Usefulness of Cepstral Acoustic Index for Estimating Objective Dysphonia Severity

Miok Yu^a, Seong Hee Choi^b, Chul-Hee Choi^b, Kyungjae Lee^b

^aDepartment of Audiology & Speech-Language Pathology, Graduate School of Daegu Catholic University, Gyeongsan, Korea

^bDepartment of Audiology & Speech-Language Pathology, Daegu Catholic University, Gyeongsan, Korea

Correspondence: Seong Hee Choi, PhD
Department of Audiology and Speech-Language Pathology, Institute of Biomimetic Sensory Control, College of Medical Sciences, 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: June 11, 2017
Revised: August 6, 2017
Accepted: August 16, 2017

-We would like to thank Dr. Byeongheun Choi in Myung ENT Voice Center and National Research Foundation of Korea to support this study (NRF-2013S1A8024543).

-This work is based on a part of the first author's master's thesis from Daegu Catholic University (2016).

Objectives: Traditional perturbation analysis using MDVP (Multi-Dimensional Voice Program) often fails on acoustic measures, making it difficult to analyze connected speech in clinical practice. The purposes of this study were to compare the differences of perturbation and cepstral measures based on auditory-perceptual severity and to identify useful acoustic indices for determining dysphonia severity, which is highly correlated with auditory-perceptual severity of both sustained vowels and connected speech. **Methods:** Perturbation and cepstral analyses were performed with /a/ vowel prolongation and connected speech voice samples. Samples of a sustained vowel /a/ and connected speech productions were elicited from 147 patients (55 males and 92 females) with dysphonia. 'G' in GRBAS (grade, roughness, breathiness, asthenia, and strain) scale, the dysphonia severity was rated by two certified speech-language pathologists. Three time-based analyses parameters—jitter (%), shimmer (%), and noise-to-harmonic ratio (NHR)—in perturbation analysis and two frequency-based analyses parameters—low/high spectral ratio [L/H ratio], cepstral peak prominence [CPP]—in cepstral analysis were compared among different 'G' groups. **Results:** Measures of jitter, shimmer, and NHR significantly increased with dysphonia severity, whereas measures of L/H ratio and CPP (dB) significantly decreased with dysphonia severity. With respect to CPP (dB), significant differences were found between the normal and mild severity group in both /a/ vowels and connected speech ($p < .05$), whereas no differences were observed between the normal and mild severity group in jitter, shimmer, NHR, and L/H ratio in /a/ vowels ($p > .05$). **Conclusion:** With cepstral-based acoustic measures, CPP can be a useful acoustic index for differentiating dysphonia severity and is a feasible objective acoustic measure for evaluating voice quality in both sustained vowels and connected speech.

Keywords: Dysphonia, Cepstral analysis, Dysphonia severity, Voice sample

음성장애를 주소로 내원한 환자의 음성을 분석할 때는 음성 분석기기를 사용하여 대상자의 음성을 객관적 수치로 측정하는 객관적 평가와 평가자의 청각 인상(perception)에 기인한 주관적 평가를 함께 시행한다. 양적 평가의 경우에는 음성장애인의 음성 특징을 수치로 나타내어 정상치와 비교할 수 있고 많은 데이터를 간편하게 정리할 수 있는 장점이 있다. 그러나 수치화할 수 없는 음성 특징의 분석을 위해서는 질적인 평가도 함께 이루어져야 한다. 청지각적 평가의 경우에는 평가자의 주관성 개입에 따른 객관성 결함으로 인하여 타당도, 신뢰도 측면에서 부족함이 있다고 볼 수 있다. 이

러한 점을 최소화하기 위하여 결과 분석에 있어서는 객관적인 평가와 주관적인 평가를 연관 지어 해석해야 한다.

현재 음성언어치료 현장에서는 장애음성에 대한 청지각적 평가와 음향학적 평가의 차이를 종종 보고하곤 한다. 선행연구에 따르면 현재 음향학적 분석 도구로 널리 사용하는 Computerized Speech Lab. (CSL), Kay PENTAX Co.의 Multi-Dimensional Voice Program (MDVP)은 기본주파수의 추출을 기반으로 분석이 이루어지므로 기본주파수가 검출이 안 될 정도로 심각한 음질의 경우에는 분석 실패가 발생할 수 있다고 한다(Choi & Choi, 2014; Kang &

Seong, 2012). 기본주파수를 추적하기 힘들 정도로 음성장애 정도가 나쁜 경우 검사결과의 신뢰도가 떨어지게 되는 것이다(Choi & Choi, 2014; Kang & Seong, 2012).

또한 MDVP 분석은 기본주파수와 성대진동양상의 가변성으로 인하여 변동을 측정이 난해한 연결발화를 분석하기에는 상당한 제약이 있다. Shim 등(2016)은 모음의 연장발성이 연결발화에 비해 음질의 비정상성이 경미하게 관찰된다고 보고하였고 현재 일부 임상 현장에서는 음성장애를 평가할 때, 환자의 음성특성을 더 잘 반영하는 문장읽기나 회화와 같은 연결발화를 추가적인 음성표본으로 선택하기도 한다(Choi, 2013). 이러한 기존의 음질 분석 문제에 대안을 제시하는 새로운 음성 분석 방법의 하나가 바로 ‘캡스트럼(Cepstrum) 분석’이다.

캡스트럼은 로그 스펙트럼의 푸리에 변환 결과물로 캡스트럼의 분석치인 캡스트럼최고정점(cepstral peak prominence, CPP)은 기본주파수의 정확한 측정을 요하지 않아 음성 문제가 심각한 경우나 문장에서 음질 분석이 가능하다(Moers et al., 2012). 캡스트럼최고정점(CPP)은 전체 캡스트럼의 회귀선과 캡스트럼정점(cepstral peak) 간 강도의 차이를 말하며, 조화음의 구조가 잘 나타나는 정상음성에서 높은 캡스트럼최고정점(CPP) 값을 나타내고 반대로 음질 문제가 심각한 경우에는 낮은 값을 보인다(Choi & Choi, 2014; Shin, 2014). 저주파수대고주파수 스펙트럼비(low/high spectral ratio, L/H ratio)는 스펙트럼 상에서 4 kHz 주파수 영역을 기준으로 위아래의 평균 에너지의 비율을 의미하며, 정상 음성에서는 고주파수에 비해 저주파수 영역 에너지가 많아 저주파수대고주파수비(L/H ratio)가 높은 값을 보이고 기식성(breathiness)이 높은 병리적 음성의 경우 낮은 값을 보인다(Shin, 2014).

캡스트럼 분석은 소음이 많이 포함되어 있는 음성 신호도 신뢰성 있게 분석할 수 있고(Choi & Choi, 2014), 음질의 청지각적 평정과도 상관성이 높아(Heman-Ackah et al., 2014; Awan et al., 2010; Choi & Choi, 2014, 2016) 장애음성 분석에 많이 사용되고 있으며 특히, 불안정한 음성 신호를 포함하는 심각한 장애음성을 분석하는데 유용하다. 하지만 이러한 캡스트럼 분석은 계산 방식이 복잡하고, 분석기기의 사용 방법이 잘 알려지지 않아 연구가 미비한 실정이다. 아울러 선행연구의 대부분은 기존의 변동을 분석의 측정치와 비교한 연구들로 주로 모음의 연장발성표본만을 이용하여 평가가 이루어졌으므로(Choi & Choi, 2014) 연결발화와 같은 다양한 말표본을 이용한 음향 분석이 필요하다.

따라서, 음성장애를 평가하기 위해 모음연장발성과 연결발화의 세분화된 음성표본에 대하여 음질의 중증도에 따른 음성 변화의 정도가 어떠한지 정량화가 필요하고, 더 나아가 심각한 음질에서

분석실패를 보고하는 기존의 음향학적 지표보다 더 객관적이고 신뢰할 수 있는 평가 지표가 요구된다. 본 연구에서는 이비인후과 음성센터에서 음성장애로 진단받은 환자를 대상으로 모음연장발성과 연결발화의 말표본을 사용하여 청지각적 평가의 음성 중증도 집단에 따라 CSL의 MDVP를 이용한 변동을 분석과 Multi-Speech의 Analysis of dysphonia in Speech and Voice (ADSV)를 이용한 캡스트럼 분석의 측정치에 대해 논하고, 이들 음향 분석 중 어떠한 음향측정치가 음성장애의 중증도를 변별하는데 임상 현장에서 적절하게 활용될 수 있을지 살펴보고자 한다.

연구방법

연구대상

2013년 1월부터 2016년 6월까지 대구 소재의 이비인후과에 음성 문제를 주소로 내원한 환자 147명을 대상으로 하였다. 대상자는 이비인후과 전문의에 의해 후두 스트로보스코피를 시행받고 음성언어치료사에게 음성검사를 받은 환자들로 대상자의 질환군은 성대결절 44명, 폴립 10명, 성대구증 23명, 성대마비 10명, 근긴장성 발성장애 28명, 연축성 발성장애 14명, 그 외 18명(성대부종, 후두횡격막증, 변성발성장애, 실성증 등)으로 Table 1과 같다. 대상자의 연령은 만 15-87세로 평균 연령은 49.5 (SD = 17.87)세였으며 남성이 55명, 여성 92명이었다.

평가도구 및 실험방법

음성자료는 대구 소재의 이비인후과 음성언어치료실에서 CSL을 통해 녹음하였다. 음성언어치료사는 대상자에게 편안한 음도와 강도로 발성하도록 안내하였다. 녹음의 내용은 모음의 경우 /아/를

Table 1. Patients' information (N=147)

Characteristic	Value
Age (yr)	49.5±17.87
Gender	
Male	55
Female	92
Diagnosis	44
Vocal fold nodules	10
Vocal polyp	23
Sulcus vocalis	10
Vocal cord paralysis	28
Muscle tension dysphonia	14
Spasmodic dysphonia	18
Etc. (edema, laryngeal web, mutational falsetto, aphonia)	44

Values are presented as mean ± SD or number.

3초 이상 발성하도록 하였고 연결발화는 대상자 본인의 이름을 넣어 “안녕하세요? 저는 대구광역시에 사는 ○○○입니다”를 말하도록 하였다. 음성표본추출률(sampling rate)은 44,100 Hz로 하였고, 마이크는 Shure사의 SM48-LC 다이내믹 마이크를 사용하였다.

분석방법

음성장애 중증도 평가

음성장애 중증도를 평가하기 위하여 7년 이상 음성장애와 치료에 경험이 있는 1급과 2급의 언어재활사 2명이(평균 년수=11.5) 각각 음성파일을 듣고 실시하였다. 음성장애 중증도는 선행연구(Pyo et al., 1999)에 근거하여 GRBAS척도를 사용하여 평가자의 숙련성에 큰 영향을 받지 않고 대부분의 음성치료 전문가들이 신뢰성 있는 항목으로 사용하는 음성의 전반적인 신 정도인 ‘G’척도만을 이용하여 0, 1, 2, 3의 4점 척도로 평가하도록 하였다(0점=정상 음질, 1점=경도, 2점=중도, 3점=심도).

본 연구의 청지각적 중증도 평가는 음성 표본의 종류에 따라 평정이 달라질 수 있으므로(Choi, 2013; Choi & Choi, 2015, 2016) 모음연장발성과 연결발화의 표본을 분리하여 각각 평가하였으며 모음의 청지각적 중증도 집단은 G0 (정상) 18명, G1 (경도) 64명, G2 (중도) 48명, G3 (심도) 17명이었고, 연결발화의 청지각적 중증도 집단은 G0 (정상) 18명, G1 (경도) 67명, G2 (중도) 52명, G3 (심도) 10

명으로 나타났다.

검사자 간 신뢰도는 모음 96%, 연결발화 92%였으며, 일치하지 않는 부분은 두 사람이 다시 듣고 동의한 최종 측정치를 사용하였다. 검사자 내 신뢰도를 위하여 무작위로 음성표본의 20%를 일주일 후 동일한 조건에서 재검사하였으며, 검사자 간 신뢰도는 연구자와 이비인후과 음성센터에서 음성장애 진단 및 치료에 대해 8년 이상의 경험이 있는 1급 언어치료사 1인이 모음과 연결발화 음성표본의 각 20%를 무작위로 검사하여 결과를 비교하였다. 적률상관관계분석(Pearson correlation analysis) 결과, 검사자 내, 검사자 간 신뢰도는 모두 일치하였다($p < .001$).

음향학적 평가

수집된 모음연장발성 자료는 KayPentax사의 CSL의 MDVP를 이용하여 변동률 분석을 하였고 multi-speech의 ADSV를 이용하여 캡스트럼 분석을 실시하였다. 모음연장발성의 분석은 /아/를 3초간 발성한 음성표본의 처음과 마지막 50 ms를 제외한 나머지 구간 중 가장 안정적인 2초를 편집하여 변동률 측정치 주파수변동률(%jitter), 진폭변동률(%shimmer), 소음대배음비(NHR)와 캡스트럼 측정치인 저주파수대고주파수비(L/H ratio), 캡스트럼최고정점(CPP)을 산출하였다.

한편, 연결발화의 분석은 MDVP에서는 분석이 불가능하므로

Figure 1. Cepstral analysis using connected speech.

ADSV에서 캡스트럼 분석만 실시하였다. 연결발화는 녹음된 모든 구간을 분석 대상으로 하되 묵음구간의 포함 정도가 분석 결과치에 영향을 미칠 수 있으므로 녹음된 음성자료의 200 ms 이상 침 구간은 제외하였다. 또한 자음 및 모음의 종류에 따라 음향학적 평가의 소음의 정도가 차이가 있으므로(Choi & Choi, 2015) 녹음된 문장의 “안녕하세요? 저는 대구광역시에 사는 ○○○입니다”에서 대상자의 이름부분은 제외하고 편집하여 동일한 연결발화에서 캡스트럼 측정치인 저주파수대고주파수비(L/H ratio)와 캡스트럼최고정점(CPP)을 산출하였다(Figure 1).

통계분석

통계처리는 SPSS version 19.0 (IBM Co., New York, NY, USA)을 이용하였으며 모든 측정값은 평균 $\pm$ 표준편차로 표기하였다. GRBAS 척도의 ‘G’척도 중증도에 따라 변동률 분석과 캡스트럼 분석의 차이를 알아보기 위하여 각각 일원배치 분산분석(one-way ANOVA)을 실시하였고, 각 집단 간 차이가 있는지를 알아보기 위하여 Tukey 사후분석을 실시하였다. 또한 장애음성에 대한 청지각적 평가의 중증도와 음향학적 분석 측정치들 간의 상관성을 살펴보기 위해 위해 스피어만 순위상관계수(Spearman rank correlation coefficients)를 산출하였다. 통계적 유의수준은 .05 수준에서 검정하였다.

연구결과

모음의 청지각적 중증도 집단(G)에 따른 변동률, 캡스트럼 측정치

147명의 대상자의 음성표본 중 모음연장발성 과제에서 CSL의 MDVP를 이용한 변동률 분석에서 분석 실패가 2예 발생하여 모음에 대한 변동률 분석은 145명의 자료로 실시하였다. 분석 실패는 모두 심도장애 음절집단에서 나타났고 분석 실패의 원인은 기본주파수를 추출하기에 음성 신호가 충분하지 않았던 성대마비 환자의

음성 1예와 실성증으로 인하여 MDVP에서 분석 가능한 진동수(pulses) 범주에서 벗어난 환자의 음성 1예였다. 변동률 분석에서 실패가 있었던 모음의 음성자료는 Multi-speech의 ADSV를 이용한 캡스트럼의 분석이 모두 가능하였다.

본 연구 참여자의 모음연장발화의 청지각적 중증도 집단(G)에 따른 변동률과 캡스트럼 분석의 측정치는 Table 2에 제시하였다.

변동률 분석치인 주파수변동률(%jitter)에서는 G0집단은 0.72% $\pm$ 0.52%, G1집단은 1.05% $\pm$ 0.50%, G2집단은 3.20% $\pm$ 2.81%, G3집단은 6.02% $\pm$ 3.80%로 ‘G’ 점수가 높아질수록 증가하였다. 진폭변동률(%shimmer)에서는 G0집단은 2.32% $\pm$ 0.74%, G1집단은 3.60% $\pm$ 1.60%, G2집단은 9.07% $\pm$ 5.57%, G3집단은 11.90% $\pm$ 7.93%로 G점수가 높아질수록 증가하였다. 이와 마찬가지로, 소음대배음비(NHR)에서는 G0집단은 0.12 $\pm$ 0.01, G1집단은 0.13 $\pm$ 0.03, G2집단은 0.21 $\pm$ 0.12, G3집단은 0.28 $\pm$ 0.21로 G점수가 증가할수록 증가하였다.

한편, 캡스트럼 분석치인 저주파수대고주파수비(L/H ratio)에서는 G0집단은 33.99 $\pm$ 7.07, G1집단은 32.23 $\pm$ 4.94, G2집단은 28.17 $\pm$ 5.39, G3집단은 22.69 $\pm$ 7.18로 G점수가 증가할수록 감소하였다. 캡스트럼최고정점(CPP)에서는 G0집단은 13.26 $\pm$ 1.78, G1집단은 11.91 $\pm$ 1.25, G2집단은 8.21 $\pm$ 2.56 G3집단은 3.44 $\pm$ 2.12로 G점수가 증가할수록 감소하였다.

모음의 청지각적 중증도에 따른 집단 간 음향 측정치의 차이를 살펴 보기 위하여 일원배치분산분석을 실시한 결과, 주파수변동률(%jitter) ($F_{(3, 143)} = 31.587, p < .001$), 진폭변동률(%shimmer) ($F_{(3, 143)} = 30.908, p < .001$), 소음대배음비(NHR) ($F_{(3, 143)} = 15.221, p < .001$)에서 유의미한 차이를 보였다.

또한 캡스트럼 분석의 측정치 저주파수대고주파수비(L/H ratio) ($F_{(3, 143)} = 18.085, p < .001$)와 캡스트럼최고정점(CPP) ($F_{(3, 143)} = 123.617, p < .001$)도 모음의 청지각적 중증도에 따른 집단 간에 유의미한 차이가 있었다.

Table 2. Acoustic measures according to G scale in /a/ vowel

Acoustic measures	G scale group					F
	G0 (N=18)	G1 (N=64)	G2 (N=47)	G3 (N=18)	Total (N=145)	
Jitter (%)	0.72 $\pm$ 0.52	1.05 $\pm$ 0.50	3.20 $\pm$ 2.81	6.02 $\pm$ 3.80	2.30 $\pm$ 2.70	31.587***
Shimmer (%)	2.32 $\pm$ 0.74	3.60 $\pm$ 1.60	9.07 $\pm$ 5.57	11.90 $\pm$ 7.93	6.21 $\pm$ 5.45	30.908***
NHR	0.12 $\pm$ 0.01	0.13 $\pm$ 0.03	0.21 $\pm$ 0.12	0.28 $\pm$ 0.21	0.17 $\pm$ 0.11	15.221***
L/H ratio	33.99 $\pm$ 7.07	32.23 $\pm$ 4.94	28.17 $\pm$ 5.39	22.69 $\pm$ 7.18	29.97 $\pm$ 6.59	18.085***
CPP (dB)	13.26 $\pm$ 1.78	11.91 $\pm$ 1.25	8.21 $\pm$ 2.56	3.44 $\pm$ 2.12	9.86 $\pm$ 3.57	123.617***

Values are presented as mean $\pm$ SD.

NHR=noise to harmonic ratio; L/H ratio=low/high spectral ratio; CPP=cepstral peak prominence.

*** $p < .001$.

연결발화의 청지각적 중증도 집단(G)에 따른 캡스트럼 측정치

연결발화의 경우 변동률 분석은 불가능하므로 캡스트럼만 분석하였으며 분석치는 Table 3과 같다. 연결발화의 캡스트럼 분석치인 저주파수대고주파수비(L/H ratio)에서는 G0집단은 30.14 ± 3.96 , G1집단은 29.74 ± 3.10 , G2집단은 26.77 ± 3.62 , G3집단은 21.44 ± 4.58 로 G점수가 증가할수록 감소하였다. 캡스트럼최고정점(CPP)에서는 G0집단은 7.48 ± 1.01 , G1집단은 6.33 ± 1.03 , G2집단은 4.67 ± 1.48 , G3집단은 1.62 ± 1.88 로 G점수가 증가할수록 감소하였다.

연결발화의 청지각적 중증도에 따른 캡스트럼 측정치 저주파수대고주파수비(L/H ratio) ($F_{(3, 143)} = 21.339, p < .001$)와 캡스트럼최고정점(CPP) ($F_{(3, 143)} = 67.025, p < .001$)에서도 유의미한 차이가 있는 것으로 나타났다.

변동률, 캡스트럼 측정치의 청지각적 중증도 집단 간 변별가능성

청지각적 중증도에 따른 각 집단의 음향학적 측정치 차이를 살펴보고 이 중 어떠한 측정치가 'G'집단 중증도를 잘 변별하는지 살

Table 3. Acoustic measures according to G scale in connected speech

Acoustic measures	G scale group					F
	G0 (N=18)	G1 (N=67)	G2 (N=52)	G3 (N=10)	Total (N=147)	
L/H ratio (dB)	30.14 ± 3.96	29.74 ± 3.10	26.77 ± 3.62	21.44 ± 4.58	28.17 ± 4.18	21.339***
CPP (dB)	7.48 ± 1.01	6.33 ± 1.03	4.67 ± 1.48	1.62 ± 1.88	5.56 ± 1.88	67.025***

Values are presented as mean $\pm$ SD.

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

*** $p < .001$.

Acoustic measures	G scale			
	0	1	2	3
jitter (%)	0	1	2	3
		$p=.941$	$p<.001^{***}$	$p<.001^{***}$
shimmer (%)	0	1	2	3
		$p=.675$	$p<.001^{***}$	$p<.001^{***}$
NHR	0	1	2	3
		$p=.998$	$p<.001^{***}$	$p=.053$
L/H ratio_V	0	1	2	3
		$p=.652$	$p<.01^{**}$	$p<.01^{**}$
L/H ratio_C	0	1	2	3
		$p=.961$	$p<.001^{***}$	$p<.001^{***}$
CPP_V (dB)	0	1	2	3
		$p<.05^*$	$p<.001^{***}$	$p<.001^{***}$
CPP_C (dB)	0	1	2	3
		$p<.01^{**}$	$p<.001^{***}$	$p<.001^{***}$

Figure 2. Differential diagnosis of the acoustic measures between G scale groups.

NHR=noise to harmonic ratio; L/H ratio=low/high spectral ratio; L/H ratio_V=L/H ratio in /a/vowel; L/H ratio_C=L/H ratio in connected speech; CPP=cepstral peak prominence; CPP_V=CPP in /a/vowel; CPP_C=CPP in connected speech. * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 4. Spearman rank correlation coefficients between perceptual and acoustic measures of /a/vowels

Acoustic measures	G scale group
Jitter (%)	.715**
Shimmer (%)	.715**
NHR	.562**
L/H ratio (dB)	-.480**
CPP (dB)	-.796**

NHR=noise to harmonic ratio; L/H ratio=low/high spectral ratio; CPP=cepstral peak prominence.

** $p < .01$.

퍼 보고자 Tukey 사후 분석을 실시하였으며 결과는 Figure 2와 같다. 변동률 분석 측정치 주파수변동률(%jitter)은 G1-G2 ($p < .001$), G2-G3 ($p < .001$) 집단 간 차이가 있었고 진폭변동률(%shimmer)은 G1-G2 ($p < .001$), G2-G3 ($p < .001$) 집단 간 차이가 있었다. 소음대배음비(NHR)는 G1-G2 ($p < .001$) 집단 간 차이만을 나타내었다. 캡스트럼 측정치 저주파수대고주파수비(L/H ratio)는 /아/모음연장발성에서 G1-G2 ($p < .01$), G2-G3 ($p < .01$) 집단 간 차이가 있었고 연결발화에서는 G1-G2 ($p < .001$), G2-G3 ($p < .001$) 집단 간 차이가 있었다. 한편, 캡스트럼최고정점(CPP)은 /아/모음연장발성에서 G0-G1 ($p < .05$), G1-G2 ($p < .001$), G2-G3 ($p < .001$)의 모든 집단 간 차이가 있었고 연결발화에서도 G0-G1 ($p < .01$), G1-G2 ($p < .001$), G2-G3 ($p < .001$)의 모든 집단 간 차이가 있었다.

청지각적 중증도와 음향학적 측정치 간 상관성

모음의 청지각적 중증도와 음향학적 측정치 간 상관성은 Table 4와 같다. 청지각적 중증도와 변동률 분석의 측정치 주파수변동률(%jitter), 진폭변동률(%shimmer), 소음대배음비(NHR) 간에 유의한 정적 상관관계를 보였다. 청지각적 중증도와 캡스트럼 분석의 측정치 저주파수대고주파수비(L/H ratio), 캡스트럼최고정점(CPP) 간에 유의한 부적 상관관계를 보였다. 모음의 청지각적 중증도와 가장 높은 상관관계를 보인 모음의 음향학적 측정치는 캡스트럼최고정점(CPP)이었다.

한편, 연결발화의 청지각적 중증도와 음향학적 측정치 간 상관성은 Table 5와 같다. 청지각적 중증도와 연결발화의 캡스트럼 분석의 측정치 저주파수대고주파수비(L/H ratio)와 캡스트럼최고정점(CPP)은 모두 유의한 부적 상관관계를 보였다. 이 중에서 청지각적 중증도와 가장 높은 상관성을 보인 연결발화의 음향학적 측정치는 캡스트럼최고정점(CPP)이었다.

Table 5. Spearman rank correlation coefficients between perceptual and acoustic measures of connected speech

Acoustic measures	G scale group
L/H ratio (dB)	-.471**
CPP (dB)	-.717**

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

** $p < .01$.

논의 및 결론

장애음성의 청지각적 중증도 집단(G)에 따른 음향학적 측정치

장애음성의 청지각적 중증도가 증가할수록 변동률 분석의 측정치 주파수변동률(%jitter), 진폭변동률(%shimmer), 소음대배음비(NHR)이 증가하는 경향을 보였고 각각의 변동률 분석의 측정치는 장애음성의 청지각적 중증도에 따라 통계적으로 유의미한 차이를 보였다. 이는 변동률 분석의 측정치가 음성장애 환자의 음질을 판단하는데 유용한 수단이라는 Ko (2003)의 선행연구와 일치하였다.

본 연구에서는 147예의 음성표본 중 심도의 음성장애 집단에서 2예의 변동률 분석 실패가 발생하였다. 분석 실패가 발생한 사례는 장애음성이 매우 기식적이고 비주기적이기 때문에 기본주파수의 검출이 어려웠던 것으로 해석할 수 있다. Titze (1995)는 장애 음성 신호를 효과적으로 분석하기 위하여 음성 신호를 3가지 체계로 분류하였으며, Sprecher, Olszewski, Jiang과 Zhang (2010)은 음성 신호를 4가지로 더 확대하여 분류하여 Type1은 주기적인 신호로, Type2는 비주기적인 신호, Type3은 제한된 차원을 가진 혼돈 신호, Type4는 무한차원을 가진 음성신호로 정의하고 변동률 분석을 위한 음성 분석 시에 신뢰성 있는 주파수변동률(%jitter)과 진폭변동률(%shimmer) 수치를 얻기 위해서는 주기적이거나 거의 주기적인 Type1신호만을 변동률 분석 시에 사용하도록 권고하였다. 국내의 Choi와 Choi (2014)의 연구에서도 장애 음성을 신호별로 분류하고 TF32를 이용하여 변동률분석을 실시하였을 때, 주파수변동률(%jitter)과 진폭변동률(%shimmer) 측정치는 Type1신호에서 가장 낮게 나타났고 Type4에서 가장 높게 나타났으며, Type1을 제외한 Type2, 3, 4의 신호는 변동률 분석이 부적합한 것으로 나타났다. 이러한 연구 결과는 심각한 음질의 경우 변동률 분석에서 음향학적 분석의 실패가 발생할 수 있다는 Kang과 Seong (2012)의 연구와 연관성이 있는 것으로 보인다. 장애 음성은 대부분 비주기적인 음성이 많은 경우를 감안할 때, 이러한 변동률 분석의 단점을 보완하기 위하여 기본주파수 추출에 상관없이 음질을 객관적으로 측정할 수 있는 음향학적 분석이 캡스트럼 분석이다.

본 연구에서 캡스트럼 분석의 경우에는 모음 연장발성 및 연결 발화 모두에서 장애음성의 청지각적 중증도가 증가할수록 저주파수대고주파수비(L/H ratio), 캡스트럼최고정점(CPP)값이 감소하는 경향을 보였고 각각의 측정치는 모음 연장발성 및 연결발화의 청지각적 중증도에 따라 통계적으로 유의미한 차이를 보였다. 이것은 나쁜 음질의 경우 조조성과 기식성으로 인한 소음 에너지 비율이 높고 음성의 조화음 정도가 낮으며, 비주기적인 특성을 보인 결과라는 Choi와 Choi (2014)의 연구와 Shim, Jang, Shin과 Ko (2014)의 연구 결과와 일치하였다. 또한 Awan과 Roy (2006)의 선행 연구에서는 캡스트럼 분석 측정치와 다른 음향학적 분석 측정치를 비교한 결과, 음성장애의 중증도를 결정하는데 캡스트럼 기반 분석의 측정치가 가장 예측력 있는 지표라고 하였다. Choi와 Choi (2014)의 연구에서도 장애 음성 신호별로 GRBAS척도를 이용한 청지각적 평정을 실시한 결과, Type1은 정상 음질을 보였고, Type4로 갈수록 음질이 나빠져 음성신호 분류체계가 장애 음성의 중증도를 효과적으로 반영할 수 있는 것으로 보였다.

본 연구에서는 기존의 전통적인 음향학적 분석방법인 변동률 분석이 모든 장애음성을 평가하고 기술하는데 제한점이 있음을 확인하였다. 또한 이를 보완하기 위한 분석 방법인 캡스트럼 분석이 음성의 음질을 평가하는 데 타당성이 있음을 확인하였고, 이러한 결과는 추후 연구 및 임상 현장에서 음성장애 환자의 병리적 음성특성을 객관적으로 설명할 수 있는 분석 방법이 될 수 있을 것이다.

음향학적 측정치의 장애음성의 청지각적 중증도 집단(G) 간 변별 가능성

본 연구에서 변동률 및 캡스트럼 측정치가 장애음성의 청지각적 중증도 'G' 집단 간 변별이 가능한지 알아본 결과, 변동률 측정치 주파수변동률(%jitter), 진폭변동률(%shimmer)는 경도장애 음성집단과 중도장애 음성집단, 중도장애 음성집단과 심도장애 음성집단 간 차이를 나타내었다. 소음대배음비(NHR)은 경도장애 음성집단과 중도장애 음성집단 간의 차이를 보였다. 임상현장에서 주로 사용하는 변동률 측정치 주파수변동률(%jitter), 진폭변동률(%shimmer), 소음대배음비(NHR)는 음성장애를 평가할 때 가장 민감하게 구분해야 하는 정상과 경도장애음성의 차이를 변별하지 못하는 문제를 보였다.

모음과 연결발화의 캡스트럼 측정치 저주파수대고주파수비(L/H ratio)는 모두 경도장애 음성집단과 중도장애 음성집단, 중도장애 음성집단과 심도장애 음성집단 간 차이를 나타내었으나, 모음과 연결발화 모두에서 정상 음질과 경도 장애 음성집단을 구분하지 못하였다. 한편, 모음과 연결발화의 캡스트럼최고정점(CPP)은 정

상 음질집단과 경도장애 음성집단, 경도장애 음성집단과 중도장애 음성집단, 중도장애 음성집단과 심도장애 음성집단 모두에서 차이를 나타내어 캡스트럼 분석의 캡스트럼최고정점(CPP) 변수가 정상 음질과 장애음성의 음질을 변별하는 데 유용한 변수임을 확인할 수 있었다.

Wang, Kent, Kent, Duffy와 Thomas (2009)은 모음의 안정적인 구간을 분석한 음향학적 연구는 음성장애 환자의 일반적 음성 특성을 반영하지 못할 수 있다고 하였다. 본 연구에서는 모음과 연결 발화의 청지각적 중증도 'G'의 상관성을 살펴본 결과, 높은 양의 상관성을 보여 음질 평가 시에 모음과 연결발화 모두 음성의 중증도를 평가하기 유용하였으나, 모음의 청지각적 중증도 집단은 G0(정상) 18명, G1(경도) 64명, G2(중도) 48명, G3(심도) 17명이었고, 연결발화의 청지각적 중증도 집단은 G0(정상) 18명, G1(경도) 67명, G2(중도) 52명, G3(심도) 10명으로 나타나 모음과 연결발화 간 청지각적 평정에 차이를 보였다. 따라서, 음질 평가 시에 모음과 연결 발화를 모두 사용하는 것이 필요하다.

청지각적 중증도와 객관적 음향학적 측정치와의 상관성

모음의 청지각적 중증도와 모음의 객관적 음향학적 측정치와의 상관성을 확인한 결과, 청지각적 중증도와 변동률 분석 측정치 주파수변동률(%jitter), 진폭변동률(%shimmer), 소음대배음비(NHR)은 모두 통계적으로 유의한 정적 상관관계를 나타내었다. 청지각적 중증도와 모음과 연결발화의 캡스트럼 분석 측정치인 저주파수대고주파수비(L/H ratio), 캡스트럼최고정점(CPP)은 모두 통계적으로 높은 부적 상관관계를 나타내었다. 모음의 청지각적 중증도와 높은 상관성을 나타낸 음향학적 측정치는 캡스트럼최고정점(CPP), 주파수변동률(%jitter)과 진폭변동률(%shimmer), 소음대배음비(NHR), 저주파수대고주파수비(L/H ratio) 순으로 나타났다. 본 연구를 통해 새로운 음향학적 측정치인 캡스트럼 분석의 캡스트럼최고정점(CPP)이 장애음성 평가에 주요한 정보를 제공하는 기존의 변동률 측정치 및 청지각적 평가 결과와 높은 상관성이 있음을 살펴볼 수 있었다. 한편, 연결발화의 청지각적 중증도와 연결발화의 캡스트럼 분석 측정치 저주파수대고주파수비(L/H ratio), 캡스트럼최고정점(CPP) 간 유의한 부적 상관관계를 보였다.

Lowell, Colton, Kelly와 Mizia (2013)의 연구에서 캡스트럼 분석의 측정치는 GRBAS와 CAPE-V의 청지각적 평정과 높은 상관을 보였고 정상음성과 장애음성을 감별할 때 가장 높은 진단예측력을 갖는 지표라고 하였다. 국내에서는 Choi와 Choi (2014) 연구에서 모음의 캡스트럼최고정점(CPP)을 측정하였으며, 캡스트럼최고정점(CPP)이 GRBAS척도의 'G', 'R', 'S'와 가장 높은 상관을 보여

음질을 예측하는 가장 민감한 판별값이었으며, Seo와 Seong (2013)은 마비말장애 환자를 대상으로 연결발화의 캡스트럼최고정점(CPP)을 측정하였으며, 캡스트럼최고정점(CPP)이 중증도 및 GRBAS척도의 'G', 'R', 'S'와 높은 상관을 보여 캡스트럼 측정치들의 임상적 유용성에 대하여 제시하였다.

본 연구의 결과를 토대로 본 연구의 제한점과 임상적 제언은 다음과 같다.

첫째, 본 연구의 연결발화 과제는 자기소개로서 “안녕하세요? 저는 대구광역시에 사는 ○○○입니다”에서 ○○○에 해당하는 자기 이름을 제외한 문장을 분석 대상으로 하였다. 자음 및 모음의 종류에 따라 음향학적 평가의 소음의 정도가 차이가 있기 때문에 (Choi & Choi, 2015) 이름 부분은 제외하고 분석하였으나 일반적인 표준 문장과는 차이가 있을 수 있다. 현재 임상현장에서 음성장애를 평가할 때 가을과 산책의 표준화된 문단, 자체제작 문단, 자기소개하기 등을 사용하는데(Choi, 2013) 앞으로 이러한 다양한 문장들을 분석을 한다면 음성평가에 대한 통일된 말자료를 구축하는 기초가 되는 자료 확립에 기여할 것이다.

둘째, 본 연구에서는 청지각적 평가와 기존의 음향학적 분석인 변동률 분석과의 비교를 통하여 새로운 음향학적 분석인 캡스트럼 분석의 유용성 및 민감성을 확인하였다. Awan과 Roy (2006)는 음성장애의 중증도를 결정하는데 있어서 캡스트럼 기반의 분석 결과가 가장 중요한 정보라고 하였으나 이러한 유용성에도 불구하고 현재 캡스트럼 측정치는 기존의 변동률 분석처럼 표준화된 객관적인 수치를 제공하지 못한다는 문제점이 있다. 최근 Choi와 Choi (2016)의 연구에서는 20-30대 정상 화자를 대상으로 한국어 단모음 7개(/아/, /에/, /이/, /오/, /우/, /으/, /어/)와 표준 문단(‘가을’, ‘산책’)을 이용하여 성별에 따른 캡스트럼 정상치를 제시하였다. 앞으로 후속 연구를 통해 더 많은 대상자와 연령군을 대상으로 캡스트럼의 표준화된 객관적 수치를 제공할 수 있다면 임상현장에서 장애 음성을 평가하는 데 유용성이 있을 뿐 아니라 음성 치료 전후의 효과를 비교하는 데도 유용한 지표로 쓰일 수 있을 것이다.

마지막으로, 본 연구에서는 주관적인 음성장애 중증도를 평가하기 위하여 청지각적 평가도구인 GRBAS 중 ‘G’ 척도만을 사용하였다. 따라서, 추후의 연구에서는 GRBAS척도 중 다른 ‘R (조조성)’이나 ‘B (기식성)’ 혹은 ‘S (긴장성)’와 연관이 있는 객관적 음향적 지표를 찾는 연구를 확대하고자 한다.

또한 MDVP 측정치와 GRBAS척도 외에도 다른 음성평가와 청지각적 평가의 관계를 연구한다면 캡스트럼 분석 결과에 대한 신뢰도와 특성에 관한 더 자세한 정보를 얻을 수 있을 것이다. Awan, Roy와 Cohen (2014)은 음성장애 환자가 본인의 음성 문제로 인해

느끼는 불편함의 정도를 VHI (voice handicap index) 설문지로 살펴본 결과, 통계적으로 유의미하게 높은 상관성을 보였으며 특히 음성장애 환자들의 평가 측면에서 객관적 평가의 결과물인 캡스트럼 분석은 음성 증상에 대해 환자들이 느끼는 정도와도 밀접한 관련이 있다고 하였다.

결론적으로, 임상현장에서 주로 사용하는 기존의 변동률 측정치는 심도의 장애음성 분석에서 분석 실패가 발생하였을 뿐만 아니라 음성장애를 평가할 때 가장 민감하게 구분해야 하는 정상과 경도장애음성의 차이를 변별하지 못하는 문제를 보였다. 이에 대해 대안으로 제시된 캡스트럼 분석은 음성 중증도에 따른 집단을 모두 변별할 수 있었으며 음성장애 환자의 중증도와 상관성이 가장 높아 음성장애를 진단하는데 유용한 변수임을 확인할 수 있었다. 또한 장애음성을 평가할 때 모음과 함께 대상자의 실제적 음성 특징을 더 잘 반영하는 연결발화 표본을 사용하는 데 본 연구에서는 연결발화가 음성장애 중증도를 평가하기 적절하였으나, 청지각적 평정 시 두 말표본 간에 차이를 보여 음질 평가 시 모음뿐 아니라 연결발화를 함께 사용하는 것이 권고된다. 따라서 본 연구에서는 음성장애 환자의 음성을 평가할 때 캡스트럼최고정점(CPP)이 청지각적 음성장애 중증도를 객관적으로 평가할 수 있는 음향학적 지표임을 확인하였고, 모음과 연결발화에서 모두 음성장애 중증도를 잘 판별하였다. 이러한 연구 결과를 통해 캡스트럼최고정점(CPP)이 장애 음성을 객관적으로 평가하는 데 유용하게 사용될 수 있을 것이다.

REFERENCE

- Awan, S. N., & Roy, N. (2006). Toward the development of an objective index of dysphonia severity: a four-factor acoustic model. *Clinical Linguistics & Phonetics*, 20, 35-49.
- 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., 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.
- Choi, S. H., & Choi, C. H. (2016). The effect of gender and speech task on cepstral- and spectral-measures of Korean normal speakers. *Audiology & Speech*, 12, 157-163.
- 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 and Disorders*, 18, 473-485.
- 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. (2015). The stability and variability based on vowels in voice quality analysis. *Phonetics and Speech Sciences*, 7, 79-86.
- Heman-Ackah, Y. D., Sataloff, R. T., Laureyns, G., Lurie, D., Michael, D. D., Heuer, R., ... & Lyons, K. (2014). Quantifying the cepstral peak prominence, a measure of dysphonia. *Journal of Voice*, 28, 783-788.
- Kang, Y., & Seong, C. (2012). A cepstral analysis of breathy voice with vocal fold paralysis. *Phonetics and Speech Sciences*, 4, 89-94.
- Ko, D. H. (2003). A study of extracting acoustic parameters for individual speakers. *Phonetics and Speech Sciences*, 10, 129-143.
- Lowell, S. Y., Colton, R. H., Kelley, R. T., & Mizia, S. A. (2013). Predictive value and discriminant capacity of cepstral-and spectral-based measures during continuous speech. *Journal of Voice*, 27, 393-400.
- Moers, C., Mobius, B., Rosanowski, F., Noth, E., Eysholdt, U., & Haderlein, T. (2012). Vowel- and text-based cepstral analysis of chronic hoarseness. *Journal of Voice*, 26, 416-424.
- Pyo, H. Y., Choi, S. H., Lim, S. E., Sim, H. S., Choi, H. S., & Kim, K. M. (1999). The correlation between GRBAS scales and MDVP parameters on the pathologic voices of the patients with vocal polyps. *The Journal of the Korean Society of Logopedics and Phoniatics*, 10, 154-163.
- Seo, I. H., & Seong, C. J. (2013). Voice quality of dysarthric speakers in connected speech. *Phonetics and Speech Sciences*, 5, 33-41.
- Shim, H. J., Jang, H. R., Shin, H. B., & Ko, D. H. (2014). Spectral and cepstral analyses of esophageal speakers. *Phonetics and Speech Sciences*, 6, 47-54.
- 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.
- Sprecher, A., Olszewski, A., Jiang, J. J., & Zhang, Y. (2010). Updating signal typing in voice: addition of type 4 signals. *Journal of Acoustical Society of America*, 127, 3710-3716.
- Titze, I. R. (1995). *Workshop on acoustic voice analysis: summary statement*. Iowa City, IA: National Center for Voice and Speech.
- Wang, Y. T., Kent, R. D., Kent, J. F., Duffy, J. R., & Thomas, J. E. (2009). Acoustic analysis of voice in dysarthria following stroke. *Clinical Linguistics & Phonetics*, 23, 335-347.

국문초록

객관적인 음성장애 중증도 측정을 위한 캡스트럼 음향 지표의 유용성

유미옥¹ · 최성희^{1,2} · 최철희^{1,2} · 이경재^{1,2}

¹대구가톨릭대학교 일반대학원 언어청각치료학과, ²대구가톨릭대학교 언어청각치료학과

배경 및 목적: CSL의 MDVP를 이용한 변동률 분석은 심각한 장애음질의 경우 분석 실패가 발생할 수 있고 연결발화를 분석하기에도 어려움이 있다. 이에 본 연구에서는 청지각적 중증도에 따른 각 집단의 변동률 분석과 캡스트럼 분석치를 비교하고 청지각적 중증도를 잘 변별하고 모음뿐 아니라 연결발화의 청지각적 중증도와 상관이 높은 음향 변수를 찾고자 한다. **방법:** 본 연구는 총 147명을 대상으로 모음과 연결발화의 말표본을 사용하여 GRBAS의 'G'척도를 이용하여 청지각적 중증도를 각각 평정하였고 음향학적 분석으로는 변동률 분석치(%jitter, %shimmer, NHR), 캡스트럼 분석치(L/H ratio, CPP)를 측정하였다. **결과:** 주파수변동률(%jitter), 진폭변동률(%shimmer), 소음대배음비(NHR), 저주파수대고주파수비(L/H ratio)는 음성평가 시 가장 민감하게 구분해야 하는 정상과 경도장애 음성의 차이를 변별하지 못하였고, 심도의 장애음성에서는 분석실패가 2에 발생하였다. 모음과 연결발화의 캡스트럼측정치인 캡스트럼최고정점(CPP)은 청지각적 중증도에 따른 모든 집단을 변별할 수 있었으며 청지각적 중증도와도 높은 상관관계를 나타내었다. **논의 및 결론:** 음성장애 환자의 음성을 평가할 때 장애음성의 중증도 변별에 캡스트럼 측정치인 캡스트럼최고정점(CPP)이 가장 유용한 음향지표가 될 수 있으며, 모음과 연결발화를 이용한 음성장애 환자의 음질 평가 시 객관적 음향학적 측정치로서 유용하게 사용될 수 있을 것이다.

핵심어: 음성장애, 캡스트럼 분석, 음성장애 중증도, 음성표본

- 본 연구를 위해 지원해 주신 명예비인후과 최병훈 원장님과 한국연구재단에 감사드립니다(NRF-2013S15A802453).

- 본 연구는 제1저자의 석사학위 논문의 일부를 요약한 것임.

참고문헌

- 강영애, 성철재(2012). 성대마비로 인한 기식 음성에 대한 Cepstral 분석. *말소리와 음성과학*, 4, 89-94.
- 고도홍(2003). 개별화자의 음성파라미터 추출에 관한 연구: 음성파라미터의 상관관계를 중심으로. *말소리와 음성과학*, 10, 129-143.
- 서인효, 성철재(2013). 연결발화에서 마비말화자의 발화 특성. *말소리와 음성과학*, 5, 33-41.
- 신유정(2014). 캡스트럼 분석과 공기역학적 분석 방법을 활용한 갑상선 절제술 전·후 음성언어변화 연구. 전북대학교 대학원, 박사학위논문.
- 심희정, 장효령, 신희백, 고도홍(2014). 식도발성화자 음성의 spectral & cepstral 분석. *말소리와 음성과학*, 6, 47-54.
- 심희정, 정훈, 이수안, 최병훈, 허정화, 고도홍(2016). 내전형연축성 발생장애 음성장애에 대한 캡스트럼과 스펙트럼 분석. *말소리와 음성과학*, 8, 73-80.
- 이영아, 김형태, 김재욱(2010). 청지각적 음성평가 훈련이 음성평가 신뢰도에 미치는 효과. *언어청각장애연구*, 15, 526-536.
- 최성희(2013). 음성언어재활사의 음성평가와 음성치료의 임상실제: 표준 임상지침과 증거기반증제를 위한 설문조사. *언어청각장애연구*, 18, 473-485.
- 최성희, 최철희(2014). 음성 신호 분류에 따른 장애음성의 변동률 분석, 비선형 동적 분석, 캡스트럼 분석의 유용성. *말소리와 음성과학*, 6, 63-72.
- 최성희, 최철희(2015). 음질 분석에 있어서 모음 연장 발생에 따른 안정성과 변이성. *말소리와 음성과학*, 7, 79-86.
- 최성희, 최철희(2016). 한국 정상 화자의 캡스트럼과 스펙트럼 측정치의 성별과 발화 과제 효과. *Audiology & Speech Research*, 12, 157-163.
- 표화영, 최성희, 임성은, 심현섭, 최홍식, 김광문(1999). 성대 폴립 환자를 대상으로 한 GRBAS척도와 MDVP 측정치 간의 상관관계 연구. *대한후두음성언어의학회지*, 10, 154-163.