자의 음성적 특징을 잘 포착할 수 있는 문장을 선정하였다. 본 실험을 위하여 구축된 문장은 5개이며, 발화 횟수는 모방발화 5회와 모방자 본인발화 2회로 실시하였다.

실험에 참여한 피실험자는 모두 20대의 서울 출신 남성 화자 4명이다. 녹음은 자연스러운 음성모방 자료를 구축하기 위하여 피실험자에게 모방 대상발화자의 음성 자료와 실험 문장을 사전에 알려주어 음성모방을 훈련한 뒤 녹음에 참여하도록 하였다. 녹음 후에 실험자들은 우선 청각적인 판단으로 피실험자들의 모방 능력의 정도에 따라 우수집단과 보통집단으로 분류하였는데, 피실험자들의 모방발화를 계층적으로 분류한 이유는 최종적으로 음성모방의 유사성 여부를 기술하기 위함이다.

녹음은 일방향 마이크 C414 B-ULS와 TASCAM DA-20MKII DAT 녹음기기를 이용하여 고려대학교 음성언어정보연구실의 녹음실에서 이루어졌다. 실험문장에 대한 운율 분석을 위해서 전사 기준안은 K-ToBI(ver 3.1; Sun-ah, Jun: 2000)를 따랐으며, 분석 프로그램은 WaveSurfer 1.5를 사용하였다.

III. 분석 결과 및 논의

1. 기본주파수 비교

본 연구의 모방발화자들은 음성모방을 할 때에 자신의 본인발화 기본주파수의 대역폭을 다음의 <그림 1>과 같이 변화시킨 것을 알 수 있다. 화자별로 차이가 있지만, 음성모방시 본인발화보다 약 20~120Hz까지 변화시키는 것으로 나타났다. 그러나 <그림 1>에서도 알 수 있듯이, 모방발화자들은 기본주파수의 최저값의 변화보다 최고값의 변화폭을 더 크게 하였다.

본 연구는 각 강세구 및 억양구에서 실현된 음높낮이 값을 측정하고, 평균값을 각 집단별로 분류하여 도표로 비교, 제시하였다. 음높낮이의 최고값과 최저값의 변화를 보기 위하여 강세구의 끝성조 La, Ha와 억양구의 경계성조 L%, H%, 음높낮이 실현의 위치 비교를 위해서 억양구에 후행하는 강세구 초 성조인 L와 H의 기본주파수값을 측정하였다. 억양구 후행 강세구 초 H는 그 실현 환경의 분절음이 기식음인 경우와 기식음이 아닌 경우로 나누어 측정하였다.

2) 모방화자들의 설문을 통하여 '김대중 전 대통령'의 음성모방이 가장 용이한 것으로 나타났다.

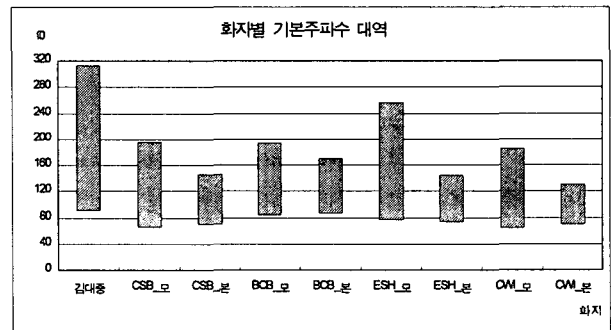


그림 1. 화자별 기본주파수 대역폭

<그림 2>와 <그림 3>은 모방화자의 우수집단과 보통집단이 실현한 강세구 끝성조 중 La의 기본주파수 평균값이다. 대상발화인 '김대중 전 대통령'의 강세구 끝성조 La는 120~140Hz 정도로 상대적으로 높게 실현되는데 모방화자들의 La는 85~100Hz 정도로 낮게 나타났다. 또한 본인발화시나 모방발화시에 La 값의 변화폭은 -5~5Hz 정도로 매우 작은 것을 알 수 있다.

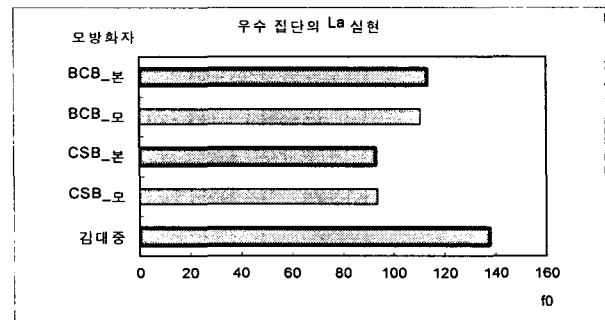


그림 2. 우수집단의 La 실현

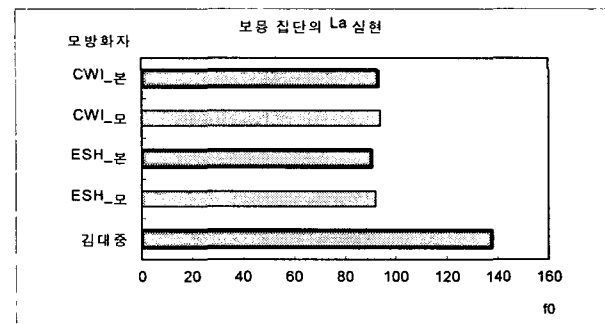


그림 3. 보통집단의 La 실현

다음의 <그림 4>와 <그림 5>는 각 화자별로 실현된 모든 억양구 경계성조 L%의 기본주파수 값의 평균으로 강세구 끝성조 La의 값보다 더 최저값으로 실현된 것을 알 수 있다. 이것이 모방화자들의 주파수 대역폭 중 최저값을 대신할 수 있는데, 대상발화의 L%는 100~110Hz대에서 주로 실현되었고, 모방화자들의 L%는 각 화자별로 100Hz, 80Hz, 80Hz, 90Hz대에서 분

인발화와 모방발화가 모두 실현되어 변화폭이 크지 않음을 알 수 있다.

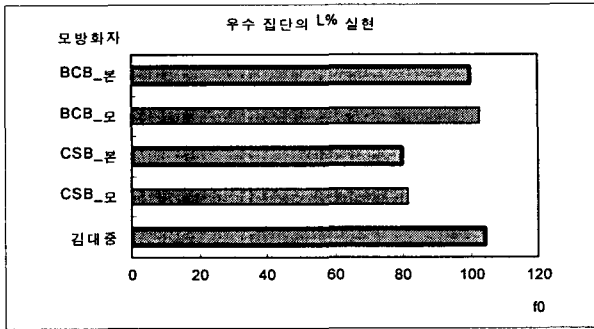


그림 4. 우수집단의 L% 실현

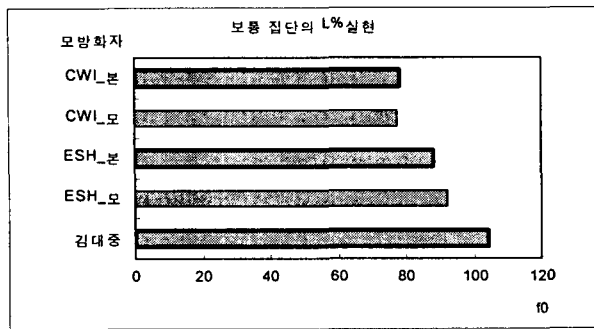


그림 5. 보통집단의 L% 실현

<그림 6>, <그림 7>과 <그림 8>, <그림 9>는 각각 강세구 끝성조 Ha와 억양구 경계성조 H%의 실현에 대한 화자별 평균 기본주파수값으로 앞서 살펴본 La와 L%의 기본주파수값보다 그 모방발화에서 변화폭이 매우 큼을 알 수 있다.

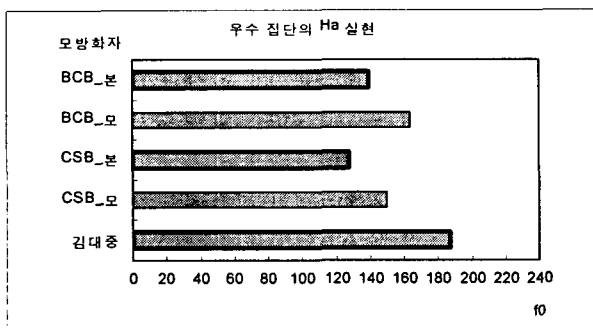


그림 6. 우수집단의 Ha 실현

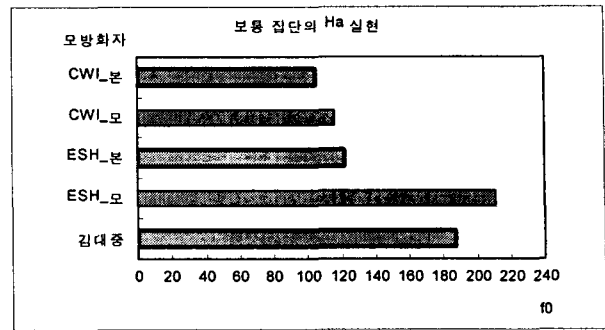


그림 7. 보통집단의 Ha 실현

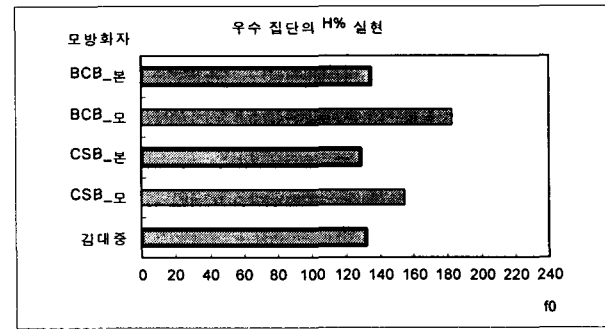


그림 8. 우수집단의 H% 실현

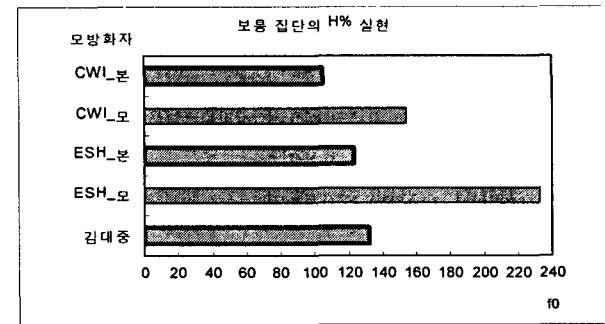


그림 9. 보통집단의 H% 실현

우수집단이나 보통집단이나 강세구 끝성조 Ha의 실현에서는 약 20~100Hz 정도 변화하였다. 억양구 경계성조 H%의 기본주파수는 화자별로 모방발화시에 25~110Hz까지 변화시켰다. 특히 모방발화에서 억양구 경계성조의 H%는 대상발화의 132Hz보다 더 높은 값으로 실현하여 높은 성조의 모방을 극대화하려는 노력을 한 것으로 볼 수 있다.

이상에서 살펴본 것처럼 강세구 끝성조와 억양구 경계성조의 기본주파수값은 강세구와 억양구 유형에서 낮은 성조보다 높은 성조로 대상발화의 음높낮이 수준이나 그 이상의 값으로 실현시키려고 노력하였으나, 강세구 및 억양구의 최저값은 자신의 수준에서 실현시켰음을 알 수 있다. 이것은 음성을 모방하는 데에 강세구 및 억양구에 배당된 각각의 음높낮이에서 높은 성조로의 변화는 가능하지만, 화자의 낮은 성조에서 기본주파수값의 변화는 불가능한 것임을 알 수 있다.

다음은 강세구 끝성조와 억양구 경계성조들과의 위치에 따른 기본주파수 변화 양상 비교를 위해서 억양구에 후행하는 강세구 초 성조의 변화 차이를 살펴 보았다. <그림 10>과 <그림 11>은 억양구에 후행하는 강세구 초 L의 실현에 관한 것이다. 모방화자 모두 본인발화의 음높낮이 기본주파수값과 -5~5Hz 수준에서 차이를 보일 뿐이다.

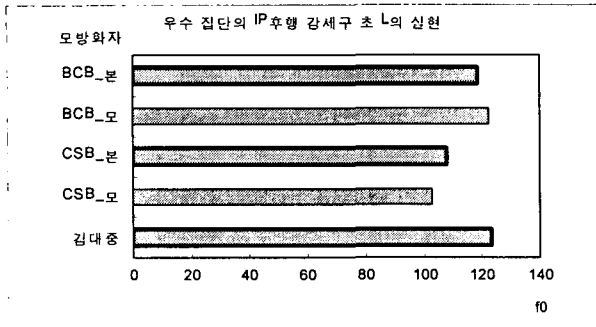


그림 10. 우수 집단의 억양구 후행 강세구 초 L 실현

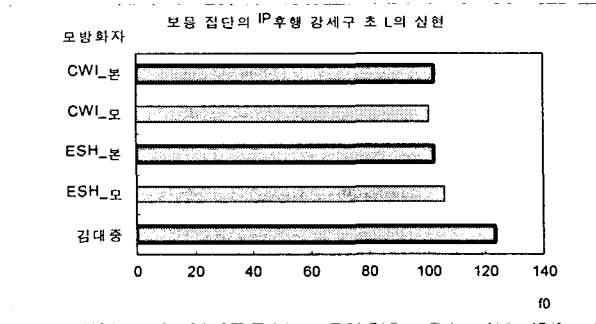


그림 11. 보통 집단의 억양구 후행 강세구 초 L 실현

그러나 다음의 <그림 12>와 <그림 13>에서 나타난 억양구 후행 강세구 초 H의 기본주파수값의 변화는 분절음의 기식성 여부에 따라 변화의 양상이 다르다.

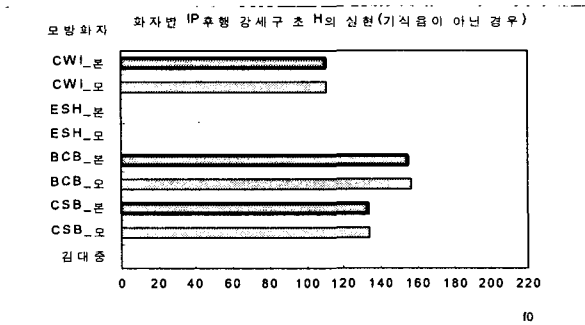


그림 12. 화자별 억양구 후행 강세구 초 H 실현(비-기식음)

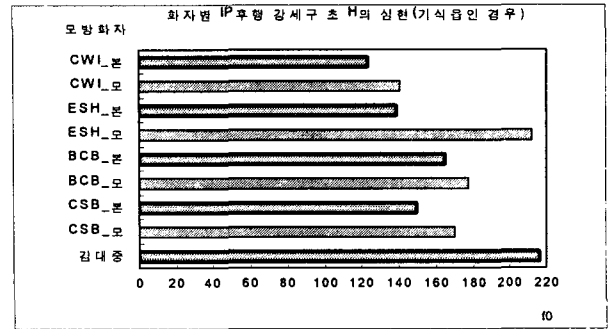


그림 13. 화자별 억양구 후행 강세구 초 H 실현(기식음)

<그림 12>에서 강세구 초 분절음이 기식음이 아닐 때 음높낮이 변화폭이 두드러지지 않는다³⁾. 그러나 <그림 13>에서처럼 기식음인 경우인데 기본주파수의 변화폭이 다른 높은 성조의 실현만큼 가능하였음을 알 수 있다. 이것은 대상 화자의 특징에 대하여 모방 화자의 음성모방은 강세구 초가 기식음 환경일 때 높은 기준주파수로의 실현 가능한 것으로 해석할 수 있다.

2. 강세구 및 억양구 유형 비교

대상발화자의 억양구 유형은 H+HL%, L+HL+L%, H+HL+L%, H+HL+>HL%가 빈번하게 사용되고 있다. 우수집단은 본인발화의 전반적인 억양구 형성 유형과는 달리 대상발화의 억양구 유형을 모방하는 데에 대체적으로 성공하였지만, 보통집단은 실패하였다.

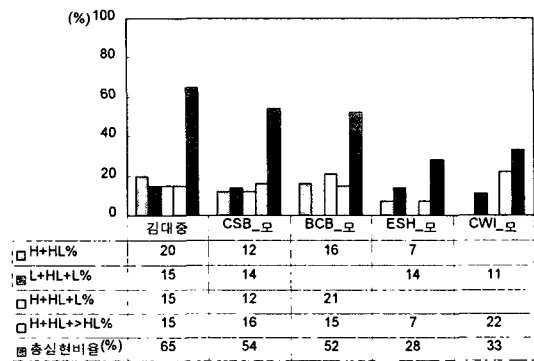


그림 14. 화자별 억양구 유형 모방 비율 30.00

3) 도표에서 여백은 자료가 발견되지 않은 예이다.

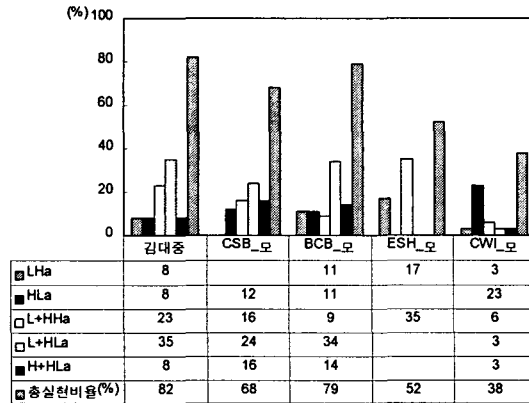


그림 15. 화자별 강세구 모방 유형 비율

강세구 형성 유형도 대상발화는 L+HLa, L+HLA가 보편적으로 나타났는데 우수집단이 강세구 유형을 성공적으로 모방한 것으로 나타났다. 음성모방시 보통집단은 강세구와 억양구의 굴곡 형성에서 대상발화만큼 자연스럽지 못하고 평탄하게 나타나지만, 우수집단은 자연스러운 강세구와 억양구 곡선을 형성하였다.

IV. 결론

음성모방시 모방화자는 음높낮이 변화에서 기본주파수를 높이는 것은 가능하지만, 본인발화보다 기본주파수를 낮추어 실현하는 것은 불가능하다. 강세구 및 억양구 형성 유형의 특징을 포착하는 것도 음성모방에 있어서 중요하다.

참고문헌

- [1] Sun-ah Jun, *K-ToBI labelling conventions, ver. 3.1*, 2000.