

한국어 발성 유형의 음향음성학적 연구

An Acoustic Study of Korean Phonation Types

박 한 상*
(Hansang Park*)

*홍익대학교 영어교육과

(접수일자: 2005년 7월 4일; 수정일자: 2005년 7월 18일; 채택일자: 2005년 7월 25일)

발성유형지수 k (PTI k)는 F0와 모음 음가의 영향을 배제한, 단순하고 단일화된 스펙트럼 기울기의 척도다. 본 논문은 한국인 20대 남성화자 10명으로부터 얻은 모음 /이, 에, 아, 오, 우/에 나타난 PTI k 를 연구한다. 환경, 발성유형, 모음, 그리고 화자를 요인으로 하여 각 집단간에 PTI k 의 값에 유의미한 차이가 있는지 알아본다. 연구 결과 환경, 발성유형, 모음, 그리고 화자 각각에 집단간에 유의미한 차이가 있었다.

핵심용어: 한국어, 발성유형, 스펙트럼의 기울기, PTI k , 환경, 모음, 화자

투고분야: 말소리 생성 및 인지분야 (12.1)

Phonation type index k (PTI k) presents a single and simplified measure of the spectral tilt, which is free from the effects of fundamental frequency and vowel quality. This study investigates PTI k with vowels /i, e, a, o, u/ obtained from 10 Korean male subjects. Specifically, this study tests the significance of differences in PTI k across positions, phonation types, vowels, and speakers, respectively. The results showed that there was a significant difference in PTI k across positions, phonation types, vowels, and speakers.

Keywords: Korean, Phonation Type, Spectral Tilt, PTI k , Position, Vowel, Speaker

ASK subject classification: Speech Production and Perception (12.1)

1. 서론

발성유형의 한 특성인 스펙트럼의 기울기 (spectral tilt)는 $H1^*-A1[1]$, $H1-F2[2][3]$, $H1^*-A3^*[1]$ 등으로 측정되어 왔다. 그러나 기존의 측정단위들은 기본주파수 (F0)와 모음의 영향을 전부 혹은 일부 고려하지 않았다. [4]에서 기존의 측정단위에서 고려되지 않았던 기본주파수와 모음의 영향을 모두 배제한 발성유형지수 k (phonation type index k , 줄여서 PTI k)를 음원여과기 이론 (Source Filter Theory)을 확장하여 새로이 제시한 이후 후속연구에서 PTI k 의 타당성을 검증하였다[5-9]. [4]에서는 3명의 남성 화자를 대상으로 문두의 CV와 어중의 VCV 연쇄체에서 한국어의 치조파연음, 마찰음, 그리고 경구개 파찰음 다음에 따라오는 모음 /a/에서 실현

되는 PTI k 만을 대상으로 연구하였다. 본 논문은 10명의 한국인 남성이 발화한 모음 /i, e, a, o, u/를 대상으로 환경 (어두 및 모음간), 발성유형 (격음, 연음, 경음), 모음 (/이, 에, 아, 오, 우/), 그리고 화자별로 각 집단간에 PTI k 의 값에 유의미한 차이가 있는지 알아본다.

PTI k 는 음성학에서 스펙트럼의 기울기를 측정하는데 쓰일 뿐만 아니라 화자의 특성을 추출하는 음성공학 분야와 음성의학 등에 두루 쓰일 수 있을 것으로 기대한다. 특히 PTI k 는 화자에 따라 그 값의 통계적 지표들이 뚜렷이 달라지는 점을 고려할 때 음성을 이용하여 범인을 확인하는 범의학이나 화자인식과 같은 분야에서 화자의 특성을 보여주는 지표로 쓰일 수 있을 것으로 기대한다.

II. 이론적 배경

음원여과기이론에 따르면 음압 (sound pressure)의 주파수반응 (frequency response)은 다음의 식 (1)과 같이 음원특성과 방출특성, 그리고 여과기특성의 곱으로 나타낼 수 있다[10].

$$|P(f)| = |U(f)||R(f)||H(f)| \quad (1)$$

위의 식 (1)에서 좌변은 음압의 주파수반응을 우변은 왼쪽에서부터 순서대로 음원특성, 방출특성, 그리고 성도전달함수를 각각 나타낸다. 그 중 음원특성과 방출특성의 곱은 다음의 식 (2)와 같이 나타낼 수 있다[10].

$$|U(f)||R(f)| = P_k \frac{(f/100)}{1+(f/100)^2} \quad (2)$$

위의 식 (2)에서 P_k 는 특정의 음압수준을 결정하는 상수를 나타낸다. 위의 식 (2)의 결과 나타나는 스펙트럼은 대략 6 dB/octave의 비율로 감소한다. 방출특성은 대략 주파수에 비례하고 음원특성은 주파수의 곱의 역수에 비례한다[10].

위의 식 (2)에서 음원특성과 방출특성의 곱은 주파수의 함수로 나타나 있는데 모음과 같은 소리에서 배음의 주파수는 기본주파수의 정수배임을 감안하면 위의 식 (2)는 다음의 식 (3)과 같이 기본주파수의 함수로 변환할 수 있다.

$$|U(nf_0)||R(nf_0)| = P_k \frac{(nf_0/100)}{1+(nf_0/100)^2} \quad (3)$$

나아가 위의 식 (3)은 아래와 같이 음원특성을 나타내는 식 (4)와 방출특성을 나타내는 식 (5)로 나눌 수 있다.

$$|U(nf_0)| = P_k \frac{(f_0/100)}{1+(nf_0/100)^2} \quad (4)$$

$$|R(nf_0)| = n \quad (5)$$

음원특성을 나타내는 식 (4)는 아래와 같이 기본주파수구성요소를 나타내는 식 (6)과 음원비율을 나타내는 식 (7)로 나눌 수 있다.

$$|U(f_0)| = P_k \frac{(f_0/100)}{1+(f_0/100)^2} \quad (6)$$

$$Source \quad Ratio = \frac{1+(f_0/100)^2}{1+(nf_0/100)^2} \quad (7)$$

배음 (harmonic)의 진폭은 기본음에 바탕을 두고 있어서 특정배음의 음원특성의 기여분은 음원비율에 의해 결정된다. 위의 식 (7)에서 알 수 있듯이 음원비율은 상수가 아니라 기본주파수의 함수이다. 이는 기본음과 배음의 음원비율은 기본주파수에 따라 변한다는 것을 의미한다.

비록 음원비율이 기본주파수에 따라 변한다고 하더라도 발생유형에 따라 달라지는 스펙트럼의 기울기의 변화는 여전히 설명하지 못한다. 기본주파수가 동일할 경우 다른 발생유형의 음원스펙트럼 기울기가 -12 dB/octave 보다 클 수도 있고 작을 수도 있다는 것을 의미한다. 스펙트럼 기울기의 변화를 설명하기 위해서는 스펙트럼의 기울기를 통제하는 요소의 설정이 필요하여 다음의 식 (8)로 표현되는 요소를 설정한다.

$$\left(\frac{1}{n}\right)^k \quad (8)$$

위의 식 (8)에서 n 은 배음번호이며 지수 k 는 "발성유형지수 (phonation type index k : PTI k)"라고 명명한다.

PTI k 는 스펙트럼 기울기의 변화를 보여주는 음성매개변수이다. "보통 목소리 (modal voice)를 "가장 빈번히 나타나는 발생유형"이라고 정의한다면, PTI k 의 분포에서 보통 목소리는 최민값 (mode)을 보일 것이고 짜내는 목소리 (creaky voice)와 숨쉬인 목소리 (breathy voice)는 최소값과 최대값을 보일 것이다."

지금까지의 논의를 바탕으로 음원특성을 식 (9)로 다시 쓸 수 있다.

$$|U(nf_0)| = P_k \frac{(f_0/100)}{1+(f_0/100)^2} \frac{1+(f_0/100)^2}{1+(nf_0/100)^2} \left(\frac{1}{n}\right)^k \quad (9)$$

결국 음원특성은 기본음, 음원비율, 그리고 스펙트럼의 기울기 통제 요소로 구성되어 있다. 따라서 배음번호 n 인 배음의 진폭은 방출특성 $|R(nf_0)|$ 와 여과기특성인 성도전달함수 $|H(nf_0)|$ 를 포함하여 다음의 식 (10)과 같이 나타낼 수 있다.

$$H_n = 20 \log_{10} P_k \frac{(f_0/100)}{1 + (f_0/100)^2} \frac{1 + (f_0/100)^2}{1 + (nf_0/100)^2} \left(\frac{1}{n}\right)^k n |H(nf_0)| \quad (10)$$

위의 식 (10)에서 알 수 있듯이 어떤 배음의 진폭은 음원특성, 방출특성, 그리고 여과기특성의 기여분으로 이루어져 있다.

PTI k 는 선형회귀분석법을 이용하여 구할 수 있다. 위의 식 (10)에서 성도전달함수 및 방출특성 그리고 음원 비율은 측정된 여러 가지 매개변수들의 값을 대입하면 상수가 된다. 그 결과 기본주파수 구성요소와 스펙트럼 기울기 통제요소를 미지수로 가지는 다음의 식 (11)과 같은 선형 방정식을 얻을 수 있다.

$$H_n - FILTER - RADIATION - SR = FC + k 20 \log_{10} \left(\frac{1}{n}\right) \quad (11)$$

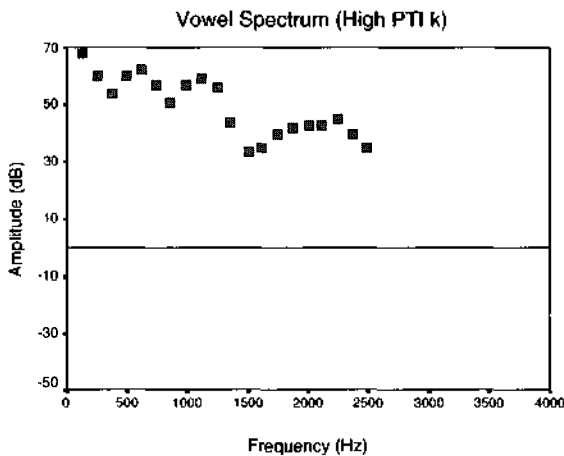


그림 1. 높은 PTI k 를 가진 음성샘플의 스펙트럼. 각각의 점은 배음의 진폭을 나타낸다

Figure 1. A spectrum of a speech sample with a high PTI k . Each dot represents the amplitude of a harmonic.

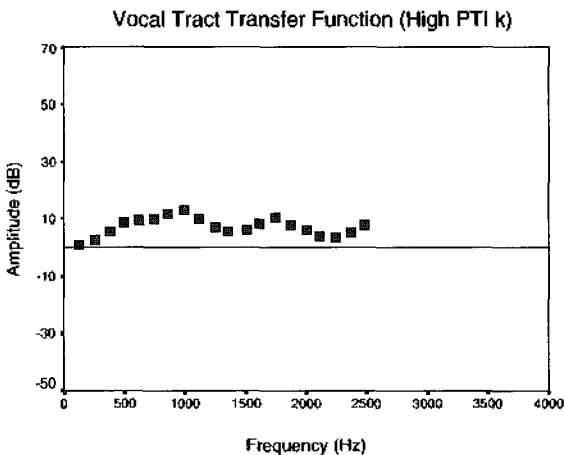


그림 3. 높은 PTI k 를 가진 음성샘플의 성도전달함수 기여분
Figure 3. Contributions of the vocal tract transfer function of the spectrum of the speech sample with a high PTI k .

위의 식 (11)에서 암시하듯이 성도전달함수 (FILTER), 방출특성 (RADIATION), 그리고 음원비율 (SR)을 배음의 진폭으로부터 뺀 값으로 구성된 스펙트럼으로부터 기울기의 추이를 보여주는 최적선 (best-fit line)을 구할 수 있다. 바로 이 최적선의 기울기가 PTI k 이다.

PTI k 산출의 구체적인 예를 다음의 일련의 그림으로 보여줄 수 있다. 높고 낮은 PTI k 값을 가진 두 음성샘플의 스펙트럼이 그림 1과 그림 2에 나타나 있다. 먼저 성도전달함수의 기여분을 계산하면 그림 3 및 그림 4와 같이 나타낼 수 있다. 다음으로 방출특성의 기여분을 계산하면 그림 5 및 그림 6과 같이 나타낼 수 있다. 마지막으로 음원비율을 계산하면 그림 7 및 그림 8과 같이 나타낼 수 있다.

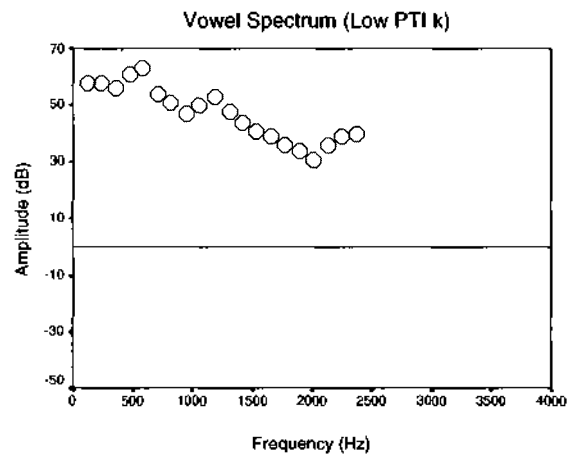


그림 2. 낮은 PTI k 를 가진 음성샘플의 스펙트럼. 각각의 점은 배음의 진폭을 나타낸다

Figure 2. A spectrum of a speech sample with a low PTI k . Each dot represents the amplitude of a harmonic.

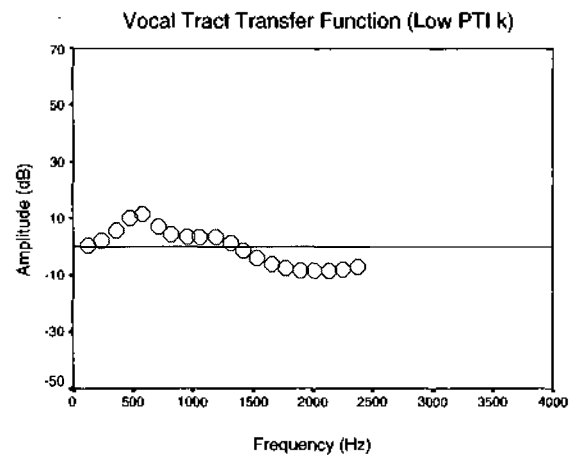


그림 4. 낮은 PTI k 를 가진 음성샘플의 성도전달함수 기여분
Figure 4. Contributions of the vocal tract transfer function of the spectrum of the speech sample with a low PTI k .

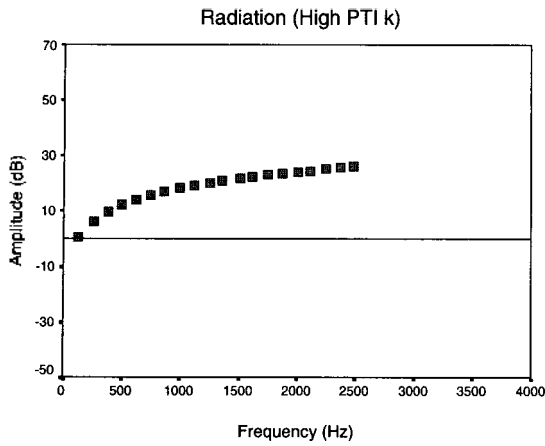


그림 5. 높은 PTI k 를 가진 음성샘플의 방출 특성 기여분
Figure 5. Contributions of the radiation characteristics of the spectrum of the speech sample with a high PTI k .

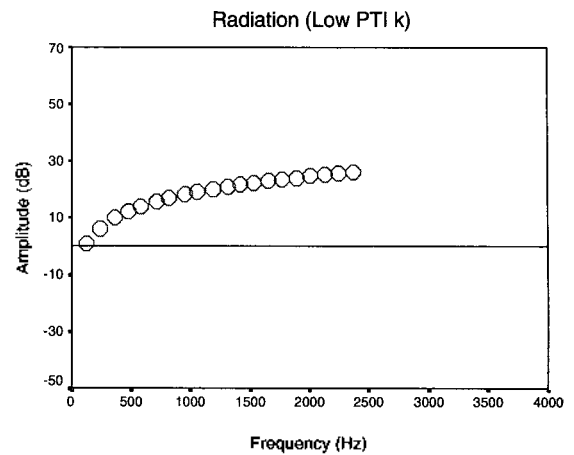


그림 6. 낮은 PTI k 를 가진 음성샘플의 방출 특성 기여분
Figure 6. Contributions of the radiation characteristics of the spectrum of the speech sample with a low PTI k .

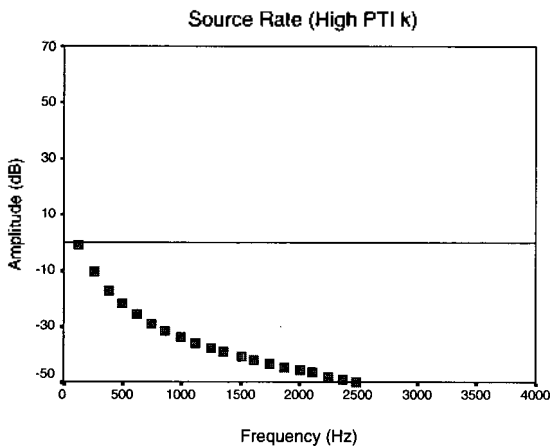


그림 7. 높은 PTI k 를 가진 음성샘플의 음원 비율 기여분
Figure 7. Contributions of the source rate of the spectrum of the speech sample with a high PTI k .

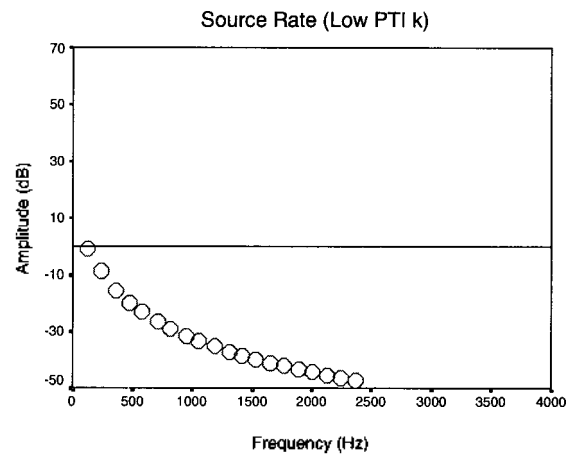


그림 8. 낮은 PTI k 를 가진 음성샘플의 음원 비율 기여분
Figure 8. Contributions of the source rate of the spectrum of the speech sample with a low PTI k .

그림 1의 모음 스펙트럼에서 그림 3, 5, 7의 성도전달 함수, 방출특성, 그리고 음원비율의 기여분을 빼면 그림

9를 얻을 수 있고, 그림 2의 모음 스펙트럼에서 그림 4, 6, 8의 성도전달함수, 방출특성, 그리고 음원비율의 기

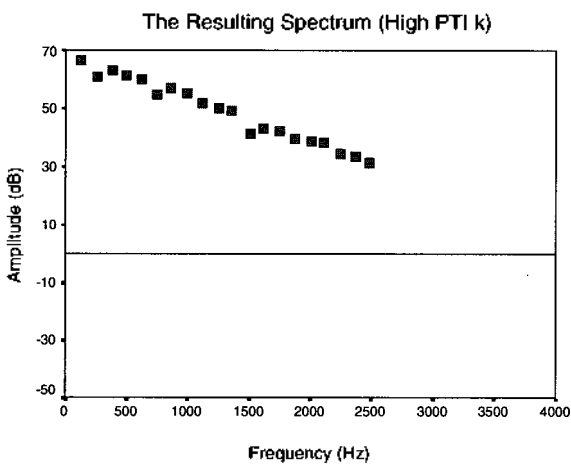


그림 9. 성도전달함수, 방출특성, 그리고 음원비율의 기여분을 제외한 높은 PTI k 를 가진 음성샘플의 스펙트럼
Figure 9. Spectrum of the speech sample with a high PTI k exclusive of the contributions of vocal tract transfer function, radiation characteristics, and source rate.

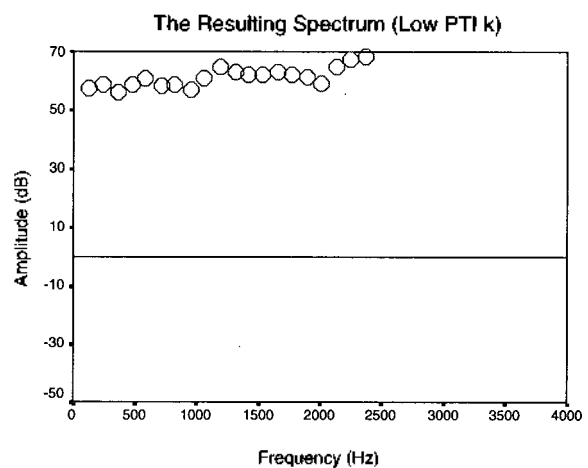


그림 10. 낮은 PTI k 를 가진 음성샘플의 음원 비율 기여분
Figure 10. Spectrum of the speech sample with a low PTI k exclusive of the contributions of vocal tract transfer function, radiation characteristics, and source rate.

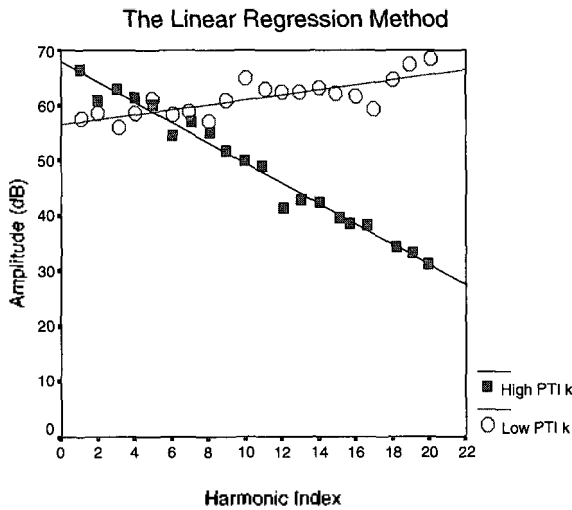


그림 11. 상도전달함수, 방출특성, 그리고 음원비율의 기여분을 제외한 스펙트럼의 선형 추이와 최적선
 Figure 11. Linear trends and best-fit lines derived from the spectra of the speech samples exclusive of the contributions of vocal tract transfer function, radiation characteristics, and source rate.

여분을 빼면 그림 10과 같은 스펙트럼을 얻을 수 있다. 그림 9와 그림 10에서 보는 바와 같이 두 음성샘플의 상도전달함수, 방출특성, 그리고 음원비율의 기여분을 뺀 결과 나타나는 스펙트럼은 분명한 선형추이 (linear trend)를 보여주고 있다. 각각의 샘플에 대하여 최적선 (best-fit line)을 그리면 그림 11과 같다.

III. 연구 방법

본 연구에 사용된 음성자료는 한국어 치조파열음 [t^h, t^{*}, t] 및 마찰음 [s^h, s^{*}] 그리고 경구개 파찰음 [t^h, t^{*}, t^s] 다음에 따라오는 모음 "이, 에, 아, 오, 우"를 포함한 [CVda]형의 2음절 단어이다. 이와 같이 자음을 선택한 이유는 본 연구의 관심사 중 하나가 장애음 조음방법에 따른 변이를 살펴보는 것이어서 파열음, 파찰음, 마찰음을 포함시켜야 하기 때문이다. 양순음과 연구개음은 파찰음과 마찰음이 없고 치조 및 경구개음 (coronal)에만 세 가지 조음방법이 모두 나타난다. 조음위치의 영향을 피하기 위하여 양순음과 연구개음은 생략하고 치조 및 경구개음만 선택하였다. 선택된 단어들 중 일부는 한국어의 동사 원형이고 나머지는 무의미 단어이다. 각 단어는 " __가 아니라 __다"라는 틀문장 (frame sentence)에 넣어서 발음하였다. 틀문장의 두 번째 빈 자리에는 모음간 환경에서 발성유형의 특성을 살펴보기 위하여 단

어의 두 음절을 서로 바꾸어 넣었다.

위에서 제시한 두 자리에 나타나는 모음 5개와 자음 8개의 조합으로 구성된 40개의 문장을 10명의 화자가 10번씩 반복하여 모두 4,000개의 문장을 녹음하였다. 화자는 모두 서울 출신의 20대 남자들이었고, 녹음 당시에 발음 이상이 없었으며 이전에 언어병력이 없었다. 녹음은 고려대학교 민족문화연구원 음성언어정보연구실에 있는 녹음실에서 이루어 졌다. 녹음된 음성자료의 표본추출율 (sampling rate)은 22,050 Hz로 하였고 양자화 (quantization)는 16 비트 (bit)로 하였다.

PTI *k*를 계산하기 위해서는 많은 음향매개변수의 값을 측정해야 한다. 먼저 상도전달함수의 기여분을 계산하기 위해서는 가장 낮은 네 포먼트의 주파수와 대역폭 (bandwidth)의 값이 필요하고, 최적선을 구하기 위해서는 배음 20개의 주파수와 대역폭의 값이 필요하며, 음원비율의 계산을 위해서는 기본주파수의 값이 필요하다. Microsoft Visual C++ 6.0을 이용하여 PTI *k*를 얻는 프로그램을 짰다. 음성매개변수 중 기본주파수는 자기상관 피치검출법을 이용하여 구하였고 포먼트의 주파수와 대역폭은 MFCC를 이용하여 구하였다. 포먼트 주파수와 대역폭 값은 LPC 스펙트럼에서 구했고, 배음 20개의 주파수와 진폭은 FFT 스펙트럼에서 구했다. LPC 스펙트럼 (LPC 계수: 25)과 FFT 스펙트럼 (FFT point: 1024) 및 기본주파수는 20 ms Hamming 창을 이용하여 구하였다. 이 값들은 모음의 한 특정 부분에서 구한 것이 아니라 모음 전체에 걸쳐 처음부터 끝까지 10 ms씩 중복하여 이동하면서 구하였다. 그래서 100 ms의 길이를 가진 모음에서는 모두 9개의 FFT와 LPC 스펙트럼을 얻었다. 창의 크기가 20 ms 이하일 경우 음성매개변수들의 값을 구하지 않았다.

이와 같은 방법으로 얻은 기본주파수와 포먼트 주파수 및 대역폭 값에 오류가 발견된 경우 PTI *k*의 계산에서 배제 하였다. 먼저 기본주파수와 관련해서는 기본주파수의 값이 90 Hz 이하일 경우와 200 Hz 이상일 경우 피치 반감 (pitch halving)과 피치배증 (pitch doubling)이 있는 음성샘플은 PTI *k*의 계산에서 배제하였다. 측정된 기본주파수 값과 FFT를 통해 얻은 기본주파수구성요소의 주파수 값의 차이가 20 Hz 이상인 경우 PTI *k*의 계산에서 배제하였다.

다음으로 배음과 관련해서는 측정된 기본주파수 값의 10배와 10번째 배음의 주파수 값의 차이와 첫 번째 배음의 주파수 값의 10배와 10번째 배음의 주파수 값의 차이

의 합이 500 Hz 이상일 때는 PTI k 의 계산에서 배제하였다. 그리고 20번째 배음의 주파수가 4,000 Hz 이상일 경우 PTI k 의 계산에서 배제하였다.

다음으로 포먼트의 주파수와 관련해서는 F1의 값이 1500 Hz 이상일 경우 또는 F2의 값이 3000 Hz 이상일 경우 PTI k 의 계산에서 배제하였다.

본 연구에서는 모음 "오"와 "우"를 다른 모음과 동일한 LPC 계수를 적용하여 구하였다. 이는 사실과 다른 F1과 F2의 값을 얻을 수 있음을 시사한다. 본 연구에서는 다른 모음과 동일한 LPC 계수를 이용했을 경우 어떤 차이점이 있는지 살펴보는 것도 일단 하나의 연구 내용으로 설정하였다. 하지만 이를 최소화하기 위하여 모음 "오"와 "우"의 경우 F2가 1900 Hz가 넘으면 PTI k 의 계산에서 배제하였다. F1의 값이 900 Hz인 이상인 경우에도 PTI k 의 계산에서 배제하였다. 우리가 일반적으로 알고 있는 모음 "오"와 "우"의 포먼트 값보다 훨씬 높은 상한선을 부여한 이유는 포먼트 전이구간에서는 일반적으로 알고 있는 포먼트 값보다 높은 값이 나올 가능성이 있기 때문이었다. 그러나 사실과 다른 포먼트 값이 기록되었을 가능성은 여전히 배제할 수 없다. 따라서 연구결과에서도 모음 "오"와 "우"는 다른 모음과 다른 행동을 보일 것으로 예측된다. 그러나 이 자체도 PTI k 와 관련된 화자 특성을 보여줄 수 있는 한 특징으로 보고 본 연구에서는 그대로 포함시켜 제시할 것이다.

마지막으로 포먼트의 대역폭과 관련해서는 각 포먼트의 대역폭이 1000 Hz 이상일 경우 PTI k 의 계산에서 배제하였다. 전체 음성샘플의 수 44,645중에서 오류가 있는 음성샘플의 수 16,851 (37.7 %)를 배제한 결과 유효한 음성샘플의 수는 27,794 (62.3 %)였다.

IV. 연구 결과

[4]에서 화자에 따라 PTI k 의 평균 (mean)과 표준편

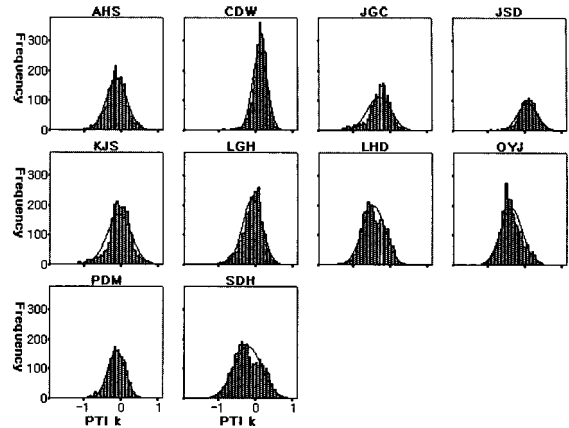


그림 12. 각 화자의 PTI k 의 분포를 보여주는 히스토그램
Figure 12. Histograms of individual speakers showing the distribution of PTI k .

차 (standard deviation), 그리고 첨도 (kurtosis)와 영역 (range)에서 특징적인 차이가 있음을 지적하였다. 뿐만 아니라 환경과 발생유형에 따라 PTI k 의 값이 특징적인 차이를 보여 주었다. 여기에서는 환경 (#CV, VCV), 발생유형 (격음, 연음, 경음), 모음 (/이, 에, 아, 오, 우 /), 화자 등의 요인에 따른 PTI k 의 기술통계량을 제시하고 각 요인별로 나타나는 특징적인 차이를 기술할 것이다. 먼저 화자별 차이를 보기 위해 PTI k 의 분포를 살펴보았다. 화자별 PTI k 의 분포를 보여주는 히스토그램이 그림 12에 제시되어 있다.

그림 12에서 볼 수 있듯이 각 화자의 히스토그램은 대체로 정규분포에 가까우며 화자별로 차이를 보여주고 있다. 표 1에 화자별 PTI k 의 기술통계가 제시되어 있다.

표 1에서 평균, 표준편차, 첨도, 영역 등을 포함한 기술통계 값에 대해 화자별로 차이를 보여주고 있다. 이러한 기술통계에 대한 화자별 차이를 보기 위하여 평균과 표준편차를 대상으로 한 산점도가 그림 13에, 첨도와 영역을 대상으로 한 산점도가 그림 14에 각각 제시되어 있다.

그림 13과 그림 14에서 볼 수 있듯이 평균과 표준편차 그리고 첨도와 영역을 보여주는 각각의 산점도에서 화자별로 차이를 보여주고 있다.

다음으로 PTI k 가 환경, 발생유형, 모음에 따라 어떤

표 1. 화자별 PTI k 의 기술통계

Table 1. Descriptive statistics of PTI k across speakers.

	AHS	CDW	JGC	JSD	KJS	LGH	LHD	OYJ	PDM	SDH
Mean	-0.13	0.109	-0.302	0.077	-0.054	-0.094	-0.455	-0.368	-0.118	-0.207
Variance	0.078	0.04	0.101	0.072	0.107	0.071	0.094	0.086	0.056	0.142
St. Dev.	0.279	0.2	0.318	0.269	0.328	0.267	0.307	0.293	0.236	0.376
Minimum	-1.385	-1.149	-1.616	-1.015	-1.617	-1.119	-1.924	-1.341	-1.2	-1.488
Maximum	0.928	0.933	0.676	1.102	1.132	0.795	0.326	0.834	0.505	0.991
Range	2.313	2.082	2.293	2.118	2.749	1.914	2.25	2.175	1.706	2.479
Skewness	-0.131	-1.207	-0.889	-0.356	-0.574	-0.53	-0.156	0.389	-0.511	0.126
Kurtosis	0.664	4.182	1.316	1.416	1.757	0.59	0.065	0.305	0.685	-0.386

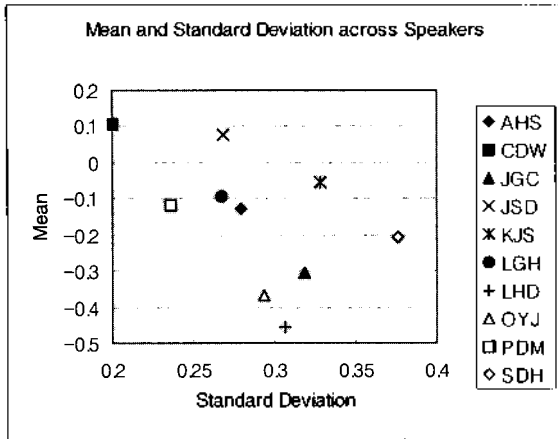


그림 13. 화자간의 차이를 보여주는 평균과 표준편차의 산점도
Figure 13. Scatterplot of mean and standard deviation illustrating differences across speakers.

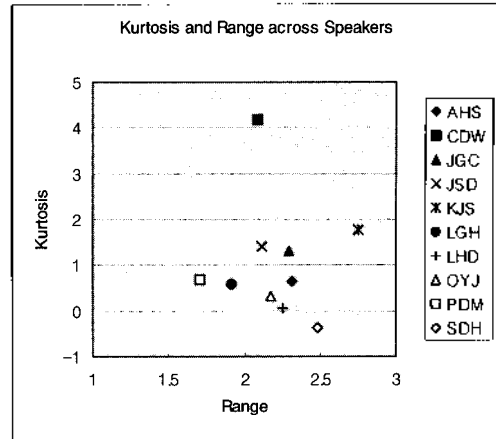


그림 14. 화자간의 차이를 보여주는 영역위와 침도의 산점도
Figure 14. Scatterplot of range and kurtosis illustrating differences across speakers.

차이를 보이는지 살펴본다. PTI k의 환경별, 발성유형별, 모음별 평균과 표준편차가 표 2에 제시되어 있다.

표 2에서 보여 주는 집단간의 평균이 유의미한 차이인지 알아보기 위하여 PTI k를 종속변수로, 환경, 발성유형, 모음, 화자를 각각 요인으로 하는 일변량4원배치분산분석을 실시하였다. 유의도는 0.05로 하였다. 분산분석의 결과가 표 3에 제시되어 있다.

표 3에 나타나 있듯이 분산분석의 결과 모든 요인에 대하여 유의미한 주효과 (main effect)가 있었다. Tukey's HSD를 이용하여 발성유형, 모음, 화자에 대한 사후 분석을 실시하였다. 그 결과 발성유형에 대해서는 연음과 경음이 격음보다 유의미하게 큰 PTI k 값을 가지고 있었으며 연음과 경음 사이에는 유의미한 차이가 없었다. 모음에 대해서는 /u/ > /o/ > /i/ > /a/, /e/의 순으로 유

표 3. 분산분석의 결과
Table 3. Results of analysis of variance.

계제-전 효과 검정					
종속변수 PTI k					
수소	제 배 유형 세입항	자유도	평균 제곱	F	유의확률
모형	2804.099 ^a	300	9.347	204.588	.000
position	.218	1	.218	4.769	.03 [*]
phonotype	12.790	2	6.395	139.971	.000
vowel	408.036	4	102.008	2232.751	.000
speaker	427.811	9	47.535	1040.441	.000
position * phonotype	2.667	2	1.334	29.189	.000
position * vowel	19.298	4	4.821	105.598	.000
phonotype * vowel	31.201	8	3.900	85.366	.000
position * phonotype * vowel	2.458	8	.307	6.726	.000
position * speaker	16.381	9	1.820	39.838	.000
phonotype * speaker	32.323	18	1.796	39.306	.000
position * phonotype * speaker	20.062	18	1.115	24.396	.000
vowel * speaker	153.161	36	4.254	93.123	.000
position * vowel * speaker	16.224	36	.451	9.864	.000
phonotype * vowel * speaker	23.735	72	.330	7.215	.000
position * phonotype * vowel * speaker	16.815	72	.234	5.112	.000
오차	1256.116	27494	.046		
합계	4060.215	27794			

a. R 제곱 = .691 (수정된 R 제곱 = .687)

표 2. PTI k의 환경, 발성유형, 모음별 평균과 표준편차
Table 2. Mean and standard deviation of PTI k across positions, phonation types, and vowels.

Phonation Type	Vowel	#CV			VCV			Total		
		Mean	St. Dev.	N	Mean	St. Dev.	N	Mean	St. Dev.	N
Aspirated	/i/	-0.201	0.324	382	-0.326	0.283	452	-0.269	0.309	834
	/e/	-0.386	0.417	1002	-0.379	0.296	1063	-0.382	0.36	2065
	/a/	-0.438	0.443	820	-0.402	0.306	1081	-0.418	0.372	1901
	/o/	0.06	0.267	734	0.023	0.211	715	0.042	0.242	1449
	/u/	0.173	0.303	456	0.055	0.266	486	0.112	0.291	942
	Total	-0.206	0.442	3394	-0.248	0.341	3797	-0.228	0.392	7191
Lenis	/i/	-0.253	0.288	441	-0.283	0.277	525	-0.269	0.282	966
	/e/	-0.261	0.303	773	-0.252	0.277	826	-0.256	0.289	1599
	/a/	-0.305	0.338	602	-0.213	0.259	967	-0.249	0.295	1569
	/o/	0.079	0.177	589	0.034	0.2	630	0.056	0.191	1219
	/u/	0.138	0.216	337	0.094	0.222	521	0.111	0.221	858
	Total	-0.147	0.329	2742	-0.142	0.291	3469	-0.144	0.308	6211
Fortis	/i/	-0.19	0.289	1095	-0.309	0.277	976	-0.246	0.29	2071
	/e/	-0.244	0.317	2016	-0.296	0.284	1518	-0.266	0.304	3534
	/a/	-0.253	0.347	2123	-0.263	0.274	1732	-0.257	0.316	3855
	/o/	0.048	0.27	1706	-0.024	0.219	1136	0.019	0.253	2842
	/u/	0.138	0.262	1187	0.041	0.243	903	0.096	0.258	2090
	Total	-0.122	0.343	8127	-0.191	0.297	6265	-0.152	0.326	14392
Total	/i/	-0.207	0.297	1918	-0.306	0.279	1953	-0.257	0.292	3871
	/e/	-0.285	0.349	3791	-0.311	0.29	3407	-0.297	0.323	7198
	/a/	-0.305	0.377	3545	-0.29	0.289	3780	-0.297	0.335	7325
	/o/	0.057	0.254	3029	0.004	0.213	2481	0.033	0.238	5510
	/u/	0.146	0.265	1980	0.059	0.245	1910	0.103	0.259	3890
	Total	-0.147	0.368	14263	-0.194	0.311	13531	-0.17	0.342	27794

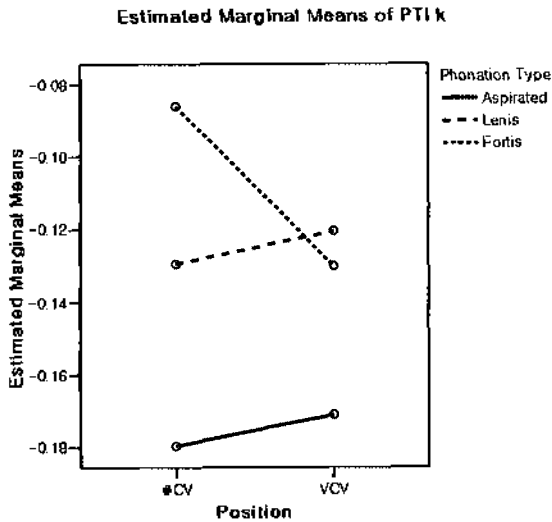


그림 15. 환경과 발성유형의 교호작용을 보여주는 PTI k의 추정주변평균
 Figure 15. Estimated marginal means of PTI k illustrating a significant interaction effect between position and phonation type.

의미한 차이가 있었으며 /a/와 /e/ 사이에는 유의미한 차이가 없었다. 화자에 대해서는 화자 AHS와 PDM 사이에만 유의미한 차이가 있었고 모든 화자간에 유의미한 차이가 있었다. 한편 분산분석의 결과 요인의 조합에 대하여 유의미한 교호작용 (interaction effect)이 있었다. 이들 중 환경, 발성유형, 모음 등 세 요인의 조합에 대하여 추정된 주변평균을 보여주는 프로파일 도표를 살펴 보았다. 먼저 환경과 발성유형에 대한 PTI k의 추정된 주변 평균이 그림 15에 제시되어 있다.

그림 15에 나타나 있듯이 환경과 발성유형간의 교호작용은 환경에 따른 발성유형 상호간의 차이가 다른 데서 비

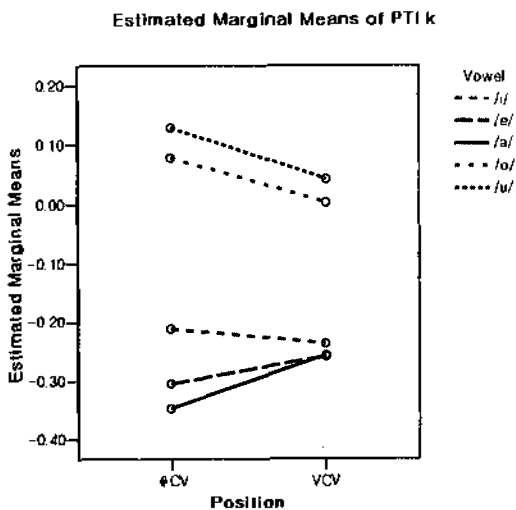


그림 16. 환경과 모음의 교호작용을 보여주는 PTI k의 추정주변평균
 Figure 16. Estimated marginal means of PTI k illustrating a significant interaction effect between position and vowel.

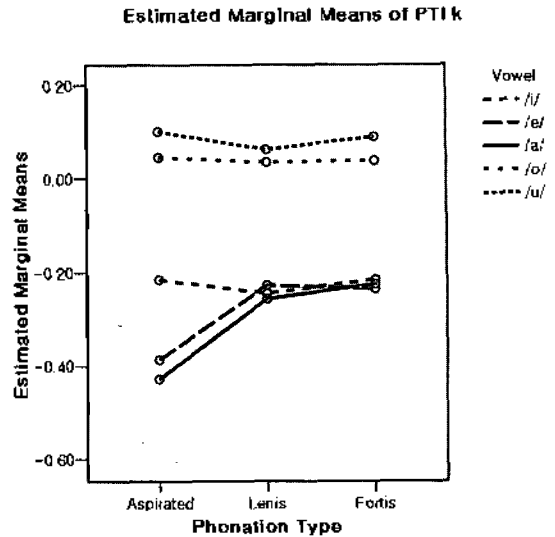


그림 17. 발성유형과 모음의 교호작용을 보여주는 PTI k의 추정주변평균
 Figure 17. Estimated marginal means of PTI k illustrating a significant interaction effect between phonation type and vowel.

롯된 것이다. #CV 환경에서는 경음, 연음, 격음 순으로 PTI k의 값이 크지만 VCV 환경에서는 연음, 경음, 격음 순으로 PTI k의 값이 크다. 다음으로 환경과 모음에 대한 PTI k의 추정된 주변 평균이 그림 16에 제시되어 있다.

그림 16에 나타나 있듯이 환경과 모음간의 교호작용은 환경에 따른 모음 상호간의 값의 폭이 다른 데서 비롯된 것이다. #CV 환경에서는 모음간 평균값의 폭이 넓지만 VCV 환경에서는 모음간 평균값의 폭이 #CV에 비해 좁아 평행한 모양을 보여주지 못한다. 마지막으로 발성유형과 모음에 대한 PTI k의 추정된 주변 평균이 그림 17에 제시되어 있다.

그림 17에 나타나 있듯이 발성유형과 모음간의 교호작용은 발성유형에 따른 모음 상호간의 값의 폭이 다른 데서 비롯된 것이다. 격음에서는 모음간 평균값의 폭이 넓지만 연음이나 경음에서는 모음간 평균값의 폭이 격음에 비해 좁아 평행한 모양을 보여주지 못한다. 특히 모음 /a/와 /e/가 연음이나 경음에서보다 격음에서 두드러지게 낮은 PTI k 값을 보여주고 있다.

요약하면 환경, 발성유형, 모음, 화자 등 각 요인에 대하여 유의미한 주효과가 있었고 환경, 발성유형, 모음 등 요인의 조합에 대해 유의미한 교호작용이 있었다. 이는 환경, 발성유형, 모음, 화자에 따라 PTI k값이 유의미한 차이를 보여준다는 것을 의미하며 각 요인간에서 서로 다른 값의 추이를 보여준다는 것을 의미한다.

V. 토론 및 결론

본 연구에서는 기본주파수와 모음의 영향을 배제한 가운데 전체적인 스펙트럼의 기울기를 하나의 지표로 나타내는 PTI k 를 한국어 모음 /i, e, a, o, u/를 포함한 음성샘플을 대상으로 살펴보았다. 본 연구의 결과 VCV 환경보다 #CV 환경에서 PTI k 의 값이 크게 나타났고, 격음에서보다 연음과 격음에서 PTI k 의 값이 크게 나타났으며, 모음 /o, u/가 모음 /i, e, a/보다 PTI k 의 값이 크게 나타났다. 이는 F1과 F2가 인접해 있어 배음의 진폭이 상승된 결과로 보인다. 평균과 표준편차 및 범위와 첨도를 대상으로 한 산점도에서 화자별로 특징적인 차이를 보임으로써 PTI k 가 화자간의 발성유형의 차이를 보여주는 의미 있는 음성 파라미터임을 확인하였다.

본 연구의 결과 모음 "오"와 "우"의 경우 다른 모음과 뚜렷이 구분되는 행동을 보여주었다. 이는 기존의 Park (2002)[4]에서 모음 "아"만을 대상으로 한 연구에서 한 걸음 더 나간 결과이다. 모음 "오"와 "우"의 포먼트 값을 정확하게 찾을 수 있는 방법은 별도의 연구과제로 남겨둔다. "오"와 "우"의 포먼트 주파수뿐만 아니라 음성매개변수를 신뢰성 있게 추출할 수 있는 방법을 개발한 후 다시 살펴보아야 할 것이다.

본 연구에서 확인된 바와 같이 PTI k 는 발성유형의 특성을 측정할 수 있는 의미 있는 음성매개변수로 보인다. PTI k 는 측정상의 문제를 보완하고 보다 많은 자료를 바탕으로 검증할 수 있을 것으로 보인다. PTI k 는 기본적으로 음원의 특성 중 스펙트럼의 기울기를 나타내는 매개변수이고 이는 성대가 달리는 속도와 관계가 있다. 성대의 진동시 성대가 달리는 속도가 특이한 경우 이 매개변수를 이용하여 화자 혹은 환자의 특성을 살펴볼 수 있으므로 음성의학에서 사용할 수 있을 것으로 기대된다. 뿐만 아니라 SenSyn과 같은 음성매개변수를 이용한 음성합성에서도 제3포먼트의 진폭이 스펙트럼의 기울기가 6dB/octave로 감소할 때와 얼마나 다른가를 TL이라는 변수로 설정하여 스펙트럼의 기울기를 통제한다. PTI k 는 특정주파수에서 추가로 어떤 값을 주기보다는 스펙트럼의 기울기 자체를 하나의 단일한 값으로 제시한다는 점에서 더욱더 간편한 방법이다. 이런 점을 고려하면 음성공학에서도 이용할 수 있을 것으로 기대된다.

본 연구에서는 양순음과 연구개음은 파찰음과 마찰음이 없고 치조 및 경구개음에만 세 가지 조음방법이 모두

있다. 조음위치의 영향을 피하기 위하여 양순음과 연구개음은 선택하지 않았다. 보다 포괄적인 연구를 위해서는 다른 조음위치에 나타나는 파열음도 포함되어야 한다. 비음과 유음은 독특한 특성을 가지고 있어서 이 부류의 음성에 대한 충분한 연구 후에 포함시킬 예정이다. 본 연구에서는 모음만을 대상으로 하여 PTI k 를 계산하였다. 하지만 다른 필터 특성이 있는 비음이나 유음 그리고 기식음의 연구에서는 가능한 한 정확한 음원특성을 얻기 위하여 필터 특성을 적절히 제거한 후 PTI k 를 계산하여야 할 것이다. 또 본 연구에서는 남성만을 대상으로 연구하였는데 남성과 여성의 발성유형상의 특성을 비교하기 위해서는 여성의 음성을 녹음하여 연구하여야 할 것이다. 이러한 연구는 다음의 과제로 남겨둔다.

참고 문헌

1. H. M. Hanson, "Glottal characteristics of female speakers: acoustic correlates," *Journal of the Acoustical Society of America*, **101** (1), 466-481, 1997.
2. P. Kirk, P. Ladefoged, J. Ladefoged, "Linguistic use of different phonation types," *UCLA Working Papers in Phonetics*, **59**, 102-113, 1984.
3. P. Ladefoged, "The linguistic use of different phonation types," *Vocal fold physiology: contemporary research and clinical issues* College Hill Press, San Diego, 1983, 351-360, 1983.
4. H. Park, *Temporal and Spectral Characteristics of Korean Phonation Types*, Ph.D. Dissertation, (The University of Texas at Austin, 2002.)
5. 박한상, "발성유형지수 k," 2002 대한음성학회 가을 학술대회 논문집, 77-80, 서울 11월, 2002.
6. H. Park, "Phonation Type Index k : A New Measure of Spectral Tilt," *Eoneohag Journal of the Linguistic Society of Korea* **36**, 101-111, 2003.
7. 이후동, 강선미, 박한상, 장문수, "녹음 환경의 차이에 따른 화자의 음원 특성 비교: 발성유형지수 k 를 중심으로," *음성과학* **10** (3), 213-224, 2003.
8. H. Park, "Problems in Source Characteristics," *Proceedings of the 2nd International Conference on Speech Sciences*, 107-114, Seoul, May, 2004.
9. H. Lee, S. Kang, M. Chang, H. Park, "A Study of Phonation Type Index k as a Parameter for Speaker Verification," *Proceedings of the 2nd International Conference on Speech Sciences*, pp.255-260, Seoul, May, 2004.
10. G. Fant, *Acoustic theory of speech production* (Mouton, The Hague, 1960.)

감사의 글

이 논문은 2002년도 한국학술진흥재단의 지원에 의하여 연구 되었음. (KRF-2002-037-A00055)

저자 약력

• 박 한 상 (Hansang Park)



2002년: The University of Texas at Austin (언어학 박사)

2003-2004: 대구한의대학교 국제어문학부 교수

2004- 현재: 홍익대학교 영어교육과 교수

※주관심분야: phonation type, experimental phonetics, acoustic phonetics