

한국 성인 음성의 음도인식에 관한 연구  
A Study on Pitch Perception of Normal Korean

정 옥 란\* · 김 형 순\*\* · 김 영 태\*\*\* · 서 장 수\*\*\*\*  
(Ok-Ran Jeong · Hyung Soon Kim · Young Tae Kim · Jang Su Suh)

ABSTRACT

This study attempts to determine the fundamental frequency level of male and female voices that Koreans perceive as normal. Seventy-three college students majoring in Speech Pathology participated in the study on a voluntary basis. The subjects listened to a male voice with fundamental frequency of 60 Hz, 80 Hz, 100 Hz, 120 Hz, 140 Hz, 160 Hz, 180 Hz, and 200 Hz, and a female voice with fundamental frequency of 140 Hz, 160 Hz, 180 Hz, 200 Hz, 220 Hz, 240 Hz, 260 Hz, and 280 Hz. The PSOLA (Pitch Synchronous Overlap) method and harmonic modeling method of speech signal were used to change pitch in the 20 Hz interval. The voices were presented in a random order to prevent listener bias. The results were as follows ; Firstly, 46.6% judged male voice with 120 Hz as normal, and 19.2% judged 140 Hz as normal, and another 19.2% judged 160 Hz as normal. Secondly, 50.7% perceived female voice with 220 Hz as normal, and 32.9% and 30.1% responded to 200 Hz and 240 Hz, respectively. The problems and recommendations for a future investigation are discussed.

I. 서 론

최근들어 국내에서도 음성장애의 진단 및 치료에 대한 관심이 높아져 가고 있다. 기질적인 문제를 수술로 해결하고 나서도 음성 자체에는 커다란 효과를 보지 못하는 경우가 많기 때문에, 최근에는 이비인후과적 수술 후 언어치료사의 음성치료를 받아 그 치료 효과를 높이는 전문인 사이의 협력체계가 이루어져 가고 있다. 초기 성대결절의 경우에는 음성치료만으로도 병리조직이 감소 또는 제거되어 굳이 수술을 하지 않아도 되는 경우가 많다. 그러나 이러한 협력체계가 일부 임상현장에만 이루어지는 이유는 음성치료를 할 수 있는 전문인이 부족하다는 점과 한국인 음성에 대한 기초자료가 부족하여 객관적인 진단 및 치료의 방향설정이 어려운 실정이기 때문이다.

\* 대구대학교 재활과학대학 언어치료학과

\*\* 부산대학교 공과대학 전자공학과

\*\*\* 이화여자대학교 사범대학 특수교육학과

\*\*\*\* 영남대학교 의과대학 이비인후과

따라서 이 연구에서는 한국 남녀 정상 성인의 음도에 대한 인식을 살펴보고자 하였다. 즉 어느 정도의 음도를 정상적인 음도로 인식하고 있는지 연구하였다. 음도장애, 즉 개인이 자신의 성과 연령에 부적절한 음높이로 발성할 때, 예컨대 남성이 여성과 같이 높은 음도로 말하거나 여성이 지나치게 낮은 음도로 말하는 경우, 그 사회에서 어느 정도의 음도를 '정상' 혹은 '비정상'으로 인식하는 지는 음도장애의 판단에 큰 변인으로 작용한다. 따라서 한국인 정상 성인의 음도 인식에 대한 이 실험연구는 음도장애의 진단 및 치료에 필요한 기초자료를 제공한다는 점에서 그 의의를 찾을 수 있다.

Boone에 따르면, 정상 성인 남성의 음도는 128 Hz, 정상 성인 여성의 음도는 213 Hz로 이는 미국의 기준치이다(정옥란, 1996). 그러나 미국인과 한국인의 후두구조가 다소 차이가 있을 수 있다는 점을 감안해 볼 때, 한국인의 음도가 미국인의 음도에 비하여 약간 높을 것이라는 추측을 쉽게 할 수 있다. 이는 최홍식 등(1994)의 연구에서 뒷받침되었다. 이들의 연구에서는 정상 성인 남성의 음도가 143-155 Hz, 정상 성인 여성의 음도가 273-287 Hz인 것으로 밝혀진 바 있다.

그렇다면 음도 인식에 있어서도 미국인들의 정상 음도 기준치보다 높은 음도를 정상 음도로 인식할 것이라는 흥미로운 추론을 이끌어낼 수 있다. 그러므로 이 연구에서는 여러 가지 기본 주파수를 가진 음성을 연구대상에게 들려주고 음도 인식에 대한 체계적인 실험을 시도하였다.

본 연구의 구체적인 연구문제는 다음과 같다 :

- 첫째, 한국 정상 성인은 어느 정도의 기본주파수를 정상적인 남성의 음성으로 인식하고 있는가?
- 둘째, 한국 정상 성인은 어느 정도의 기본주파수를 정상적인 여성의 음성으로 인식하고 있는가?

## II. 연구방법

### 1. 연구대상

시각 청각 및 음성에 병력이 없으면서 대학에서 언어병리학을 전공하는 20-30세 사이의 한국 남녀 대학생 73명이 연구에 자발적으로 참여하였다.

### 2. 실험도구

#### 1) 음성샘플을 담은 오디오 테이프

한국 정상 성인 남성의 60 Hz 80 Hz 100 Hz 120 Hz 140 Hz 160 Hz 180 Hz 200 Hz 주파수에 해당하는 /아/ 장발성 음성샘플과 한국 정상 성인 여성의 140 Hz 160 Hz 180 Hz 200 Hz 220 Hz 240 Hz 260 Hz 280 Hz 주파수에 해당하는 /아/ 연장발성 음성샘플이 담겨 있는 normal bias의 HF-S 60 Sony 녹음용 테이프가 사용되었다. 20 Hz 인터벌로 남녀 음성의 음도 주기를 변경시킨 절차는 다음과 같다.

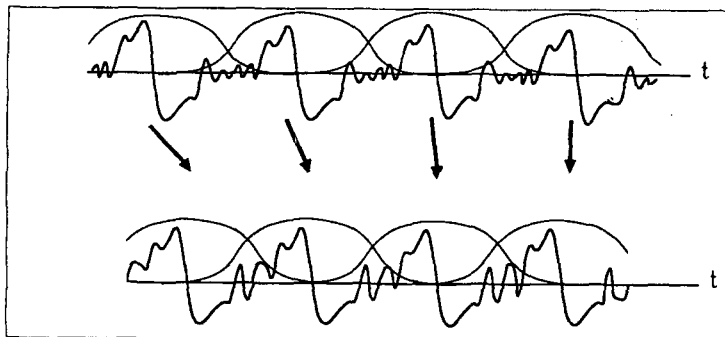
<음성신호의 음도주기 변경방법>

음성신호의 음도주기를 변경시키는 방법으로는 음성신호의 시간축 압축/신장 방법(Oppenheim & Schaffer, 1989), LPC(Linear Predictive Coding) 방법(Markel & Gray, 1976), PSOLA(Pitch Synchronous Overlap and Add) 방법(Hamon, Moulines, & Charpentier, 1989), 그리고 음성신호의 harmonic modeling 방법(Quayieri & McAulay, 1992). 등이 있다. 그 중 첫 번째와 두 번째 방법은 음도 변경된 음의 스펙트럼 왜곡이 상대적으로 심한 문제점이 있다. 따라서 이 연구에서는 비교적 양호한 음성을 얻을 수 있는 세 번째와 네 번째 방법 중에서 계산량이 적고 구현이 용이한 PSOLA 방법을 사용하여 음성신호의 음도 주기를 변경시켰다.

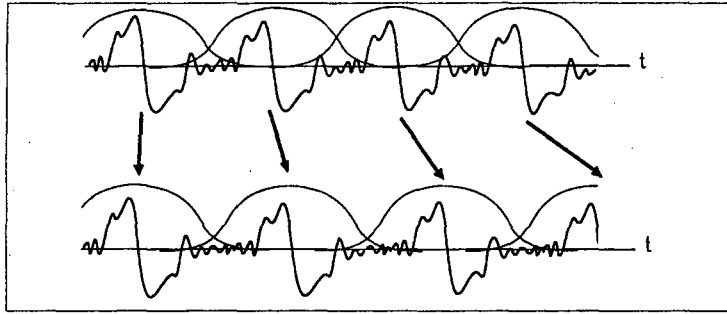
PSOLA 방법에서는 먼저 입력음성을 중첩(overlap)되는 단구간 음성 구간들로 분할하는데, 이 과정에서 분할된 음성구간들은 미리 음성분석을 통하여 구해진 음성신호의 음도위치(pitch marker)와 동기화(synchronization)시킨다. 그 다음으로 변경할 음도 주기에 의해 시간축 상에서의 음도위치를 재구성하고, 이에 따라 분할된 음성구간들을 스펙트럼 왜곡이 최소화되도록 중첩가산(overlap and add)시킴으로써 음도주기가 변경된 음성신호를 만들게 된다. 그림 1에 PSOLA 방법에 의해 음성신호의 음도주기를 변경하는 예가 제시되어 있다. 그림 1(a) 및 (b)에서 위의 그림은 원래의 음성신호, 그리고 아래 그림은 변화된 음성을 보여주고 있으며, 원래의 음성신호가 음도 위치에 따라 중첩되어 분할되는 모양이 나타나 있다. 그림에서 (a)는 음성신호의 음도를 높히는 경우를, 그리고 (b)는 음성신호의 음도를 낮추는 경우를 각각 표현하고 있다.

이 연구에서는 PSOLA 방법을 이용하여 평균 음도가 120 Hz인 남성 음성의 데이터로부터 60 Hz에서 200 Hz까지 20 Hz 단위로 증가하는 음도를 가지는 8 개의 음성을 생성시켰으며, 평균 음도가 220 Hz인 여성 음성의 데이터에 대해서도 140 Hz에서 280 Hz까지 역시 20 Hz 단위로 증가하는 음도를 가지는 8 개의 음성을 생성시켰다. 보다 객관적인 청취평가 결과를 유도하기 위하여 이들 음성 데이터들은 음높이에 대한 순서를 무작위(random)로 배열시켰다.

<그림 1> PSOLA 방법에 의해 음성신호의 음도주기를 변경시키는 예  
(a) 음도를 높히는 경우 (b) 음도를 낮추는 경우



(a)



(b)

## 2) 닥터위콤 녹음기(model 92060579)

남녀 음성 각각 8 개를 연구대상에게 소음이 통제된 실험공간에서 들려주는 데에 사용된 녹음기이다.

## 3) Visi-Pitch (model 6097)

120 Hz의 기본주파수를 가진 한국 정상 성인 남성의 /아/ 초기음성과 220 Hz의 기본주파수를 가진 한국 정상 성인 여성의 /아/ 초기음성을 샘플링하는 데에 사용된 기구이다.

## 4) 설문지

설문지는 남성(여성)의 음성으로는 '지나치게 낮다(-3),' '중간정도 낮다(-2),' '약간 낮다 (-1),' '보통이다(0),' '약간 높다(+1),' '중간정도 높다(+2),' '지나치게 높다(+3),' 등 7 개 등급으로 구성되어 연구대상의 음도에 대한 인식을 측정하였다. 실험에 사용된 설문지가 <부록 1>에 제시되어 있다.

## 3. 실험절차

소음이 통제된 공간에서 73명의 연구대상은 36명과 37명 각각 A·B 집단으로 나뉘어 실험에 참여하였다. A집단은 20 Hz 인터벌로 조절된 8개의 남성 음성에 대한 청취를 한 다음 이어서 8개의 여성 음성에 대한 청취를 하였다. B집단은 여성 음성을 먼저 청취하고 이어서 남성 음성을 청취하였다. 혼돈 변인을 제거하기 위하여 8개 음성의 청취순서도 무작위로 제시되었다. 연구대상들은 제시되는 개개의 음성을 듣고 설문지에 제시되어 있는 7개의 등급 중 어떤 범주에 속하는지 표시하도록 실험에 앞서 연구자가 구두로 설명한 후 실험이 진행되었다.

III. 연구결과

<표 1>과 <표 2>에 남녀 음성의 음도인식에 대한 등급 척도별 반응빈도수가 제시되어 있다.

<표 1> 8개 주파수를 가진 남성 음성의 음도인식에 대한 등급 척도별 반응빈도

| 남성의 음성<br>기본주파수 | 매우낮다      | 중간정도<br>낮다 | 약간낮다     | 보통이다     | 약간높다     | 중간정도<br>높다 | 매우높다     |
|-----------------|-----------|------------|----------|----------|----------|------------|----------|
| 60Hz            | 46(63.0)* | 22(30.1)   | 5(6.8)   | 0(0)     | 0(0)     | 0(0)       | 0(0)     |
| 80Hz            | 40(54.8)  | 28(38.4)   | 13(17.8) | 1(1.4)   | 0(0)     | 1(1.4)     | 0(0)     |
| 100Hz           | 9(12.3)   | 38(52.1)   | 20(27.4) | 5(6.8)   | 1(1.4)   | 0(0)       | 0(0)     |
| 120Hz           | 3(4.1)    | 10(13.7)   | 21(28.8) | 34(46.6) | 4(5.5)   | 1(1.4)     | 0(0)     |
| 140Hz           | 0(0)      | 6(8.2)     | 17(23.3) | 14(19.2) | 27(37.0) | 9(12.3)    | 0(0)     |
| 160Hz           | 1(1.4)    | 3(4.1)     | 3(4.1)   | 14(19.2) | 31(42.5) | 19(26.0)   | 2(2.7)   |
| 180Hz           | 0(0)      | 3(4.1)     | 2(2.7)   | 10(13.7) | 32(43.8) | 24(32.9)   | 2(2.7)   |
| 200Hz           | 1(1.4)    | 2(2.7)     | 2(2.7)   | 2(2.7)   | 17(23.3) | 34(46.6)   | 15(20.5) |

\* ( )의 수치는 반응빈도수에 대한 백분율이다.

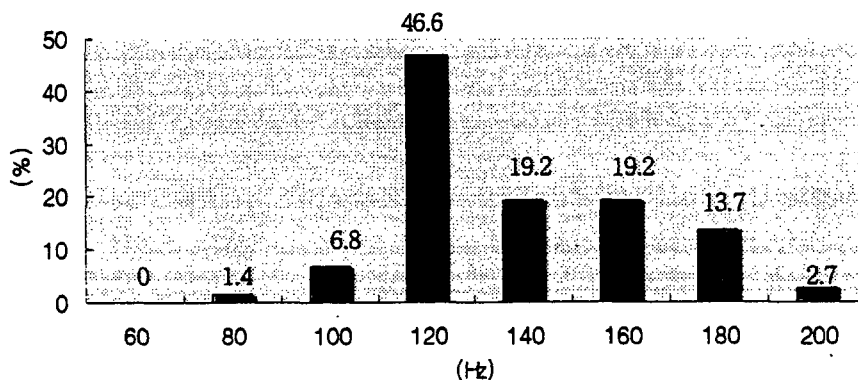
<표 2> 8개 주파수를 가진 여성 음성의 음도인식에 대한 등급 척도별 반응빈도수

| 여성의 음성<br>기본주파수 | 매우낮다      | 중간정도<br>낮다 | 약간낮다     | 보통이다     | 약간높다     | 중간정도<br>높다 | 매우높다     |
|-----------------|-----------|------------|----------|----------|----------|------------|----------|
| 140Hz           | 38(52.1)* | 22(30.1)   | 10(13.7) | 0(0)     | 2(2.7)   | 0(0)       | 1(1.4)   |
| 160Hz           | 11(15.1)  | 21(28.8)   | 14(19.2) | 11(15.1) | 15(20.5) | 1(1.4)     | 0(0)     |
| 180Hz           | 1(1.4)    | 33(45.2)   | 31(42.5) | 4(5.5)   | 3(4.1)   | 1(1.4)     | 0(0)     |
| 200Hz           | 1(1.4)    | 15(20.5)   | 26(35.6) | 24(32.9) | 4(5.5)   | 2(2.7)     | 1(1.4)   |
| 220Hz           | 0(0)      | 1(1.4)     | 24(32.9) | 37(50.7) | 11(15.1) | 0(0)       | 0(0)     |
| 240Hz           | 5(6.8)    | 0(0)       | 6(8.2)   | 22(30.1) | 24(32.9) | 16(21.9)   | 0(0)     |
| 260Hz           | 0(0)      | 1(1.4)     | 6(8.2)   | 17(23.3) | 32(43.8) | 15(20.5)   | 2(2.7)   |
| 280Hz           | 0(0)      | 0(0)       | 1(1.4)   | 3(4.1)   | 23(31.5) | 30(41.1)   | 16(21.9) |

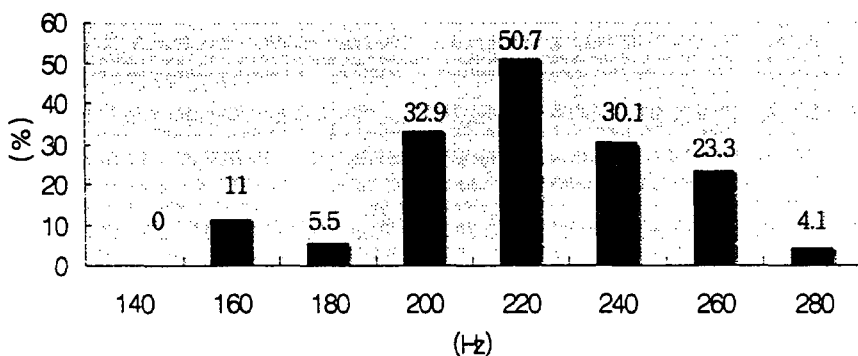
\* ( )의 수치는 반응빈도수에 대한 백분율이다.

<그림 2>와 <그림 3>에 연구대상들이 20Hz 인터벌로 조정된 다양한 음성을 듣고 정상적인 음도라고 인식하는 반응율이 제시되어 있다.

<그림 2> 다양한 기본주파수의 음성을 듣고 정상적인 남성의 음도라고 인식하는 반응율이 제시되어 있다.



<그림 3> 다양한 기본주파수의 음성을 듣고 정상적인 여성의 음도라고 인식하는 반응율이 제시되어 있다.



<표 1>과 <그림 2>에서 보여지듯이 한국 성인남성의 정상 음도로 가장 많은 인식을 받은 기본주파수는 120 Hz로 46.6%(73명중 43명)가 이에 응답하였다. 그 다음으로 높은 반응율을 보인 기본주파수는 140 Hz와 160 Hz로 각각 19.2%(73명중 14명씩)의 반응율을 보였다.

<표 2>와 <그림 3>에서 보여지듯이 한국 성인여성의 정상음도로 가장 많은 인식을 받은 기본주파수는 220 Hz로 50.7%(73명중 37명)가 이에 응답하였다. 그 다음으로 32.9%(73명중 24명)가 200 Hz를, 그리고 30.1%(73명중 22명)가 240 Hz를 각각 정상 음도로 판단하였다.

#### IV. 해석 및 토의

이 연구의 결과는 다음과 같이 요약할 수 있다:

첫째, 한국 정상 성인은 120 Hz(46.6%)를 정상적인 남성의 음도로 인식하였다. 그러나 19.2%가 140 Hz를, 또 19.2%가 160 Hz를 남성의 정상 음도로 인식하였다.

둘째, 한국 정상 성인은 220 Hz(50.7%)를 정상적인 여성의 음도로 인식하였다. 그러나 32.9%가 200 Hz를, 그리고 30.1%가 240 Hz를 여성의 정상 음도로 인식하였다.

한국 정상 성인 남성의 음도 평균치를 143 - 155 Hz로 볼 때 120 Hz의 음성에 가장 많은 반응을 보인 것은 약간의 괴리가 있는 것으로 보인다. 그러나 140 Hz와 160 Hz에도 각각 약 20%에 해당하는 연구대상들이 반응을 보인 점은 한국 남성의 음도가 미국 남성의 음도(128 Hz)보다 다소 높으며 따라서 음도의 인식에 있어서도 미국인의 음도보다 다소 높은 음도를 정상적인 것으로 인식한다는 점과 무관하지 않은 것 같다.

한국 정상 성인 여성의 음도 평균치를 273-287 Hz로 볼 때 220 Hz의 음성에 가장 많은 반응을 보인 것은 흥미로운 사실이다. 더구나 32.9%의 반응자가 220 Hz보다 20 Hz 낮은 정상적인 여성의 음도로 판단하였다. 그리고 30.1%가 240 Hz를 정상 음도로 판단하였다. 반면, 선행연구에서 밝혀진 여성의 음도 평균치 273 - 287 Hz 범주에 해당하는 280 Hz에는 단 4.1%의 응답자만이 이를 정상적인 음도로 판단하였으며, 오히려 280 Hz의 음도를 “여성의 음성으로는 약간 높다 (31.5%)” “여성의 음성으로는 중간정도 높다 (41.1%)” “여성의 음성으로는 매우 높다 (21.9%)”로 인식하였다. 이와 같이 산출과 인식간의 큰 차이가 있는 것은 아마도 다음과 같은 이유 중 하나에서 혹은 여러 이유가 복합적으로 작용한 것 같다.

첫째, 최홍식 등(1994)의 연구에서 밝혀진 한국 정상 성인 여성의 음도 평균치가 다소 높게 측정되었을 수 있다. 이들의 연구에 참여한 피험자 수가 제한적이라는 점을 감안해보면 273 - 287 Hz를 한국 정상 성인 여성의 음도 표준치로 해석하는 것은 고려의 여지가 있을 수 있다.

둘째, 산출과 인식은 반드시 일치하지 않을 수 있다. 그러나 남성음성의 인식에서는 비교적 일치율을 보이면서 여성음성의 인식에서는 상당한 차이를 빚는 것은 어떤 이유에서인지 또 다른 의문이 제기된다.

셋째, 이 연구에 사용된 음성 샘플이 적절하지 않은 것일 수 있다. 실험에 사용된 음성은 /아/의 연장 발성으로 이는 반드시 개인의 대표적인 음성을 말한다고 할 수는 없다. /아/보다는 가령 /아 애 이 오 우/를 연장 발성하거나 더 대표적인 음성이라면 발화시 음성을 말한다. 그러나 각각의 자·모음이 지니는 주파수가 다르기 때문에 음도주기를 20 Hz 인터벌로 변경시키는 데에 많은 문제점이 따른다.

이러한 문제 때문에 편의상 /아/를 사용하였으나 반응자들이 자신의 음도에 대한 인식을 정확히 판단내리기에는 최상의 샘플이 아니었을 수 있다. 따라서 후속 연구에서는 보다 객관적인 인식도를 측정하기 위하여 /아/대신 /아 애 이 오 우/를 연결한 연장발성이나 짧은 발화를 음성 샘플로 사용하여 본 연구의 결과를 재검증해보는 것이 바람직하리라 생각된다. 또한 주파수 변

경을 20 Hz 인터벌로 하기보다 보다 작은 인터벌로 조정하여 실험하는 것도 고려해 봄직하다.

### 참 고 문 헌

- [1] 정옥란. 1996. *음성과 음성치료*. 도서출판 원미사.
- [2] 최홍식 · 장미숙 · 이정준. 1994. "정상인과 후두폴립환자에서의 음성학적 측정." *대한음성언어의학회지* 5(1), 38-43.
- [3] Hamon, C., Moulines, E., & Charpentier, F. 1989. "A Diphone Synthesis System Based on Time-domain Prosodic Modification of Speech." in Proceedings of IEEE International Conference on Acoustics. *Speech and Signal Processing*, 238-241.
- [4] Markel J. D. & Gray, A. H. 1976. *Linear Prediction of Speech*. Springer-Verlag.
- [5] Oppenheim, A. V. & Schafer, R. W. 1989. *Discrete-Time Signal Processing*. Prentice-Hall.
- [6] Quayieri, T. F. & McAulay, R. J. 1992. "Shape Invariant Time-scale and Pitch Modification of Speech." *IEEE Transactions on Signal Processing* 40(3), 497-510.

접수일자 : '97. 1. 27.

게재결정 : '97. 2. 26.

#### ▲ 정옥란

대구광역시 남구 대명동 2288  
 대구대학교 재활과학대학 언어치료학과 (우편번호 : 705-033)  
 Tel : (053) 650-8274 FAX : (053) 650-8259  
 e-mail : oj@biho.taegu.ac.kr

#### ▲ 김형순

부산광역시 금정구 장전동 30  
 부산대학교 전자공학과 (우편번호 : 609-735)  
 Tel : (051) 510-2452(O) FAX : (051) 515-5190  
 (051) 516-0353(H)  
 e-mail : kimhs@hyowon.pusan.ac.kr

#### ▲ 김영태

서울시 서대문구 대현동 11-1  
 이화여대 특수교육학과 (우편번호 : 120-750)  
 Tel : (02) 360-2410(O) 512-6423(H)

#### ▲ 서장수

대구시 남구 대명동 317-1  
 영남대학병원 ENT (우편번호 : 705-030)  
 Tel : (053) 620-3782(O) FAX : (053) 628-7884



<부록 1> 음도인식을 평가하기 위한 설문지

| 성명:        | 아<br>주<br>낮<br>다 | 중(中)<br>간<br>정<br>도<br>낮<br>다 | 약<br>간<br>낮<br>다 | 표<br>준<br>이<br>다 | 약<br>간<br>높<br>다 | 중(中)<br>간<br>정<br>도<br>높<br>다 | 아<br>주<br>높<br>다 |
|------------|------------------|-------------------------------|------------------|------------------|------------------|-------------------------------|------------------|
| 1. 남자음성으로는 | -3               | -2                            | -1               | 0                | +1               | +2                            | +3               |
| 2. "       | -3               | -2                            | -1               | 0                | +1               | +2                            | +3               |
| 3. "       | -3               | -2                            | -1               | 0                | +1               | +2                            | +3               |
| 4. "       | -3               | -2                            | -1               | 0                | +1               | +2                            | +3               |
| 5. "       | -3               | -2                            | -1               | 0                | +1               | +2                            | +3               |
| 6. "       | -3               | -2                            | -1               | 0                | +1               | +2                            | +3               |
| 7. "       | -3               | -2                            | -1               | 0                | +1               | +2                            | +3               |
| 8. "       | -3               | -2                            | -1               | 0                | +1               | +2                            | +3               |
| 1. 여자음성으로는 | -3               | -2                            | -1               | 0                | +1               | +2                            | +3               |
| 2. "       | -3               | -2                            | -1               | 0                | +1               | +2                            | +3               |
| 3. "       | -3               | -2                            | -1               | 0                | +1               | +2                            | +3               |
| 4. "       | -3               | -2                            | -1               | 0                | +1               | +2                            | +3               |
| 5. "       | -3               | -2                            | -1               | 0                | +1               | +2                            | +3               |
| 6. "       | -3               | -2                            | -1               | 0                | +1               | +2                            | +3               |
| 7. "       | -3               | -2                            | -1               | 0                | +1               | +2                            | +3               |
| 8. "       | -3               | -2                            | -1               | 0                | +1               | +2                            | +3               |