

## 음성분석의 최근 연구 동향에 관한 문헌적 고찰

박찬규 · 박영재 · 박영배

경희대학교 학과간협동과정 한방인체정보의학과

---

### Abstract

#### A Review on Recent Study of Acoustic Analysis

Chan-Kyu Park, Young-Jae Park, Young-Bae Park

*Dept. of Human Informatics of Oriental Medicine, Interdisciplinary Programs, Kyung Hee University*

**Objective :**

To review studies about acoustic analysis and to study its medical application.

**Method :**

We investigate the research results through internet search engines, such as Pubmed.

**Conclusions :**

I would like to conclude that there should be continued study, along with the objectification and quantification of diagnosis and application of acoustic analysis in the wide variety of oriental medicine.

**Key Words :** Acoustic Analysis

## I. 서 론

음성은 사람의 감정을 표현하는 중요한 수단 중의 하나로서 최근에 들어 음성의 음질이 치료 전후 혹은 치료 과정에서 어떻게 변화했는가를 객관적으로 기술하기 위해 음성 분석이 많이 이루어지고 있다. 즉, 음성장애 및 후두질환의 정확한 진단과 치료를 위해 음성분석이 시행되고 있는 것이다.

음성을 검사하는 방법에는 공기 역학적 검사, 음향 검사, 성대 진동 검사, 근 신경 검사, 청각 심리 검사 등이 있다. 최근까지는 청각인상에 기인한 평가법으로서 GRBAS척도를 표준으로 정했지만, 주관적인 평가법에 의한 평가의 불안정이라는 문제가 있었다<sup>1)</sup>. 이로 인하여 최근에는 컴퓨터 기술 발달에 힘입어 음성분석을 객관화하고 정량화하여 활용하는 경향을 증가하고 있다.

음성을 객관적으로 분석하는 연구는 지난 수십년간 꾸준히 이어져오고 있다. 특히 컴퓨터 프로그램의 개발로 인해 음성분석의 정량화 연구는 활발히 진행되어 왔는데, 이는 1991년부터 1995년까지의 음성과 관련된 논문 263편들 중에서 약 42.2%가 음성분석에 관한 것이라는 것에서 알 수 있다<sup>2)</sup>. 논문에서 언급된 음성분석방법에는 Cspeech, CSRE, ILS-PC, Mac-SpeechLab, Signalyse 등이 있는데 이 중에서 CSL(Computerized Speech Lab, Kay Elemetrics Co.)의 MDVP(Mutidimensional Voice Program)가 가장 많이 사용되어왔다(17.1%). 하지만 최근 들어 Dr. Speech(Tiger Electronics, Neu-Anspach, Germany)가 낮은 가격의 장점으로 인해 많이 사용되고 있다<sup>3)</sup>.

이러한 다양한 음성분석방법으로 인해 각각의 음성분석결과에 대한 비교연구가 진행되었다. Read<sup>4)</sup> 등은 대부분의 분석방법이 유의성은 있지만 과정이 어떻게 진행되느냐에 따라서 세세한 부분에서 차이

가 날 수 있다고 하였다. 따라서 그들은 보다 향상된 알고리즘의 필요성을 주장하였다. Karnell<sup>5)</sup> 등과 Bielamowicz<sup>6)</sup> 등은 음성분석방법들을 비교연구한 결과, 서로 다른 알고리즘을 이용하여 얻은 측정값은 연관성이 없다고 하였다. 또한 유<sup>7)</sup> 등과 Ilse<sup>3)</sup> 등의 연구에서도 음도와 같은 일반적인 알고리즘을 사용하는 경우는 상관관계가 높은 것으로 나타났지만 그렇지 않은 경우는 상관관계가 낮은 것으로 나타났다.

최근 음성 분석기기를 이용하여 특정 집단의 음성 기준 데이터에 대한 연구들이 많이 실시되어져 오고 있다. 이에 본 논문에서는 최근 연구들에 관한 문헌적 고찰을 통해 음성분석에 대한 연구가 어떻게 진행되었는지 살펴보고 앞으로의 활용 방안에 대해 서술하고자 한다.

## II. 본 론

### 1. 음성분석의 방법

음향학적 분석을 위한 음성 표본은 피험자들을 마이크에서 10cm 떨어져서 일정 시간 동안 가장 편안한 목소리로 발성시킨 음성에서, 음성의 떨림(perturbation)의 차이를 가장 잘 나타낸다고 보고되었으며<sup>8)</sup>, /a/모음 또는 기타 다른 모음을 발성시킨 것을 실험하기도 하고, 일정 길이의 단락의 글을 읽는 것 또는 평상시의 대화를 녹음하기도 하였다.

또한 음성의 객관적인 평가를 위해서는 음성분석 검사를 시행함에 있어 일관성 있고 안정된 결과를 얻을 수 있는 발성조건을 찾는 것이 중요하다. 손<sup>9)</sup> 등은 정상인을 대상으로 음도 및 강도의 차에 따라 발성의 조건을 주고 그에 따른 결과를 비교분석하였다. 그 결과, 편안한 발성에 비해 강도가 10dB 작을

때는 jitter 값이 증가되는 불안정한 양상을 보였고, 반대로 10dB 클 때는 shimmer 값이 편안한 발성에 비해 감소되는 안정된 발성양상을 보였다. 따라서 안정된 발성에 영향을 주는 것은 발성의 주파수보다 발성의 강도임을 주장하였고 음성분석검사를 시행함에 있어서도 편안한 발성보다는 좀 더 큰 소리로 발성하게 하는 것이 일관성있는 결과를 얻을 수 있는 조건이라고 하였다.

음성 sample의 길이도 음성분석에서 변수가 될 수 있다. 표<sup>10)</sup> 등의 연구에서는 음성 sample의 길이 변화에 따른 MDVP 측정치와 GRBAS 척도간의 상관관계 변화에 대한 비교를 실시하였다. 이 연구에서 음성의 전체 지속시간인 2.75초 중에서 0.5초, 1초, 1.5초, 2초, 2.5초 길이의 음성을 각각 편집해서 분석하였다. 그 결과, 1.5초의 음성 sample이 주관적인 평정과 객관적인 분석 결과간의 일치도가 가장 높은 것으로 나타났다. 또한 가장 상관성이 적었던 길이가 1초인 점은 음성 sample 길이 0.5초의 차이가 전혀 다른 결과를 가져올 수 있다는 것으로 보여주었으며 이는 음성분석에 있어 sample 길이 선정이 중요한 문제임을 시사해준다고 할 수 있다.

## 2. 음성분석의 지표

음성 분석기기를 통해 얻을 수 있는 음성 분석 지표들은 여러 가지가 있지만 일반적으로 많이 쓰이는 지표는 기본주파수(Fundamental frequency, F0), Jitter, Shimmer, NHR(Noise-harmony ratio) 등이 있다<sup>11)</sup>.

F0는 기능학적으로나 해부학적으로 후두를 평가하는데 있어 중요한 지표이며 이는 성대의 진동에 의해서 결정된다. F0는 음성 분석 지표들 중에서 가장 일반적으로 사용되고 있는 지표이다.

jitter는 성대진동의 주기마다 주파수의 변이를 나타

내는 것으로 거친 목소리(harsh voice quality)를 반영하고 shimmer는 성대진동의 주기마다 진폭의 변이를 나타내는 것으로 애성(hoarse voice quality)을 반영한다. 그리고 NHR은 잡음 대 정상음 간의 비율을 나타내는 지표라고 할 수 있다.

## 3. 음성분석의 임상적용

1) 연령 및 성별에 따른 차이에 관한 음성분석 연구  
연령에 따른 발성기관의 변화는 많은 저자들에게 의하여 보고되어 Segre<sup>12)</sup>는 연령증가에 따라 발성을 담당하는 기관의 변화가 음성의 변화에 영향을 미친다고 하였고, Endres<sup>13)</sup>, Hollien<sup>14)</sup>, Wilcox<sup>15)</sup> 등은 노인 음성 분석의 결과에 있어서 F0와 jitter 등이 변화를 보였다고 하였다.

진<sup>16)</sup> 등은 노인남녀와 성인남녀 각각 20명씩을 대상으로 실험한 결과, 성인남자에 비하여 노인남자에 있어서 jitter와 shimmer가 의미있게 증가하였고 여자의 경우 jitter만 의미있게 증가하였다고 보고하였다. 이는 후두기능에 미묘한 변화를 반영하는것이라고 하였으며 발성의 안정성을 나타낼 수 있는 지표인 NHR을 측정된 결과 여자노인에서 의미있게 증가되어 연령증가에 따른 발성의 불안정성을 증명하고 있다고 하였다.

김<sup>17)</sup> 등은 성인남녀와 노인남녀를 대상으로 한국어 모음의 제 1음형대와 제 2음형대 주파수를 측정하였다. 그 결과, 연령이 증가할수록 제 2음형대 주파수는 남녀 모두 유의하게 낮았고 이는 제 2음형대 주파수의 형성에 관여하는 성도의 신장에 의한것이라고 추측하였다. 그러나 제 1음형대 주파수는 노년군 남성의 경우 성인군에 비해 오히려 높았고 노년군 여성의 경우는 낮게 나타났으나 유의하지는 않았다. 따라서 제 1음형대의 형성에 관여하는 성대에서 혀의

최고 수축점까지를 이루고 있는 성도는 연령증가에 별다른 영향을 받지 않는다고 하였다.

M.M. Gorham-Rowan<sup>18)</sup> 등은 노인남녀와 성인남녀의 음성을 평가자들에게 들려준 후 이들로 하여금 숨이 찬 듯 약한 음성(breathiness)과 거친 음성(hoarseness)에 대해 주관적인 평가를 하도록 하였다. 음성분석기기의 지표값에 있어서는 남녀모두 연령의 증가에 따른 변화가 나타났지만, 주관적인 평가의 경우 여성은 연령이 증가할수록 breathiness와 hoarseness가 유의한 차이가 있었으나 남성은 둘 다 유의한 차이가 없었다. 이는 음성분석의 객관적인 평가와 주관적인 평가의 결과가 다를수 있다는 것을 보여준다.

한편 성별에 따라서도 음성분석의 결과가 달라지는데, 표<sup>19)</sup> 등의 연구에서 남녀 대상자들에게 모음을 발성하게 한후 비교한 결과, 남자보다 여자가 F0가 유의하게 증가하는 것으로 나타났다. jitter의 경우는 남자가 여자보다 유의하게 낮았으며 shimmer는 남녀간의 유의한 차이는 나타나지 않았다. NHR은 남자가 여자보다 유의하게 높은 것으로 나타났다. 특히 여성의 경우, 폐경기에 이르면 성대 자체에 변화를 유발하여 F0가 저하된다는 보고도 있다<sup>20-21)</sup>.

## 2) 각종 스트레스에 의한 음성분석에 관한 연구

P.Wittels<sup>22)</sup> 등은 인간의 감정상태와 음성에 관한 연구를 실시하였다. 실험자들로 하여금 정신적 스트레스를 가한 후 음성을 측정하였고, 30분 후 다시 스트레스를 유발한 다음 음성을 측정하였으며 마지막으로 3일 후에 똑같은 방법으로 스트레스를 가한 후 음성을 측정하였다. 그 결과, 첫 번째 실험에서 F0 값이 기본상태에서 측정한 값보다 유의하게 증가하였다. 하지만 재차 실험한 결과에서는 F0 값이 감소는 하였으나 유의한 결과는 아닌 것으로 나타났다.

E.Mendoza<sup>23)</sup> 등도 정신적 스트레스를 받는 환경

에서 음성 데이터를 수집하여 분석하였다. 실험자들은 /a/ 모음을 5초간 발음하여 음성을 녹음하였고 분석결과, F0와 jitter, shimmer 값 모두 기본상태보다 유의하게 증가한 것으로 나타났다.

한편 후두피로는 목소리를 과다사용하거나 잘못 사용하여 발생하는 병리적, 신경학적인 이상을 말하는데 J.C.Stemple<sup>24)</sup> 등은 후두피로(laryngeal fatigue)가 음성에 어떤 영향을 미치는지를 실험하였다. 실험자들은 2시간동안 소설의 한 부분을 정해진 목소리 크기로 읽도록 하였다. 그 결과, 실험전에 비해서 F0 값이 유의하게 증가한 것으로 나타났다.

## 3) 기타 다른 연구들

음성분석의 결과를 이용하여 중의학과 연관성에 대한 연구들도 진행되어왔다. C.C.Chiu<sup>25)</sup> 등은 음성분석의 지표값들이 정상인과 기허 및 음허인 사람들을 변증하는데 유의성이 있다는 것을 밝혀냈다. 이는 중의학에서 이용하는 진단방법인 望, 聞, 問, 切의 四診 중에서 聞診에 있어 새로운 가능성을 제시했다고 할 수 있다.

그리고 Barros<sup>26)</sup> 등은 음성분석 지표와 중의학의 오음인 宮, 商, 角, 徵, 羽가 연관성이 있다는 것을 밝혀냈고 특히 F0가 중의학의 진단방법을 보완할 수 있을 것이라고 주장하였다. 따라서 F0는 오행이나 오장과의 관련성도 크다고 할 수 있다.

한편 흡연과 음성에 관한 연구들 또한 많이 있었는데 기존 연구들에서, 흡연은 성대의 진동패턴에 변화를 가져오게 되고 이는 결국 F0 값을 저하시키고 jitter 값이 증가한다고 하였다<sup>20, 27-29)</sup>. I. Guimaraes<sup>30)</sup> 등은 최근 연구에서 흡연자들이 비흡연자들에 비해서 조직학적으로 후두의 병변이 더 많이 나타나고 F0의 평균값이 떨어지게 된다고 하였다.

이<sup>1)</sup> 등은 정상 및 후두질환 환자의 음성분석을

Dr. Speech를 이용하여 비교분석하였다. 후두질환 환자는 성대 결절 및 용종으로 수술을 받은 환자들을 대상으로 하였다. 그 결과, 성대 결절환자는 수술전 후의 jitter 및 shimmer 값의 의미있는 변화는 없었으며 용종환자의 경우는 수술전후로 jitter 및 shimmer 값이 유의성있게 감소되는 것으로 나타났다. 이는 Dr. Speech 를 이용한 음성분석검사가 수술전후의 음성변화의 객관적인 평가를 위한 방법으로 사용이 가능하다는 것을 보여준다고 할 수 있다.

### Ⅲ. 고 찰

음성은 사람의 감정을 표현하는 중요한 수단 중의 하나로서 최근에 들어 음성분석을 통하여 임상적으로 여러 분야에서 이용을 하고 있다. 음성분석의 중요한 목적의 하나는 음성의 음질이 치료 전후 혹은 치료 과정에서 어떻게 변화했는가를 객관적으로 기술하는 것이다. 병적음성의 음성분석에 관한 연구는 1960년대 Liberman<sup>31)</sup>이 음성 기본 주기의 불규칙성의 연구에서 시작되었다. 정상인의 경우 성대가 규칙적인 진동을 보이지만 병변이 있는 경우는 주기마다 다른 진동을 나타내어 이들의 목소리는 거칠고 곱고럽다고 표현된다고 하였다. 이러한 음성의 떨림을 객관적으로 측정하기 위해 Horn<sup>32)</sup>는 jitter 와 shimmer를 언급하였다. 그 후 기술의 발달에 힘입어 음성 분석에 관한 연구가 활발하게 진행되어왔고 여러 음성분석기기가 개발되어 음성과형 분석이 보다 용이해지고 있다.

음성분석방법에 있어 객관적인 분석외에도 일본 음성 언어학회 등에서는 주관적인 방법인 GRBAS 척도를 표준으로 정했으나 평가의 애매성이나 불완전성의 문제가 야기되었다. 하지만 표 등의 연구에

따르면 주관적 음성분석 방법인 GRBAS 척도와 객관적 음성분석 방법인 MDVP 측정치와의 상관성이 높은 것을 밝혀냈으며 1.5초정도의 음성 sample이 주관적인 평정과 객관적인 분석의 결과 간의 일치도가 가장 높은 것이라고 설명하였다.

한편 유<sup>7)</sup> 등은 객관적 음성분석기기인 MDVP, Dr.Speech, Praat 간의 음향학적 측정치에 관한 상관 연구를 시행하였는데, 각 기기간의 FO 값과 shimmer 값은 매우 높은 상관성을 나타내었다. 하지만 jitter 값은 상관성이 없는 것으로 나타났는데 이는 각 기기 간 알고리즘의 차이에 의한 것이라고 의심을 하였다. 따라서 향후 보다 객관적인 음성분석을 위해서는 여러 음성분석기기에 상관관계의 분석을 통해서 적절한 지표를 찾아내고 이를 활용하여야 할 것이다.

임상적으로 음성분석에 관한 연구는 주로 성별, 연령, 스트레스, 흡연유무 등과 연관되어 이루어져 왔다. 연령이나 성별에 따라 성대의 구조 등의 차이가 발생하기 때문에 음성의 평가에 있어서도 차이가 생기는 것으로 보여지며 지금까지의 연구들은 주로 성인과 노인그룹의 비교연구가 많았는데 향후 소아와 성인 및 노인그룹에 대한 비교연구가 이루어져야 할 것으로 사료된다. 또한 정신적 스트레스나 흡연유무와 음성의 관한 연구도 보다 많은 인원을 대상으로 한 연구가 필요할 것으로 보인다.

특히 한의학에서 이용하는 望, 聞, 問, 切의 四診 중에 聞診과 음성분석과의 비교연구는 향후 임상적으로도 의의가 있을 것으로 사료되며 보다 객관적인 진단근거를 뒷받침해줄 것으로 생각된다.

### Ⅳ. 결 론

최근 임상적으로 후두질환 등의 진단 및 치료를

위해 음성분석에 관한 연구가 활발히 진행되어왔다. 향후 음성분석의 객관성을 유지하기 위해 여러 음성 분석기기들의 상관성을 제고시킬 필요가 있으며 이를 한의학적인 진단방법에 응용한다면 임상적으로 다양하게 활용을 할 수 있을 것으로 사료된다.

## 참 고 문 헌

- 이형석, 태경, 장경진, 김경우, 김경래, 박철원. Dr. Speech Science를 이용한 정상 및 후두질환 환자의 음향분석. 대한음성언어의학회지. 1997; 8(2):166-172.
- De Bodt MS. A framework for voice assessment: the relation between subjective and objective parameters in the judgement of normal and pathological voice. Doctoral dissertation. University of Antwerp;1997:23.
- Ilse Smits, Piet Ceuppens, De Bodt MS. A comparative study of acoustic voice measurements by means of Dr. Speech and Computerized Speech Lab. Journal of Voice. 2005;19(2):187-196.
- Read C, Buder EH, Ken RD. Speech Analysis Systems: an evaluation. J Speech Hear Res. 1992; 35:314-332.
- Karnell MP, Hall KD, Landahl KL. Comparison of fundamental frequency and perturbation measurements among three analysis systems. J Voice. 1995;9:383-393.
- Bielamowicz S, Kreiman J, Gerratt BR, Dauer MS, Berke GS. Comparison of voice analysis system for perturbation measurement. J Speech Hear Res. 1996;39:126-134.
- 유재연, 정옥란, 장태엽, 고도홍. MDVP와 Praat, Dr. Speech간의 음향학적 측정치에 관한 상관연구. 음성과학. 2003;10(3):29-36.
- Lehman JJ, Bles DM, Brandenburg JH : An objective assessment : of voice production after radiation therapy for stage 1 squamous cell carcinoma of the glottis. Otolaryngol H & N Surg. 1998;98:121-129.
- 손영익, 윤영선, 권중근, 추광철. 발성시 음도 및 강도의 변화가 음성분석검사 결과에 미치는 영향. 대한음성언어의학회지. 1997;8(1):12-17.
- 표화영, 심현섭, 임성은. 음성 Sample의 길이 변화에 따른 MDVP 측정치와 GRBAS 척도간의 상관관계 변화 비교. 음성과학. 2000;7(2):71-81.
- De Felipe ACN, Martelletti Grillo MHM, Grecbi TH. Standardization of acoustic measures for normal voice patterns. Brazilian Journal of Otorhinolaryngology. 2006;72(5):659-664.
- Segre R. Senescence of the voice. EENT Monthly. 1971;50:223-227.
- Endres W, Bambach W, Flosser G. Voice spectrograms as a function of age, voice disguise and voice imitation. J of Acoust Soc of Am. 1971;49:1842-1848.
- Hollien H, Shipp T. Speaking fundamental frequency and chronological age in males. J of Speech and Hearing Research. 1972;15:155-159.
- Wilcox KA, Horii Y. Age and changes in vocal jitter. J Gerontol. 1980;35:194-198.
- 진성민, 권기환, 강현국. 연령증가에 따른 정상 노인의 음향분석학적 특징. 대한음성언어의학회지. 1997;8(1):44-48.
- 김찬우, 차홍익, 장일환, 김선태, 오승철, 석윤식,

- 이영숙. 한국어 모음에서 연령증가에 따른 제2 음형대의 변화양상. *대한음성언어학회지*. 1999; 10(2):119-123
18. Gorham-Rowan MM, Laures-Gore J. Acoustic-perceptual correlates of voice quality in elderly men and women. *Journal of Communication Disorders*. 2006;39:171-184.
  19. 표화영, 심현섭, 송윤경, 윤영선, 이은경, 임성은, 하현령, 최홍식. 한국 성인의 정상 음성에 관한 기본 음성 측정치 연구. *음성과학*. 2002;9(2): 179-192
  20. Stoicheff ML. Speaking fundamental frequency characteristics of nonsmoking females. *J Speech Hear Res*. 1981;24:437-441.
  21. Hollien HP, K. Massey. Longitudinal research on adolescent voice change in males. *J Acoust Soc Am*. 1994;96(5):2646-2654.
  22. P. Wittels, B. Johannes, R. Enne, K. Kirsch, HC. Gunga. Voice monitoring to measure emotional load during short-term stress. *Eur J Appl Physiol*. 2002;87:278-282.
  23. E. Mendoza, G. Carballo. Acoustic analysis of induced vocal stress by means of cognitive workload tasks. *Journal of Voice*. 1998;12(3): 263-273.
  24. Stemple JC, Stanley J, Lee L. Objective measures of voice production in normal subjects following prolonged voice use. *Journal of Voice*. 1995;9(2):127-133.
  25. Chiu CC, Chang HH, Yang CH. Objective auscultation for traditional Chinese medical diagnosis using novel acoustic parameters. *Computer Methods and Programs in Biomedicine*. 2000;62:99-107.
  26. Dotta de Barros FC, De Barros SE. The fundamental frequency of voice and its correlations with the five sounds of the traditional Chinese medicine. *Acupuncture and Electro-Therapeutics Res*. 2007;32:211-221.
  27. Gilbert H, Weismer G. The effects of smoking on the fundamental frequency of adult women. *J Psycholinguist Res*. 1974;3:225-331.
  28. Sorensen D, Horii Y. Cigarette smoking and voice fundamental frequency. *J Commun Disord*. 1982;15:135-144.
  29. Damborenea Tajada J, Fernández Liesa R, Llorente Arenas E, Naya Glávez MJ, Marín Garrido C, Rueda Gormedino P, et al. Efecto del consumo de tabaco en el análisis acústico de la voz. *Acta Otorrinolaring Esp*. 1999;50(6): 448-25.
  30. I. Guimarães, E. Abberton. Health and voice quality in smokers: an exploratory investigation. *Logopedics Phoniatrics Vocology*. 2005; 30:185-191.
  31. Liberman P. Some acoustic measures of the fundamental periodicity of normal and pathologic larynges. *J Acoust Soc Am*. 1963;35: 344-353.
  32. Horri Y. Vocal shimmer in sustained phonation. *J Speech and Hearing Research*. 1980;23: 202-209.