

음성 정서표현의 음향프로파일 및 재인

남경식 · 민경환

서울대학교 심리학과

본 연구는 음성 정서표현의 음향프로파일을 조사하고 재인의 정확성을 알아보기 위해 수행되었다. 만족, 환희, 슬픔, 비통, 불안, 공포, 짜증, 분노, 혐오, 경멸, 지루함, 죄책감, 자괴심, 흥미의 14 정서를 대상으로 하였으며, Banse와 Scherer(1996)의 절차를 따라서 연구를 수행하였다. 실험 1에서는 이들 정서에 대한 다양한 음향특정치를 구하여 정서별로 음향프로파일을 만들었다. 실험 1의 음향프로파일은 공포를 제외하고는 대체로 Banse와 Scherer의 결과와 일치하였다. 실험 2에서는 음성자료만을 제시하였을 때 얼마나 정서를 잘 추정할 수 있는가를 알아보았다. 재인율은 50%내외로 Banse와 Scherer의 결과와 비슷했으나 개별정서의 재인율에서는 일치도가 낮았다. 재인율이 높은 정서로는 비통, 분노, 지루함 등이었고 낮은 정서에는 만족, 환희, 불안, 혐오 등이 포함되었다. 동일방향의 정서를 지닌 정서들간에 혼동률이 높았다. 여성음성자극에 대한 재인율이 높았고, 동성의 음성자극에 대한 재인율이 높았으며, 남녀 피험자간에 재인율에서 차이는 발견되지 않았다.

주제어: 음성 정서 표현, 음향 프로파일, 재인율

정서표현에 관한 연구는 Darwin(1872)으로 거슬러 올라간다. Darwin은 진화론을 검증하기 위한 방법으로 다양한 문화권의 사람들과 다양한 종의 동물들을 대상으로 정서표현, 특히 얼굴표정의 유사성을 확인하기 위한 연구를 하였다. Darwin이후 정서표현에 관한 연구는 거의 단절되었다가 1970년대부터 Ekman(1971, 1982), Izard(1971, 1982) 등을 중심으로 얼굴표정에 관한 연구들이 활발히 이루어져 왔다. 그러나 음성표현에 관한 연구는 심리학내에서 정서에 대한 관심이 급속히 확대된 이후에도 대단히 드물게 수행되었다. 이러한 연구부재의 원인은 주로 방법론적인 문제에 기인한다. 우선 음성표현을 녹음하고, 적절한 음성특정치를 구하고, 자료를 분석하며 그래픽으로 제시하는 등의 작업이 제대로 이루어지기 위해서는 공학기술과 컴퓨터 프로그래밍 등이 뒷받침되어야 하는데 이러한 기반이 어느 정도 갖추어진 것은 비교적 최근의 일이다. 더욱이 대부분의

심리학자들은 음성연구를 수행하기 위해 요구되는 음향학적 지식을 갖추고 있지 못하다. 한편 음성이 얼굴표정만큼 섬세하고 정확하게 정서를 전달하지 못한다는 점도 방법론적인 어려움과 함께 음성연구를 소홀히 하게 되는 원인으로 작용하여 왔다.

그러나 최근에는 Scherer를 중심으로 정서의 음성표현에 관한 연구들이 수행되기 시작하고 있다. 한국에서도 조은경과 조철우 등(조은경, 조철우와 민경환, 1997; 조철우, 조은경과 민경환, 1997)에 의해서 정서의 음성표현을 분석하는 연구가 수행된 바 있다.

음성 표현의 연구는 크게 특정 정서에 대한 특정 음성 표현 패턴이 무엇인가를 밝히는 입력연구와 청자들이 음성 단서로부터 정서를 추론할 수 있나를 밝히는 해독연구의 두 갈래로 나눌 수 있는데, 기존의 연구들은 대부분 입력연구를 중심으로 이루어져 왔다.

그런데, Banse와 Scherer(1996)의 연구 이전에 수행

된 입력 연구들은 다음과 같은 문제점들을 안고 있다. 첫째, 매우 제한된 숫자의 정서들만이 대상이 되었기 때문에 좀 더 다양한 정서들을 표집할 필요가 제기되었다. 둘째, 활성화(activation)와 정서가(valence) 단서의 역할을 분리하여 보지 않았다. 셋째, 과거의 연구는 대개 기본주파수와 에너지 측정치들을 사용하였는데, 이는 비특정적인 생리학적 각성을 반영하기가 쉽다. 즉 음향 측정치의 선택의 폭이 좁았기 때문에 정서 특정적 음향 목록을 만들기가 힘들었다. 넷째, 이 영역에서의 연구가 이론이 없기 때문에 경험적 발견의 축적이나 검증 가능한 가설의 발전이 불가능했다.

Banse와 Scherer(1996)는 이러한 네 가지 측면의 제한점들을 극복하려고 했다. 첫째, 대상 정서의 숫자를 늘렸는데, 만족, 환희, 혐오, 경멸, 슬픔, 비통, 불안, 공포, 짜증, 분노, 지루함, 죄책감, 자존심, 흥미 등이 그것이다. 특히 이중 4쌍은 정서가차원 (valence dimension)에서는 유사하지만 활성화차원 (activation dimension)에서는 상이한 정서들이다 (분노/짜증, 공포/불안, 비통/슬픔, 환희/만족). 따라서 정서와 활성화 단서들의 영향을 분리, 비교할 수 있게 되었다.

둘째, 음향측정치를 다양화하였다. 음향의 에너지를 나타내는 기본주파수와 강도차원 이외에도 발음속도, 공학적 측정치, 유성음과 무성음 영역의 대역별 평균 등을 추가하였다.

셋째, Scherer (1984, 1986)는 정서의 음성연구에 이론적 기반을 제공하기 위하여 성분과정이론(component process theory)을 제안하였다. 이 이론에서는 정서가 5개의 성분(인지, 생리적 조절, 동기, 운동표현, 검색-느낌) 각각의 하위체계들이 협응함으로써 발생한다고 한다. 즉 비정서적 상태에서 독립적으로 작용하던 하위체계들이 정서상태에서 협응한다는 것이다. 이 협응과정은 유기체에 중요한 외부 혹은 내부 자극들을 평가하는 이른바 “자극평가검사”(Stimulus Evaluation Checks; SECs)의 결과로 유발된다. 대부분의 정서 이론가들은 세분화된 정서적 경험이 유기체의 생존과 안녕의 관점에서 자극이나 사건을 평가하는데서 기인한다고 본다 (Plutchik, 1980; Frijda, 1986; Lazarus, 1991). 정서는 이와같이 자극평가검사의 결과로 협응과정이 시작되면서 발생하고 하위체계의 독립적 기능으로 돌아가면서

끝난다. 자극평가가 각 하위체계에 직접적으로 영향을 미치고 또한 모든 체계들이 정서에피소드동안 깊이 상호 관련되어 있기 때문에 복잡한 조절과 다면적인 상호 작용이 수반된다. 이 이론적 틀에 근거하여 정서특정적인 운동표현(음성표현을 포함하여)의 패턴을 예언할 수 있다. 첫째, 주요 정서들에 대하여 평가결과(즉 특정 SECs의 결과)에 따른 전형적인 프로파일을 명세화할 수 있다. 과거의 연구 및 음성산출의 생리적 기제에 관한 지식에 근거하여 특정 SECs의 결과가 음성산출(특히 호흡과 발성)에 미치는 영향을 음향신호에서의 변화의 속성을 명세화하는 것을 통하여 예언한다.

즉, 발성은 다른 모든 운동 표현과 마찬가지로, 근육 조직의 움직임에 의해서 생성된다. SECs의 각 결과는 체신경계와 자율신경계에 영향을 미치고 이는 음성근육의 변화, 호흡 및 점액분비의 변화를 야기하여 음성표현을 변화시킨다.

만일 자극이 위협하고 행동을 필요로 한다고 평가된다면 (생존의 목적에 중대한 장애), 체신경계의 기능적 반응 때문에 (증가된 근육 긴장) 목소리의 기본주파수가 상승할 것이다. 동시에 자율신경계에서의 변화 때문에 (교감신경계의 우세) 타액의 분비가 감소할 것이다. 두 가지의 효과는 목소리를 높게 하기 쉽다 (목젖 진동과 성대 공명에서의 변화). 만일 잠시 후에 그 상황이 장난이었다고 재평가되면, 재빠르게 반응하는 체신경계는 다시 변할 것이고, 근육의 긴장 감소로 목소리는 낮아질 것이다. 자율신경계는 좀 더 느리게 반응하기 때문에 성대 공명에서의 감소된 타액의 효과는 좀더 지속될 것이다. 아래의 표 1은 정서가 주요 음향학적 지표들에 미치는 영향을 예언한 것이다. Scherer의 성분과정이론은 음성 정서표현의 특징들을 예언하는데 유용한 이론적 틀을 제공해 주지만 아직 경험적 증거는 별로 없다.

음성 의사소통의 해독(decoding)에 관한 연구들은 아직 소수에 불과하다(Van Bezooijen, 1984; Scherer, London, & Wolf, 1973; Wallbot & Scherer, 1986). 아마도 음성정보를 다루는데 요구되는 기술상의 어려움이 주된 이유일 것이다. 몇 안되는 선행 연구들을 고찰해보면, 청자들이 음성 표현으로부터 화자의 감정적 상태와 태도를 추정하는데 능숙함을 알 수 있다. 정답률은 일반적으로 50% 내외의 수준인데 (제한된 숫자의 선지

표 1. 주요 음향학적 지표에 미치는 정서의 영향 예언

정서	지표	기본주파수 평균	평균 에너지	발성구간의 길이
만족	속	낮음	낮음	길음
환희	회	높음	높음	짧음
협오	오	높음	높음	-
경멸	멸	보통	매우높음	-
슬픔	픔	보통	매우낮음	길음
비통	통	높음	높음	짧음
불안	안	높음	-	-
공포	포	매우높음	높음	매우짧음
짜증	증	보통	높음	-
분노	노	보통	높음	짧음
지루함	루	낮음	보통	-
죄책감	책	높음	-	-

주. Banse & Scherer(1996, p.617)로부터 발췌.

로부터 추정하는 효과를 조정한 이후), 이는 우연에 의해 기대되는 확률을 훨씬 상회한다 (Pittam & Scherer, 1993). 이 분야의 대부분의 연구들은 분노, 두려움, 슬픔, 기쁨, 혐오 등 적은 숫자의 정서들에 초점을 맞추고 있다. 음성 정서표현의 해독 연구에서 다음과 같은 질문들이 관심을 끈다.

첫째, 개별 정서들의 재인율은 각각 어느 정도일까? 선행연구들에 따르면 정서들의 평균 재인율은 약 50% 정도이고(Van Bezooijen, 1984; Scherer, Banse, Wallbott & Goldbeck, 1991), 개별 정서들도 대체로 비슷한 수준의 재인율을 보였다. 개별 정서들 중 슬픔이나 분노와 같은 정서는 높은 재인율을 보였고, 혐오는 낮은 재인율을 보였다.

둘째, 화자의 성과 청자의 성은 재인과 어떤 관계가 있을까? 여자가 남자보다 정서적인 존재이며 정서적 능력이 더 뛰어나다고 알려져 있다. 그렇다면 여자가 남자보다 재인율이 더 높을까? 그리고 남자의 음성은 남자가, 여자의 음성은 여자가 더 잘 재인할 것인가? 이러한 남녀 차이가 모든 정서들에서 발견될 것인가 등이다.

셋째, 활성화 단서와 정서가 단서들이 음성표현의 재인(그리고 혼동)에 어떤 영향을 미칠까? 정서가 동일한 방향의 정서표현들간(즉, 정적 정서와 정적 정서간, 그리고 부정 정서와 부정 정서간) 혼동률이 높을 것으로 예상되며, 활성화의 경우에도 활성화정도가 비슷한 정서표

현들간(즉, 활성화정서와 활성화정서간, 그리고 비활성정서와 비활성정서간) 혼동률이 높을 것으로 예상된다.

국내에서는 조은경과 조철우등 (조은경, 조철우와 민경환, 1997; 조철우, 조은경과 민경환, 1997)이 정서의 음성표현의 입력패턴을 연구한 논문이 있다. 그런데 이 연구에서는 기쁨, 슬픔, 두려움, 화의 4가지 정서만을 조사하였고 분석도 피치, 강도, 발연속도의 세 지표에 국한했다.

본 연구의 목적은 기존의 연구들보다 정서의 종류와 분석지표를 다양화하여 음성 정서표현의 음향학적 프로파일을 구하고(실험 1), 정서와 성 및 정서기/활성화 차원에 따른 재인율의 차이를 알아보는 것(실험 2)이다. 본 연구는 Banse와 Scherer의 1996년 연구의 절차를 따라서 수행되었다.

실험 1

실험 1은 주요 정서들의 음성표현을 분석하여 음향프로파일을 발견하기 위해서 수행되었다. 기존의 입력연구에서 사용된 정서의 종류는 다양하지 못하였고, 음향 측정치 또한 한정되어 있었으며 이론적인 배경이 없었다. 본 실험에서는 그러한 한계점들을 극복하기 위하여 노력하였다. 우선 정서를 다양화하여 14개의 정서를 대상으로 하였다. 특히 4쌍의 정서 (만족/환희, 슬픔/비통, 불안/공포, 짜증/분노)는 유사한 정서가를 지니고 있지만 활성화 차원에서 상이한 정서들로 이들 쌍간에 음성 패턴에서의 차이가 있는지를 보았다. 또한 다양한 음향 측정치를 분석하였다.

방법

연기자

일년 이상 성우 전문교육과정을 거친 남자 3명, 여자 3명의 성우가 실험에 참가하였다. 이들에게는 실험참가에 대한 사례비가 지급되었다.

정서

만족, 환희, 슬픔, 비통, 불안, 공포, 짜증, 분노, 혐오, 경멸, 지루함, 죄책감, 자괴심, 흥미의 14가지의 대표적인 정서들이 포함되었다. 이들중 만족/환희, 슬픔/비통,

불안/공포, 짜증/분노는 유사한 정서가를 지니고 있지만 상이한 활성화 강도를 지닌 정서 쌍들이다.

시나리오

14개 각각의 정서에 대하여 두 개씩의 시나리오가 개발되었다. 각 시나리오는 해당 정서를 환기시키는데 적절한 내용으로 정서 연구실 대학원생들에 의하여 작성되었다. 이렇게 만들어진 총 28개의 시나리오를 심리학과 대학원생 22명에게 제시하였다. 그리고 각 시나리오들이 만족, 환희, 혐오, 경멸, 슬픔, 비통, 불안, 공포, 짜증, 분노, 지루함, 죄책감, 자괴심, 흥미의 14개의 정서 중 어떤 정서를 유발하는 상황인지를 답하게 하였다. 그 결과 27개의 시나리오가 80%이상의 정답률을 보여서 채택되었고, 정답률이 80% 이하인 시나리오 1개를 대체하였다. 대체한 시나리오를 심리학과 대학원생 12명에게 평정시킨 결과 100%의 정답률을 나타냈다.

표준 문장

말레이시아어로 이루어진 다음과 같은 무의미 문장을 사용하였다. “아나 썸뽕시니 썸뽕 찬틱”

설계

본 연구의 설계는 정서(14) X 시나리오(2) X 성우(6) X 문장(1) X 시행(3) 이었다. 원래 Banse와 Scherer의 연구는 정서(14) X 시나리오(2) X 성우(12) X 문장(2) X 시행(2)의 설계였다. 본 연구에서 성우의 숫자가 열두명에서 여섯명으로 줄어든 것은 시간과 비용을 절감하기 위해서였다. 또한 문장이 두 개에서 한 개로 줄어든 것은 두 문장간에 음성표현에서의 차이가 미미하였기 때문이었다. 원래 본 실험에서도 각 정서당 두 개씩의 문장을 사용하여 음성을 채록했는데, 녹음기사가 자의로 녹음강도를 조정하였기 때문에 재녹음을 하면서 문장을 하나로 통일하였다. 이로써 분석이 좀더 용이해지는 장점이 있는 반면에 일반화 가능성은 조금 낮아지는 단점도 있다.

음성녹음 절차

연구자들은 녹음시작 한 시간 전에 스튜디오로 오도록 하여 시나리오와 표준문장을 충분히 숙지하고 여러

차례 연습을 하도록 한 후에 녹음을 하였다. 녹음은 전문적인 스튜디오의 디지털 녹음실에서 이루어졌다. 시나리오별로 정서를 환기시킨 후 각 문장을 세 차례씩 읽도록 하여 녹음하였다. 음성 강도를 일정하게 유지하기 위하여 남자는 마이크에서 수직으로 내려온 지면에서 30cm 앞, 여자는 15cm 앞에 발을 두고 녹음하였다. 녹음 수준은 처음에 조정된 후 중간에 변화시키지 않았다. 디지털 테이프(120분)가 녹음에 사용되었다.

음성표현의 선정

심리학과 대학원 학생 6명의 평정에 의하여 시나리오별로 성우 당 3개씩 얻어진 음성표현 중 가장 정서표현이 잘된 녹음을 한 개씩 선정하여 분석하였다. 따라서 분석에 사용된 음성표현은 시나리오(28) X 성우(6)로 총 168개였다.

음향측정치

음향측정치로는 기본주파수 평균(MF0), 기본주파수 평균분산(SDF0), P25F0, P75F0, 평균에너지(Melog), 유성음 구간의 길이(DurVo), 발성구간의 길이(DurArt), Hamml, DO1000, PE500, PE1000등의 11가지를 얻었다. 이들 측정치들에 대한 간략한 소개는 표 2에 제시되어 있으며 보다 상세한 설명을 위해서는 Banse와 Scherer(1996)을 참조하기 바란다.

표 2. 분석에 사용한 음향 측정치

측정치	의 미	단위
MF0	발성 내 유성음 구간의 기본주파수 평균	Hz
SDF0	발성 내 유성음 구간 기본주파수의 표준편차	Hz
P25F0	발성 내 백분율로 25번째 피치값	Hz
P75F0	발성 내 백분율로 75번째 피치값	Hz
Melog	발성 내 평균에너지 (10*log)	dB
DurVo	발성 내 유성음 구간의 총 길이	msec
DurArt	무발성 구간을 제외한 음성의 총길이	msec
Hamml	0 - 2kHz 내 최대 에너지와 2 - 5kHz 내 최대에너지의 차이	dB
DO1000	1000Hz 이상 에너지 envelop의 기울기 (dB scale : log에서 구함)	
PE500	500Hz 이상과 이하의 에너지 비율	
PE1000	1000Hz 이상과 이하의 에너지 비율	

결과 및 논의

각 정서별로 두 개의 시나리오를 사용하였기 때문에 두 시나리오간 음향측정치들간에 차이가 있는지를 조사하였다. 그 결과 불안이 유성음 구간의 길이(DurVo)에서 ($t(10)=5.201, p<.01$), 분노가 발생구간의 길이(DurArt)에서($t(10)=2.590, p<.05$), 환희가 PE500($t(10)=2.666, p<.05$)과 PE1000($t(10)=2.284, p<.05$)에서 각각 유의미한 차이를 보였을 뿐으로 대체로 동일정서의 시나리오간에 음성표현의 차이가 없다고 결론을 내릴 수 있었다. 따라서 두 시나리오의 음향측정치들의 평균값을 제시하고 분석에 사용하였다.

표 3은 각 정서별로 6명의 성우들의 음향측정치 평균값들을 제시하고, 쌍을 이루는 정서들간을 비교한 표이다. 표 3에서 보면 만족/환희는 기본 주파수 평균(MF0), P75F0, 평균 에너지(Melog), Hamml, PE500, PE1000에서 유의미한 차이를 보였고, 슬픔/비통은 기본 주파수 평균(MF0), P25F0, 평균 에너지(Melog), PE500, PE1000에서 유의미한 차이를 보였다. 짜증/분노는 P25F0, 평균 에너지(Melog), Hamml, PE500, PE1000에서 유의미한 차이를 보였고, 불안/공포는 어떤 측정치에서도 유의미한 차이를 보이지 않았다. 대체로 쌍을 이루는 정서들

중에서 활성화 차원이 높은 정서들(환희, 비통, 공포, 분노)이 낮은 정서들(만족, 슬픔, 불안, 짜증)보다 기본 주파수 평균(MF0), P25F0, P75F0, 평균 에너지(Melog)에서 더 높았고, 반대로 공학적 측정치인 Hamml, PE500, PE1000에서는 더 낮았다. 이는 활성화 차원이 높은 정서가 더 발음하기가 힘들음을 나타낸다.

그림 1에서 그림 9까지는 각 성우의 정서별 음향측정치들의 프로파일을 제시하고 있다. 표 3에 제시된 음향측정치들 중에서 P25와 P75는 기본 주파수 평균(MF0)과 거의 유사하였기 때문에 그림에서는 생략하였다.

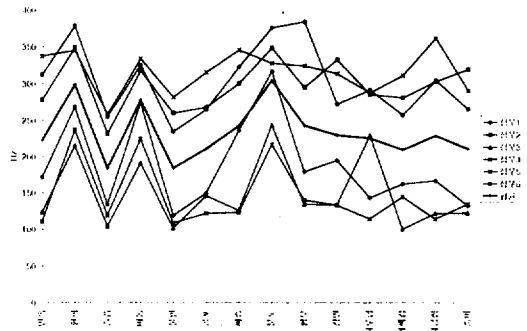


그림 1. 성우별 기본주파수 평균 (MF0)

표 3. 정서별 음향측정치의 평균

	MF0 (Hz)	SDF0 (Hz)	P25F0 (Hz)	P75F0 (Hz)	Melog (dB)	DurVo (msec)	DurArt (msec)	Hamml	DO1000	PE500	PE1000
만족	222.3*	4.3	210.3	238.8**	58.89**	1358	1936	19.00*	-0.024	0.11**	0.21**
환희	298.8*	6.2	256.6	345.3**	70.53**	1391	1903	10.00*	-0.025	0.08**	0.18**
슬픔	183.9**	4.3	175.5**	194.6	49.86**	1278	2256	20.83	-0.025	0.11*	0.21*
비통	278.0**	4.6	319.3**	233.4	62.25**	1463	2311	19.33	-0.021	0.09*	0.19*
불안	184.4	2.7	198.4	172.4	50.76	1132	2246	18.75	-0.024	0.10	0.20
공포	211.1	3.2	226.0	196.9	53.48	1192	2221	17.33	-0.023	0.09	0.19
짜증	242.7	8.8	271.3*	224.1	59.84**	1295	1861	18.08**	-0.024	0.10**	0.20**
분노	305.4	6.9	363.4*	259.6	72.97**	1253	1884	10.75**	-0.025	0.09**	0.18**
혐오	243.4	5.6	267.8	227.2	59.05	1284	1947	17.17	-0.025	0.10	0.19
경멸	230.4	4.7	253.0	213.9	61.19	1422	1945	16.58	-0.028	0.10	0.20
지루함	225.8	5.9	247.0	216.8	57.55	1740	2461	20.42	-0.022	0.10	0.21
죄책감	210.1	5.3	230.8	191.6	50.89	1225	2196	22.92	-0.024	0.11	0.21
자극심	229.4	7.2	249.7	218.1	60.01	1262	1910	17.17	-0.025	0.10	0.19
흥미	211.7	7.0	231.1	198.1	55.06	1374	2118	20.42	-0.021	0.10	0.20

** p<.01, *p<.05

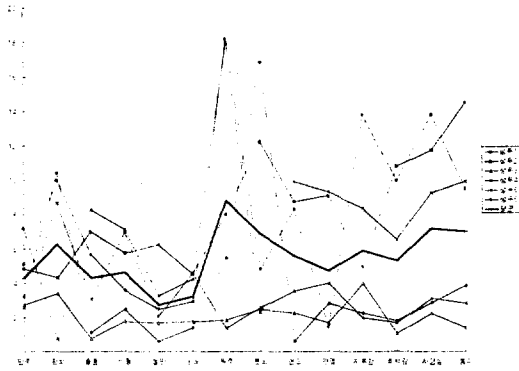


그림 2. 성우별 유성음 구간 내 기본주파수 평균 변산

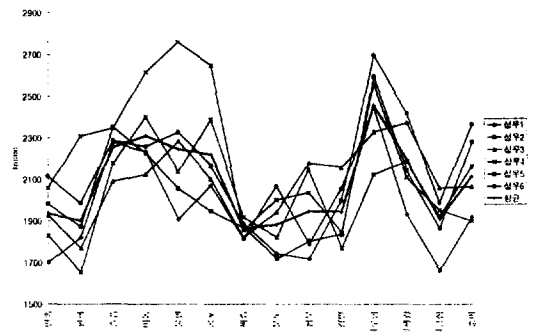


그림 5. 성우별 발성구간의 길이 (DurArt)

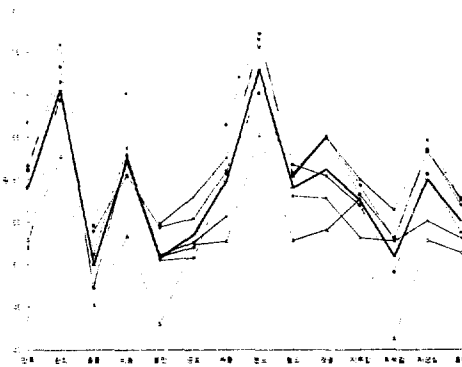


그림 3. 성우별 평균에너지 (Melog)

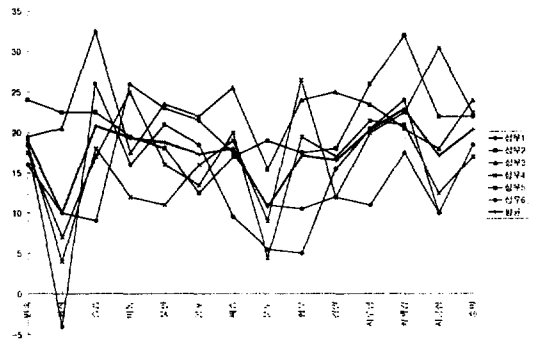


그림 6. 성우별 Hammi

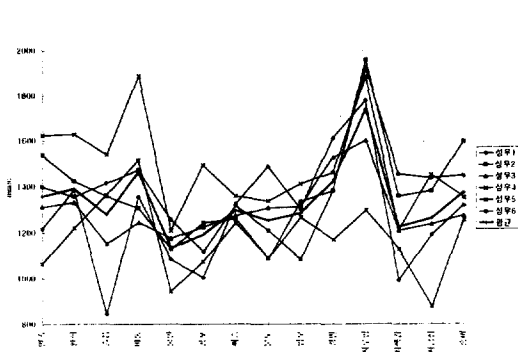


그림 4. 성우별 유성음 구간의 길이

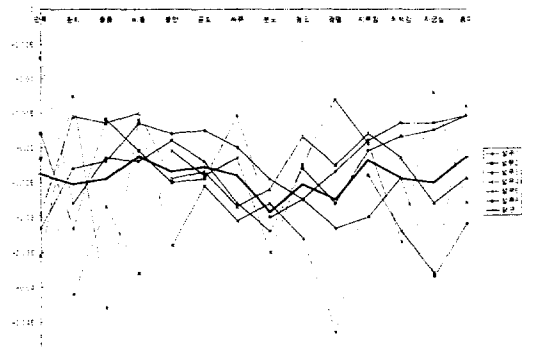


그림 7. 성우별 DO1000

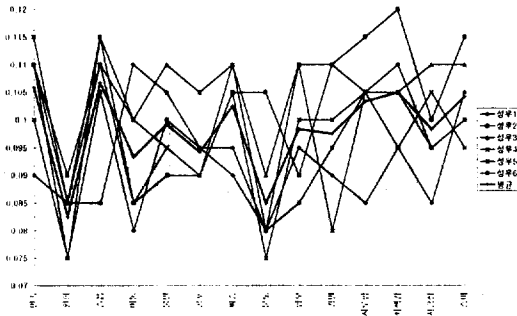


그림 8. 성우별 PE500

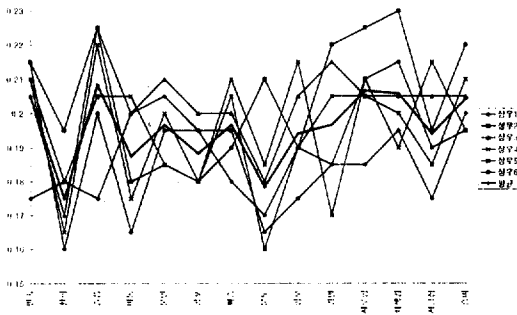


그림 9. 성우별 PE1000

실험 1의 결과를 Scherer의 정서의 성분과정이론을 토대로 표 1에서 예언한 것들과 비교해보면 다음과 같다. 표 1에서는 기본주파수 평균(MF0), 평균에너지(Melog), 발생구간의 길이(DurArt)에 대해서만 예언을 하였기 때문에 이 세 측정치에서만 예언과 비교하였다.

- (1) 만족은 예언했던 대로 낮은 기본주파수 평균, 낮은 평균에너지가 나타났다.
- (2) 환희는 예언했던 대로 높은 기본주파수 평균, 높은 평균에너지, 짧은 발생구간의 길이가 나타났다.
- (3) 혐오는 예언했던 대로 결과가 나타나지 않았다. 보통의 기본주파수 평균, 보통의 평균에너지가 나타났다.
- (4) 경멸은 복합적인 결과가 나타났다. 기본주파수 평균은 예언했던 대로 보통의 수준이 나타난 반면, 평균 에너지는 예언과는 달리 보통의 수준이 나타났다.
- (5) 슬픔도 복합적인 결과가 나타났다. 기본주파수 평균은 예언과는 달리 낮은 수준이 나타난 반면 평균 에너지는 예언했던 대로 매우 낮은 수준이 나타났고, 발생

구간의 길이도 예언했던 대로 길게 나타났다.

(6) 비통도 복합적인 결과가 나타났다. 기본주파수 평균은 예언대로 높은 수준이, 평균에너지도 예언했던 대로 높은 수준이 나타난 반면, 발생구간의 길이는 예언과는 다르게 길었다.

(7) 불안은 예언과는 다르게 낮은 기본주파수 평균이 나타났다.

(8) 공포는 예언과는 완전히 다른 결과가 나타났다. 낮은 기본주파수 평균, 낮은 평균 에너지가 나타났고, 발생구간의 길이도 길었다.

(9) 짜증은 복합적인 결과가 나타났다. 예언했던 대로 보통의 기본주파수 평균이 나타났지만, 평균에너지는 예언과는 달리 보통의 수준이 나타났다.

(10) 분노도 복합적인 결과가 나타났다. 예언과는 달리 매우 높은 기본주파수 평균이 나타난 반면, 평균 에너지는 예언했던 대로 높은 수준이 나타났다.

(11) 지루함은 예언했던 대로 낮은 기본주파수 평균과 보통의 평균 에너지가 나타났다.

(12) 죄책감은 예언과는 달리 보통의 기본주파수 평균이 나타났다.

이상의 결과들을 종합하면 Scherer의 정서의 성분과정이론은 본 연구에서 일관성 있게 모두 지지되지는 않았고 그 중 일부가 지지되었다. 이러한 결과는 Scherer의 이론이 좀 더 이론적인 정교화와 실험연구를 통한 검증이 거쳐야 하는 초기 단계의 이론임을 시사한다.

한편, 주요한 음향측정치인 기본주파수 평균, 평균에너지, 발생구간의 길이 등에서 6명의 성우들이 상당히 유사한 프로파일을 보여주었다. 이것은 주요한 음향측정치들에서 상당히 신뢰로운 값을 얻었음을 의미한다. 그림1의 기본주파수 평균의 경우에는 1, 3, 5의 남자성우와 2, 4, 6의 여자성우로 프로파일이 구분될 수 있음을 보여주었다. 이것은 기본주파수 평균이 남녀의 음성 차이에 민감한 측정치임을 시사한다. 특정 정서를 잘 대표하는 지표도 발견되었다. 예를 들면 지루함은 발생 구간의 길이에서 다른 정서들과 구분되는 특징을 나타냈으며, 분노나 환희는 공학적 측정치들(Hamml, PE500, PE1000 등)에서 일관성 있는 패턴을 보여주었다.

그림 10부터 그림 12까지는 주요 음향측정치들을 선행연구와 비교한 것이다. 대부분의 음향측정치들은

피험자

남자 18명, 여자 18명의 심리학과 학부생이 실험 2에 참가하였다.

음성자극 선정

실험 1의 입력연구에서 분석되었던 음성표현중 시나리오별로 남녀 음성표현 각각 1개씩, 즉 정서(14)×시나리오(2)×남여(2)의 총 56개의 음성표현을 자극으로 선정하였다.

자극제시

음성표현은 CD로 제작되었고, CD플레이어로 무선적으로 제시되었다. 각 음성표현이 제시된 후 약 10초간의 간격을 주어 피험자에게 정서를 추정케 하였다. 피험자들은 남녀가 쌍을 이루어서 두 명씩 실험에 참가하였다.

결과 및 논의

표 4의 왼쪽의 비합산 조건은 각 정서의 정답률(재인율)을 나타낸 것이다. 이 조건에서의 정답률은 Scherer의 연구가 47.9%로 실험 2의 정답률 46.3%보다 조금 높았다. 표 4의 오른쪽의 합산 조건은 쌍을 이루는 정서중 대응하는 다른 하나의 정서로 대답한 반응을 정답으로 간주한 것이다. 예를 들면 만족에 해당하는 문장을 제시하였을 때 만족과 환희를 모두 정답으로 간주하여 이 두가지 응답비율을 합산하여 정답률을 구한 것이다. 이 합산조건에서의 정답률도 Scherer 연구가 54.9%로 실험 2의 52.2%보다 높았다. 그러나 합산과 비합산 두 조건 모두 Scherer의 실험과 실험 2의 차이는 크지 않았다.

개별 정서별로 볼 때, 만족과 환희의 두 정서는 Scherer 연구에 비하여 매우 낮은 정답률을 나타냈다. 그리고 혐오의 경우는 실험 2에서는 우연수준(7.1%)과 비슷한 정도의 정답률(7.6%)을 보였다. 실험 2에서 낮은 정답률(비합산조건에서 .25이하)을 보인 정서들은 만족, 환희, 불안, 혐오 등이었으며, 높은 정답률(비합산조건에서 .75 이상)을 보인 정서들은 비통, 분노, 지루함 등이었고 슬픔, 공포, 짜증, 경멸, 죄책감, 자긍심, 흥미 등은 중간 정도의 정답률(비합산조건에서 .25에서 .75사이)을 보였다.

표 4. 주요 정서들의 재인율

정서 (비합산)	실험 2	Scherer	정서 (합산)	실험 2	Scherer
만 족	0.229	0.520	만 족	0.257	0.540
환 회	0.174	0.380	환 회	0.194	0.390
슬 픔	0.611	0.520	슬 픔	0.806	0.730
비 통	0.813	0.470	비 통	0.931	0.550
불 안	0.236	0.420	불 안	0.319	0.550
공 포	0.417	0.360	공 포	0.778	0.630
짜 증	0.465	0.340	짜 증	0.472	0.510
분 노	0.813	0.780	분 노	0.826	0.880
혐 오	0.076	0.150	혐 오	0.076	0.150
경 멸	0.403	0.600	경 멸	0.403	0.600
지루함	0.896	0.760	지루함	0.896	0.760
죄책감	0.444	0.220	죄책감	0.444	0.220
자긍심	0.299	0.430	자긍심	0.299	0.430
흥 미	0.611	0.750	흥 미	0.611	0.750
평 균	0.463	0.479	평 균	0.522	0.549

동일한 기준을 적용했을 때 Scherer의 연구에서는 혐오와 죄책감의 정답률이 낮았고, 분노, 지루함, 흥미의 정답률이 높았으며, 만족, 환희, 슬픔, 비통, 불안, 공포, 짜증, 경멸, 자긍심 등의 정답률이 중간 수준이었다. 따라서 여러 정서들을 망라한 평균 정답률은 두 연구간에 비슷했지만 개별 정서의 정답률은 높은 일치도를 보이지는 않음을 알 수 있다.

각 정서별로 오답패턴을 분석해보면 각 정서의 음성표현이 어떤 정서로 자주 오인되는지를 알 수 있다. 특히 만족, 환희, 불안, 혐오의 정서들은 정답률보다 다른 특정정서로 오인한 비율이 더 높았다. 만족의 경우에는 정답률이 22.9%인데 비해 자긍심으로 응답한 비율이 29.9%였으며, 환희의 경우에는 정답률이 17.4%인데 비해 분노로 응답한 비율이 18.8%, 비통으로 응답한 비율이 17.4%였다. 또 불안의 경우에는 정답률이 23.6%였으나 슬픔과 죄책감으로 응답한 비율이 각각 28.5%였으며, 혐오의 경우에는 정답률이 7.6%에 불과하여 죄책감, 불안, 슬픔, 지루함으로 응답한 비율이 모두 정답률을 상회하였다. 그밖의 정서들은 정답률이 특정정서로 응답한 비율보다 높았다. 참고로, 슬픔은 비통으로 오답한 비율이 가장 높았고 비통 역시 슬픔으로 오답한 비

율이 가장 높았다. 또한 공포는 불안, 짜증은 경멸, 분노는 경멸, 경멸은 혐오, 지루함은 경멸, 죄책감은 슬픔, 자긍심은 만족, 흥미는 지루함으로 오답한 비율이 각각 가장 높았다.

오답패턴을 정서가와 활성화의 두 차원상에서 해석하기 위해서 표 5에서 보듯이 차원별로 혼동수를 분석하였다. 첫 번째 칸은 혼동수를 나타낸다(예, 정 - 정은 정적 정서를 제시했을 때 제시된 정서 이외의 다른 정적 정서를 대답한 혼동수이다). 두 번째 칸은 혼동할 수 있는 정서의 개수이다. 14개의 정서중 정적 정서는 4개(만족, 환희, 자긍심, 흥미)이고 부적 정서는 10개(슬픔, 비통, 불안, 공포, 짜증, 분노, 혐오, 경멸, 지루함, 죄책감)이므로 “정 - 정”(즉, 정적 정서를 제시했을 때 다른 정적 정서로 오답할 가능정서수)는 3이고 마찬가지로 “부 - 부”는 9, “정 - 부”는 10, “부 - 정”은 4이다. 혼동수를 혼동가능정서수로 나누면 세 번째 칸의 혼동지수를 구할 수 있다. 활성화차원을 보면 비활성정서(만족, 슬픔, 불안, 짜증)와 활성정서(환희, 비통, 공포, 분노)가 각각 4개씩 짝을 이룬다. 따라서 두 번째 행의 혼동가능정서수가 “활 - 활”은 3, “비활 - 비활”은 3, “활 - 비활”은 4, “비활 - 활”은 4가 된다.

표 5의 정서가 차원에서의 혼동 수 분석을 보면 부적 정서를 다른 부적 정서로 혼동한 “부 - 부”의 혼동지수가 가장 높고, 그 다음이 “정 - 정”, 그리고 “정 - 부”, “부 - 정”의 순서였다. 따라서 동일방향의 정서가를 가진 정서들간에 음성표현에 따른 변별이 어렵고, 그중에서도 정적 정서들보다는 부적 정서들간에 변별이 더 어렵다고 말할 수 있다. 상이한 방향의 정서가를 가진 정서들간에는 비교적 변별이 쉬웠으며 특히 부적 정서를 정적 정서로 혼동하는 일은 아주 드물었다.

표 5의 활성화 차원에서의 혼동수 분석을 보면 활성 정서를 비활성 정서로 혼동한 “활 - 비활”의 혼동지수가 가장 높고, 차례로 “활 - 활”, “비활 - 비활”, “비활 - 활” 순으로 나타났다. 동일한 방향에서의 혼동지수가 높았던 정서가 차원과 달리 “활 - 비활”에서의 혼동지수가 가장 높게 나온 것은 공포 - 불안의 혼동수가 “활 - 비활”의 혼동수 96개중 60개로 압도적인 영향을 미쳤기 때문이며 이를 제외하면 역시 같은 활성화집단에서의 혼동지수가 더 높았다.

표 5. 정서가/활성화 차원상 혼동수 분석

	혼동수	혼동가능 정서수	혼동수/혼동 가능정서수
부 - 부	601	9	66.8
정 - 정	123	3	41.0
정 - 부	259	10	25.9
부 - 정	70	4	17.5
활 - 비활	96	4	24.0
활 - 활	65	3	21.7
비활 - 비활	63	3	21.0
비활 - 활	63	4	15.8

아래 표 6에서는 음성표현을 제공하는 연기자와 재인 과제를 하는 해독자(피험자)를 남녀로 구분하여 평균 재인율을 비교, 제시하였다. 여기서 남녀 평균 재인율은 피험자별로 두 개의 시나리오에 대한 재인율의 평균값을 구한 뒤 다시 남자집단과 여자집단의 평균값을 구한 것이다. 맨 아래칸의 평균은 14개 정서의 재인율을 평균한 값이다. 연기자의 성별을 피험자 내 변인으로, 해독자의 성별을 피험자 간 변인으로 하는 변량분석을 실시하였다.

표 6. 남녀간 정서별 재인율 비교

정서	성우남		성우여		성우효과 (F)	해독자 효과(F)	성우×해독자효과 (F)
	남 해독	여 해독	남 해독	여 해독			
평균	0.454	0.391	0.490	0.522	26.889**	0.415	7.693**
만족	0.084	0.028	0.500	0.306	31.158**	3.732+	1.246
환희	0.278	0.194	0.111	0.111	4.135*	0.552	0.459
슬픔	0.528	0.444	0.722	0.778	12.975**	0.26	0.899
비통	0.778	0.834	0.834	0.806	0.046	0.046	0.410
불안	0.194	0.056	0.334	0.334	13.235**	1.177	1.471
공포	0.361	0.334	0.500	0.472	5.484*	0.158	0.000
짜증	0.361	0.139	0.639	0.722	27.550**	1.341	3.469+
분노	0.806	0.861	0.806	0.778	0.515	0.033	0.515
혐오	0.056	0.028	0.111	0.111	2.297	0.124	0.092
경멸	0.472	0.500	0.334	0.306	5.693*	0.000	0.158
지루함	0.944	0.861	0.834	0.944	0.069	0.074	3.400+
죄책감	0.334	0.416	0.389	0.639	2.471	3.755+	0.890
자긍심	0.584	0.194	0.139	0.278	7.871**	3.577+	16.814**
흥미	0.578	0.584	0.611	0.722	1.772	1.423	0.111

**p<.01, *p<.05. +p<.08

표 6에서 발견된 중요한 결과들을 요약하면 다음과 같다. 우선 정서별 재인율을 평균한 맨 아래 칸을 보면 남성 음성자극보다는 여성 음성자극이 더 잘 재인 되었음을 알 수 있다. 개별 정서별로 보면 만족, 슬픔, 불안, 공포, 짜증 등에서 남성음성 보다는 여성음성이 더 잘 해독되었고, 환희, 경멸, 자긍심 등에서는 남성음성이 여성음성보다 더 잘 해독되었다.

연기자의 성과 해독자의 성의 상호작용 효과가 정서 평균에서 나타났는데, 남성음성은 남성이, 그리고 여성음성은 여성이 더 잘 해독하였다. 이러한 상호작용 효과는 개별정서별로는 자긍심에서 나타났고, 짜증과 지루함에서도 약하게 나타났다.

해독자의 성은 재인율과 거의 관계가 없었다. 개별 정서들중 만족과 자긍심 및 죄책감에서 약한 정도로 남녀간 차이를 보였을 뿐이었다. 이들 중 만족과 자긍심은 남자가, 그리고 죄책감은 여자가 더 잘 재인하였다.

전 체 논 의

본 연구의 목적은 Scherer의 이론과 연구에 기반을 두고 한국인을 대상으로 주요정서들의 음성표현의 특징적인 지표들을 수집하고(실험 1의 입력연구), 더불어 이들 음성표현들을 사람들이 얼마나 잘 재인할 수 있나를 조사하였다(실험 2의 재인연구).

본 연구에서 확인된 주요 결과들과 그 의미는 다음과 같이 정리될 수 있다.

첫째, 실험 1의 결과는 Scherer의 정서의 성분과정 이론 중 일부를 지지하는 것으로 나타났다. Scherer의 이론을 검증한 연구는 아직 별로 없으며 이론 자체도 정교한 체계를 갖추고 있지 못하다. 따라서 Scherer 이론은 본 실험 1과 같은 연구들이 축적됨으로써 타당성이 검증될 수 있고 또한 보다 정교한 이론으로 발전될 수 있다고 본다.

둘째, 실험 1의 결과는 Scherer의 연구(Banse & Scherer, 1996) 결과와 대체로 일치하였다. 대부분의 음향측정치들이 Scherer의 연구와 비슷한 패턴을 나타냈는데 공포의 경우에만 두 연구에서 뚜렷한 차이를 보였다. Scherer의 연구에서 공포가 강하고 높고 짧은 음성으로 표현되는데 반하여 실험 1에서는 약하고 낮고 긴 음

성으로 표현되었다. 실험 1의 공포의 음성표현이 잘못 되었을 가능성은 실험 2의 재인연구에서 공포의 재인율이 41.7%로 Scherer연구의 36.0%보다 높다는 점에서 배제될 수 있다. 이 결과는 서양과 동양 사이에 공포의 음성표현에서 문화적 차이가 존재할 가능성을 시사해준다. 요약하면, 정서의 음성표현들이 문화적 보편성을 갖고 있으나 일부 정서(예, 공포)에서는 문화적 차이도 존재한다는 사실이 확인되었다.

셋째, 활성화 차원에서 쌍을 이루는 정서들의 경우, 활성화와 관련된 지표들(기본주파수 평균, 평균에너지 등)에서 차이를 보였다. 또한 특정 정서를 잘 대표하는 지표들도 발견되었다(예, 지루함은 발성구간의 길이 등). 앞으로 특정정서나 특정차원을 대표하는 음향측정치들을 탐색하는 후속연구들이 기대된다. 이러한 연구들은 실용적인 가치도 높을 것으로 생각된다.

넷째, 실험 2의 재인연구를 Scherer의 연구와 비교해 볼 때 개별정서에 따른 재인율의 높고 낮음에서는 두 연구간에 일치도가 낮았으나 전체적인 재인율(즉 평균 재인율)은 .5내외로 아주 근접하였다. 그러므로 Scherer 연구의 피험자들과 실험 2의 피험자들은 음성으로 표현된 정서에 대해 비슷한 정도의 재인 능력을 갖고 있다고 말할 수 있다.

다섯째, 실험 2에서 재인율이 높았던 정서는 비통, 분노, 지루함 등이며 재인율이 낮았던 정서는 만족, 환희, 불안, 혐오 등이었다. 특히 혐오는 우연수준에 근접할 정도로 재인율이 낮았으며 Scherer의 연구에서도 재인율이 가장 낮았다. 혐오의 경우에는 죄책감, 불안, 슬픔, 지루함으로 오답한 비율이 모두 정답율을 상회하여 음성표현의 재인이 대단히 어려운 정서임을 보여 주었다.

여섯째, 오답패턴의 분석에서 동일방향의 정서들을 가진 정서들간에 혼동율이 높았고 특히 부적 정서들간의 변별이 정적 정서들간보다 더욱 어려웠음이 나타났다. 상이한 방향의 정서들을 가진 정서들간에는 비교적 변별이 쉬웠고 특히 부적 정서를 정적 정서로 혼동하는 일은 극히 드물었다. 한편 활성화 차원의 경우는 공포-불안의 혼동수가 너무 많아서 결과를 해석하기가 어려웠다.

일곱째, 재인율에서의 성차를 분석해보면, 전체적으로 여성 음성자극이 남성 음성자극보다 더 잘 재인되었으

나 개별정서에 따라 남성 음성자극이 더 잘 재인되기도 하였다. 또한 남성음성은 남성이, 그리고 여성음성은 여성이 더 잘 재인하는 경향을 보였으며, 재인율에서 남녀간 성차는 발견되지 않았다.

정서의 음성표현에 관한 연구는 쉽지 않고 여러 가지 제약이 따른다. 우선 첨단 녹음시설이 갖추어져야 하고, 적절한 음성표현을 얻을 수 있는 연기자들을 충분히 확보하여야 하고, 철저하게 통제된 상황에서 녹음이 이루어져야 하며, 채록된 자료를 음성공학에 정통한 프로그래머가 분석을 하여야 한다. 이러한 이유 때문에 한국은 물론 외국에서도 음성표현의 연구는 드물다. 본 연구에서도 이러한 조건들이 완벽하게 충족되었다고 말할 수 없다. 우선 소수의 성우에 의한 정서의 연기가 과연 그 정서의 대표적인 음성표현을 제공할지 의문이다. 성우의 역량도 문제이거니와 개인차도 간과할 수 없다. 특히 본 연구에서는 시간과 경비의 문제 때문에 6명의 성우(Scherer의 연구에서는 12명)만의 음성표현들을 채록하였기 때문에 측정치의 안정성이나 일반화의 면에서 취약하다는 점을 인정할 수밖에 없다. 또한 이러한 “연기”에 의한 음성표현이 자연스러운 정서적 의사소통 상황에서의 음성표현을 잘 반영했다고 보기 힘들다. 이를 보완하기 위하여 자연스러운 사회적 상황에서의 음성자료를 수집하여 분석하는 연구도 이루어져야 한다.

이러한 이유 때문에 본 연구의 결과를 14개의 정서에 대한 한국인의 음성표현의 대표적 음향특징치들을 제시한 연구라고 주장하기는 어렵다. 특히 비언어 측정치들은 높은 신뢰도를 확보하기 어려운 만큼 아무리 잘 수행된 연구라도 하나의 연구에 의존하여 결론을 내리는 것은 위험하다. 앞으로 음성 정서표현에 대한 후속연구들이 필요한 이유이다.

본 연구가 갖는 의미는 다음 세 가지로 정리할 수 있다. 첫째, 본 연구에서는 다양한 정서들을 대상으로 중요한 음향특징치들의 프로파일을 얻었다. 비록 이들 측정치들의 신뢰도를 현재로서는 자신할 수 없지만 후속 연구들에서 확인되거나 비교될 수 있는 기준을 확보했다는 점은 이 연구의 중요한 성과이다. 둘째, 본 연구는 정서의 음성표현의 입력 및 재인에 관하여 많은 중요한 통찰들을 제공하였다. 이들은 위에서 일곱가지로 정리하였으며 모두 후속연구들에서 검증이 가능한 중요한

가설들을 구성한다. 음성표현과 같은 역사가 일천한 연구분야에서는 적절한 가설의 존재가 연구를 활성화시키는데 중요한 역할을 한다. 셋째, 본 연구는 앞으로 음성표현연구에 방법론상으로 많은 참고가 될 것이라 생각한다. 음성표현연구가 드물며 기술적인 어려움이 수반되기 때문에 선행연구의 존재는(훌륭한 연구이든 아니든) 후속연구에 많은 도움을 줄 수 있기 때문이다.

참 고 문 헌

- 조은경 · 조철우 · 민경환. (1997). 자전적 회상을 통한 자연스런 정서음성정보 수집방법에 관한 연구. *한국음향학회지*, 16권 2호, 66-70.
- 조철우 · 조은경 · 민경환. (1997). 정서정보의 변화에 따른 음성신호의 특성 분석에 관한 연구. *한국음향학회지*, 16권 3호, 33-37.
- Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, 70, 614-636.
- Darwin, C. (1872/1965). *The expression of the emotions in man and animals*. Chicago: University of Chicago Press.
- Ekman, P. (1971). Universals and cultural differences in facial expressions of emotion. In J. K. Cole (Ed.), *Nebraska Symposium on Motivation*, 19 (pp. 207-282). Lincoln: University of Nebraska Press.
- Ekman, P. (Ed.). (1982). *Emotion in the human face*. London: Cambridge University Press.
- Frijda, N. H. (1986). *The emotions*. Cambridge: Cambridge University Press.
- Izard, C. E. (1971). *The face of emotion*. New York: Appleton-Century-Crofts.
- Izard, C. E. (1982). *Measuring emotions in infants and children*. London: Cambridge University Press.
- Lazarus, R. S. (1991). *Emotion and adaptation*. New York: Oxford University Press.
- Pittam, J. & Scherer, K. R. (1993). Vocal expression and communication of emotion. In M. Lewis &

- J. M. Haviland (Eds.), *Handbook of Emotions* (pp. 185-198). New York: Guilford Press.
- Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In R. Plutchik & H. Kellerman(Eds.), *Emotion: Theory, research, and experience, Vol. 1*(pp. 3-33). New York: Academic Press.
- Scherer, K. R. (1984). On the nature and function of emotion: A component process approach. In K. R. Scherer & P. Ekman (Eds.), *Approches to emotion* (pp. 293-318). Hillsdale, NJ: Erlbaum.
- Scherer, K. R. (1986). Vocal affect expression: A review and a model for futrure research. *Psychological Bulletin*, 99, 143-165.
- Scherer, K. R., Banse, R., Wallbott, H. G., & Goldbeck, T. (1991). Vocal cues in emotion encoding and decoding. *Motivation and Emotion*, 15, 123-148.
- Scherer, K. R., London, H., & Wolf, J. (1973). The voice of confidence: Paraliguistic cues and audience evaluation. *Journal of Research in Personality*, 7, 31-44.
- Van Bezooijen, R. (1984). *The characteristics and recognizability of vocal expressions of emotion*. Dordrecht, The Netherlands: Foris.
- Wallbott, H. G., & Scherer, K. R. (1986). Cues and channels in emotion recognition. *Journal of Personality and Social Psychology*, 51, 690-699.

Acoustic Profiles and Recognition in Vocal Emotion Expression

Kyung-Sik Nam and Kyung-Hwan Min

Department of Psychology, Seoul National University

The present study replicated Banse & Scherer(1996) to obtain acoustic profiles and estimate recognition accuracy in vocal emotion expression for Korean people. In experiment 1 professional actors' portrayals of 14 emotions(joy, elation, sadness, grief, anxiety, fear, irritation, anger, disgust, contempt, boredom, guilt, pride, interest) were subjected to digital acoustic analysis to obtain profiles of vocal parameters for different emotions. The Korean profiles in general replicated those obtained in Banse & Scherer with the exception of fear. In experiment 2 vocal emotional portrayals were presented to judges to investigate the accuracy of recognition. The overall recognition rate was around 50%, close to the one in Banse & Scherer, but correspondence in recognition rates for the individual emotions was not high between the two studies. Vocal expressions of grief, anger, boredom were easily recognized while those of joy, elation, anxiety, disgust were not. Emotions with the same valence tended to be confused with each other. Female vocal expressions and vocal expressions of the same sex were more easily recognized than male ones and those of the different sex respectively. However differences in recognition accuracy between male and female judges were not found.

Keywords : acoustic profile, recognition rate, vocal emotion expression

원고 접수: 2000년 1월 14일
심사 통과: 2000년 2월 7일