

# Predicting Normal and Pathological Voice using a Cepstral Based Acoustic Index in Sustained Vowels versus Connected Speech

Miok Yu<sup>a</sup>, Seong Hee Choi<sup>a,b</sup>, Chul-Hee Choi<sup>a,b</sup>, Byeongheun Choi<sup>c</sup>

<sup>a</sup>Department of Audiology & Speech-Language Pathology, Graduate School of Daegu Catholic University, Gyeongsan, Korea

<sup>b</sup>Department of Audiology & Speech-Language Pathology, Daegu Catholic University, Gyeongsan, Korea

<sup>c</sup>Myung ENT Clinic, Daegu, Korea

**Correspondence:** Seong Hee Choi, PhD  
Department of Audiology & Speech-Language Pathology, Institute of Biomimetic Sensory Control & Catholic Hearing Voice Speech Center, Daegu Catholic University, 13-13 Hayang-ro, Hayang-eup, Gyeongsan 38430, Korea  
Tel: +82-53-850-2542  
Fax: +82-53-850-2540  
E-mail: shgrace@cu.ac.kr

Received: October 5, 2018  
Revised: November 20, 2018  
Accepted: November 20, 2018

**Objectives:** In recent years, the use of cepstral measures for acoustic evaluation of voice has increased. The objective of this study is to evaluate the diagnostic value of spectral/cepstral measures to differentiate dysphonia from normal voice and to determine what type of voice sample (sustained vowel /a/ or connected speech) is the most sensitive in differentiating normal and pathological voice. **Methods:** Two hundred and eighty-eight individuals (99 men, 189 women) from 214 dysphonia patients and 74 normal speakers recorded connected speech and a sustained vowel /a/. One laryngologist and two speech-language pathologists performed visual and auditory-perceptual rating of voice samples in terms of the degree of dysphonia/normality. Recorded voices were analyzed with two spectral/cepstral measures. The cutoff value for positivity that has the highest specificity for discriminating between normal and dysphonia voices was determined based on receiver operating characteristic (ROC) analyses. **Results:** Measures of cepstral peak prominence (CPP) and ratio of low- to high-frequency spectral energies (L/H ratio) were significantly different between groups in both speech conditions. ROC analysis demonstrated CPP had high sensitivity and specificity for the classification of dysphonia versus controls in the both speech conditions (area under curve [AUC] = .815 in vowel, AUC = .91 in connected speech); and CPP, in particular, showed higher discrimination accuracy. **Conclusion:** CPP is a good predictable acoustic measure to detect dysphonic speakers in both vowel prolongation and connected speech from normal voice. Therefore, this study suggested cepstral-based acoustic measures should be included for clinical evaluation of dysphonia.

**Keywords:** Cepstral analysis, Dysphonia, Voice sample, Diagnostic predictability

음성언어재활사는 대상자의 음성평가 시 청지각적 평가와 더불어 음향의 파형 분석에서 얻는 객관적 측정치를 포함하는 도구적 평가를 함께 시행하여 타당하고 신뢰로운 평가를 하기 위해 노력한다. 음성장애 환자의 음질을 평가하기 위해 사용하는 음성표본은 모음연장발성이나 연결발화를 사용하는데, 일반적으로 음향학적 분석과 지각적 검사에서 주로 쓰이고 있다(Choi, 2013). 모음연장발성은 음성의 지각적인 판단에 영향을 주는 개인의 발화 특성과 문맥이나 강세 같은 요인의 변동을 줄여주기 때문에(Parsa & Jamieson, 2001) 더욱 적절하게 평가할 수 있다고 기대된다. 하지만

Revis 등(1999)은 연결발화에 비해 모음연장발성에서 음성장애 환자가 힘을 주어 성대 내전을 높이므로 단시간에 더 좋은 음성을 산출할 수 있어 음질의 비정상성이 실제 대화보다 경미하게 측정된다고 하였다. 이에 반해 연결발화의 경우 모음연장발성과는 달리 일상의 발화 패턴을 반영하는 환자 음질 특성을 포함하므로 보다 실제적인 음성 평가가 이루어질 수 있다.

음향 분석은 객관적인 음질 기준을 제공하지만, 현재 임상현장에서 주로 사용하는 음향 측정치 중 많은 수는 상당한 한계를 가진다. 변동을 분석과 같은 전통적 음향 분석의 경우, 환자의 음질에

대한 중요한 정보를 제공하더라도 일반적으로 지속적인 음성을 분석할 때만 그 값이 유효하므로 화자의 일상적인 음성 산출 특성을 반영하는 연결발화를 분석하지 못해 그 신뢰성이 저하될 수 있다. 또한 전통적 시간 기반 측정은 시작하거나 끝나는 주기의 경계 또는 주기 내 최고 진폭과 같은 특성을 분석하는 데 한계가 있다. 주기를 만들어 낼 수 없을 정도로 심각한 음질인 불안정한 음성 신호에서 그 유효성이 의심스럽게 되는 것이다(Roark, 2006). 반면에 스펙트럼 및 캡스트럼 분석은 음성 신호 내에서 변동률의 정도를 측정하기 위한 메커니즘인 기본주파수 추적에 의존하지 않으므로 매우 비주기적인 음성 신호에서도 유효하게 적용된다(Awan, Roy, & Dromey, 2009; Choi & Choi, 2014, 2016; Yu, Choi, Choi, & Lee, 2017). 또한 스펙트럼 기반 측정은 음성신호에 대하여 일정한 음도 및 강도를 가정하지 않으므로 연결발화도 분석이 가능하다.

캡스트럼 분석은 음파의 스펙트럼에서 기본주파수를 추출하는 절차로 Noll (1964)에 의해 고안되었으나 과학적 제한으로 인해 최근에는 임상적으로 이용 가능하게 되었다. 캡스트럼은 로그파워스펙트럼의 역푸리에 변환으로 얻어지는 데, 배경소음 레벨에서 벗어나는 정도를 그래픽으로 표시하는데 사용할 수 있다(Baken & Orlikoff, 2000). 캡스트럼의 주요 분석치인 캡스트럼최고정점(cepstral peak prominence, CPP)은 캡스트럼정점(cepstral peak)과 캡스트럼의 소리에너지 평균을 나타내는 선형회귀직선(linear regression)과의 사이 거리를 의미한다. 주기적 신호에서 조화음의 구조가 잘 나타나는 정상음성의 경우에는 높은 캡스트럼최고정점(CPP) 값을 나타내며, 반대로 음질 문제가 심각한 경우 낮은 측정치를 보인다. 저주파수대고주파수 스펙트럼 에너지비율(ratio of low to high-frequency spectral energies, L/H ratio)은 4 kHz 주파수 영역을 기점으로 스펙트럼상에서 위아래 평균 에너지비율을 계산하는데, 정상음성은 고주파수에 비해 저주파수 영역 에너지가 많으므로 저주파수대고주파수 스펙트럼 에너지비율(L/H ratio) 값이 더 크게 나타난다(Shin, 2014; Watts & Awan, 2011).

최근 국내외에서 스펙트럼 및 캡스트럼 기반 음향학적 분석에 대한 연구에서 캡스트럼 측정치가 모음연장발성 및 연결발화 음성 표본에서 장애음성의 중증도를 강력하게 예측하는 것으로 나타났다(Awan et al., 2010; Awan & Roy, 2005; Choi & Choi, 2014, 2016; Heman-Ackah, Michael, & Goding, 2002; Maryn, Roy, De Bodt, Van Cauwenberge, & Corthals, 2009; Peterson et al., 2013; Seo & Seong, 2013; Yu et al., 2017). 또한 캡스트럼 분석 결과는 음성의 청지각적 평가와 높은 상관성을 보일 뿐만 아니라(Awan et al., 2010; Choi & Choi, 2014, 2016; Hasanvand, Salehi, & Ebrahimipour, 2017; Heman-Ackah et al., 2002; Yu et al., 2017), 음성장애지수(Voice

Handicap Index, VHI)와 같은 주관적 자가평가와도 상관성을 보였다(Awan, Roy, & Cohen, 2014).

그러나 현재까지 캡스트럼에 대한 연구는 정상음성과 장애음성의 진단적 결정을 위한 기준값을 설정하기에 충분한 음성표본을 사용한 연구가 부족하며 국내의 연구는 대부분이 모음으로 음성표본을 제한하여(Choi & Choi, 2014) 음성장애 및 중증도의 확실한 예측인자인 캡스트럼 측정치가 어떤 음성표본에서 더욱 민감성을 보이는지에 대한 충분한 근거가 제시되어 있지 않다. 따라서 본 연구에서는 모음연장발성 및 연결발화의 세분화된 음성표본과 음성장애 유무 집단에 따른 스펙트럼 및 캡스트럼의 측정치에 대해 논하고 이들 음향학적 측정치가 음성장애의 유무를 평가하는 지표로서 객관적이고 신뢰로운지 알아보고자 한다. 나아가 음성장애의 진단을 위한 스펙트럼과 캡스트럼의 기준값과 그 값의 민감도 및 특이도를 정의하고, 스펙트럼과 캡스트럼을 분석하기 위해 사용한 두 가지 음성표본을 비교하여 음성장애를 감별하는 데 가장 민감한 표본이 무엇인지 살펴보고자 한다.

## 연구방법

### 연구대상

본 연구의 참여대상자는 2013년 1월부터 2018년 6월까지 대구 소재 이비인후과에 목소리 문제를 주소로 내원한 환자 214명과 최근 3개월 내 상기도감염이나 후두 질환이 없었고 음성 문제를 보고하지 않은 일반인 실험지원자 74명이었다. 대상자는 모두 1명의 이비인후과 전문의와 2명의 음성언어재활사에게 후두내시경(스트로보스코피 또는 하이스피드 이미징)과 음향학적 평가 및 청지각적 평가를 시행받았고, 한국판 음성장애지수(Korean version of Voice Handicap Index, K-VHI) 설문지를 작성하였다. 실험에 참여한 대상자에 대한 음성장애 유무에 대한 판단은 후두내시경과 음향학적 평가 및 청지각적 평가를 기준으로 이비인후과 전문의에 의해 후두 내시경으로 확인된 의학적 진단을 받고 음성언어재활사의 청지각적 음성 평가에 의해 음성에 문제가 있다고 판단하였을 경우로 하였다. 대상자의 평균 연령은  $46.56 \pm 18.87$ 세였으며 남자 99명 여자 189명이었다. 대상자에 대한 정보는 Tables 1과 2에 나타내었다.

### 말과제 및 음성자료 수집

음성 녹음은 소음이 40 dB 이하로 통제된 이비인후과 음성검사실과 대학교 음성검사실에서 이루어졌다. 녹음은 90° 각도로 고정된 Shure SM48-LC 다이내믹 마이크를 이용하여 Computerized Speech Lab (CSL; KayPENTAX Inc., Lincoln Park, NJ, USA)으로

**Table 1.** Demographic information of the normal and pathological voice samples (N = 288)

	Normal voice		Pathological voice	
	Male	Female	Male	Female
Number of samples	27	47	72	142
Age (yr)	31.160 ± 13.088 (21-59)	38.214 ± 16.662 (21-67)	52.254 ± 20.47 (11-87)	48.908 ± 17.16 (15-87)

Values are presented as mean ± SD (range).

**Table 2.** Disease of the pathological voice samples (N = 214)

Disease	Male	Female
Nodule	14	71
Polyp	7	3
Cyst	2	1
Vocal Palsy	6	13
Sulcus	21	15
Scar	12	4
Edema	1	1
MTD	3	19
Etc.	6	15

하였으며 음성표본추출률(sampling rate)은 44,100 Hz, 16 bit 양자화로 하였다.

모음연장발성 과정은 편안하고 자연스러운 음도와 강도로 모음 /아/를 3초간 발성하도록 하였고 연속구어 과정은 본인의 이름을 넣어 “안녕하세요. 저는 OOO입니다”를 1회 말하도록 하였다.

### 분석 방법

#### 청지각적 평가

청지각적 평가를 위하여 10년 이상 음성장애의 평가 및 치료에 경험이 있는 2명의 1급 언어재활사(평균연수 11년)가 실시하였다. 모음연장발성과 연결발화의 두 가지 음성표본은 실험대상자의 정보 없이 WAV 형식으로 각각의 평가자에게 제공되었고 평가자는 대상자의 음성을 듣고 음성의 전반적인 신 정도(rough, hoarse)만으로 장애음성(+)/정상음성(-)으로 평가하였다. 검사자 간 신뢰도는 92.014% (265/288)였으며, 판정이 일치하지 않는 음성표본의 경우에는 제1저자와 교신저자가 함께 음성을 듣고 동의한 최종 측정치를 사용하였다. 검사자 내 신뢰도는 제1저자와 교신저자가 음성표본의 20%를 무작위로 선정하여 재검사하였고 이를 음성장애 평가 및 치료에 10년 이상 경험이 있는 1급 언어재활사 1인의 검사 결과와 비교하였다. 청지각적 평가를 통하여 일반인 실험지원자의 음성 중 장애음성으로 평가되거나 후두내시경 검사에서 양성(+)으로 판단된 7예와 음성 문제를 주소로 병리적인 음성으로 수집된 음성 중 후두내시경 소견상 의학적 진단이 없고 음성언어재활사에 의해

정상음성으로 판정된 2예는 분석에서 제외하였다. 총 288예의 표본 음성 중 279예가 분석에 사용되었다.

#### 음향학적 평가

녹음된 음성자료는 Analysis of Dysphonia in Speech and Voice (ADSV, Model 5109; KayPENTAX Inc.)의 프로그램을 사용하여 캡스트럼최고정점(CPP), 저주파수대고주파수 스펙트럼 에너지비율(L/H ratio)을 측정하였다. 모음연장발성의 분석은 /아/모음을 3초간 발성한 음성표본의 처음과 끝부분의 50 ms를 제외한 구간 중 2초를 분석하였다. 연결발화는 본인의 이름 말하기 문장 “안녕하세요. 저는 OOO입니다”에서 분석 결과에 영향을 미칠 수 있는 대상자의 이름 부분과 200 ms 이상의 쉬 구간을 분석에서 제외하였다.

#### 통계 분석

음성장애 집단과 정상집단의 모음연장발성과 음성연결발화 음성표본의 스펙트럼 및 캡스트럼 분석치는 기술 통계를 이용하여 평균과 표준편차를 산출하였다. 음성장애 유무 집단과 음성표본에 따라 스펙트럼 및 캡스트럼 측정치의 차이가 있는지 비교하기 위해 개체 간 독립변수(집단-음성장애군 vs. 대조군)와 개체 내 독립변수(음성표본-모음연장발성 vs. 연결발화)로 하는 이원혼합분산분석(two-way mixed ANOVA)을 실시하였다. 유무 집단과 음성표본 간 상호작용이 나타나는 경우는 각 음성표본에서 음성장애 유무 집단 간 스펙트럼 및 캡스트럼 측정치의 차이가 있는지 알아보기 위해 독립표본 t-검정을 실시하였다. 유의수준은 Bonferroni의 방법으로 교정하여 .025로 하였다. 더불어 캡스트럼 측정치의 음성장애 유무 예측 가능성을 알아보기 위해 수신자 조작 특성 곡선 분석(Receiver-Operating Characteristic analysis)을 실시하였고, Youden의 Index를 이용하여 각 측정치의 민감도와 특이도, 그리고 음성장애를 판단하는 기준값(cutoff value)을 확인하였다. 또한 음성에 대한 객관적인 음향학적 분석치와 화자의 주관적인 음성장애지수 간의 상관성을 살펴보기 위해 적률상관관계(Pearson correlation analysis)분석을 실시하였다. 통계 프로그램으로는 SPSS version 21.0 (IBM Co., New York, NY, USA)을 사용하였다.

연구결과

음성장애 유무 집단 간 음성표본에 따른 캡스트럼 및 스펙트럼 측정치 차이

ADSV 프로그램을 이용하여 분석한 모음연장발성과 연결발화에 대한 음성장애군과 대조군의 캡스트럼 분석치의 평균과 표준편차는 Table 3과 Figures 1, 2와 같다. 캡스트럼최고정점(CPP)은 모음연장발성에서 음성장애군(9.024 ± 3.898 dB)이 정상대조군(12.994 ± 1.820 dB)보다 낮았고, 연결발화에서도 음성장애군(5.076 ± 1.867 dB)이 정상대조군(7.605 ± 1.031 dB)보다 낮았다. 저주파수대 고주파수 스펙트럼 비율(L/H ratio)은 모음연장발성에서 음성장애군(28.405 ± 6.176 dB)이 정상대조군(31.266 ± 5.319 dB)보다 더 낮은 값을 보였고, 연결발화에서도 음성장애군(27.022 ± 4.105 dB)이 정상대조군(29.831 ± 3.369 dB)보다 더 낮은 값을 보였다.

음성장애 유무 집단과 음성표본에 따라 캡스트럼최고정점의 차

**Table 3.** Cepstral & spectral measures between normal versus dysphonia group in different voice samples

	Control group	Dysphonic group	t
Sustained vowel /a/			
CPP (dB)	12.994 ± 1.820	9.024 ± 3.898	11.409***
L/H ratio (dB)	31.266 ± 5.319	28.405 ± 6.176	3.413**
Connected speech			
CPP (dB)	7.605 ± 1.031	5.076 ± 1.867	14.073***
L/H ratio (dB)	29.831 ± 3.369	27.022 ± 4.105	5.085***

Values are presented as mean ± SD.

CPP = cepstral peak prominence; L/H ratio = low-high spectral ratio.

\*\*p < .01, \*\*\*p < .001.

이가 있는지 알아보기 위하여 이원혼합분산분석을 시행한 결과, 음성장애 유무 집단과 음성표본에 따라 캡스트럼최고정점은 유의한 주효과가 있었으며( $F_{(1, 277)} = 768.117, p < .001$ ), 상호작용 또한 유의하였다( $F_{(1, 277)} = 18.276, p < .001$ ). 따라서, 각 음성표본에 따라 집단 간 독립표본 t-검정을 실시한 후 Bonferroni의 방법으로 교정하여 살펴본 결과, 캡스트럼최고정점은 정상집단과 음성장애환자 집단 간 모음연장발성에서 유의한 차이를 보였다( $p < .001$ ). 이와 마찬가지로, 연결발화에서도 두 집단 간 스펙트럼과 캡스트럼 측정치에 유의한 차이를 보였다( $p < .001$ ) (Table 3).

저주파수대 고주파수 스펙트럼 비율(L/H ratio)은 음성장애 유무 집단과 음성표본에 따라 상호작용은 없었으나, 통계적으로 유의한 음성표본의 주효과가 있었으며( $F_{(1, 277)} = 20.245, p < .001$ ), 집단 간 주효과가 있었다( $F_{(1, 277)} = 19.807, p < .001$ ). 즉, 저주파수대 고주

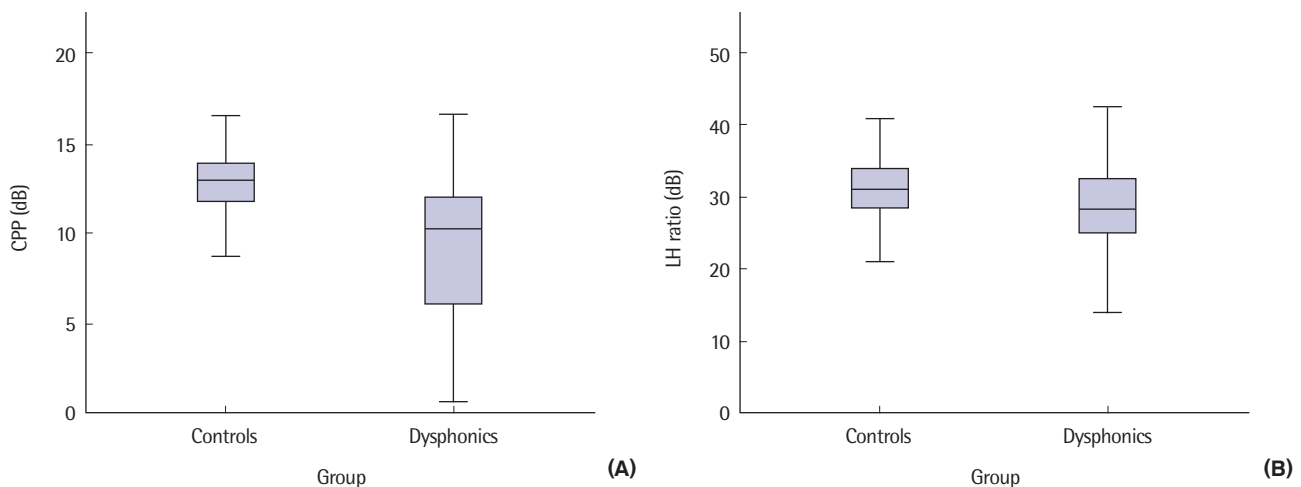
**Table 4.** Results of CPP and L/H ratio as computed from the ROC analysis in sustained vowels and connected speech

Acoustic variable	Cutoff value	Sensitivity (%)	Specificity (%)	AUC (SEM)	95% CI
Sustained vowel /a/					
CPP	12.088	70.6	75.8	.815*** (.026)	.763-.866
L/H ratio	29.241	69.1	57.8	.637** (.037)	.565-.710
Connected Speech					
CPP	7.076	80.9	90.0	.910*** (.020)	.871-.949
L/H ratio	28.284	66.2	58.8	.693*** (.035)	.652-.761

Cutoff values were determined via Youden's Index.

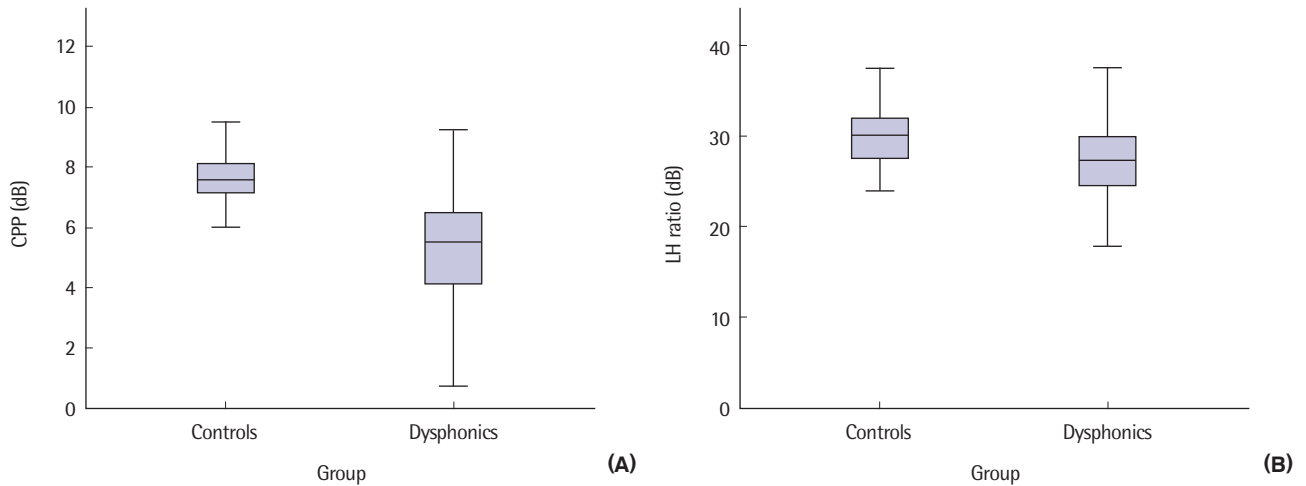
CPP = cepstral peak prominence; L/H ratio = low-high spectral ratio; ROC = receiver operating characteristic; AUC = area under the curve; SEM = standard error of measurement; CI = confidence interval.

\*\*p < .01, \*\*\*p < .001.

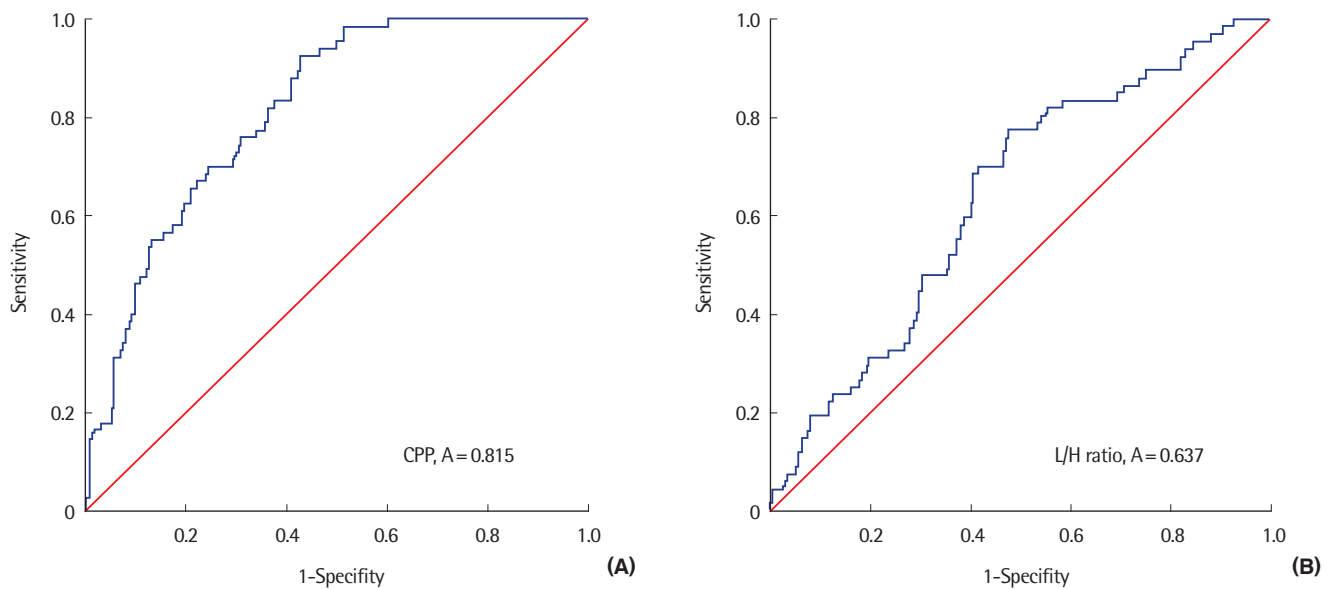


**Figure 1.** Comparison of (A) CPP and (B) L/H ratio between the control and dysphonic group in sustained vowel. CPP = cepstral peak prominence; L/H ratio = low-high spectral ratio.





**Figure 2.** Comparison of (A) CPP and (B) L/H ratio between the control and dysphonic group in connected speech. CPP=cepstral peak prominence; L/H ratio=low-high spectral ratio.



**Figure 3.** Receiver operating characteristic for (A) CPP and (B) L/H ratio in sustained vowel. CPP=cepstral peak prominence; L/H ratio=low-high spectral ratio; A=area under the curve.

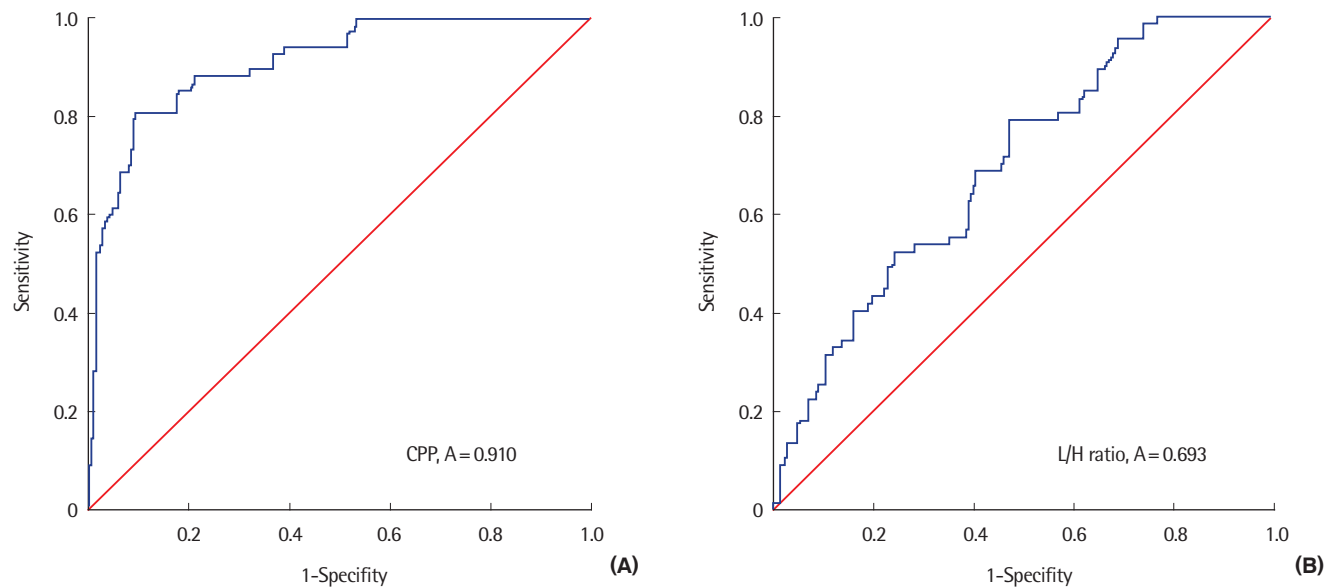
파수 스펙트럼 비율(L/H ratio)도 모음연장발성 및 연결발화 음성 표본 모두에서 음성장애군과 정상대조군에 따라 유의미한 차이를 보였으며, 음성장애 집단에서 더 유의하게 낮았다.

**캡스트럼 측정치의 음성장애 예측 가능성**

음향학적 측정치가 음성장애 환자의 음질 문제에 대한 감별 진단적 가능성이 있는지 알아보기 위하여 수신자 조작 특성 곡선 분석을 실시하였고 그 결과값과 Youden의 Index를 이용하여 각각

측정치의 민감도 및 특이도 및 그 진단 기준값(cutoff value)을 Table 4와 Figures 3, 4에 나타내었다.

모음연장발성의 스펙트럼, 캡스트럼 측정치의 수신자 조작 특성 곡선 아래의 면적(area under curve, AUC)값을 살펴보면 캡스트럼 최고정점(CPP)은 .815 (SEM=.026), 저주파수대고주파수 스펙트럼 비율(L/H ratio)은 .637 (SEM=.037)로 나타났고 두 측정치 중 캡스트럼최고정점이 장애음성의 음질 감별진단의 가능성이 더 높은 것으로 나타났다. Youden의 Index에 따르면 모음연장발성 표본



**Figure 4.** Receiver operating characteristic for (A) CPP and (B) L/H ratio in connected speech. CPP = cepstral peak prominence; L/H ratio = low-high spectral ratio; A = area under the curve.

**Table 5.** Pearson correlation analysis between acoustic variables and K-VHI on /a/ vowel and connected speech

Acoustic variable	K-VHI
Sustained vowel /a/	
CPP	-.497**
L/H ratio	-.204**
Connected speech	
CPP	-.468**
L/H ratio	-.285**

CPP = cepstral peak prominence; L/H ratio = low-high spectral ratio; K-VHI = Korean version of Voice Handicap Index.

\*\* $p < .01$ .

에서 캡스트럼최고정점은 70.6%의 민감도, 75.8%의 특이도로 진단 기준값은 12.088 (dB)로 나타났고, 저주파수대고주파수 스펙트럼 비율은 69.1%의 민감도, 57.8%의 특이도로 그 진단 기준값은 29.241 (dB)로 나타났다.

연결발화의 스펙트럼, 캡스트럼 측정치의 수신자 조작 특성 곡선 아래의 면적(AUC)값을 살펴보면, 캡스트럼최고정점(CPP)은 .910 (SEM = .020), 저주파수대고주파수 스펙트럼 비율(L/H ratio)은 .693 (SEM = .035)으로 나타났고 두 측정치 중 캡스트럼최고정점이 장애음성의 음질 감별진단의 가능성이 더 높은 것으로 나타났다. Youden의 Index에 따르면 연결발화 표본에서 캡스트럼최고정점은 80.9%의 민감도, 90.0%의 특이도로 그 진단 기준값은 7.076 (dB)로 나타났고, 저주파수대고주파수 스펙트럼 비율은 66.2%의 민감도, 58.8%의 특이도로 그 진단 기준값은 28.284 (dB)로 나타났다.

### 음향학적 측정치와 음성장애지수 간 상관성

모음연장발성과 연결발화의 음향학적 측정치와 음성장애지수 간 상관성은 Table 5와 같다. 음향학적 측정치 캡스트럼최고정점(CPP)과 저주파수대고주파수 스펙트럼 비율(L/H ratio)은 모음연장발성과 연결발화 표본 모두에서 한국판 음성장애지수(K-VHI)와 유의한 부적 상관관계를 보였다.

### 논의 및 결론

본 연구의 목적은 음성장애환자와 정상화자를 대상으로 음성장애를 평가하기 위한 객관적 음향학적 지표인 스펙트럼 및 캡스트럼 측정치의 차이를 비교하고, 정상음성과 병리적 음성을 구분하는 스펙트럼 및 캡스트럼 측정치의 진단적 기준점을 확인하는 것이다. 또한 음성평가 시 일반적으로 사용하는 두 음성표본을 비교하여 비교하여 음성장애를 측정하는 데 가장 민감한 표본이 무엇인지 알아보는 것이다.

스펙트럼 및 캡스트럼의 측정치 캡스트럼최고정점(CPP)과 저주파수대고주파수 스펙트럼 비율(L/H ratio)은 병리적 음성에서 정상음성보다 그 값이 낮았으며, 각 측정치는 모음연장발성과 연결발화의 두 가지 말하기 상황 모두에서 통계적으로 유의한 음성장애 유무 집단 간 차이를 보였다. 이는 주기적인 음향신호를 나타내는 정상음성은 스펙트럼 내에서 조화음의 구조를 잘 나타내므로 현저한 캡스트럼 정점을 나타낸다는 이전 연구와 일치하였다(Choi &

Choi, 2014; Shim et al., 2016; Watts & Awan, 2011; Yu et al., 2017). 캡스트럼최고정점(CPP)은 기식성으로 인한 소음에너지의 비율에 가장 큰 영향을 받으나, 음성장애의 일반적인 척도로 사용하는 거친 소리나 쇠 목소리에 의한 평정에서도 유용하였다. 저주파수대고주파수 스펙트럼 비율(L/H ratio)은 성문에서 발생하는 난기류 소음과 관련이 있는데 이 기식성의 음질 특징은 음성장애군의 낮은 저주파수대고주파수 스펙트럼 비율로 관찰되었다. Awan, Roy, Zhang과Cohen (2016)은 332명을 대상으로 임상가에 의한 청지각적 평정, 대상자에 의한 음성장애지수(VHI), 그리고 후두내시경적 진단의 세 가지 기준으로 정상음성과 병리적 음성을 감별하는 객관적 음향음성학적 캡스트럼 지수인 Cepstral Spectral Index of Dysphonia (CSID)의 유용성을 제시하였다. 캡스트럼최고정점(CPP), 저주파수대고주파수 스펙트럼 비율(L/H ratio), 표준편차, 그리고 성별에 다중회기방식으로 가중치를 두어 계산한 CSID는 세 가지 기준 가운데 특히 음성평가 시 일차 기준인 청지각적 판단과 높은 정확성을 보였다. 또한 Sauder, Bretl과 Eadie (2017)는 정상화자 70명과 장애음성 화자 100명을 대상으로 Praat과 ADSV에서 분석한 캡스트럼 측정치의 음성장애 진단정확도를 비교하였고 그 결과, 캡스트럼최고정점(CPP)은 두 음성프로그램 모두에서 높은 진단정확도를 나타내었다. Maryn 등(2009)은 음성의 음향적 평가에서 사용되는 모음연장발성과 연결발화 두 가지 유형의 음성표본과 관련하여 25편의 연구를 메타분석을 실시한 결과, 음질을 평가하는 음향학적 지표 중 이 두 가지 음성에서 동시에 유효성을 가지는 지표는 캡스트럼최고정점이 유일하다고 보고하였다. 본 연구에서도 캡스트럼 분석이 음질을 평가하는 데 타당성이 있음을 확인하였고, 이러한 결과는 캡스트럼 최고정점(CPP)이 임상현장에서 정상과 병리적 음성을 객관적으로 감별하는 데 유용하게 사용될 수 있을 것을 시사한다.

본 연구에서 수신자 조작 특성 곡선 분석 결과, 캡스트럼최고정점(CPP) 값은 모음연장발성과 연결발화의 음성표본에 관계없이 높은 정확도로 정상음성과 장애음성을 변별할 수 있었다. 특히 모음연장발성 과제에 비해 연결발화 과제에서 캡스트럼최고정점(CPP)은 더욱 높은 변별정확도를 나타내었다. Watts와 Awan (2011)은 수신자 조작 특성 분석을 통해 정상화자 16명, 음성장애 16명을 대상으로 캡스트럼최고정점(CPP)을 정상음성과 병리적 음성을 감별진단하는 매개변수의 가능성을 제시하였는데, 분석을 위해 사용한 두 가지 음성표본 중 모음연장발성에서 더욱 높은 변별정확도를 보인다고 보고하여 그 결과가 본 연구와 차이가 있었다. 한편 Heman-Ackah 등(2002, 2003)은 연결발화에서 캡스트럼최고정점(CPP)과 청지각적 평가의 높은 상관관계를 보고하였다. 국내에서 Shim 등

(2016)은 내전형연축성 발성장애환자 30명을 대상으로 발화과제(모음연장발성, 문장읽기)에 따른 스펙트럼 및 캡스트럼 측정치를 살펴보고, 그 결과 캡스트럼최고정점(CPP) 값이 연결발화에서 음성 중증도를 잘 반영한다고 하였다. 이는 특정 장애군 집단에 따라 특정 음성표본에서 음성장애환자의 음성 특징이 더욱 두드러질 수 있고, 동일한 발화자 내에서도 발성장애의 중증도에 따라 모음연장발성과 연결발화에서 음성의 특징이 다르게 나타날 수 있는 것으로 추측된다. 특히 연결발화 분석에서의 캡스트럼최고정점(CPP) 값은 다양한 무무성음과 음향학적 영향과 발화 속도, 억양, 특유의 조음 습관 등과 같은 개인적인 발화 특성, 또는 읽기나 말하기 유창성의 영향이 나타날 수 있다. 따라서 환자의 두드러진 음성 증상이 잘 나타날 수 있도록 음성평가 시 음성과제의 문맥은 주의 깊게 고려되어야 한다.

모음연장발성과 연결발화의 객관적 음향 지표인 스펙트럼 및 캡스트럼 측정치와 주관적인 음성장애지수의 상관성을 확인한 결과, 스펙트럼 및 캡스트럼 측정치 캡스트럼최고정점(CPP), 저주파수대고주파수 스펙트럼 비율(L/H ratio)은 화자의 주관적인 음성장애지수와 통계적으로 유의한 부적 상관관계를 나타내었다. Solomon 등(2013)은 음성평가 시 음향학적 평가 및 청지각적 평가와 더불어 자가평가 방식의 음성장애지수와 같은 평가가 유용하다고 하였다. 본 연구에서 객관적이고 신뢰로운 음향 지표인 스펙트럼 및 캡스트럼의 측정치는 화자의 주관적인 음성평가와도 상관성을 보였으며 이는 캡스트럼 측정치가 음성장애지수와 상관성을 보인다고 보고한 선행연구와 일치하였다(Awan et al., 2014).

연구의 결과를 토대로 본 연구의 제한점과 임상적 제언은 다음과 같다. 첫째, 본 연구에서 채택한 연결발화 과제는 자기소개였으며 음성분석 시 연구참여자의 이름을 제외한 전체 문장을 대상으로 하였다. Choi와 Choi (2016)는 20-30대의 정상화자를 대상으로 두 표준문단 '가을'과 '산책'의 스펙트럼 및 캡스트럼 측정치를 비교하였는데, 두 연결발화는 서로 다른 운율과 자음 및 모음의 음소 결합에도 불구하고 두 문단은 유사한 캡스트럼최고정점(CPP)이나 저주파수대고주파수 스펙트럼 비율(L/H ratio) 값을 나타내었다. 그러나 본 연구에서는 음향학적 측정치에 영향을 미칠 수 있는 참여자의 이름을 제외하고 분석하였으나 음성장애를 평가하기 위해 사용하는 일반적인 표준 문장과는 차이가 있을 수 있으므로 연구의 결과를 객관적인 진단 기준값으로 사용하기에는 무리가 있다. 따라서 추후에는 음성장애환자의 음질을 비교하기 위해 개발된 표준 문장을 이용하여 집단 간 차이를 살펴보는 것이 필요하다(Choi, 2018).

둘째, 본 연구는 많은 대상자와 연령군을 대상으로 스펙트럼 및

캡스트럼의 표준화된 객관적 수치를 제공하고자 하였으나 대상자의 성별, 연령별 분포를 고려하지 못하였다. Choi와 Choi (2016)는 여성의 기식성(breathiness) 음질 특성으로 인해 저주파수대고주파수 스펙트럼 비율(L/H ratio)에서의 성별 차이가 있다고 하였고 Lee, Pyo와 Choi (2018)도 과제와 성별에 따른 스펙트럼 및 캡스트럼 측정치 차이를 보고하였다. 특히 스펙트럼 및 캡스트럼 분석의 경우 기식성으로 인한 소음에너지의 비율에 가장 큰 영향을 받으므로 성별이나 대상자의 연령별 분포에 따른 고려가 필요할 것이다.

음성에 대한 음향학적 분석은 정상음성과 병리적인 음성의 특징을 설명하는 데 유용하며, 특히 캡스트럼 분석치인 캡스트럼최고정점(CPP)은 정상음성과 병리적인 음성을 구분하는 높은 감별정확도를 보였다. 본 연구에서는 의학적 진단에 의한 음성장애 유무에 따라 스펙트럼 및 캡스트럼 측정치를 비교하였으며, 더 나아가 각 측정치의 절단점을 도출하여 음성장애환자를 판별하는 진단 기준치를 제시하였다. 따라서, 기존의 모음 분석 위주의 시간 기반 음향학적 분석인 변동률 분석보다 비주기적인 음성신호를 가진 음성장애환자의 음성을 임상에서 보다 더 신뢰롭고 정확하게 진단할 수 있는 스펙트럼 및 캡스트럼 측정치에 대한 기준치를 제시하였다는 것에 의의가 있다. 또한, 본 연구의 결과는 병리적 음성의 객관적 검사뿐만 아니라 주관적 검사에서 어떠한 음성 자료를 대상으로 분석할 것인지에 대한 참고 자료가 될 것으로 기대된다.

## REFERENCES

- Awan, S. N., & Roy, N. (2005). Acoustic prediction of voice type in women with functional dysphonia. *Journal of Voice*, 19, 268-282.
- Awan, S. N., Roy, N., & Cohen, S. M. (2014). Exploring the relationship between spectral and cepstral measures of voice and the Voice Handicap Index (VHI). *Journal of Voice*, 28, 430-439.
- Awan, S. N., Roy, N., & Dromey, C. (2009). Estimating dysphonia severity in continuous speech: application of a multi-parameter spectral/cepstral model. *Clinical Linguistics & Phonetics*, 23, 825-841.
- Awan, S. N., Roy, N., Jetté, M. E., Meltzner, G. S., & Hillman, R. E. (2010). Quantifying dysphonia severity using a spectral/cepstral-based acoustic index: comparisons with auditory-perceptual judgements from the CAPE-V. *Clinical Linguistics & Phonetics*, 24, 742-758.
- Awan, S. N., Roy, N., Zhang, D., & Cohen, S. M. (2016). Validation of the cepstral spectral index of dysphonia (CSID) as a screening tool for voice disorders: development of clinical cutoff scores. *Journal of Voice*, 30, 130-144.
- Baken, R. J., & Orlikoff, R. F. (2000). *Clinical measurement of speech and voice* (2nd ed.). San Diego, CA: Singular Thomson Learning.
- Choi, S. H. (2013). Speech-language pathologists' voice assessment and voice therapy practices: a survey for standard clinical guideline and evidence-based practice. *Communication Sciences & Disorders*, 18, 473-485.
- Choi, S. H. (2018). Development of Korean standardized sentences on voice quality evaluation for dysphonia. *Audiology and Speech Research*, 14, 128-142.
- Choi, S. H., & Choi, C. H. (2014). The utility of perturbation, non-linear dynamic, and cepstrum measures of dysphonia according to signal typing. *Phonetics and Speech Sciences*, 6, 63-72.
- Choi, S. H., & Choi, C. H. (2016). The effect of gender and speech task on cepstral-and spectral-measures of Korean normal speakers. *Audiology and Speech Research*, 12, 157-163.
- Hasanvand, A., Salehi, A., & Ebrahimipour, M. (2017). A cepstral analysis of normal and pathologic voice qualities in Iranian adults: a comparative study. *Journal of Voice*, 31, 508-e17.
- Heman-Ackah, Y. D., Michael, D. D., Baroody, M. M., Ostrowski, R., Hillenbrand, J., Heuer, R. J., ... & Sataloff, R. T. (2003). Cepstral peak prominence: a more reliable measure of dysphonia. *Annals of Otolaryngology & Rhinology*, 112, 324-333.
- Heman-Ackah, Y. D., Michael, D. D., & Goding Jr, G. S. (2002). The relationship between cepstral peak prominence and selected parameters of dysphonia. *Journal of Voice*, 16, 20-27.
- Lee, S. J., Pyo, H. Y., & Choi, H. S. (2018). Normative data of cepstral and spectral measures in Korean adults using vowel phonation and passage reading tasks. *Communication Sciences & Disorders*, 23, 208-217.
- Maryn, Y., Roy, N., De Bodt, M., Van Cauwenberge, P., & Corthals, P. (2009). Acoustic measurement of overall voice quality: a meta-analysis. *The Journal of the Acoustical Society of America*, 126, 2619-2634.
- Noll, A. M. (1964). Short-time spectrum and "cepstrum" techniques for vocal-pitch detection. *The Journal of the Acoustical Society of America*, 36, 296-302.
- Parsa, V., & Jamieson, D. G. (2001). Acoustic discrimination of pathological voice: sustained vowels versus continuous speech. *Journal of Speech, Language, and Hearing Research*, 44, 327-339.
- Peterson, E. A., Roy, N., Awan, S. N., Merrill, R. M., Banks, R., & Tanner, K. (2013). Toward validation of the cepstral spectral index of dysphonia (CSID) as an objective treatment outcomes measure. *Journal of Voice*, 27, 401-410.
- Revis, J., Giovanni, A., Wuyts, F., & Triglia, J. M. (1999). Comparison of dif-



- ferent voice samples for perceptual analysis. *Folia Phoniatica et Logopaedica*, 51, 108-116.
- Roark, R. M. (2006). Frequency and voice: perspectives in the time domain. *Journal of Voice*, 20, 325-354.
- Sauder, C., Bretl, M., & Eadie, T. (2017). Predicting voice disorder status from smoothed measures of cepstral peak prominence using Praat and Analysis of Dysphonia in Speech and Voice (ADSV). *Journal of Voice*, 31, 557-566.
- Seo, I., & Seong, C. (2013). Voice quality of dysarthric speakers in connected speech. *Phonetics and Speech Sciences*, 5, 33-41.
- Shim, H. J., Jung, H., Lee, S. A., Choi, B. H., Heo, J. H., & Ko, D. H. (2016). Cepstral and spectral analysis of voices with adductor spasmodic dysphonia. *Phonetics and Speech Sciences*, 8, 73-80.
- Shin, Y. (2014). *Cepstral and Aerodynamic characteristics of subjects with thyroidectomy in pre and post operation* (Doctoral dissertation). Chonbuk National University, Jeonju, Korea.
- Solomon, N. P., Helou, L. B., Henry, L. R., Howard, R. S., Coppit, G., Shaha, A. R., & Stojadinovic, A. (2013). Utility of the voice handicap index as an indicator of postthyroidectomy voice dysfunction. *Journal of Voice*, 27, 348-354.
- Watts, C. R., & Awan, S. N. (2011). 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*, 54, 1525-1537.
- Yu, M., Choi, S. H., Choi, C. H., & Lee, K. (2017). Usefulness of cepstral acoustic index for estimating objective dysphonia severity. *Communication Sciences & Disorders*, 22, 587-596.

## 국문초록

### 모음과 연결발화에서 캡스트럼 음향 지표의 정상 및 음성장애 예측

유미옥<sup>1</sup>(학생, 제1저자) · 최성희<sup>1,2</sup>(교수, 교신저자) · 최철희<sup>1,2</sup>(교수) · 최병훈<sup>3</sup>(센터장)

<sup>1</sup>대구가톨릭대학교 일반대학원 언어청각치료학과, <sup>2</sup>대구가톨릭대학교 언어청각치료학과, <sup>3</sup>명이비인후과 음성센터

**배경 및 목적:** 본 연구에서는 모음연장발성과 연결발화에서 음성장애 진단을 위한 기준값을 정회하고, 두 음성표본 중 음성장애를 분석하기 위한 민감한 표본이 무엇인지 알아보았다. **방법:** 74명의 정상화자와 214명의 음성장애 환자를 대상으로 모음연장발성과 연결발화의 캡스트럼 및 스펙트럼 변수를 측정하였다. 음성장애 유무와 음성표본에 따라 각 측정치를 비교하고, 각 과제의 캡스트럼 및 스펙트럼 측정치의 진단 기준값을 정의하였다. **결과:** CPP와 L/H ratio는 두 음성표본 모두에서 음성장애 유무 집단 간 유의미한 차이를 보였다. CPP 값은 음성표본에 관계없이 높은 정확도로 정상음성과 병리적 음성을 변별할 수 있었고 특히 모음연장발성 과제에 비해 연결발화 과제에서 CPP는 더욱 높은 변별 정확도를 나타내었다. **논의 및 결론:** 캡스트럼 및 스펙트럼 측정치는 정상음성과 병리적인 음성을 구분하는 높은 감별 정확도를 보였다. 따라서, 본 연구는 장애음성을 평가할 때 진단적 예측으로서 캡스트럼 측정치를 포함시켜야 할 것을 제안한다.

**핵심어:** 캡스트럼 분석, 음성장애, 음성표본, 진단적 예측가능성

## 참고문헌

- 서인호, 성철재(2013). 연결발화에서 마비말화자의 음질 특성. *말소리와 음성과학*, 5, 33-41.
- 신유정(2014). 캡스트럼 분석과 공기역학적 분석 방법을 활용한 갑상선 절제술 전·후 음성언어변화 연구. 전북대학교 대학원 박사학위논문.
- 심희정, 정훈, 최병훈, 허정화, 고도홍(2016). 내전형연축성 발성장애 음성장애에 대한 캡스트럼과 스펙트럼 분석. *말소리와 음성과학*, 8, 73-80.
- 유미옥, 최성희, 최철희, 이정재(2017). 객관적인 음성장애 중증도 측정을 위한 캡스트럼 음향 지표의 유용성. *Communication Sciences & Disorders*, 22, 587-596.
- 이승진, 표화영, 최홍식(2018). 모음 발성 및 문단 읽기 과제 시 한국 성인의 캡스트럼 및 스펙트럼 측정치 정상 기준 연구. *Communication Sciences & Disorders*, 23, 208-217.
- 최성희(2013). 음성언어재활사의 음성평가와 음성치료의 임상실제: 표준 임상지침과 증거기반증재를 위한 설문조사. *Communication Sciences & Disorders*, 18, 473-485.
- 최성희(2018). 음성장애환자의 음질 평가를 위한 표준문장 개발. *청능재활*, 14, 128-142.
- 최성희, 최철희(2014). 음성 신호 분류에 따른 장애 음성의 변동률 분석, 비선형 동적 분석, 캡스트럼 분석의 유용성. *말소리와 음성과학*, 6, 63-72.
- 최성희, 최철희(2016). 한국 정상 화자의 캡스트럼과 스펙트럼 측정치의 성별과 발화 과제 효과. *청능재활*, 12, 157-163.