

운율이식을 통해 나타난 감정인지 양상 연구

이서배(부산대)

<차 례>

- | | |
|--------------------|------------------------|
| 1. 서론 | 4. 결과 및 논의 |
| 2. 기존연구의 한계와 대안 | 4.1. 이식유형에 따른 감정인지 |
| 3. 감정음성의 분석과 인지실험 | 4.2. 감정유형에 따른 감정인지 |
| 3.1. 감정음성 데이터베이스 | 4.3. 감정별 이식유형에 따른 감정인지 |
| 3.2. 감정음성의 운율 | 4.4. 문장길이에 따른 감정인지 |
| 3.3. 감정음성의 음가 및 음색 | 4.5. 운율 대비 음가 및 음색 |
| 3.4. 감정음성의 인지실험 | 5. 결론 |

<Abstract>

A Study on the Perceptual Aspects of an Emotional Voice Using Prosody Transplantation

So Pae Yi

This study investigated the perception of emotional voices by transplanting some or all of the prosodic aspects, i.e. pitch, duration, and intensity, of the utterances produced with emotional voices onto those with normal voices and vice versa. Listening evaluation by 24 raters revealed that prosodic effect was greater than segmental & vocal quality effect on the preception of the emotion. The degree of influence of prosody and that of segments & vocal quality varied according to the type of emotion. As for fear, prosodic elements had far greater influence than segmental & vocal quality elements whereas segmental and vocal elements had as much effect as prosody on the perception of happy voices. Different amount of contribution to the perception of emotion was found among prosodic features with the descending order of pitch, duration and intensity. As for the length of the utterances, the perception of emotion was more effective with long utterances than with short utterances.

1. 서 론

의사소통에 있어서는 말하고자 하는 내용만큼이나 화자의 의도가 중요하다. 특히 음성언어에 있어서는 음성으로부터 나오는 운율과 분절과 음색과 같은 음향적 특질들을 종합해서 해석할 때, 보다 정확한 의사전달이 가능해 지는 것이다. 상대방을 보지 않고 전화음성만으로도 기분을 알 수 있다는 사실은 인간의 이러한 능력을 말해주는 단적인 예이다. 최근 음성기술은 감정인식을 자동화하려는 단계에 이르렀다. 음성을 듣고 사용자의 기분에 따라 서비스의 수준을 결정하는 예라든지, 주로 전화로 고객을 대하는 회사에서 사원들의 친절도 평가를 위해 방대한 양의 대화 내용을 녹음했다가 감정인식 시스템을 통해 문제가 될 만한 대화만 선택해서 평가할 수 있게 한 사례는 감정음성에 대한 연구가 실생활에 응용될 수 있는 가능성이 크다는 것을 보여준다.

본 연구는 2장에서 기존연구의 한계와 대안을 논하고 3장에서 감정음성의 분석과 인지실험을 설명하였다. 4장에서 결과 제시 및 논의를 하고 5장에서 결론을 맺었다.

2. 기존연구의 한계와 대안

음성을 분절적 요소(segmental elements)와 초분절적 요소(suprasegmental elements)로 나눌 수 있다면 본 연구에서는 편의상 분절적요소를 포먼트의 값으로 나타나는 음가(phonetic value)로 보았고 초분절적 요소를 피치(pitch), 지속시간(duration), 강도(intensity)로 구성되는 운율(prosody)과 스펙트럼의 기울기(spectral tilt), HNR(Harmonic to Noise Ratio) 등등의 변화가 반영되어 나타나는 음색(vocal quality)으로 세분화 하였다. 감정음성(emotional voice)을 분석한 다수의 연구들은 감정의 영향에 따른 음가, 운율 그리고 음색의 변화를 보고하고 있다[1][2][3][4]. 아울러 이러한 자연음의 분석에 이어 포먼트 합성기를 사용한 감정인지에 대한 연구도 있어왔다[5][6]. 자연음 분석에서 나타나는 감정에 따른 운율의 변화[1][2]는 운율을 변화시킨 합성음성이 감정인지에 영향을 준다는 연구를 뒷받침해준다[6]. 자연음에서 나타나는 음가와 음색의 변화가 감정인지에 영향을 준다는 보고 역시 음성합성기를 통한 포먼트와 음색의 변화가 감정인지에 영향을 준다는 보고와 맥을 같이한다[3][4][5].

그러나 이러한 포먼트 합성기를 이용한 연구들은 한 영역에 국한된 자질의 제한된 변화(예, F0 평균, F3, 스펙트럼 기울기 조정) 안에서만 이루어졌고 자연음이 아닌 기계적인 신호음을 대상으로 한 간접 연구를 벗어나지 못했다. 이러한 제약은 대용량 TTS(Text-to-Speech) 합성기를 이용한 연구에서 자연음들을 이어붙이는

방법으로 극복할 수 있었으나 음색과 분절음의 변화를 반영하기 어려운 구조 때문에 운율을 변화시키는 연구에만 국한되었다. 그러므로 자연음의 운율, 분절음, 그리고 음색의 변화 모두를 동시에 고려한 감정인지의 연구는 어려웠다.

최근 음성 인식기술의 발달로 기계에 의해 감정을 인식하려는 시도들 중 한 연구에 의하면 운율정보를 이용하는 것보다는 분절음의 변화정보를 이용하는 것이 기계에 의한 감정음성 인식에 더 효과적이라고 한다[7]. 이러한 양상이 인간의 감정인지에서는 어떻게 나타나는지 알아보기 위해서는 자연음의 운율, 분절음 그리고 음색의 변화 모두를 고려한 감정인지에 대한 연구가 필요하다.

본 연구는 음성이 운율과 음가 및 음색으로 나누어질 수 있다고 보았다. 이를 전제로 감정인지에 대한 운율의 영향을 살펴보고자 감정음성의 운율을 운율요소 별로 나누어 감정중립음성(낭독체)에 이식(감정음성의 운율+낭독체의 음가 및 음색)시키고 감정인지에 대한 음가 및 음색의 영향을 연구하기 위해 역으로 감정중립음성의 운율을 감정음성에 이식(낭독체의 운율+감정음성의 음가 및 음색)시키는 방법[8]으로 위에서 살펴본 제약을 극복하고 운율과 음가 및 음색의 감정인지에 대한 영향을 살펴보았다.

3. 감정음성의 분석과 인지실험

3.1. 감정음성 데이터베이스

원 음성 데이터베이스[9]는 4명의 성우가 발화하였는데, 그 가운데 20명의 청취 평가 결과 감정인식 점수가 가장 높은 화자인 MYS의 음성만을 대상으로 하였다. 음절수의 증가로 인한 문장길이의 변화가 감정인지에 어떤 영향을 주는지 알아보고자 녹음된 문장들 가운데 음절수로 3배 이상 차이가 나는 두 문장 a(장문), b(단문)를 실험자료로 삼았다. 장문과 단문은 각각 보통 목소리에 해당하는 ‘낭독체’ 음성과 ‘기쁨’, ‘화남’, ‘두려움’, ‘슬픔’의 4가지 감정음성을 포함한 5가지의 음성으로 구성되었다.

- a. 난 가지 말라고 하면서 문을 닫았어
- b. 나도 몰라

3.2. 감정음성의 운율

본 연구에 사용된 음성의 분석결과 <표 1>에서와 같이 감정의 종류에 따라 운율이 각각 다르게 나타나는 것을 볼 수 있다. ‘슬픔’의 경우 강도가 낮아지고 ‘화

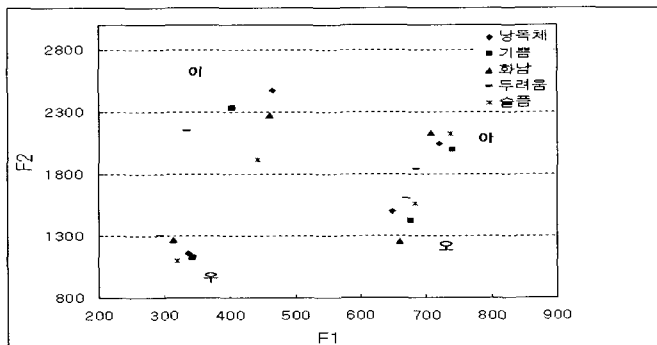
남'의 경우 피치가 크게 올라가는 것과 '두려움'에서의 높은 피치와 '기쁨'에서의 높은 강도는 기존연구[1][10]에서 나타난 것과 같은 경향을 보여준다. 특히 '화남'의 경우 피치, 강도가 가장 높고 총 휴지구간이 가장 짧은 것으로 나타났다. 반면에 슬픔은 가장 긴 휴지구간을 보여주었다.

<표 1> 감정에 따른 운율의 변화

운율유형	낭독체	기쁨	화남	두려움	슬픔
평균피치(Hz)	188.8	231.2	<u>264.9</u>	252.7	245.8
평균강도(dB)	58.6	66.0	<u>69.2</u>	56.3	49.1
평균발화시간(sec)	1.502	1.725	1.373	<u>1.365</u>	1.473
총휴지(pause)	0.544	0.87	<u>0.189</u>	3.096	5.267

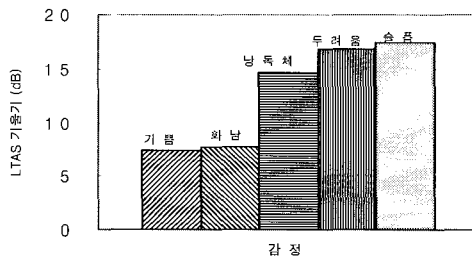
3.3. 감정음성의 음가 및 음색

감정의 영향으로 턱과 혀끝의 위치가 달라진다는 조음적 연구[11][12]에서 알 수 있듯이 감정에 의해 음가가 변한다는 것은 자연스러운 현상이라 할 수 있다. <그림 1>은 본 연구에 사용된 자연음의 감정에 따른 포먼트의 위치를 보여준다. 분절음들의 F1과 F2가 감정의 변화에 따라 달라지는 것을 볼 수 있다. 예를 들어, <그림 1>에서 낭독체를 기준으로 할 때 F1의 경우 두려움과 슬픔에서 '이'와 '우'의 값이 내려가는 것을 볼 수 있다. F2의 경우 슬픔의 '이'와 '오', 두려움의 '이', '우', '아', '오'가 중앙화하는 경향이 있는 것으로 나타났다. 이는 심리적 위축으로 인해 입을 작게 벌린 결과로 보인다. 특히 두려움의 경우 빠른 발화속도(<표 1> 참조)로 인하여 조음목표점에 도달하지 못한 결과(undershoot)도 영향을 미친 것으로 보여진다.



<그림 1> 감정에 따른 포먼트의 변화

감정별 LTAS(Long Term Average Spectrum)의 기울기를 보면 <그림 2>와 <표 2>에서와 같이 낭독체를 기준으로 두려움과 슬픔은 기울기가 큰 반면 기쁨과 화남은 작은 것으로 나타났다. 기쁨과 화남의 경우 낭독체보다 높아진 강도(<표 1> 참조)에서도 알 수 있듯이 큰 목소리에서 나타나는 고주파 성분의 증가로 인해 전체 스펙트럼의 기울기가 작아진 반면에 낭독체보다 강도가 작은 두려움과 슬픔은 고주파가 감소하는 작은 목소리의 특징 때문에 스펙트럼의 기울기가 증가한 것으로 보인다[13]. <표 2>에서 HNR(Harmonic to Noise Ratio)의 경우, 감정중립음성인 낭독체에 비해 전반적으로 감정음성에서 노이즈가 증가한 것을 볼 수 있다. 특히 기쁨과 슬픔의 경우 화남과 두려움에서 보다 노이즈의 비율이 더 크게 나타났다.



<그림 2> 감정유형별 LTAS 기울기

<표 2> 감정에 따른 음색의 변화 (평균값)

운율유형	낭독체	기쁨	화남	두려움	슬픔
LTAS 기울기(dB)	14.7	7.5	7.8	16.8	17.4
HNR (dB)	15.0	10.5	12.2	12.1	10.6

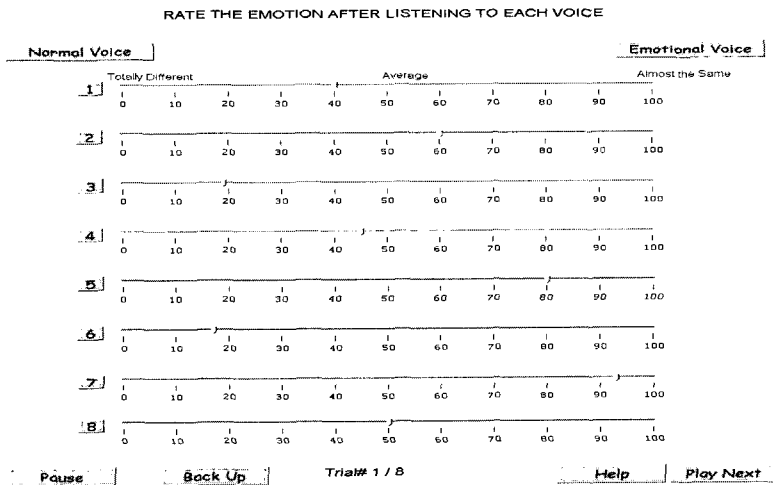
3.4. 감정음성의 인지실험

감정음성의 운율을 낭독체 음성에 이식하고 낭독체의 운율을 감정음성에 이식해서 <표 3>과 같이 8가지의 이식이 이루어졌고 총 80개(2×5×8)의 음성을 얻었다.

<표 3> 이식유형과 설명

유형	설 명	약어
nI	낭독체에 감정음성의 강도를 이식	n: normal e: emotion D: duration P: pitch I: intensity
nD	낭독체에 감정음성의 지속시간을 이식	
nDI	낭독체에 감정음성의 지속시간&강도를 이식	
nP	낭독체에 감정음성의 피치를 이식	
nPI	낭독체에 감정음성의 피치&강도를 이식	
nDP	낭독체에 감정음성의 지속시간&피치를 이식	
nDPI	낭독체에 감정음성의 지속시간&피치&강도를 이식	
eDPI	감정음성에 낭독체의 지속시간&피치&강도를 이식	

청취실험에는 28명(남15명, 여13명)의 대학생이 참여하였고 이들의 평균 나이는 26세였다. 80개의 합성음에 대해 Alvin[14]을 사용해 아날로그 스케일로 평가하도록 구성하였다. <그림 3>에서 왼쪽에 있는 1번부터 8번까지의 버튼을 클릭하면 동일한 문장이면서 동일한 감정음성에서 이식된 8개(<표 3> 참조)의 음성을 들을 수 있게 된다. 이들 8가지 발화의 제시 순서는 무작위처리 되었다. 오른쪽의 'Emotional Voice'의 버튼을 클릭하면 동일한 문장의 감정음성 원음을 듣게 된다.



<그림 3> 감정 청취평가에 사용된 컴퓨터 인터페이스

평가자들은 감정음성 원음의 감정 표현을 100%로 보았을 때 8개의 음성을 듣고 느껴지는 감정이 어느 정도인지를 슬라이드 위에 있는 슬라이드바를 마우스로 움직여 표시하도록 지시받았다. 예를 들어, '1번'과 'Emotional Voice'를 들어 보고 느껴지는 감정의 정도가 원음과 비슷할수록 오른쪽(100%)에 가깝게, 상이할수록 왼쪽(0%)에 가깝게 슬라이드바를 움직이는 것이다. 슬라이드바의 위치는 수치화되

어 자동으로 입력되었다.

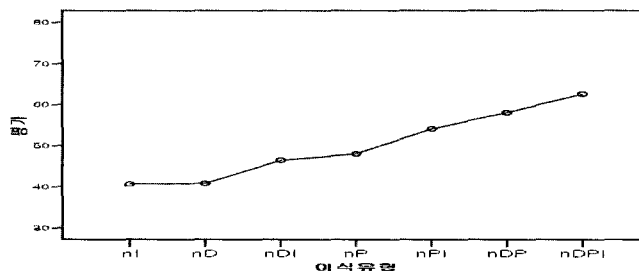
평가자들이 원하는 만큼 발화를 들을 수 있게 하였고 지나간 발화도 되돌아가 다시 들을 수 있게 하였다. 청취실험에서 이산 스케일보다 아날로그 스케일이 평가자의 일관성 확보에 유리하고 복수의 자극을 한 화면에서 동시에 비교하는 것이 신뢰도와 일관성을 더욱 높여주는 것으로 보고되고 있다[15].

4. 결과 및 논의

평가자들의 신뢰도를 알아보기 위해 그룹내 상관관계(intraclass correlation)를 구해 0.886을 얻었다. 평가자 28명 중 4명의 평가점수를 제외하면 크론바하의 알파(Cronbach's alpha)값이 전체값보다 올라가므로 그 4명의 평가자는 신뢰도 확보를 위해 제외하였고 나머지 24명의 평가자에 대해 그룹내 상관관계를 다시 구한 결과 0.894가 나와 비교적 높은 신뢰도를 확보하였다. 이식유형(7가지)¹⁾, 감정유형(4가지), 장단문(2가지)의 요인들을 혼합디자인 분산분석(ANOVA)한 결과 이식유형 [F(6,138)=32.279, P<.05], 감정유형[F(3,69)=8.479, P<.05], 장단문[F(1,23)=24.474, P<.05], 이식유형&감정유형[F(18,414)=6.486, P<.05], 감정유형&장단문[F(3,69)=4.338, P<.05], 이식유형&감정유형&장단문[F(18,414)=4.348, P<.05]이 유의한 것으로 나왔다.

4.1. 이식유형에 따른 감정인지

이식유형에 따른 감정인지의 전반적 경향은 <그림 4>에서 볼 수 있는데 운율 요소들의 유기적인 결합이 개개의 운율보다 더 큰 영향을 미치는 것으로 나타났다.



<그림 4> 이식유형에 따른 감정인지

1) 이식유형에 따른 분석에서 eDPI는 별도의 분석을 위해 제외하였다.

<표 4> 이식유형별 감정인지 평가 평균과 중요도 (*P < 0.05)

유형 (평균)	nI (41.3)	nD (41.4)	nDI (46.4)	nP (48.2)	nPI (53.5)	nDP (57.3)	nDPI (62.0)
nI		0.924	0.002*	0.004*	0.000*	0.000*	0.000*
nD	0.924		0.000*	0.001*	0.000*	0.000*	0.000*
nDI	0.002*	0.000*		0.435	0.002*	0.000*	0.000*
nP	0.004*	0.001*	0.435		0.002*	0.000*	0.000*
nPI	0.000*	0.000*	0.002*	0.002*		0.076	0.001*
nDP	0.000*	0.000*	0.000*	0.000*	0.076		0.001*
nDPI	0.000*	0.000*	0.000*	0.000*	0.001*	0.001*	

개개의 요소로 볼 때 <표 4>에서와 같이 nI는 nD를 제외한 모든 유형과 유의한 차이를 보였고 nD는 nI, nDI는 nP, nP는 nDI, nPI는 nDP 그리고 nDP는 nPI를 제외한 모든 유형과 유의한 차이를 보였다. 그리고 nDPI는 모든 유형과 유의한 차이를 보였다. 다시 말해 (nI, nD), (nDI, nP), (nPI, nDP) 간에는 차이가 없었다.

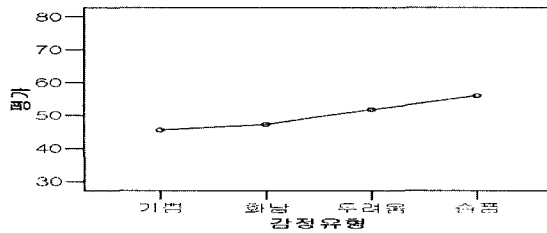
<표 4>에서 (nD, nDI), (nP, nPI), (nDP, nDPI) 간에 유의한 차이가 나는 것으로 볼 때 감정인식에 있어 강도가 일관성 있는 영향력을 미치는 것으로 보인다. 그러나 (nI, nDI), (nP, nDP) 그리고 (nI, nPI), (nD, nDP)의 관계에서 보면 지속시간과 피치의 영향력이 강도보다 월등히 더 큰 것을 알 수 있다(<그림 4> 참조). 그러므로 감정인지에 대한 운율요소들의 기여도는 ‘피치>지속시간>강도’ 순인 것으로 나타났다.

4.2. 감정유형에 따른 감정인지

<표 5>에서와 같이 기쁨은 화남을 제외한 모든 유형과 유의한 차이를 보이고 화남은 기쁨을, 두려움은 슬픔을, 슬픔은 두려움을 제외한 모든 유형과 유의한 차이를 보이는 것을 알 수 있다. <표 5>과 <그림 5>로 볼 때, 운율이식된 음성들의 감정인지 정도는 (두려움, 슬픔) > (기쁨, 화남) 순이라고 말할 수 있다. 운율적 변화가 두려움, 슬픔과 같은 소극적인 감정에서 기쁨, 화남과 같은 적극적인 감정에 서보다 감정인지에 더 크게 작용한 것으로 보인다.

<표 5> 감정유형별 감정인지 평가 평균과 중요도 (*P < 0.05)

유형 (평균)	기쁨 (46.1)	화남 (46.5)	두려움 (51.4)	슬픔 (55.5)
기쁨		0.321	0.003*	0.002*
화남	0.321		0.040*	0.002*
두려움	0.003*	0.040*		0.085
슬픔	0.002*	0.002*	0.085	



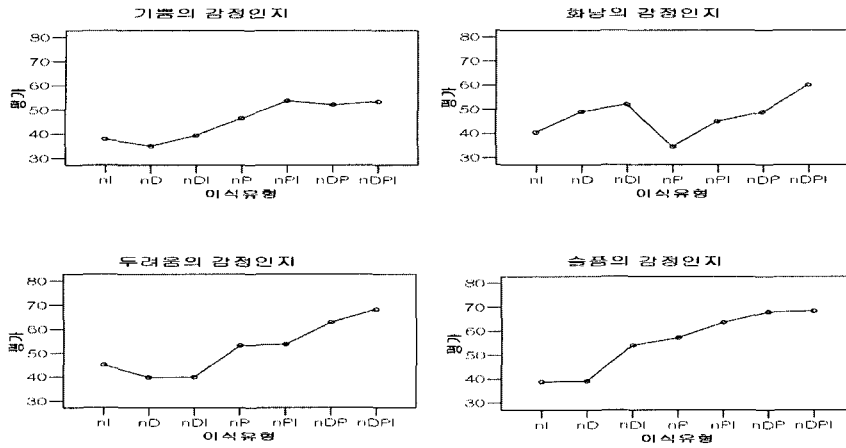
<그림 5> 감정유형에 따른 감정인지

4.3. 감정별 이식유형에 따른 감정인지

앞절에서 나타난 바와 같이 운율의 변화가 기쁨과 화남과 같은 적극적인 감정보다 두려움과 슬픔과 같은 소극적인 감정에서 감정인지에 더 크게 작용한 이유는 <그림 6>와 같이 두려움과 슬픔의 경우 기쁨과 화남의 경우보다 더 다양한 운율이식 조건에서 더 높은 평가를 받았기 때문이다. 두려움과 슬픔의 평균은 각각 51.8과 55.8이고 기쁨과 화남의 평균은 각각 45.4와 47.2이다.

<그림 6>와 같이 감정별 이식유형의 양상은 화남의 경우를 제외하고는 감정 전체에 따른 이식의 전반적 양상(<그림 4> 참조)과 비슷한 양상을 보였다(기쁨, 두려움, 슬픔의 nI vs. nD는 유의한 차이가 없었다: $p < 0.05$).

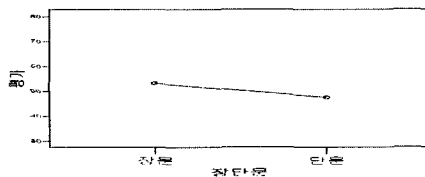
화남의 경우 다른 감정에 비해서 지속시간의 상대적 중요성이 컸다. 이것은 적극적인 화남(Hot Anger)에 나타나는 빠른 발화속도로 인한 지속시간의 변화(<표 1> 참조)가 피치의 변화보다 감정인지에 더 큰 기여를 하기 때문인 것으로 보인다. 화남에서의 피치의 상대적 영향력은 떨어진 것으로 나타났다.



<그림 6> 감정별 이식유형에 따른 감정인지

4.4. 문장길이에 따른 감정인지

단문보다는 장문에서 감정인지가 더 잘되는 것(p=0.000*)은 음절의 길이가 늘어 나면서 감정인지의 단서들이 더 많아지기 때문인 것으로 해석된다(<그림 7> 참조). 또한 감정별 분석에서도 동일한 경향이 드러났다.

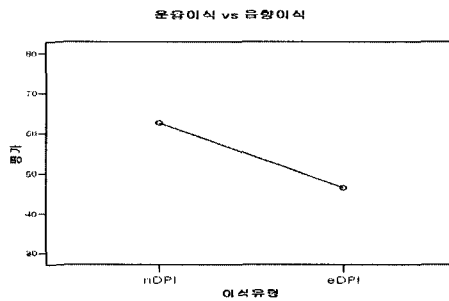


<그림 7> 문장길이에 따른 감정인지

4.5. 운율 대비 음가 및 음색

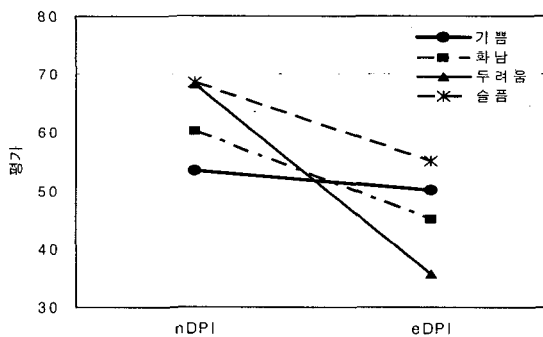
<표 3>에서와 같이 eDPI는 낭독체의 운율과 감정음성의 음가와 음색을 가지게 되므로 감정음성의 운율과 낭독체의 음가와 음색을 가지는 nDPI와 비교하면 음가와 음색으로 인해 해당 감정의 인지가 얼마나 영향을 받는가의 판단에 사용될 수 있게 된다. 즉, eDPI의 청취평가를 통해 음가와 음색이 감정인식에 어느 정도 영향을 미치는지 알 수 있고 nDPI의 청취평가를 통해 운율의 총합이 감정인식에 미

치는 영향을 가늠할 수 있게 되는 것이다. <그림 8>과 <표 6>에서와 같이 운동 대비 음가 및 음식의 평균비율은 74.3%였고 유의한 차이($p=0.000^*$)를 보였다. 그러므로 기계에 의한 감정인식에서는 분절음정보가 운동보다 더 큰 기여를 한다는 보고[7]에 반해 인간의 감정인지에는 운동의 영향이 더 크다고 말할 수 있겠다.



<그림 8> 운동 대비 음가 및 음식

<그림 9>와 <표 6>에서 각기 감정별로도 운동의 영향이 음가 및 음식보다 일관성 있게 큰 것으로 나타났다. <표 6>에 나타나듯이 감정인지에 있어서 감정별 운동에 대한 음가 및 음식의 비중은 기쁨이 제일 크고 두려움이 제일 작은 것으로 나타났다. 즉, 기쁨을 표현함에 있어서는 운동의 변화 못지않게 음가 및 음식의 변화가 중요한 반면 두려움의 표현에는 음가 및 음식보다는 운동의 변화를 더 중요하게 인지하는 것으로 보인다.



<그림 9> 운동 대비 음가 및 음식 (감정유형별)

<표 6> 운율 대비 음가 및 음색의 감정인지 평가 평균

	기쁨	화남	두려움	슬픔	평균
nDPI	53.4	60.2	68.2	68.5	62.6
eDPI	50.2	45.1	35.6	55.0	46.5
eDPI/nDPI	94.0%	74.9%	52.2%	80.3%	74.3%

주목할만한 점은 두려움과 슬픔 같은 소극적인 감정이 운율에는 비슷하게 높은 영향을 받다가 음가 및 음색의 영향에서 두려움은 최저의, 슬픔은 최고의 수치를 보인다는 것이다. 이는 두려움보다 LTAS의 기울기가 더 크고 노이즈 성분(예, 흐느끼는 음색)이 더 많은 (<표 2> 참조) 슬픔의 음색적 특징이 영향을 미쳤기 때문으로 추정된다.

5. 결 론

지금까지 자연음을 대상으로 한 운율이식 실험에 나타난 감정인지의 양상을 살펴보았다. 전반적으로 음가와 음색과 같은 자질보다는 운율적 자질이 감정인지에 더 큰 영향을 미치는 것으로 나타났다. 감정별 운율 대비 음가 및 음색의 비교 경우 두려움에서는 운율의 상대적 역할이 컸던 반면 기쁨에서는 음가 및 음색의 영향력이 운율과 대등했다. 합성기를 이용한 음색실험과 음가실험에서 음색변화만으로 혹은 음가변화만으로 감정인지가 영향 받을 수 있다는 결과[5][6][16]에 비추어볼 때 자연음에서도 음가와 음색이 감정인지에 영향을 준다는 고찰은 설득력을 가진다. 다만 자연음에서는 감정별로 음가와 음색의 영향력이 달라질 수 있다는 점을 고려해야 할 것이다.

운율이식유형에 따른 감정인지에서 전반적으로 운율요소들의 유기적인 결합이 개개의 운율보다 더 큰 영향을 미치는 것으로 나타났다. 감정인지에 대한 운율요소들의 기여도는 ‘피치>지속시간>강도’ 순이었는데 감정별로 기여도의 순서는 바뀔 수 있는 것으로 나타났다.

자연음의 운율, 음가, 음색이 감정의 종류에 따라 감정인지에 서로 다른 정도의 영향을 미친다는 결과는 자연음의 분석에 있어서도 감정의 종류에 따라 달라지는 운율, 음가, 음색이 관찰된다는 보고[1][2][3][4]와 맥을 같이한다고 볼 수 있다. 또한 효과적으로 감정을 전달하기 위해서는 운율, 음가, 음색과 같은 자질들을 목적인 감정음성의 인지적 특성에 맞게 표현해야 한다는 점을 시사하고 있다(예, 화남의 감정표현에서는 음가와 음색보다는 운율의 효과적인 표현에 중점을 두고 피치보다는 운율의 지속시간에 유의해야 하며 기쁨의 감정에서는 운율만큼 음가와 음색의 표현을 중시하는 반면에 두려움의 표현에서는 음가와 음색의 비중이 떨어

짐을 고려해야함).

향후, 음가와 음색을 분리해서 각각의 기여도를 알 수 있는 실험이 필요하리라고 본다. 이상의 연구가 감정음성 인지의 이해에 더 큰 시각을 제공해주고 효과적인 감정표현 및 인식의 기초자료가 되기를 바란다.

참 고 문 헌

- [1] S. Yildirim, "Study of acoustic correlates associated with emotional speech", *148th ASA Meeting*, San Diego, CA, November 15, 2004.
- [2] I. Murray, J. Arnott, "Toward to simulation of emotion in synthetic speech: A review of the literature on human vocal emotion", *JASA*, Vol. 93, No. 2, pp. 1097-1108, 1993.
- [3] K. Scherer, "Vocal affect expression: A review and a model for future research", *Psychological Bulletin*, Vol. 99, pp. 143-165, 1986.
- [4] K. Scherer, "Vocal measurement of emotion", In: R. Plutchik, H. Kellerman, (Eds.), *Emotion: Theory, Research, and Experience*, Academic Press, San Diego, Vol. 4. pp. 233 - 259, 1989.
- [5] C. Gobl, "The role of voice quality in communicating emotion, mood and attitude", *Speech Communication*, Vol 40, Issue 1-2, pp. 189-212, April 2003.
- [6] M. Schröder, "Emotional speech synthesis: A review", *Proceedings from Eurospeech*, Vol. 1, pp. 561-564, Aalborg, Denmark, 2001.
- [7] C. Lee, S. Yildirim, M. Bulut, A. Kazemzadeh, C. Busso, Z. Deng, S. Lee, S. Narayanan, "Emotion recognition based on phoneme classes", *Proc. of Int'l Conf on Spoken Language Processing*, Jeju island, Korea, 2004.
- [8] K. Yoon, "Imposing native speakers' prosody on non-native speakers' utterances", *WESPAC IX 2006, The 9th Western Pacific Acoustics Conference*, 2006.
- [9] 정서 음성 데이터베이스 *Emotion01*, 음성정보기술산업지원센터, 2004.
- [10] 조철우, "정서정보의 변화에 따른 피치와 지속시간의 변화," *창원대학교 산기연 논문집*, 제 11집 별쇄, 1997.
- [11] D. Erickson, O. Fujimura, B. Pardo, "Articulatory correlates of prosodic control: emotion and emphasis", *Language and Speech*, Vol. 41, pp. 395-413, 1998.
- [12] M. Nordstam, G. Svanfeldt, B. Granstrom, D. House, "Measurements of articulatory variation in expressive speech for a set of Swedish vowels", *Speech Communication*, Vol. 44, pp. 187-196, 2004.
- [13] E. Holmberg, R. Hillman, J. Perkell, "Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal and loud voice", *JASA*, Vol. 84, No. 2, pp. 511-529, 1988.
- [14] J. Hillenbrand, Alvin experiment-control software: Full install (v 1.19, 2/24/2006), <http://homepages.wmich.edu/~hillenbr/>, 2006.
- [15] S. Granqvist, "Enhancements to the Visual Analogue Scale, VAS, for listening tests", *TMH-QPSR*, 4/1996, 1996.

- [16] T. Waaramaa, P. ALKU, A. Laukkanen, "The role of F3 in the vocal expression of emotions", *Logopedics Phoniatics Vocology*, Vol. 31, pp. 153-156, 2006.

접수일자: 2007년 5월 10일

게재결정: 2007년 6월 18일

▶ 이서배(So Pae Yi)

주소: 609-735 부산시 금정구 장전동 산30번지 부산대학교 전자공학과 음성통신연구실

소속: 부산대학교 인지과학 협동과정

전화: 010-5555-6305, 055-339-9555, 051-516-4279

E-mail: sopaeyi@pusan.ac.kr