

기능성 음성장애의 진단을 위한 음향학적, 청지각적 평가

부산대학교병원 이비인후과,¹ 양산부산대학교병원 이비인후과,² 경상대학교병원 이비인후과,³ 동남권원자력의학원 이비인후과,⁴ 춘해보건대학교 언어치료과,⁵ 부산대학교 인문대학 언어정보학과⁶
김근효¹ · 이연우¹ · 배인호² · 이재석³ · 이창윤⁴ · 박희준⁵ · 이병주¹ · 권순복⁶

Acoustic Analysis and Auditory-Perceptual Assessment for Diagnosis of Functional Dysphonia

Geun-Hyo Kim, PhD¹, Yeon-Yoo Lee, MS¹, In-Ho Bae, PhD², Jae-Seok Lee, MS³,
Chang-Yoon Lee, MS⁴, Hee-June Park, PhD⁵, Byung-Joo Lee, MD, PhD¹ and Soon-Bok Kwon, PhD⁶

¹Department of Otorhinolaryngology-Head and Neck Surgery and Biomedical Research Institute, Pusan National University Hospital, Busan; and ²Department of Otorhinolaryngology-Head and Neck Surgery, Pusan National University Yangsan Hospital, Yangsan; and ³Department of Otolaryngology, Gyeongsang National University Hospital, Jinju; and ⁴Department of Otorhinolaryngology, Dongnam Institute of Radiological & Medical Sciences, Busan; and ⁵Department of Speech-Language Therapy, Choonhae College of Health Sciences, Ulsan; and ⁶Humanities, Language and Information, Pusan National University, Busan, Korea

– ABSTRACT –

Background and Objectives : The purpose of this study was to compare the measured values of acoustic and auditory perceptual assessments between normal and functional dysphonia (FD) groups. **Materials and Methods** : 102 subjects with FD and 59 normal voice groups were participated in this study. Mid-vowel portion of the sustained vowel /a/ and two sentences of ‘Sanchaek’ were edited, concatenated, and analyzed by Praat script. And then auditory-perceptual (AP) rating was completed by three listeners. **Results** : The FD group showed higher acoustic voice quality index version 2.02 and version 3.01 (AVQIv2 and AVQIv3), slope, Hammarberg index (HAM), grade (G) and overall severity (OS), values than normal group. Additionally, smoothed cepstral peak prominence in Praat (PraatCPPS), tilt, low-to high spectral band energies (L/H ratio), long-term average spectrum (LTAS) in FD group were lower than normal voice group. And the correlation among measured values ranged from -0.250 to 0.960. In ROC curve analysis, cutoff values of AVQIv2, AVQIv3, PraatCPPS, slope, tilt, L/H ratio, HAM, and LTAS were 3.270, 2.013, 13.838, -22.286, -9.754, 369.043, 27.912, and 34.523, respectively, and the AUC of each analysis was over .890 in AVQIv2, AVQIv3, and PraatCPPS, over 0.731 in HAM, tilt, and slope, over 0.605 in LTAS and L/H ratio. **Conclusions** : In conclusion, AVQI and CPPS showed the highest predictive power for distinguishing between normal and FD groups. Acoustic analyses and AP rating as noninvasive examination can reinforce the screening capability of FD and help to establish efficient diagnosis and treatment process plan for FD. (J Clinical Otolaryngol 2018;29:212-222)

KEY WORDS : Functional dysphonia · Acoustics · Praat · GRBAS · CAPE-V.

논문접수일 : 2018년 9월 14일 / 논문수정일 : 2018년 10월 15일 / 심사완료일 : 2018년 11월 4일
교신저자 : 권순복, 46241 부산광역시 금정구 부산대학교로63번길 2 부산대학교 인문대학 언어정보학과
전화 : (051) 510-2003 · 전송 : (051) 582-0572 · E-mail : sbkwon@pusan.ac.kr

서 론

기능적 음성장애(functional dysphonia, FD)는 구조적 또는 신경학적 후두 병리학적 특징이 없는 경우에 발생하는 음성장애를 말하며, 많은 임상현장에서 진료받는 환자의 10~40%정도를 차지한다.^{1,2)} 기질적인 문제가 없고 발성 중 나타나는 후두생리의 결과로 여겨진다. 후두의 좌우, 전후방 수축, 뒤쪽 성문의 열림, 경련 등과 같은 후두경 소견은 기능적 음성장애의 증상으로 간주된다. 기능적 음성장애는 일반적으로 상기도 감염 후에 자주 발생하며, 일시적이기도 하며 치료에 대한 반응이 매우 다양하게 나타난다.^{3,4)}

기능적이라는 용어는 해부학적 구조보다는 생리적 기능의 음성 문제를 의미한다.⁵⁾ 임상현장에서 ‘기능적’이라는 용어는 ‘기질적’인 것과 대조되며 종종 심인성, 심리적 요인이 추가된 의미를 전달한다.⁶⁾ 스트레스, 감정, 심리적 갈등은 기능적 문제를 유발하거나 악화시키는 것으로 알려져 있다. 전환장애, 히스테리, 근육 오용, 근긴장성 발성장애(muscle tension dysphonia) 등 명확하게 설명할 수 없는 음성장애를 포함하기 있기 때문에 기능적 음성장애라는 진단 범주에 속해있다.^{7,8)} 각각의 진단적 표지(diagnostic label)가 병인학적 이질성(etiological heterogeneity)을 내포하고 있지만, 이들 질환은 질적으로 다르고 병인학적 명확성이 불분명하다. 이러한 다양한 진단적 지표는 종종 임상가의 가정, 편향, 선호가 반영된다. 이러한 기능적 음성장애를 신뢰성 있게 구별하는 것은 쉽지 않다.⁹⁾

본 연구에서는 정상 음성과 기능적 음성장애를 구분하기 위해서 다양한 음향학적 측정들과 청지각적 평가를 실시하였다. 이를 통해서 두 집단의 차이를 확인하고 구분할 수 있는 측정 변수를 알아보고자 하였다. 본 연구의 문제는 다음과 같다.

첫째, 측정 변수들에 대한 집단간에 차이가 있는가?

둘째, 청지각적 평가 변수들과 음향학적 측정 변수들 간 상관관계는 어떠한가?

셋째, 정상과 병리적 음성을 구분하는 절사값, 진단에 측력은 무엇인가?

대상 및 방법

대 상

본 연구의 대상이 된 기능성 음성장애 환자군은 2013년 4월부터 2018년 1월까지 부산, 경남지역의 이비인후과를 통해 내원한 환자들 중 기능성 음성장애로 진단된 환자 102명(59.4±12.5세)이었다. 정상집단 59명(52.9±9.1세)은 후두내시경 상으로 관찰되는 병변이 없고 음성장애를 호소하지 않는 사람들을 대상으로 후두내시경 및 후두 스트로보스코피 검사를 시행하여 성문폐쇄부전 혹은 과내전의 소견을 보이지 않고 음성검사에서도 정상 범위의 값이 측정되고 음성장애를 호소하지 않은 대상자들로 분류하였다.

기능성 음성장애 집단은 구조적, 신경학적, 후두병리학적 특징이 없이 음성장애를 호소하는 대상자들로 분류하였다. 본 연구에 포함된 기능적 음성장애 환자들은 근긴장성 발성장애 66명, 가성대 발성 23명, 음성과다 사용(음성피로) 9명, 후두외상 4명이었다. 두 집단 모두 외과적 수술 병력 및 신경학적 질환이 없는 사람들을 대상으로 하였다.

연구의 절차

음성녹음

음성녹음은 언어치료실 내 방음부스에서 Computerized Speech Lab(CSL model 4500, Kay Electronic, USA) 과 단일지향성 마이크인 Shure SM48(Shure Inc, USA) 를 사용하여 녹음하였다. 모음 연장발성과 산책 문단 읽기과업을 편안한 음도와 강도로 수행하였다. 녹음 전 2회의 연습을 하도록 지시하였고, 모음 연장발성은 3회 반복하여 가장 안정적인 음성 샘플의 안정구간 2초를 선정하여 저장하였다. 읽기과업은 산책문단 중 2개의 문장(26음절)을 편안한 속도로 읽었다. 저장 형태는 표본 추출률 44.1 kHz, 양자화 16bit, wav이었다.

음성편집

음성샘플에 대한 음향학적, 청지각적 평가를 위해서 모음연장발성 샘플과 문장읽기 샘플을 연결하여 하나의 음성샘플(concatenated voice sample)로 구성하였다.

Praat 스크립트를 통해서 모든 음향학적 측정이 자동으로 완료되며 결과는 텍스트파일(txt)로 저장된다.

신호 대 잡음비(signal to noise ratio, SNR) 측정

음향학적 분석에 대한 높은 신뢰성을 갖기 위해서 SNR 30 db 이상의 조건을 만족하는 음성샘플을 분석할 것을 권고하였다.¹⁰⁾ 본 연구에서 SNR은 31.0~72.11이었으며 평균 44.09이었다. 이 중 30 db 이하인 6명을 제외하였다.

청지각적 평가

음성장애 평가에 있어 7년 이상의 경력이 있는 언어 재활사 3명이 기능성 음성장애의 음성샘플을 평가하였다. 평가자들은 사전 정보 없이 무작위로 제시된 음성 샘플을 통해서 청지각적 평가를 수행하였다. 청지각적 평가는 무성음, 쉼 구간이 제거되기 전의 음성샘플을 이용하여 평가하였다. GRBAS, CAPE-V를 통해서 grade, overall severity를 평가하였다. 평가자 내 신뢰도 측정을 위해서 전체 음성 샘플의 약 20%에 해당하는 20명의 음성샘플을 무작위로 선정하여 초기 평가 후 4주 내에 재평가하였다.

음향학적 측정

본 연구에서 사용된 음향학적 측정은 Praat 스크립트를 사용하여 자동으로 분석하였다. 측정 변수는 두 가지 버전의 AVQI(v.2.02, v.3.01), PraatCPPS, low-to high spectral band energies(L/H ratio), Hammarberg index (HAM), Long-term average spectrum(LTAS)이다. AVQI를 계산한 두 가지 회귀식은 (1), (2)와 같다.

$$AVQI\ v.2.02=9.072-(0.245x\ PraatCPPS)-(0.161x\ HNR)-(0.470xSL)+(6.158xSLdB)-(0.071xSlope)+(0.170xTilt)(1)$$

$$AVQI\ v.3.01=(4.152-(0.177x\ PraatCPPS)-(0.006x\ HNR)-0.037xSL)+(0.941xSLdB)+(0.01*Slope)+(0.093*Tilt)*2.8902(2)$$

Slope는 1 kHz 미만의 평균 스펙트럼 강도와 1 kHz에서 10 kHz까지 평균 스펙트럼 강도 사이에 dB 차이를 계산한다. Tilt는 1~10,000 Hz 평균 스펙트럼의 추세

선 기울기를 의미한다. Praat에서 get trend line으로 측정한다. Praat에서의 L/H ratio는 저주파수 대역과 고주파수 대역 사이의 에너지 비율을 측정하는 것이며, 저주파수 대역은 0~4,000 Hz, 고주파수 대역은 4,000~8,000 Hz를 포함한다. 병리적인 음성일수록 수치가 낮다. Ham-marberg index(HAM)는 1980년 Hammarberg에 의해서 소개되었으며 0~2,000 Hz범위의 LTAS 곡선의 최대강도와 2,000~5,000Hz 영역의 최대강도 차이를 의미한다.¹¹⁾ 감정 발화(emotionally expressive speech)에 대한 지각(perception) 및 음향학적 특성을 분석할 때 주로 사용된다.^{12,13)} LTAS는 연속 구어를 분석하는데 유용한 도구이며, 전체 음성 스펙트럼 특성을 재현할 수 있다. LTAS는 고속 푸리에 변환으로 생성된 파워 스펙트럼으로, 스펙트럼 모멘트 분석(spectral moments analysis)을 사용하여 가우스 종 곡선(gaussian bell curve)과 비교할 수 있다.¹⁴⁾ Long-term average spectrum(LTAS)는 불완전 성문폐쇄, 기식성과 연관이 있으며 특히 모음연장보다는 연속 구어에서 더욱 연관성이 높았다.^{15,16)} Praat을 통해서 전체 음성 스펙트럼의 평균 에너지 값을 측정하였다.

자료 처리

음성평가에 대한 통계 분석 및 자료 처리는 R, version 3.4.0(The R Foundation for Statistical Computing, Vienna, Austria) and RStudio 1.0.143(RStudio Inc., Boston, MA, USA)을 통해서 수행되었다. 청지각적 평가에 대한 평가자 간, 내의 신뢰도 측정을 위해서 intrarater correlation coefficient, interrater correlation coefficient(ICC) 통계 분석을 수행하였다.

집단 간 AVQIv2, AVQIv3, PraatCPPS, L/H ratio, HAM, LTAS, Grade, OS의 차이는 독립표본 t-검정을 통해 검증되었다. 평가 변수들 간의 상관관계는 피어슨 상관분석을 통해서 검증하였고 유의수준은 95%였다. 정상음성과 기능적 음성장애에 대한 감별 진단적 변수를 확인하기 위해서 ROC(receiver operating characteristics) 곡선 분석을 실시하여 민감도, 특이도, 곡선하 영역을 확인하였다.

결 과

청지각적 평가의 신뢰도 측정

평가자 내 신뢰도

기능성 음성장애에 대한 청지각적 평가의 평가자 내, 간 신뢰도 결과는 Table 1, 2에 제시되었다. 모음연장과 연속구어의 평균 G에 대한 평가자 내 신뢰도는 평가자 1은 0.883(p<0.01), 평가자 2는 0.835(p<0.01), 평가자 3은 0.851(p<0.01)이었다. OS에 대한 평가자 내 신뢰도는 평가자 1은 0.824(p<0.01), 평가자 2는 0.812(p<0.01), 평가자 3은 0.784(p<0.01)이었다(Table 1).

평가자 간 신뢰도

첫 번째 평가에서 G는 0.762~0.837(mean=0.801, p<0.01), OS는 0.811~0.855(mean=0.832, p<0.01)이었다. 두 번째 평가에서는 G는 0.772~0.797(mean=0.786, p<0.01), 0.735~0.819(mean=0.765, p<0.01)이었다(Table 2).

모든 측정 변수들의 집단별 비교

모든 측정 변수들에 대한 집단별 비교 결과는 Fig. 1, Table 3에 제시하였다.

Table 1. Intra-rater reliability of the three listeners who rated G and OS

ICC	Rater 1	Rater 2	Rater 3
G	0.883**	0.835**	0.851**
OS	0.824**	0.812**	0.784**

** : p<.01. ICC : intra-rater correlation coefficient, G : mean value of grade, OS : overall severity

Table 2. Inter-rater reliability of the three listeners who rated G and OS

	Rater 2		Rater 3	
	ICC on G	ICC on OS	ICC on G	ICC on OS
Rater 1				
1st rating	0.806**	0.832**	0.762**	0.811**
2nd rating	0.772**	0.741*	0.797**	0.735**
Rater 2				
1st rating	-		0.837**	0.855**
2nd rating			0.789**	0.819**

* : p<.05, ** : p<.01. ICC : inter-rater correlation coefficient, G : grade, OS : overall severity

AVQIv2, v3, PraatCPPS의 집단별 비교

집단 간 차이를 알아본 결과, AVQIv2에서 정상집단은 2.8±0.5, 기능성 음성장애 집단은 5.2±1.5, AVQIv3에서는 정상집단은 1.2±0.5, 기능성 음성장애 집단은 3.9±1.9, PraatCPPS는 정상집단 14.7±1.0, 기능성 음성장애 집단은 11.1±2.7로 측정되었고, 모두 통계적으로 유의미한 차이(p<0.01)가 있었다. 음성의 질이 나쁠수록 AVQI 값이 증가, PraatCPPS 값이 감소하는데 기능성 음성장애 집단이 정상집단보다 더 큰 AVQI, 더 작은 PraatCPPS값으로 확인되었다.

Slope, tilt, L/H ratio, HAM, LTAS의 집단별 비교

Slope(p<0.01), tilt(p<0.01), L/H ratio(p<0.05), HAM(p<0.01), LTAS(p<0.01)에서 집단별 유의한 차이가 확인되었다(Table 3). 기능성 음성장애 집단은 정상 집단에 비해 더 큰 slope, HAM 값과 더 작은 tilt, L/H ratio, LTAS 값을 나타내었다.

청지각적 평가의 집단별 비교

집단 간 G의 차이를 알아본 결과, 정상집단은 G0(59명, 100.0%)으로 평균 0.00, 기능성 음성장애 집단은 G0(12명, 11.8%), G1(37명, 36.3%), G1.5(6명, 5.9%), G2(40명, 39.2%), G2.5(6명, 5.9%), G3(1명, 1.00%)으로 평균 1.45로 평가되었다. 기능성 음성장애 집단이 정상집단보다 더 나쁜 음성을 산출하는 것으로 평가되었고, 두 집단 간의 유의미한 차이(p<0.01)가 확인되었다.

집단 간 OS의 차이를 알아본 결과, 정상집단은 9.5±5.9, 기능성 음성장애 집단은 43.7±17.2로 평가되었다. 음성의 질이 나쁠수록 OS 값이 큰 값으로 점수화되는데

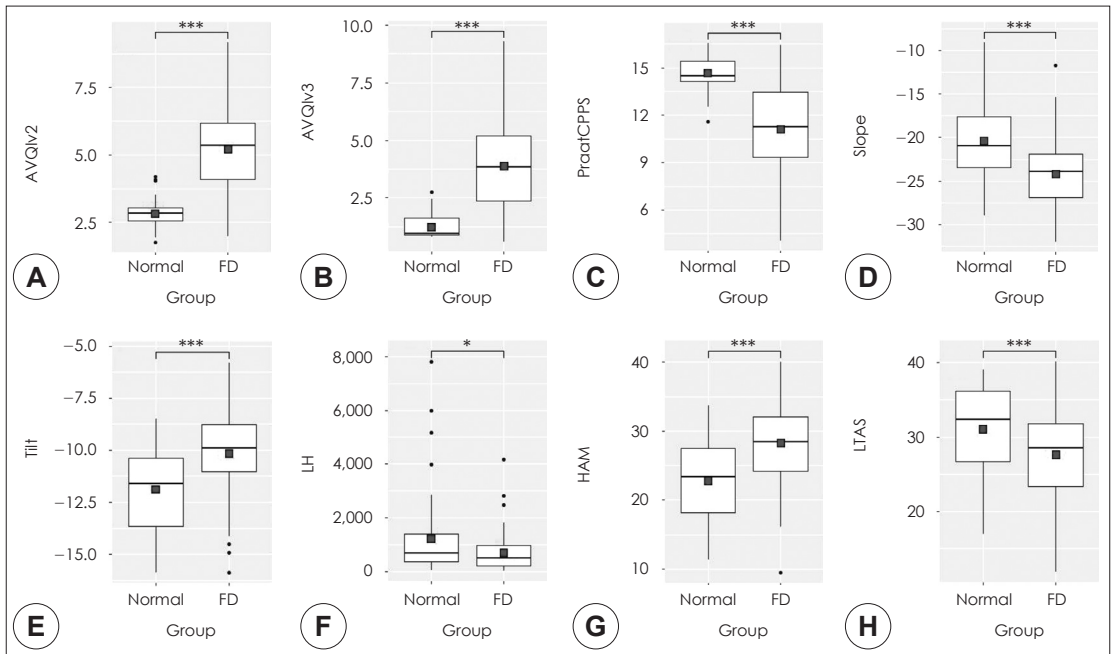


Fig. 1. Comparison of measured variables between normal and FD groups. A : AVQIv2. B : AVQIv3. C : PraatCPPS. D : Slope. E : Tilt. F : L/H ratio. G : HAM. H : LTAS. * : $p < .05$, *** : $p < .001$.

두 집단 간에도 통계적으로 유의미한 차이($p < 0.01$)가 확인되었다.

slope(-22.286), tilt(-9.754), L/H ratio(369.043), HAM (27.912), LTAS(34.523)이었다(Table 4).

측정변수들간의 상관관계

측정변수들간의 상관관계를 분석하여 Fig. 2를 제시 하였고 높은 상관관계를 나타내었다($p < 0.01$). 이 중 LTAS 는 slope, tilt, L/H ratio, HAM와 통계적으로 유의한 상관성은 확인되지 않았고, L/H ratio는 tilt를 제외한 모든 변수들과 유의한 상관관계는 없었다.

집단에 따른 측정변수들의 ROC 곡선 분석

정상집단과 기능적 음성장애 집단의 감별 진단적 특성을 확인하기 위해서 ROC 곡선 분석을 실시하여 민감도, 특이도, 곡선하 영역에 대한 결과를 Fig. 3에 제시 하였다. 곡선하 면적은 AVQIv2(0.933), AVQIv3(0.899), PraatCPPS(0.887), HAM(0.759), tilt(0.749), slope(0.731), LTAS(0.673), L/H ratio(0.605)로 우수한 진단능력을 보인 변수들을 확인하였다. 민감도와 특이도, 곡선하 면적을 확인하여 최적의 절사값을 결정하였다. 절사값은 각각 AVQIv2(3.270), AVQIv3(2.013), PraatCPPS(13.838),

고 찰

본 연구는 기능적 음성장애 환자들의 음성특징을 음향학적 분석을 통해서 확인하고 음성장애에 감별에 대한 절사값과 진단 예측력을 알아보고자 하였다. 본 연구의 결과, 기능적 음성장애 집단은 정상집단과 비교하여 큰 AVQI, slope, HAM, 작은 PraatCPPS, tilt, L/H ratio, LTAS 값을 나타내고 통계적으로 유의한 차이가 확인되었다. 그리고 L/H ratio 를 제외한 모든 변수들은 통계적으로 상관관계가 있는 것으로 분석되었다. ROC 곡선 분석에서, AVQI(v2, v3), PraatCPPS는 모두 0.89 이상의 AUC 로 우수한 진단 예측력을 보였다. 연구 결과를 통해서 다음과 같은 의미를 얻을 수 있었다.

첫째, 음향학적 측정변수들에 대한 두 집단의 결과를 비교한 결과, 두 집단 간에 통계적으로 유의한 차이가 확인되었다. 본 연구에서 적용된 AVQI의 두 가지 회귀식(v2, v3)은 선행 연구들에서 많이 사용되었다.¹⁷⁻¹⁹ 두

Table 3. Comparison of measured variables between normal and FD groups

Group	Normal (n=59)	FD (n=102)	p
AVQlv2	2.8±0.5	5.2±1.5	<0.01**
AVQlv3	1.2±0.5	3.9±1.9	<0.01**
PraatCPPS	14.7±1.0	11.1±2.7	<0.01**
Slope	-20.4±4.6	-24.2±4.0	<0.01**
Tilt	-11.9±1.8	-10.2±2.0	<0.01**
LH	1,235.1±1,492.6	699.1±653.2	<0.05*
HAM	22.8±5.5	28.3±5.6	<0.01**
LTAS	31.1±6.0	27.6±5.8	<0.01**
Grade	0.0±0.0	1.45±0.63	
0	59 (100.0%)	12 (11.8%)	
1	0 (0.0%)	37 (36.3%)	
1.5	0 (0.0%)	6 (5.9%)	
2	0 (0.0%)	40 (39.2%)	
2.5	0 (0.0%)	6 (5.9%)	
3	0 (0.0%)	1 (1.0%)	
OS	9.5±5.9	43.7±17.2	<0.01**

* : p<.05, ** : p<.01. FD : functional dysphonia, AVQlv2 : acoustic voice quality index version 2.02, AVQlv3 : acoustic voice quality index version 3.01, PraatCPPS : smoothed cepstral peak prominence in Praat, Slope : general slope of the long-term average spectrum, Tilt : tilt of the regression line through the long-term average spectrum, LH : low to high spectral band energies, HAM : Hammarberg index, LTAS : long-term average spectrum, OS : overall severity

회귀식은 포함된 변수들은 동일하지만 변수들의 가중치를 다르게 적용하여 음성장애의 심한 정도를 정량화하였다. 병리적인 음성은 큰 값의 AVQI를 산출하도록 고안되었다. AVQI v2, v3 모두 기능적 음성장애 집단이 정상 집단에 비해서 높은 값을 보였으며 통계적 유의성도 확인되었다. 선행연구들의 결과에서도 이러한 특성은 확인되었다.²⁰⁻²² 정상음성과 병리적 음성집단 전체에 대한 평균 AVQI는 초기 alfa, beta 버전에서는 3.83±1.63, 3.97±1.60으로 보고되었고, 두 버전 간의 상관관계는 0.980(p<0.0001)이었으며 이 때의 beta 버전은 AVQlv2가 되었다.²² AVQI의 내부 일관성 향상을 위해서 다양한 길이의 음성샘플을 분석, 비교하여 최적의 회귀식을 확인하였다. Voice duration(VD)-1(17음절+3초 모음), VD-2(무성음 구간 추출 후 3초에 상응하는 문장길이+3초 모음), VD-3(93개 음절을 포함하는 전체 문장+3초 모음)을 비교한 결과, VD-2에서 정상과 병리적 음성을

감별하는 최고의 정확도를 보였다. 이후 VD-2는 AVQlv3가 되었다.²³

PraatCPPS는 음성신호 내의 배음의 정도를 표현해주는 변수이기 때문에 병리적인 음성은 낮은 CPPS 값을 나타낸다. 선행연구에서 SpeechTool과 Praat으로 측정된 전체 대상자들의 CPPS 값은 각각 6.61±1.79, 11.66±2.68이었고, 두 측정방법 간의 상관관계는 0.961(p<0.0001)이었다.²² 본 연구에서도 병리적인 집단이 정상 집단에 비해서 낮은 CPPS 값으로 측정되어 선행연구의 결과와 일치하였다.^{24,25}

Slope는 1,000 Hz 이하의 평균 스펙트럼 강도와 1,000~10,000 Hz 의 평균 스펙트럼 강도의 차이를 의미한다. 기능적 음성장애 집단은 정상 집단에 비해 더 큰 차이를 보였고 선행연구에서도 유사한 결과(정상집단=-23.31±5.03, 음성장애집단=-24.86±4.99)를 확인하였다. Tilt는 1~10,000Hz 사이의 평균 스펙트럼에 대한 추세선 기울기를 의미하며 slope와는 달리 기능적 음성장애 집단이 정상 집단에 비해 더 작은 값을 나타내었고 선행연구의 결과(정상집단=-10.51±0.73, 음성장애집단=-9.45±1.38)도 일치하였다.²⁶ L/H ratio는 저주파수 대역과 고주파수 대역의 에너지 비율을 의미하며, 정상집단이 기능적 음성장애 집단보다 유의하게 높았다. 이것은 ADSV에서의 L/H spectral ratio의 결과와 일치한다.^{27,28} HAM은 선행연구에서 성악가들의 음역에 따라 다양한 값(-4.1~17.6)을 나타내었다. 또한 HAM은 하인두영역(hypopharyngeal area, HPA)와 0.54의 상관관계가 있었지만 통계적으로 유의하지는 않았다.¹² 넓은 HPA는 증가된 HAM을 산출할 수 있는 것으로 확인되었다. 모음, 발성 방식 등에 따라 달라질 수 있으며 추가적인 연구가 필요할 것으로 사료된다. 본 연구에서 LTAS는 기능적 음성장애 집단이 정상 집단에 비해 낮은 스펙트럼 강도를 보였다. 젊은 여성과 중년 여성의 음성을 비교한 선행 연구에서 젊은 여성이 중년 여성에 비해 더 큰 강도의 LTAS를 나타내었다.²⁹ 또한 정상 출생아동(term born)과 조산아동(extremely preterm, 25주 미만 출산) 아동의 음성을 비교한 연구에서 정상 출생아동이 조산아동보다 더 좋은 음성을 산출하는 것으로 확인되었다(term born : pre term ; G= 1:1.4, AVQI=3.0:5.0, CPP=7.5:5.5, slope=-16.5:-17.0, tilt=-10.6:-9.9). 이러한 결과는 본

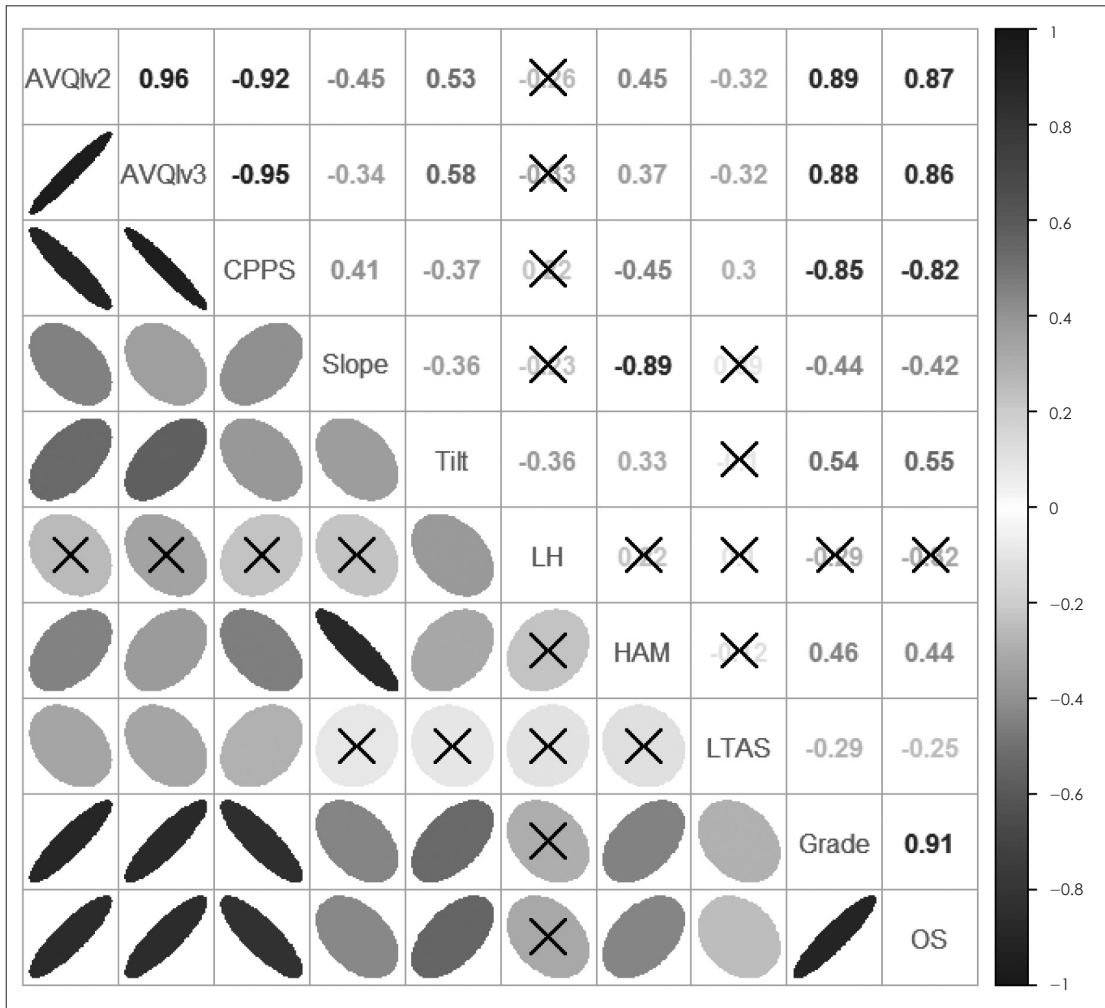


Fig. 2. Correlation plot among measured variables and p-value significance.

연구의 결과와 상응한다.³⁰⁾ 거친 음성(roughness)에 대한 음향학적 비교 결과, 정상 집단(CPPS 11.92, slope -24.78, tilt -9.62)은 거친 음성집단(CPPS 10.64, slope -24.65, tilt -9.39)과 차이를 나타내었다. 거친 음성에 대한 새로운 회귀모형(roughness model)을 만들어 GRBAS의 R 척도와 상관관계를 확인한 결과, 0.731(p=0.000)으로 높게 나타났다.³¹⁾

둘째, 음향학적 측정과 청지각적 평가 간의 상관관계는 대부분 유의한 것으로 나타났다. G는 AVQI(v2, v3)와 각각 0.88, 0.89의 상관관계가 있었고, 다른 언어권 연구에서도 0.78~0.91의 상관관계가 보고되었다. CPPS와

는 -0.85의 상관을 보였고 모음 -0.56, 연속구어 -0.73의 상관을 보인 선행연구를 뒷받침하였다.³²⁾ Slope, tilt, HAM, LTAS도 각각 -0.44, 0.54, 0.46, -0.29의 상관성이 확인되었다. 이 네 변수들은 LTAS로부터 측정 될 수 있고 이외에도 alpha ratio, H1-H2, spectral flatness, spectral centroid 등도 있다. 분석 기준이 되는 주파수 대역을 다르게 설정하여 음성의 특성을 파악하려는 연구가 진행되고 왔다.³³⁾ OS도 AVQI(v2, v3)와 0.87, 0.86의 상관을 보였고 0.88~0.92의 결과를 보인 선행연구와 일치하였다.³⁴⁻³⁸⁾ CPPS는 G, OS와 -0.85, -0.82의 상관관계가 확인되었고, Bae 등³⁹⁾의 연구에서 G는 Speech-

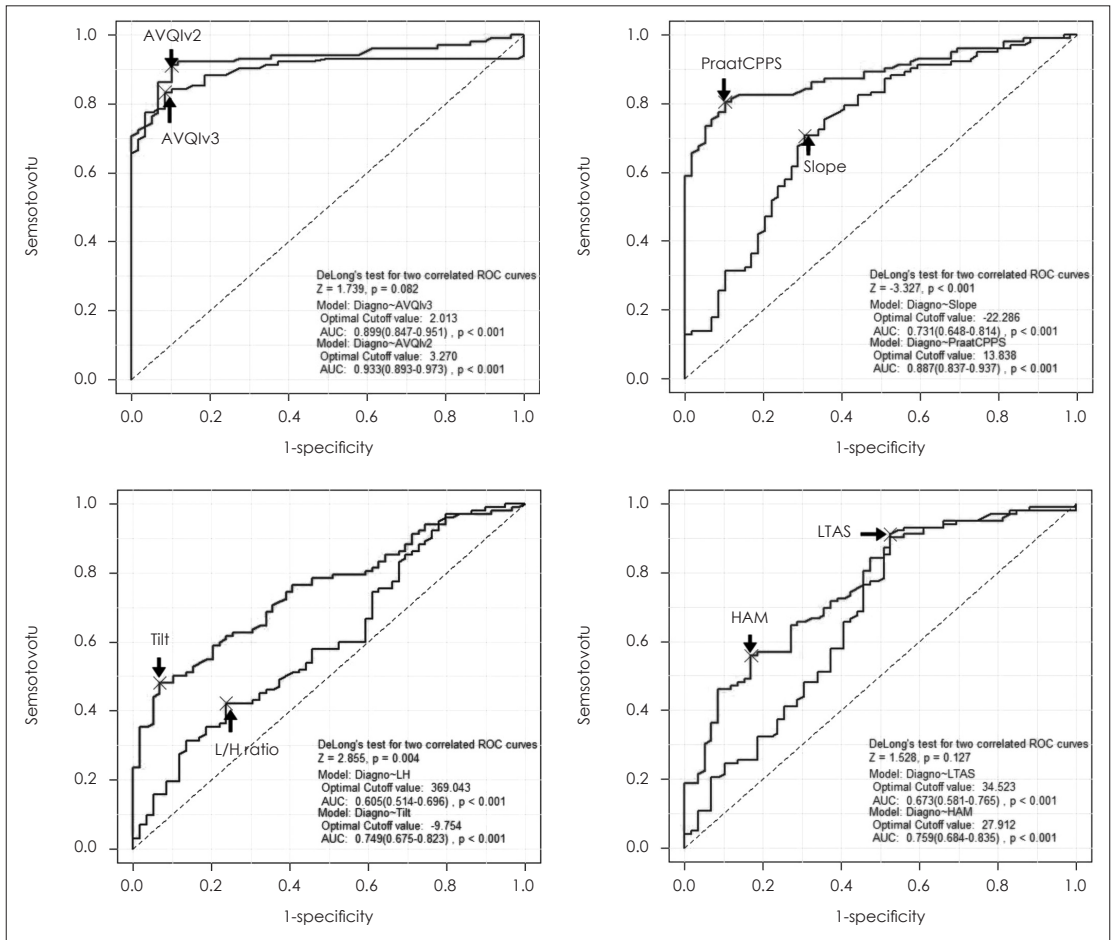


Fig. 3. ROC plot of the measured variables. The black arrow is the best AUC point.

Table 4. ROC curve analysis between normal and FD groups

Variables	Sensitivity	Specificity	Cutoff value	AUC
AVQlv2	91.2	89.8	3.270	0.933
AVQlv3	83.3	91.5	2.013	0.899
PraatCPPS	80.4	89.8	13.838	0.887
Slope	70.6	69.5	-22.286	0.731
Tilt	48.8	93.2	-9.754	0.749
LH	42.2	76.3	369.043	0.605
HAM	55.9	83.1	27.912	0.759
LTAS	91.2	47.5	34.523	0.673

AVQlv2 : acoustic voice quality index version 2.02, AVQlv3 acoustic voice quality index version 3.01, PraatCPPS : smoothed cepstral peak prominence in Praat, Slope : general slope of the long-term average spectrum, Tilt : tilt of the regression line through the long-term average spectrum, LH : low to high spectral band energies, HAM : Hammarberg index, LTAS : long-term average spectrum, AUC : area under the curve

Tool의 CPP, CPPS, ADSV의 CPP와 각각 -0.85, -0.72, -0.71, Awan 등⁴⁰⁾의 연구에서 OS는 CPP -0.68, CPPS -0.79의 상관관계가 보고되었다. 또한 음향학적 변수들 간의 상관관계에서, L/H ratio은 tilt와 유의미한 상관관계를 보였으며 나머지 변수들과는 유의하지 않았다. 선행 연구에서 AVQI는 CPPS($r=-0.71$), slope($r=0.01$), tilt ($r=0.48$)로 본 연구의 결과(CPPS, $r=-0.92$; slope, $r=-0.45$; tilt, $r=0.53$)와 유사하였다.²⁶⁾ 또한 3명의 성악가로부터 추출된 HAM과 slope의 상관관계는 -0.39으로 보고되었다.³³⁾

셋째, 측정변수들의 진단 예측력 및 절사값을 확인함으로써 효과적인 기능적 음성장애의 평가를 가능하게 하였다. 본 연구를 통해 확인된 결과들은 AVQI(v2, v3), PraatCPPS는 0.89 이상, slope, tilt, HAM은 0.73 이상, L/H ratio, LTAS는 0.61 이상의 예측력을 가졌다. 정상과 비정상 음성을 선별하는 절사값은 AVQIv2는 3.270, AVQIv3는 2.013, PraatCPPS는 13.838이며, 선행연구의 결과들과 유사하였다. 기존 연구들에서 AVQIv2의 절사값은 2.30~3.48, 진단 예측력은 0.906~0.970이었으며, v2 이전의 초기 버전의 AVQI는 2.70~3.66, 진단 예측력은 0.880~0.956으로 보고되었다.^{34-36,41-43)} AVQIv3의 절사값은 2.06~2.43, 진단예측력은 0.906~0.923이었다.^{19,44)}

캡스트럼 분석을 위한 도구는 CSL, ADSV, SpeechTool, Praat 등으로 다양하며 분석 알고리즘에 따라 CPPS의 값은 달라진다.⁴⁵⁻⁴⁷⁾ 성문암 환자 65명의 음성을 다양한 캡스트럼 분석도구로 측정된 결과, 모음과제에서 ADSV(5.87 ± 3.78), SpeechTool(4.44 ± 3.14), Praat(9.41 ± 4.02), 연속구어 과제에서 ADSV(3.87 ± 2.14), SpeechTool(3.04 ± 1.46), Praat(6.75 ± 2.18)으로 보고되었다.⁴⁵⁾ 다른 선행연구에서도 정상 음성집단의 CPPS 값은 ADSV 5.89 ± 1.00 , Praat 20.11 ± 1.27 , 병리적 음성장애 집단은 ADSV 4.15 ± 1.73 , Praat 17.49 ± 1.52 , 전체 평균 CPPS는 ADSV 4.87 ± 1.70 , Praat 18.58 ± 1.91 로 나타났다.⁴⁸⁾ 선행연구에서 Praat과 ADSV를 통한 CPPS 절사값은 각각 19.10, 5.53이며 0.910, 0.810의 진단 예측력으로 보고되었다.⁴⁸⁾ Heman-Ackah 등⁴⁷⁾의 연구에서 음성장애 감별을 위한 절사값은 CPPS 5.0이었고 이후 추가 연구에서는 Houde²⁸⁾의 알고리즘을 사용하여 CPPS 절사값 4.0, 민감도 0.924, 특이도 0.790, 진단 예측력 0.937을

보고하였다.⁴⁹⁾ 기존 연구들에서 Slope, tilt, HAM, LTAS에 대한 정상과 병리적 음성의 비교를 제시하였지만, 두 집단을 선별할 수 있는 절사값과 진단 예측력에 대해서는 보고하지 않았다. 본 연구에서는 기능적 음성장애 집단에 국한되었지만 정상 집단과 비교하여 감별할 수 있는 음향학적 기준을 제시하였다는 점에서 의미가 있다.

다양한 종류의 음성 장애 유형과 심각성을 비교할 때 모음연장보다는 연속구어가 좀 더 강한 예측력을 나타낸다고 보고하였다. 하지만 일부 환자군에서는 모음연장에서 음성장애가 더욱 두드러지기 때문에 모음연장과 연속구어를 모두 포함하는 음성샘플을 제작하여 분석해야 한다고 주장하였다.⁴⁰⁾ CAPE-V의 6개 문장은 음성장애의 특징을 찾아내기에 효과적이라고 하였지만,⁵⁰⁾ 기존 CAPE-V 문장에 대응하는 한국어판 문장이 표준화는 아직 완료되지 않았다. 6개 문장의 표준화가 이루어진다면 기존 과업 문장들과의 차이점을 확인해야 할 것이다.

본 연구의 결과를 종합하여 볼 때, 기능적 음성장애를 예측하기 위해 음향학적 분석 및 청지각적 평가는 도움이 된다고 할 수 있다. 추후 연구에서는 음향학적 측정, 청지각적 평가 외에도 후두 스트로보스코피 등과 같은 성대 진동의 정량화와 공기역학 분석 관련 요소 등도 함께 고려하여야 할 것이다.

결론

본 연구에서는 기능적 음성장애 집단을 대상으로 음향학적 측정, 청지각적 평가를 실시하였고 음성장애를 선별하는데 이러한 평가들이 유용함을 확인하였다. AVQI, CPPS가 정상과 기능적 음성장애를 구별하는데 높은 예측력을 보였다. 또한 음향학적 측정과 청지각적 평가는 높은 상관관계를 보였으며 임상현장에서 기능적 음성장애의 진단 및 평가에 유용한 도구로써 도움이 될 것으로 사료된다. 추가 연구를 통해서 기질적, 신경학적 음성장애에 대해서도 검증해야 할 것이다.

중심 단어 : 기능적 음성장애 · 음향학 · Praat · GRBAS · CAPE-V.

REFERENCES

- 1) Koufman J, Blalock P. Functional voice disorders. *Otolaryngologic Clinics of North America* 1991;24(5):1059-73.
- 2) Ramig LO, Verdolini K. Treatment efficacy: voice disorders. *Journal of Speech Language and Hearing Research* 1998;41(1):S101-S116.
- 3) Roy N, Leeper HA. Effects of the manual laryngeal musculoskeletal tension reduction technique as a treatment for functional voice disorders: perceptual and acoustic measures. *J Voice* 1993;7(3):242-9.
- 4) Roy N, Bless DM, Heisey D, Ford CN. Manual circumlaryngeal therapy for functional dysphonia: an evaluation of short-and long-term treatment outcomes. *J Voice* 1997;11(3): 321-31.
- 5) Sama A, Carding PN, Price S, Kelly P, Wilson JA. The clinical features of functional dysphonia. *Laryngoscope* 2001; 111(3):458-63.
- 6) Carding PN, Horsley IA, Docherty GJ. A study of the effectiveness of voice therapy in the treatment of 45 patients with nonorganic dysphonia. *J Voice* 1999;13(1):72-104.
- 7) Fex B, Fex S, Shiromoto O, Hirano M. Acoustic analysis of functional dysphonia: Before and after voice therapy (accent method). *J Voice* 1994;8(2):163-7.
- 8) Kollbrunner J, Seifert E. Encouragement to Increase the Use of Psychosocial Skills in the Diagnosis and Therapy of Patients With Functional Dysphonia. *J Voice* 2017;31(1): 132.e1-132.e7.
- 9) Roy N. Functional dysphonia. *Current Opinion in Otolaryngology & Head and Neck Surgery* 2003;11(3):144-8.
- 10) Deliyiski DD, Shaw HS, Evans MK. Adverse Effects of Environmental Noise on Acoustic Voice Quality Measurements. *J Voice* 2005;19(1):15-28.
- 11) Hammarberg B, Fritzell B, Gauffin J, Sundberg J, Wedin L. Perceptual and Acoustic Correlates of Abnormal Voice Qualities. *Acta Oto-Laryngologica* 1980;90(5-6):441-51.
- 12) Mainka A, Poznyakovskiy A, Platzek I, Fleischer M, Sundberg J, Murbe D. Lower Vocal Tract Morphologic Adjustments Are Relevant for Voice Timbre in Singing. *PLoS One* 2015;10(7):e0132241.
- 13) Monzo C, Alias F, Iriondo I, Gonzalvo X, Planet S. Discriminating expressive speech styles by voice quality parameterization. *Proc. of ICPhS*;2007.
- 14) Tanner K, Roy N, Ash A, Buder EH. Spectral Moments of the Long-term Average Spectrum: Sensitive Indices of Voice Change After Therapy? *J Voice* 2005;19(2):211-22.
- 15) Eadie TL, Doyle PC. Classification of Dysphonic Voice: Acoustic and Auditory-Perceptual Measures. *J Voice* 2005; 19(1):1-14.
- 16) Laukkanen A-M, Björkner E, Sundberg J. Throaty Voice Quality: Subglottal Pressure, Voice Source, and Formant Characteristics. *J Voice* 2006;20(1):25-37.
- 17) Uloza V, van Latsoszek B, Ulozaitė-Staniene N, Petrauskas T, Maryn Y. A comparison of Dysphonia Severity Index and Acoustic Voice Quality Index measures in differentiating normal and dysphonic voices. *European Archives of Otorhino-Laryngology* 2018;275(4):949-58.
- 18) Kim GH, Lee YW, Bae IH, Park HJ, Wang SG, Kwon SB. Validation of the Acoustic Voice Quality Index in the Korean Language. *J Voice* In press 2018.
- 19) Barsties B, Maryn Y. External Validation of the Acoustic Voice Quality Index Version 03.01 With Extended Representativity. *Annals of Otolaryngology and Laryngology* 2016;125(7):571-83.
- 20) Kim G-H, Lee Y-Y, Lee B-J, Kwon S-B. Acoustic and Auditory-Perceptual Evaluation as Predictor of Voice Recovery after Laryngeal Microsurgery in Patients with Vocal Polyp. *Korean Journal of Otorhinolaryngology-Head and Neck Surgery* 2018;61(7):361-9.
- 21) Hernandez JD, Gomez NML, Jimenez A, Izquierdo LM, vander Latsoszek B. Validation of the Acoustic Voice Quality Index Version 03.01 and the Acoustic Breathiness Index in the Spanish language. *Annals of Otolaryngology and Laryngology* 2018;127(5):317-26.
- 22) Maryn Y, Weenink D. Objective Dysphonia Measures in the Program Praat: Smoothed Cepstral Peak Prominence and Acoustic Voice Quality Index. *J Voice* 2015;29(1):35-43.
- 23) Barsties B, Maryn Y. The improvement of internal consistency of the Acoustic Voice Quality Index. *American Journal of Otolaryngology* 2015;36(5):647-56.
- 24) Brinca LF, Batista APF, Tavares AI, Goncalves IC, Moreno ML. Use of Cepstral Analyses for Differentiating Normal From Dysphonic Voices: A Comparative Study of Connected Speech Versus Sustained Vowel in European Portuguese Female Speakers. *J Voice* 2014;28(3):282-6.
- 25) Watts CR, Awan SN. Use of Spectral/Cepstral Analyses for Differentiating Normal From Hypofunctional Voices in Sustained Vowel and Continuous Speech Contexts. *Journal of Speech Language and Hearing Research* 2011;54(6):1525-37.
- 26) Maryn Y, Corthals P, Van Cauwenberge P, Roy N, De Bodt M. Toward Improved Ecological Validity in the Acoustic Measurement of Overall Voice Quality: Combining Continuous Speech and Sustained Vowels. *J Voice* 2010;24(5): 540-55.
- 27) Awan SN, Roy N. Outcomes Measurement in Voice Disorders: Application of an Acoustic Index of Dysphonia Severity. *Journal of Speech Language and Hearing Research* 2009;52(2):482-99.
- 28) Hillenbrand J, Houde RA. Acoustic correlates of breathy vocal quality: dysphonic voices and continuous speech. *Journal of Speech Language and Hearing Research* 1996;39(2): 311-21.
- 29) da Silva PT, Master S, Andreoni S, Pontes P, Ramos LR. Acoustic and long-term average spectrum measures to detect vocal aging in women. *J Voice* 2011;25(4):411-9.
- 30) Reynolds V, Buckland A, Bailey J, Lipscombe J, Nathan E, Vijayasekaran S, et al. Objective assessment of pediatric voice disorders with the acoustic voice quality index. *J Voice* 2012;26(5):672-e1.
- 31) Barsties B, De Bodt M, Gerrits E, Maryn Y. The Exploration of an Objective Model for Roughness With Several

- Acoustic Markers. *J Voice* 2018;32(2):149-61.
- 32) Delgado-Hernandez J, Leon-Gomez NM, Izquierdo-Arteaga LM, Llanos-Fumero Y. Cepstral analysis of normal and pathological voice in Spanish adults. Smoothed cepstral peak prominence in sustained vowels versus connected speech. *Acta Otorrinolaringologica Espanola* 2018;69(3):134-40.
 - 33) Scherer KR, Sundberg J, Tamarit L, Salomao GL. Comparing the acoustic expression of emotion in the speaking and the singing voice. *Computer Speech and Language* 2015; 29(1):218-35.
 - 34) Uloza V, Petrauskas T, Padervinskis E, Ulozaite N, Barsties B, Maryn Y. Validation of the Acoustic Voice Quality Index in the Lithuanian Language. *J Voice* 2017;31(2):257.e1-257.e11.
 - 35) Hosokawa K, Barsties B, Iwahashi T, Iwahashi M, Kato C, Iwaki S, et al. Validation of the Acoustic Voice Quality Index in the Japanese Language. *J Voice* 2017;31(2):260.e1-260.e9.
 - 36) Maryn Y, De Bodt M, Barsties B, Roy N. The value of the Acoustic Voice Quality Index as a measure of dysphonia severity in subjects speaking different languages. *European Archives of Oto-Rhino-Laryngology* 2014;271(6):1609-19.
 - 37) Barsties B, Maryn Y. The Acoustic Voice Quality Index. Toward expanded measurement of dysphonia severity in German subjects. *HNO* 2012;60(8):715-20.
 - 38) Nemr K, Simoes-Zenari M, Cordeiro GF, Tsuji D, Ogawa AI, Ubrig MT, et al. GRBAS and Cape-V Scales: High Reliability and Consensus When Applied at Different Times. *J Voice* 2012;26(6):812.e17-22.
 - 39) Bae IH, Kim GH, Lee YW, Park HJ, Wang SG, Kwon SB. Clinical Application of Cepstral Peak Prominence for Treatment Outcomes Assessment in Voice Disorders: Comparing ADSV, Speechtool, and PNU_CPP. *Journal of Speech-Language & Hearing Disorders* 2016;25(2):93-102.
 - 40) Awan SN, Helou LB, Stojadinovic A, Solomon NP. Tracking voice change after thyroidectomy: application of spectral/cepstral analyses. *Clinical Linguistics & Phonetics* 2011; 25(4):302-20.
 - 41) Barsties VLB, Ulozaite-Staniene N, Maryn Y, Petrauskas T, Uloza V. The Influence of Gender and Age on the Acoustic Voice Quality Index and Dysphonia Severity Index: A Normative Study. *J Voice* In press 2017.
 - 42) Maryn Y, Kim HT, Kim J. Auditory-Perceptual and Acoustic Methods in Measuring Dysphonia Severity of Korean Speech. *J Voice* 2016;30(5):587-94.
 - 43) Maryn Y, De Bodt M, Roy N. The Acoustic Voice Quality Index: toward improved treatment outcomes assessment in voice disorders. *Journal of Communication Disorders* 2010; 43(3):161-74.
 - 44) Hosokawa K, Barsties VLB, Iwahashi T, Iwahashi M, Iwaki S, Kato C, et al. The Acoustic Voice Quality Index Version 03.01 for the Japanese-speaking Population. *J Voice* In press 2017.
 - 45) Kim GH, Lee YW, Bae IH, Park HJ, Kwon SB. A Study of Cepstral Peak Prominence Characteristics in ADSV, SpeechTool and Praat. *Journal of Speech-Language and Hearing Disorders* 2017;26(3):99-111.
 - 46) Lowell SY, Colton RH, Kelley RT, Mizia Sa. Predictive value and discriminant capacity of cepstral- and spectral-based measures during continuous speech. *J Voice* 2013;27(4): 393-400.
 - 47) Heman-Ackah YD, Michael DD, Baroody MM, Ostrowski R, Hillenbrand J, Heuer RJ, et al. Cepstral peak prominence: a more reliable measure of dysphonia. *Annals of Otolaryngology & Laryngology* 2003;112(4):324-33.
 - 48) Sauder C, Bretl M, Eadie T. Predicting Voice Disorder Status From Smoothed Measures of Cepstral Peak Prominence Using Praat and Analysis of Dysphonia in Speech and Voice (ADSV). *J Voice* 2017;31(5):557-66.
 - 49) Heman-Ackah YD, Sataloff RT, Laureyns G, Lurie D, Michael DD, Heuer R, et al. Quantifying the cepstral peak prominence, a measure of dysphonia. *J Voice* 2014;28(6): 783-8.
 - 50) Awan SN, Roy N, Jetté ME, Meltzner GS, Hillman RE. Quantifying dysphonia severity using a spectral/cepstral-based acoustic index: comparisons with auditory-perceptual judgements from the CAPE-V. *Clinical Linguistics & Phonetics* 2010;24(9):742-58.