

감정에 따른 음성의 기본주파수 실현 연구1)

박미영* · 박미경**

고려대학교 민족문화연구원 음성언어정보연구실*, 한국과학기술원 CAD/CAM 연구센터**

A Study of F0's realization in Emotional speech

Park, MiYoung* · Park, MiKyoung**

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

Korea Institute of Science and Technology**

E-mail : pikel@unitel.co.kr*, miky@kist.re.kr**

Abstract

In this paper, we are trying to compare the normal speech with emotional speech -happy, sad, and angry states- through the changes of fundamental frequency. Based on the distribution charts of the normal and emotional speech, there are distinctive cues such as range of distribution, average, maximum, minimum, and so on. On the whole, the range of the fundamental frequency is extended in happy and angry states. On the other hand, sad states make the range relatively lessened. Nevertheless, the ranges of the f0 frequency in sad states are wider than the normal speech. In addition, we can verify that ending boundary tones reflect the information of whole speech.

I. 서론

본 연구는 동양의 전통적인 감정 '희로애락(喜怒哀樂)' 중 선행 연구들에서 '기쁨'의 2차적인 감정으로 분류된 '즐거움'을 제외한, '기쁨, 화남, 슬픔' 감정 상태와 '정상' 발화 시에 나타나는 음성의 특징을 연구하고자 한다. 감정 관련 선행 연구([4],[5])에서는 피치, 지속시간, 강도, 음질 등을 감정 변수로 보고 논의해왔다. 본 연구도 피치 정보 중 기본주파수 분포와 그 실현 양상을 발화 전체 문장과 문미역양이 포함된 어절에서의 실현으로 나누어 비교·분석하고자 한다. 감정 연구는 국내에서 여러 측면에서 연구가 시도되었지만 매우 광범위한 실험 설계가 필요하므로 총체적으로 시도된 연구는 없었다. 따라서 본 연구는 인간의 주요 감정별로

특징적인 음향 정보를 규명하는 데에 중요한 출발점이 되는 연구임에 의의가 있다 하겠다.

II. 연구 방법

감정 연구는 실험 설계에서 외부 변인 요소들이 많기 때문에 화자가 실험실에서 직접 감정을 생산하고 실제 처럼 얼마나 잘 구현했으며, 어떠한 정밀한 측정을 하였는지가 매우 중요하다. 따라서 본 연구자들은 감정 표현이 용이하도록 낭독체형 문장이 아닌, 실생활에서 자주 사용될 수 있는 구어체 문장을 선정하였는데, 각각의 감정을 표현하는 데에 가장 상관성이 높은 관용 표현들을 삽입함으로써 감정 표현의 극대화를 추구하도록 하였다(예: 이제 가슴이 벅차네.(기쁨)/이제 눈에서 불이 나네.(화남)/이제 가슴이 미어지네.(슬픔)). 이렇게 관용 표현이 들어간 문장은 총 21 문장이었다.

또한 특정한 상황을 제시하지 않아도 단독 상황에서 피험자가 감정을 변별적으로 표현하는 것이 용이할 것이고, 자연스러운 발화가 가능할 것으로 예상되는 일반 문장(예: 여보세요./누구세요) 16 문장도 본 연구의 녹음 텍스트로 구성하였다.

자료의 측정에 용이하게 감정 별로 문장 구성과 종결 어미의 양상을 통일시켰으며, 전체 문장의 음절수(±1~2음절)를 가능한 한 통일시키도록 하였다. 또한 일부 문장에는 후에 감정별 모음 포먼트 연구를 위해 국어의 7 기본모음을 문두에 넣는 것을 기본적으로 하여 실험 문장을 선정하였다.

이번 연구를 위해 분석된 음성 자료는 동일한 녹음 텍스트로 감정 연기를 수행한 전문 성우 6명(남 3, 여 3)의 발화이다. 성우 전문 녹음실 '보이스위즈(voicewiz)'를 통해 감정 표현의 전문 연기가 가능한 성우들을 섭외하였으며, 전체 녹음이 진행되었다. 실험에

1) 이 연구는 KIST의 기관고유과제인 TSI 과제 지원으로 연구되었습니다.

참여해준 전문 성우들은 다음과 같다.

	파일명	화자명	성 별	나 이
1	F1	KSH	여	22
2	F2	ACH	여	31
3	F3	KIJ	여	27
4	M1	KKS	남	34
5	M2	SMS	남	32
6	M3	SSH	남	34

<표 1. 피험자 성우 명단 및 인적 사항>

녹음에 사용된 장비는 믹서 MACKIE 1604-VLZPRO, 오디오 인터페이스 ESI WaMi Rack 192X, 컴프레서 JOEMEEK VC3Q, 마이크 RODE NT2가 사용되었으며 Cool Edit Pro 2.1에서 22050 byte, 16bit로 디지털화되었다.

전체 37 개의 실험 문장은 각각 ‘정상, 기쁨, 슬픔, 화남’의 감정마다 2회씩 반복되었는데, 실험자들은 피험자들에게 녹음 시 설정 감정이 극명하게 표현될 수 있도록 충분한 연습시간을 배려하였으며, 녹음 시 설정된 감정이 피험자 모두에게 끝까지 유지될 수 있도록 주지시켰다. 또한, 녹음 시 설정된 감정이 ‘정상’인 경우, 발화속도가 갑자기 빨라지지 않고 자연스럽게 읽히도록 주의하였다.

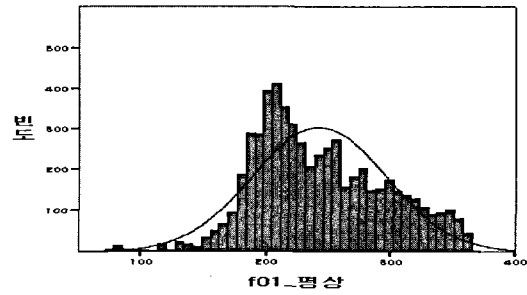
본 연구의 주요 분석 대상은 기본주파수인데, 기본주파수는 WaveSurfer 1.8.2를 이용하여 기본주파수 데이터의 전체값과 마지막 문미역양이 포함된 2음절 또는 1어절 구간의 값을 추출하였다. 예측되지 않은 기본주파수값의 오류는 모두 개별적으로 검토 후 수정하였다. 기본주파수값들의 분포 계산을 위해서는 pitch_chal(2004. 3. 자체 제작)이라는 프로그램을 이용하였으며, SPSS 12.0을 이용하여 통계 처리하였다.

III. 분석 결과

1. 감정별 기본주파수 분포도

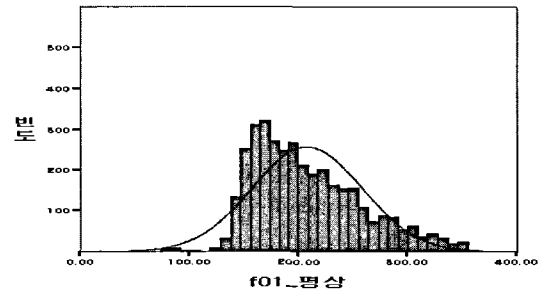
1) 감정별 전체 문장과 문미역양의 기본주파수 분포도

(1) 정상



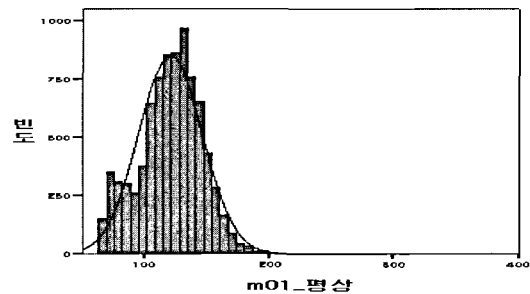
<그림 1. f01의 ‘정상’ 감정 전체 기본주파수 분포도>

여성 f01의 발화 자료의 전체 기본주파수값 분포 양상을 도출하여 <그림 1>과 같이 나타냈다. 또한 감정이나 태도를 나타내는 중요한 요소로 보는 문미역양의 양상도 보기 위해 문장의 종결어미가 포함된 발화 끝에서부터 2음절, 또는 마지막 어절의 기본주파수값 분포도도 다음의 <그림 2>와 같이 나타냈다.

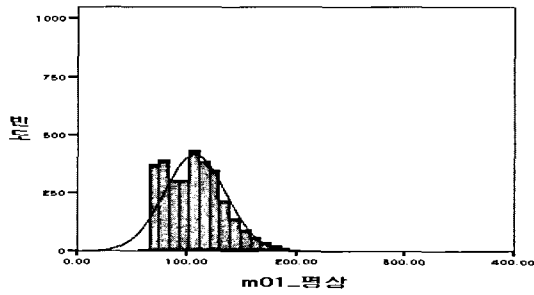


<그림 2. f01의 ‘정상’ 감정 문미 기본주파수 분포도>

아래의 <그림 3>과 <그림 4>는 남성 화자 m01의 문장 전체 기본주파수값과 문미역양이 실현된 구간의 기본주파수값 분포도이다.



<그림 3. m01의 ‘정상’ 감정 전체 기본주파수 분포도>



<그림 4. m01의 '평상' 감정 문미 기본주파수 분포도>
 '평상' 감정의 발화에서 피험자들은 모두 선행 연구 [2], [3]의 정상 발화 결과와 같이 주파수 대역 내에서 중앙값에서 왼쪽으로 치우치는 분포 양상을 띠고 있음을 알 수 있다.

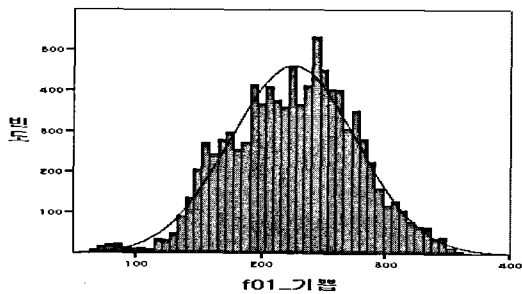
문미역량이 포함된 구간만 추출한 기본주파수 분포도도 전체 문장에서 보이는 기본주파수 분포도와 매우 유사한 양상을 띠고 있는데, 이것은 뒤이어 살펴볼 감정별 분포 양상에서도 유사한 양상을 보인다.

분석 결과가 제시되지 않은 여성 화자 f02, 03과 남성 화자 m02, 03에서도 각 성별에 따라 위의 결과들과 유사한 모습을 보이고 있다.

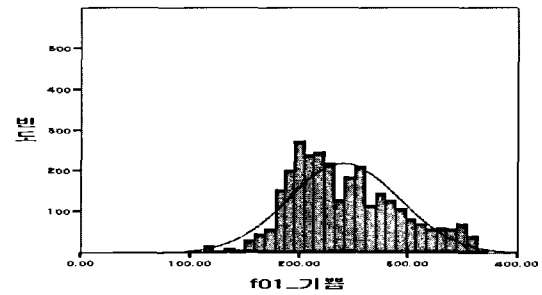
(2) 기쁨

'기쁨'의 감정으로 발화했을 시에 피험자 f01과 m01 모두 '평상' 발화에서와 달리 기본주파수 분포가 오른쪽의 고차 주파수대로 확대되었고 분포도의 양상도 주파수 대역 내에서 오른쪽으로 치우친 값들에서 고빈도를 보이는 모습을 알 수 있다.

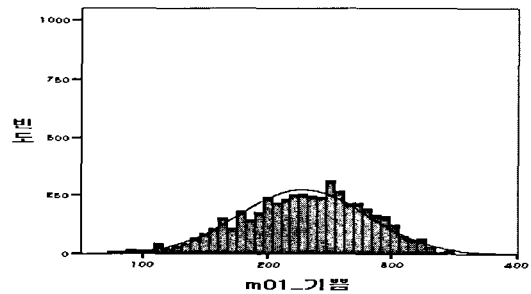
선행 연구 [4], [5]에서도 '기쁨/행복함/유쾌함' 등의 감정에서 기본주파수 대역이 증가하는 것으로 나타났던 연구의 결과가 일치하는 것임을 알 수 있다.



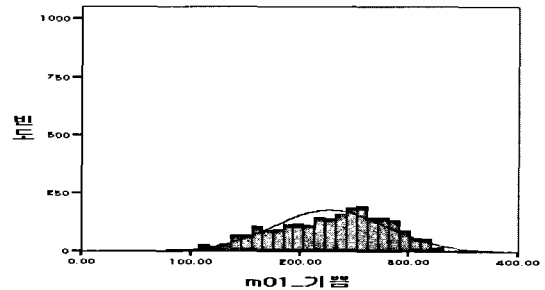
<그림 5. f01의 '기쁨' 감정 전체 기본주파수 분포도>



<그림 6. f01의 '기쁨' 감정 문미 기본주파수 분포도>



<그림 7. m01의 '기쁨' 감정 전체 기본주파수 분포도>

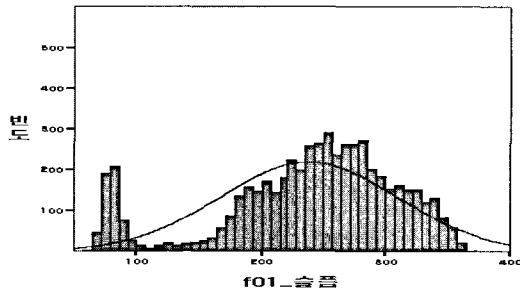


<그림 8. m01의 '기쁨' 감정 문미 기본주파수 분포도>

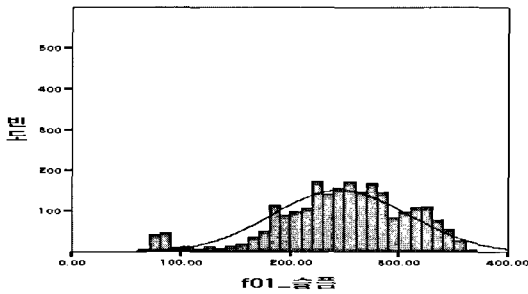
'평상' 발화에서의 전체 문장과 문미역량 구간의 기본주파수 분포와 마찬가지로 '기쁨'의 감정으로 발화 시에도 기본주파수 분포도가 두 경우에서 모두 유사한 양상으로 나타나고 있다. 이것은 문미역량에서의 기본주파수 분포만으로도 발화의 실현 양상을 알 수 있는 근거가 되는 것으로 해석된다.

(3) 슬픔

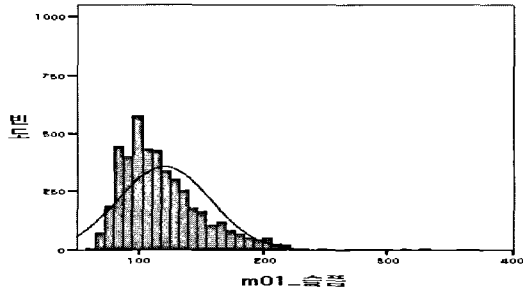
'슬픔'의 감정과 관련한 선행 연구들은 주로 '슬픔'의 기본주파수 특징으로 보통이거나 낮은 기본주파수와 좁은 기본주파수 대역을 특징으로 들었다. 그러나 본 연구에서 피험자들이 발화한 아래의 자료들은 성별에 따라 다른 결과를 보이고 있음을 확인할 수 있었다.



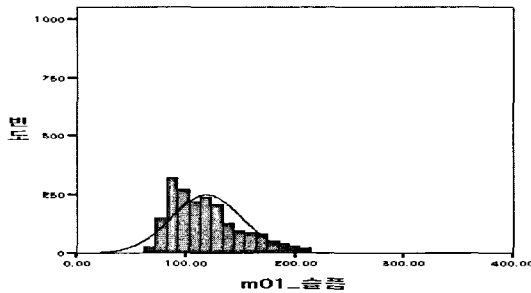
<그림 9. f01의 '슬픔' 감정 전체 기본주파수 분포도>



<그림 10. f01의 '슬픔' 감정 문미 기본주파수 분포도>



<그림 11. m01의 '슬픔' 감정 전체 기본주파수 분포도>



<그림 12. m01의 '슬픔' 감정 문미 기본주파수 분포도>

우선 여성 화자 f01의 <그림 9>의 전체 문장 기본주파수 분포도와 <그림 10>의 문미역양 기본주파수 분포도의 모습을 보면 앞선 '평상' 발화에서의 기본주파수 대역과 달리 넓은 대역에서 기본주파수값이 분포하고 있음을 알 수 있으며, 중앙값에서 오른쪽으로 치우쳐

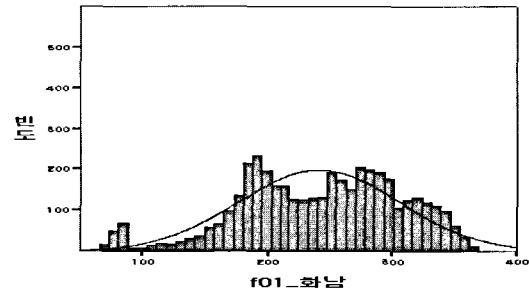
고빈도의 값들이 나타난다.

이것은 남성 화자 m01과는 다른 결과를 보이고 있는데, 남성 화자 m01의 결과 <그림 11>과 <그림 12>의 자료에서는 주파수 대역이 '평상' 발화와 유사한 수준으로 실현되었으면 고빈도의 값들도 주파수 대역 내의 중앙값으로부터 왼쪽으로 치우쳐 실현되고 있음을 알 수 있다.

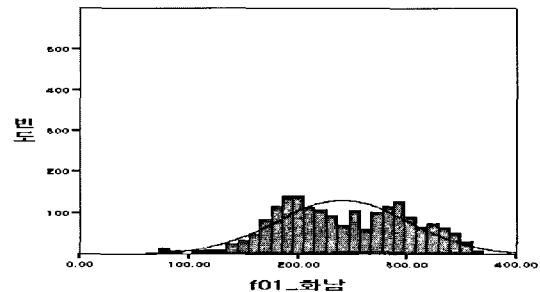
이러한 결과는 전체 여성 화자와 남성 화자들에서 대립되는 같은 결과를 보이고 있다. 이에 대해서는 다음 III-2에서 자세하게 논의하고자 한다.

(4) 화남

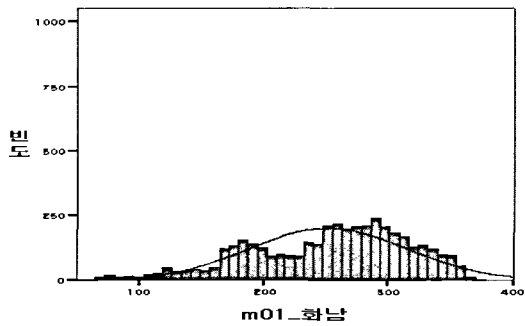
'화남'의 감정으로 발화했을 시에는 '기쁨'의 감정에서 나온 결과와 유사한 패턴을 보이고 있다. '화남'의 감정은 그동안 가장 연구가 많이 된 분야로 높은 기본주파수와 넓은 대역이 특징인데, 본 연구의 피험자들에게서도 같은 결과를 보이고 있음을 알 수 있다. 저주파수대에서도 최빈값들을 보이는 기본주파수값들이 있는데, 이것은 아마도 '화남'의 감정을 표현하는 특징으로서 불규칙적으로 급격하게 오르내리는 억양이 기본주파수값을 좌우하기 때문인 것으로 보인다.



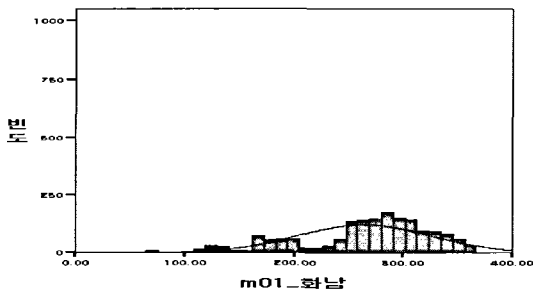
<그림 13. f01의 '화남' 감정 전체 기본주파수 분포도>



<그림 14. f01의 '화남' 감정 문미 기본주파수 분포도>



<그림 15. m01의 '화남' 감정 전체 기본주파수 분포도>



<그림 16. m01의 '화남' 감정 문미 기본주파수 분포도>

2) 감정별 기본주파수 분포도의 왜도, 첨도

다음의 <표 2>와 <표 3>은 각각 전체 발화와 문미 역양의 왜도, 첨도를 제시한 것이다.

	f01				m01			
	평상	기쁨	슬픔	화남	평상	기쁨	슬픔	화남
왜도	.224	-.119	-.793	-.363	.160	-.315	1.759	-.509
첨도	-.314	-.206	.090	-.334	2.360	-.221	5.243	-.370
	f02				m02			
	평상	기쁨	슬픔	화남	평상	기쁨	슬픔	화남
왜도	-.003	-.051	.195	-.068	-.292	-.319	1.093	.470
첨도	.166	-.333	-.508	-.662	.332	-.879	2.502	-.649
	f03				m03			
	평상	기쁨	슬픔	화남	평상	기쁨	슬픔	화남
왜도	.144	.219	.085	.299	.012	.345	2.120	.358
첨도	.527	-.192	.161	-.324	.236	-1.031	10.087	-1.005

<표 2. 각 화자 전체 문장 발화의 왜도, 첨도>

	f01				m01			
	평상	기쁨	슬픔	화남	평상	기쁨	슬픔	화남
왜도	.686	.351	-.704	-.082	1.175	-.304	.789	-.845
첨도	-.035	-.273	.477	-.634	5.51	-.555	-.013	.188
	f02				m02			
	평상	기쁨	슬픔	화남	평상	기쁨	슬픔	화남
왜도	.754	.041	.227	.092	.061	-.421	1.008	.247
첨도	.612	-.460	-.711	-.903	.500	-.838	4.125	-.966
	f03				m03			
	평상	기쁨	슬픔	화남	평상	기쁨	슬픔	화남
왜도	1.231	.232	.040	.270	.209	.077	1.815	.142
첨도	1.886	-.300	.071	-.353	-.142	-1.258	6.723	-1.279

<표 3. 각 화자 문미역양 발화의 왜도, 첨도>

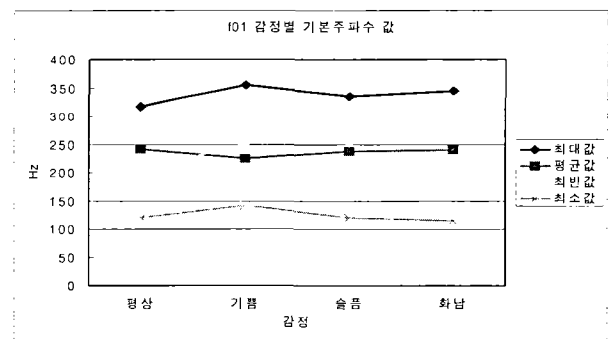
주로 왜도가 0에서 1에 위치하거나 첨도가 0에 가까

운 값일수록 '평상' 발화인데, 각 감정에서는 '평상' 발화와 다르게 나타남을 알 수 있다. 특히 각 값들이 큰 수치의 음의 값을 갖는 경우는 '평상' 발화와는 기본주파수값 분포가 많이 다름을 알 수 있다.

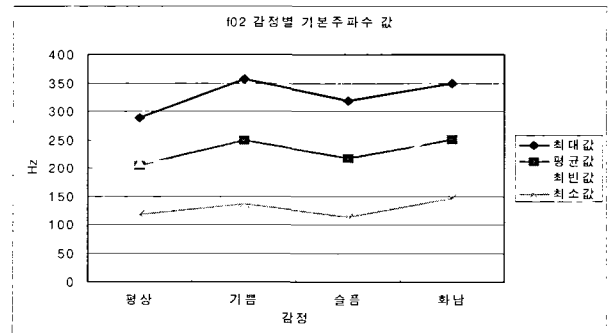
2. 감정별 기본주파수 최대값, 최소값, 평균값

실험에 참여한 전체 피험자들의 각 감정별로 실현된 기본주파수 값을 가지고 최대값과 최소값의 평균, 평균값, 최빈값을 그래프로 나타내었다.

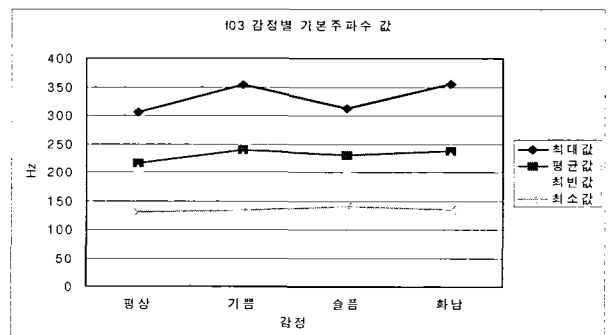
<그림 17>, <그림 18>, <그림 19>는 여성 화자 f01, f02, f03의 기본주파수값들의 그래프이며, <그림 20>, <그림 21>, <그림 22>는 남성 화자 m01, m02, m03의 것이다.



<그림 17. f01의 감정별 각 기본주파수 평균값 분포>



<그림 18. f02의 감정별 각 기본주파수 평균값 분포>



<그림 19. f03의 감정별 각 기본주파수 평균값 분포>

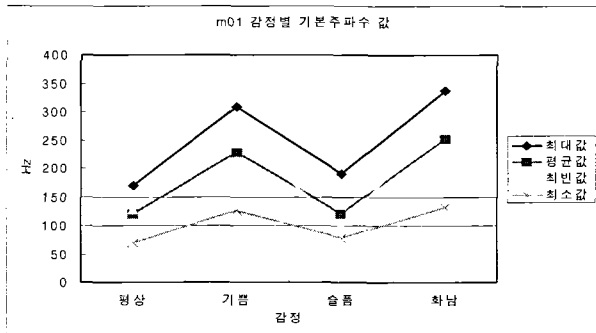
여성 화자들은 '기쁨'과 '화남'의 감정에서 기본주파수 대역이 확대되어 나타나고 '슬픔'의 감정에서도 f01, f02

는 '평상' 발화보다 기본주파수 값 대역이 확대되었다. '슬픔' 감정의 f03은 '평상' 발화의 수준과 유사한 양상으로 기본주파수 대역을 실현하였다.

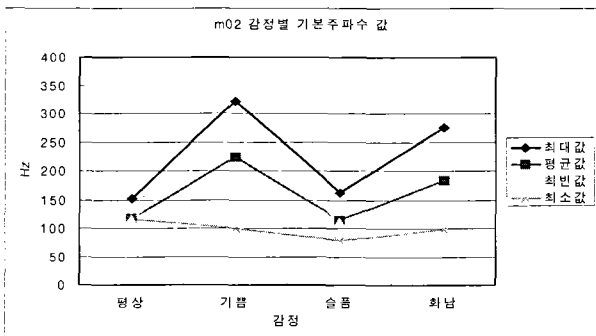
기본주파수 대역 내에서 f01의 경우를 제외하고는 f02, f03의 기본주파수 평균값들은 '평상' 발화에서의 수준보다 높게 실현되었다. f01의 경우는 고주파수대에서 최빈값이 차이가 있음을 확인할 수 있다. '기쁨, 화남'의 기본주파수 최대값들은 '평상' 발화보다 약 50Hz 내외에서 차이를 보이며 실현되고 있는데, 이것은 다음에 살펴볼 남성 화자들과는 차이가 있다.

다음의 자료들은 남성 화자들의 그래프인데, 여성 화자들의 경우와 확연하게 다름을 알 수 있다. 남성 화자들은 감정별로 연결된 기본주파수값들의 그래프 선의 경사가 급격한데, 이것은 감정 변화를 할 때, 기본주파수 정보를 남성들이 여성들보다 더 잘 이용하는 것으로 볼 수 있다. m01은 '화남'의 감정에서 기본주파수 변화가 더 컸으며, m02와 m03은 '기쁨'의 감정에서 기본주파수 변화폭이 더 크다.

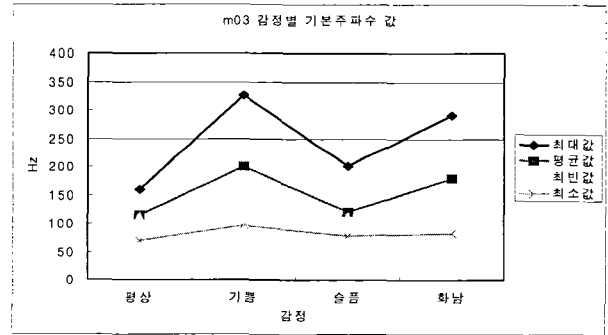
남성 화자들은 '기쁨'과 '화남'의 감정에서 기본주파수의 최대값 평균은 '평상' 발화보다 약 150Hz 내외로 증가하여 실현하는 것이 여성 화자보다 큰 차이가 있음을 알 수 있다.



<그림 20 . m01의 감정별 각 기본주파수 평균값 분포>



<그림 21. m02의 감정별 각 기본주파수 평균값 분포>



<그림 22. m03의 감정별 각 기본주파수 평균값 분포>

남성 화자들도 '슬픔'의 감정에서는 다른 감정들과 달리 '평상' 발화의 기본주파수 대역 수준으로 좁아지고 있는데, 여성 화자들에서와 마찬가지로 자신의 '평상' 발화 수준보다 좁아지는 대역폭을 보이지는 않는다. 또한 급격하게 좁은 주파수 대역에서의 발화는 평균값과 최빈값에서도 '평상' 발화와 같이 중앙값으로부터 저주파수대에 위치하는 것으로 보아 여성 화자들과는 달리 낮은 주파수 대역을 이용하여 '슬픔' 감정을 표현하는데 기본주파수가 안정적으로 실현되고 있음을 알 수 있다.

IV. 결론

이상에서 '평상' 발화와 '기쁨, 슬픔, 화남'의 감정 발화에서 나타나는 기본주파수의 실현을 살펴보았다. 문장 전체에서 추출된 기본주파수의 분석도 중요하지만, 문미역양만을 추출했을 때에도 전체 문장 발화 시와 유사한 경향을 보이며, 기존의 문미역양 실현으로 감정의 상태를 정의할 수 있었던 것도 문미역양의 해석만으로도 전체 감정을 해석할 수 있는 중요한 근거임을 뒷받침하였다.

감정 발화 시에 화자들은 모두 기본주파수 분포 실현을 변화하였는데, '기쁨, 화남'에서는 '평상' 발화보다 주파수 대역이 확대되었으며, '슬픔' 감정에서는 그 실현 양상이 성별에 따라 달랐다. 여성 화자들은 선행 연구들과 달리 '슬픔' 감정에서도 기본주파수 대역을 확대하기도 하였으나, 남성 화자들은 기본주파수 대역이 다른 감정들보다 확연하게 좁아지거나 '평상' 발화 수준으로 실현되고 있음을 알 수 있었다. 또한 남성 화자는 '기쁨'과 '화남' 감정에서의 기본주파수 대역의 확대도 여성 화자보다 더 큰 값으로 변화한 것을 알 수 있었는데, 이것은 남성이 감정을 표현할 때에 기본주파수 변화를 더 활용하는 것으로 볼 수 있다.

따라서 앞으로의 연구는 성별에 따라 감정을 표현하는 데에 사용하는 다른 특징들을 살펴봄, 기존 연구들에서 알려진 기본주파수 정보 외에 어떤 음향적 정보들이 감정 연구에 활용될 수 있는지를 알아보겠다. 또

한 궁극적으로는 실제 감정 표현을 연구하는 것이 중요하겠지만, 실제 대화상의 녹음이나 광범위한 외부 변인 요소들로 인한 연구의 어려움을 좀 더 치밀한 녹음 설계로 보완하여 정밀한 연구가 이루어져야겠다.

참고문헌

- [1] 장인창, 이태승, 박미경, 김태수, 장동식(2005), <발화 내 감정의 정밀한 인식을 위한 한국어 문미역양의 활용>, 한국통신학회 논문지 vol.30, 한국통신학회
- [2] 박미영, 신지영, 강선미(2004). 위장·모방 발화에서 나타나는 기본주파수 변화 양상에 관한 연구. 한국어학 24. 한국어학회
- [3] 안병섭, 박미영(2004). 모방발화의 음향음성학적 연구(3); 전문 성대 모사자의 자료를 중심으로. 말소리 52. 대한음성학회
- [4] Iain R. Murray and John L. Aronott(1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. J. Acoust. Soc. 93-2. Acoustical Society of America
- [5] Carl E. Williams, Kenneth N. Stevens.(1972). Emotion and Speech : Some Acoustical Correlates. J. Acoust. Soc. 52-4. Acoustical Society of America