

노인성 음성장애에 대한 켈스트럼 분석에서의 무성음 분절 영향

부산대학교병원 이비인후과,¹ 양산부산대학교병원 이비인후과,² 부산가톨릭대학교 언어청각치료학과,³ 동남권원자력의학원 이비인후과,⁴ 부산대학교 인문대학 언어정보학과⁵
김근효¹ · 이연우¹ · 배인호² · 박희준³ · 이창윤⁴ · 권순복⁵

The Effect of Unvoiced Segments on the Cepstral Analysis in Age-Related Dysphonia

Geun-Hyo Kim, PhD¹, Yeon-Yoo Lee, PhD¹, In-Ho Bae, PhD²,
Hee-June Park, PhD³, Chang-Yoon Lee, MS⁴ and Soon-Bok Kwon, PhD⁵

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 Speech and Hearing Therapy, Catholic University of Pusan, Busan; and ⁴Department of Otorhinolaryngology, Dongnam Institute of Radiological & Medical Sciences, Busan; and ⁵Humanities, Language and Information, Pusan National University, Busan, Korea

– ABSTRACT –

Background and Objectives : The purpose of this study is to investigate the effect of unvoiced segments on cepstral analysis in patients with presbyphonia. **Materials and Methods** : A total of 381 participants (263 participants with presbyphonia and 118 with normal voices participated in this study. The sustained vowel /a/ 2 seconds and one sentence from ‘San-Chaek’ were edited and analyzed using the Praat script. Auditory-perceptual rating was also performed by three raters. **Results** : All parameters except LHRatio_V and extracted continuous speech (EXT) showed a significant difference between the normal and the dysphonia groups. The normal group showed higher cepstral peak prominence (CPP) and smoothed CPP (CPPS) than the dysphonia group. Additionally, AP ratings of the normal group were lower than that of the dysphonia group. The correlation within the measured values ranged from -0.86 to -0.66. In continuous speech (CS), there was a significant difference of area under the curve (AUC) between CPP and CPPS, but there was no significant difference in EXT. **Conclusions** : Through this study, we confirmed the effect of unvoiced segments on cepstral analysis. By removing the unvoiced segments, the measured values of the cepstral analysis showed an increase. Through further studies, we hope to increase the accuracy of discriminating dysphonia by extracting the unvoiced segments. (J Clinical Otolaryngol 2019;30:68-76)

KEY WORDS : Dysphonia · Acoustics · Aging · Praat.

논문접수일 : 2019년 3월 20일 / 논문수정일 : 2019년 5월 8일 / 심사완료일 : 2019년 5월 24일
교신저자 : 권순복, 46241 부산광역시 금정구 부산대학로63번길 2 부산대학교 인문대학 언어정보학과
전화 : (051) 510-2003 · 전송 : (051) 582-0572
E-mail : sbkwon@pusan.ac.kr

서 론

노화는 신체의 모든 기관에 영향을 주며, 성대도 예외는 아니다.¹⁾ 음성은 호흡, 성대진동, 공명의 결합된 작용으로부터 생성된다. 연령이 증가할수록 음성 산출과 연관된 모든 요소들이 악화되고 노인성 음성의 기전(mechanism)은 복잡해진다. 호흡이 감소되고 성대기능이 저하되며, 입천장, 혀, 입술 등과 같은 조음 및 공명기관의 운동 기능 또한 약화된다. 성대는 얇아지고 약해져 성대위축의 원인이 된다. 성대기능 부전으로 인해 성대진동이 약해지고 음성은 거친 소리, 쉼 소리, 약한 소리 등을 포함한다.

노인성 음성장애를 정확하게 평가하기 위해 다양한 방법들을 시행한다.^{2,3)} 임상현장에서 음성장애환자의 음질을 평가하기 위해서 음향학적, 공기역학적, 청지각적 평가, 자가 설문지 등을 실시하게 된다.⁴⁾ 다양한 평가들을 종합하여 음성장애의 중증도를 파악한다. 이 중 널리 사용되는 방법 중 하나인 음향학적 평가는 음성장애 환자의 음질을 간단하게 수치화 할 수 있는 방법이다. 음향학적 분석은 섭동분석과 캡스트럼 분석 방식을 사용하고 있으며, 모음연장과 연속발화 과제를 통해서 음성샘플을 수집한다.^{5,6)}

최근 미국언어치료협회(American Speech-Language-Hearing Association, ASHA)에서는 병리적 음성에 대한 평가를 위해 캡스트럼 분석을 수행하는 가이드라인을 제시하였다.⁷⁾ 중증도가 심한 음성에서는 섭동분석의 신뢰도가 낮아 캡스트럼 분석을 권고하였다.⁸⁾ 캡스트럼 분석은 무성음, 쉼 구간에 영향을 많이 받기 때문에 모음연장 샘플보다는 연속발화 샘플에서 cepstral peak prominence(CPP), smoothed cepstral peak prominence(CPPS)의 측정치가 낮게 관찰된다. CPP, CPPS는 음성신호의 주기성 정도를 나타내는 변수이며, 캡스트럼 피크가 얼마만큼 두드러지는지를 통해 음성장애의 중증도를 정량화한다. 기존 연구들에서는 연구자가 임의로 연속발화 내의 무성음, 쉼 구간을 제거하는 것은 주관적이기 때문에 제거하지 말고 포함시켜 분석하도록 권고하였다.⁹⁾ 연속발화 녹음 시 환자들의 상태에 따라 다양한 쉼 길이를 가질 수 있다. 이러한 무성음, 쉼 길이는 상대적으로 원래의 음질보다 더 중증의 수치로 측정된다.

본 연구에서 캡스트럼 분석을 위해 Praat을 활용하였

다. 무료로 배포된 프로그램이기 때문에 임상현장에서 누구나 쉽게 사용해 볼 수 있다. Praat은 다양한 스크립트를 적용하여 많은 수의 음성샘플을 자동으로 분석할 수 있게 해준다. 또한 결과를 텍스트파일로 출력할 수 있어 결과 분석 및 연구 진행에도 많은 도움을 준다.

따라서 본 연구의 목적은 연속발화 내의 무성음 구간을 제거, 분석한 결과가 노인성 음성장애의 감별에 유용한지 확인하고 이러한 음향학적 측정치와 청지각적 평가와의 상관성을 알아보고자 하였다.

대상 및 방법

연구 대상

본 연구의 대상은 부산 소재 대학병원에서 노인성 음성장애로 진단받은 대상자 381명을 대상으로 하였다. 노인성 음성장애군은 263명(mean age : 52.9±14.3, 남자 104명, 여자 159명), 대조군인 정상 음성군은 118명(mean age : 58.0±10.3, 남자 40명, 여자 78명)이었다. 노인성 음성장애의 진단은 선행연구에서 제시한 기준을 준수하였다.¹⁰⁾ 본 연구의 대상자 포함 기준은 차트리뷰, 후두내시경, 후두스트로보스코피 검사 시 궁형성대(vocal cord bowing), 성대위축(vocal atrophy), 성대돌기 돌출(prominence of vocal process), 성문폐쇄부전과 같은 소견을 확인하였다. 또한 제외 기준은 후두기능에 영향을 주는 신경학적 질병, 만성 퇴행성 질환, 성대 병변(성대결절, 성대용종, 육아종, 성대낭종, 성대마비), 알레르기, 인후두 급성감염 등이었다.

본 연구에서 정상음성 집단은 본원을 내원한 두경부 환자 중에서 후두질환이 없고 음성장애를 호소하지 않으며 GBRAS 평가에서 G척도가 0(normal voice)으로 판단된 환자들의 음성샘플들로 결정하였다. 본 연구는 획득된 음성 샘플을 후향적으로 분석한 연구로서 부산대학교병원 연구윤리위원회(Institutional Review Board)의 심의를 면제받았다(IRB No. H-1801-012-063).

연구의 절차

음성샘플

음성녹음은 음성검사실 내의 방음부스에서 Comput-

erized Speech Lab(CSL, Model 4500, Kay Electronic, USA)과 단일지향성 마이크인 Shure SM48(Shure Inc, USA)를 이용하였다. 음성샘플의 저장형태는 표본추출률(sampling rate) 44.1 kHz, 양자화(quantization) 16 bit로 하였다. 발성 과제에서 모음연장은 편안한 음도와 강도로 /아/ 모음을, 연속발화는 산책문단의 25음절을 읽었다.

Praat

Praat은 University of Amsterdam의 Paul Boersman와 David Weenick가 공동으로 개발한 음성분석 및 합성 프로그램으로 그 기능의 다양함과 방대함에 있어서는 다른 프로그램을 능가한다.¹¹⁾ Praat은 인터넷에서 검색기능을 활용하여 손쉽게 찾아낼 수 있으며 연구 목적으로 즉시 내려 받아 사용할 수 있다. 다양한 운영체제(Windows, Mac OS, Linux)에서 사용이 가능하고 PC에 설치할 필요가 없으며 지속적인 기능 개선, 오류 수정 등의 업데이트가 이루어진다. Pitch, spectrum, formants, fundamental frequency, jitter, shimmer 등 음성 분석 외에 다양한 조합의 분석 가능하다. 또한 Praat 스크립트를 이용하여 대량의 음성샘플을 자동으로 빠른 시간 내에 편집, 분석할 수 있다. 선행 연구나 구글링을 통해서 다양한 스크립트를 적용할 수 있으며, 측정 결과에 대한 분석 및 시각화를 위해서 R, RStudio에서 PraatR 패키지를 사용할 수도 있다.

캡스트럼 분석

CSL에서 녹음한 모음연장 샘플은 사운드에디터 창에서 select를 이용하여 정확하게 2초 구간을 편집하고, 연속발화 샘플은 발화의 시작과 끝 부분을 선택하여 불필요한 부분을 편집하였다. 모음연장 샘플은 무성음 구간 제거 스크립트는 적용하지 않고 연속발화 샘플에만 적용하였다. 캡스트럼 분석은 모음연장 샘플(sustained vowel samples, SV), 연속발화 샘플(continuous speech samples, CS), 무성음 구간이 제거된 연속발화 샘플(voiceless extracted samples, EXT)을 대상으로 실시하였다. 무성음 구간 제거와 캡스트럼 분석의 스크립트는 선행 연구에서 제시한 것을 적용하였다.¹²⁻¹⁴⁾ 유성음 추출 스크립트는 1) 사운드 에너지가 전체 신호 에너지의 30%를

초과하고, 2) 제로 크로싱 속도가 1,500 Hz 미만이고, 3) 정규화된 자기 상관 피크가 0.3보다 큰 경우, 30 밀리세컨드(milliseconds)의 프레임을 유성음으로 지정했다. LHRatio는 4,000 Hz 이하의 에너지와 4,000~8,000 Hz 사이의 에너지 비율을 말하며 너무 큰 값으로 측정되어 자연로그(natural logarithm ; ln)를 취하여 측정값을 제시하였다.

청지각적 평가

무성음 구간 제거 전의 음성샘플을 통해서 청지각적 평가를 실시하였다. 음성평가 및 음성치료 경력이 7년 이상인 언어재활사 1급 3명이 음성샘플에 대해서 청지각적 평가를 수행하였다. 평가자들은 대상자에 대한 사전 정보 없이 무작위로 제시된 음성샘플을 반복 청취하여 평가하였다. 모음연장과 연속발화에 대해서 각각 평가를 수행하였고, 전반적인 음성 중증도 확인을 위해서 GRBAS의 grade(G), Consensus Auditory Perceptual Evaluation-Voice(CAPE-V)의 overall severity(OS)를 평가하였다. 평가자 내 신뢰도 측정을 위해서 전체 음성 샘플 중 약 20%에 해당하는 75명의 음성샘플을 초기 평가 후 2주 뒤에 재평가하였다.

자료 처리

캡스트럼 분석에 대한 자료 처리는 R, version 3.5.2(The R Foundation for Statistical Computing, Vienna, Austria) and RStudio 1.1.463(RStudio Inc., Boston, MA, USA)을 이용하였다. 청지각적 평가에 대한 평가자 간, 내의 신뢰도 측정을 위해서 급내상관계수(intraclass correlation coefficient, ICC ; 단일측도)를 수행하였다. 그 결과 평가자 간 신뢰도[G : 0.72~0.93(mean : 0.87), OS : 0.71~0.91(mean : 0.82)], 평가자 내 신뢰도[G : 0.74~0.91(mean : 0.87, OS : 0.74~0.86(mean : 0.81))]으로 높은 수준이었다.

정상집단과 노인성 음성장애 집단 간의 캡스트럼 측정치 및 청지각적 평가의 차이가 있는지 알아보기 위해 독립표본 t-검정(independent t test)을 실시하였다. 또한 과업별 CPP, CPPS, LHRatio 측정값의 비교를 위해서 one way ANOVA를 실시하였다. 캡스트럼 측정치와 청지각적 평가 간의 상관관계를 알아보기 위하여 피어

는 상관분석(Pearson correlation coefficient)을 시행하였다. 또한 두 집단을 감별하기 위한 receiver operating characteristic curve(ROC) 분석을 실시하여 민감도(sensitivity)와 특이도(specificity), 곡선 하 영역(area under the curve, AUC)을 산출하였다. CS, EXT에서 측정된 CPP, CPPS의 AUC를 비교하기 위해서 DeLong's test¹⁵⁾를 수행하였다.

결 과

집단 간 음향학적 측정치 및 청지각적 평정 비교

집단 간 음향학적 측정치 및 청지각적 평정을 비교한 결과는 Fig. 1, 2와 Table 1에 제시되었다. 집단 간 음향학적 측정치를 비교한 결과, LHRatio_V, LHRatio_EXT를 제외한 모든 변수들이 정상집단과 노인성 음성장애 집단 간에 통계적으로 유의한 차이를 보였다($p < 0.01$).

집단 간 청지각적 평정을 비교한 결과, 정상집단은 노인성 음성장애 집단에 비해 유의하게 낮은 평정치를 보였다($p < 0.01$).

말과제에 따른 음향학적 측정치 비교

말과제에 따른 음향학적 측정치를 비교한 결과는 Table 2, Fig. 3에 제시되었다. CPP, CPPS의 측정에서 SV와 CS, SV와 EXT간에는 유의미한 차이($p < 0.01$)가 확인되었지만 CS와 EXT간에는 유의미한 차이가 없었다. 또한 LHRatio에서는 모든 말과제 간에 유의미한 차이는 확인되지 않았다.

음향학적 측정과 청지각적 평가 간의 상관관계

음향학적 측정과 청지각적 평가 간의 상관관계는 Table 3에 제시되었다. CPP, CPPS는 청지각적 평가와 0.80 이상의 상관관계를 보였으며($p < 0.01$), LHRatio는 청지각적 평가와는 0.20 이하의 낮은 상관관계를 나타내었다.

무성음 구간 추출에 따른 병리적 음성의 감별 예측도

무성음 분절의 추출에 따른 병리적 음성의 감별 예측도는 Fig. 4에 제시되었다. CS에서 CPP, CPPS의 감별 예측도는 각각 0.873, 0.845였으며 DeLong's test 결과, 두 감별 예측도 간에는 유의미한 차이가 있는 것으로 확

인되었다($p < 0.001$). 반면 EXT에서 CPP, CPPS의 감별 예측도는 각각 0.898, 0.907였으며 DeLong's test 결과, 두 예측도 간에는 유의미한 차이가 없는 것으로 확인되었다($p = 0.093$).

고 찰

본 연구는 노인성 음성장애 환자들의 연속발화에서 유성음 구간만을 추출하여 캡스트럼 분석을 실시하였다. 이를 통해 무성음 분절이 캡스트럼 분석 상에서 어떠한 영향을 미치는지를 알아보고자 하였다. 대부분의 선행 연구들은 연속발화(CS)에서 환자들의 개별 발화 특성이 반영되는 쉽, 호흡 등의 무성음 구간을 연구자가 말, 음성분석 프로그램을 이용하여 수작업(manual)으로 제거하였다. 이러한 수작업은 음성신호의 스펙트로그램을 눈으로 확인하여 구간을 지정하고 제거하는 방식이기 때문에 연구자의 주관 혹은 편향(bias)이 개입될 수 있다. 최근 Analysis of Dysphonia in Speech and Voice (ADSV, Model 5109, KayPENTAX)를 이용하여 자동화 된 음성음 탐지가 이루어졌지만 알고리즘의 옵션은 변경할 수 없고 음성샘플을 하나씩 분석해야 한다는 제한점이 있었다.^{16,17)} 본 연구에서는 기존에 수작업 혹은 ADSV에서 하나씩 분석하던 비효율적이고 부정확한 방식에서 벗어나 Praat의 스크립트를 이용하여 유성음 구간의 추출 및 캡스트럼 분석, 결과의 출력까지 모두 자동으로 진행하였다는 점에서 의미가 있다.

기존에는 캡스트럼 분석을 위해서 모음연장과 연속발화 과업을 수행하여 음성의 중증도를 측정하였다. 하지만 연속발화에서 대상자의 발화특성(호흡길이, 억양, 쉬 구간 등)이 이러한 중증도의 정량화에 영향을 준다는 것이 밝혀졌다. 그래서 동일한 길이의 문장 읽기를 실시하고 선행연구에서 제시한 유성음 추출(무성음 구간 제거) 스크립트를 통해서 좀 더 신뢰성 높은 음성분석을 하고자 본 연구를 수행하였다.

노인성 음성장애 집단은 보통 성대위축(atrophy), 궁형성대(bowing), 성대의 탈수(dehydration of the vocal fold) 등으로 인해 성대 점막과동 및 진폭 감소가 나타나고 성문폐쇄부전이 유발된다. 이로 인해 약한 음성, 거칠고 쉼 음성이 나타나고 발성길이가 짧아지고 음도가 변

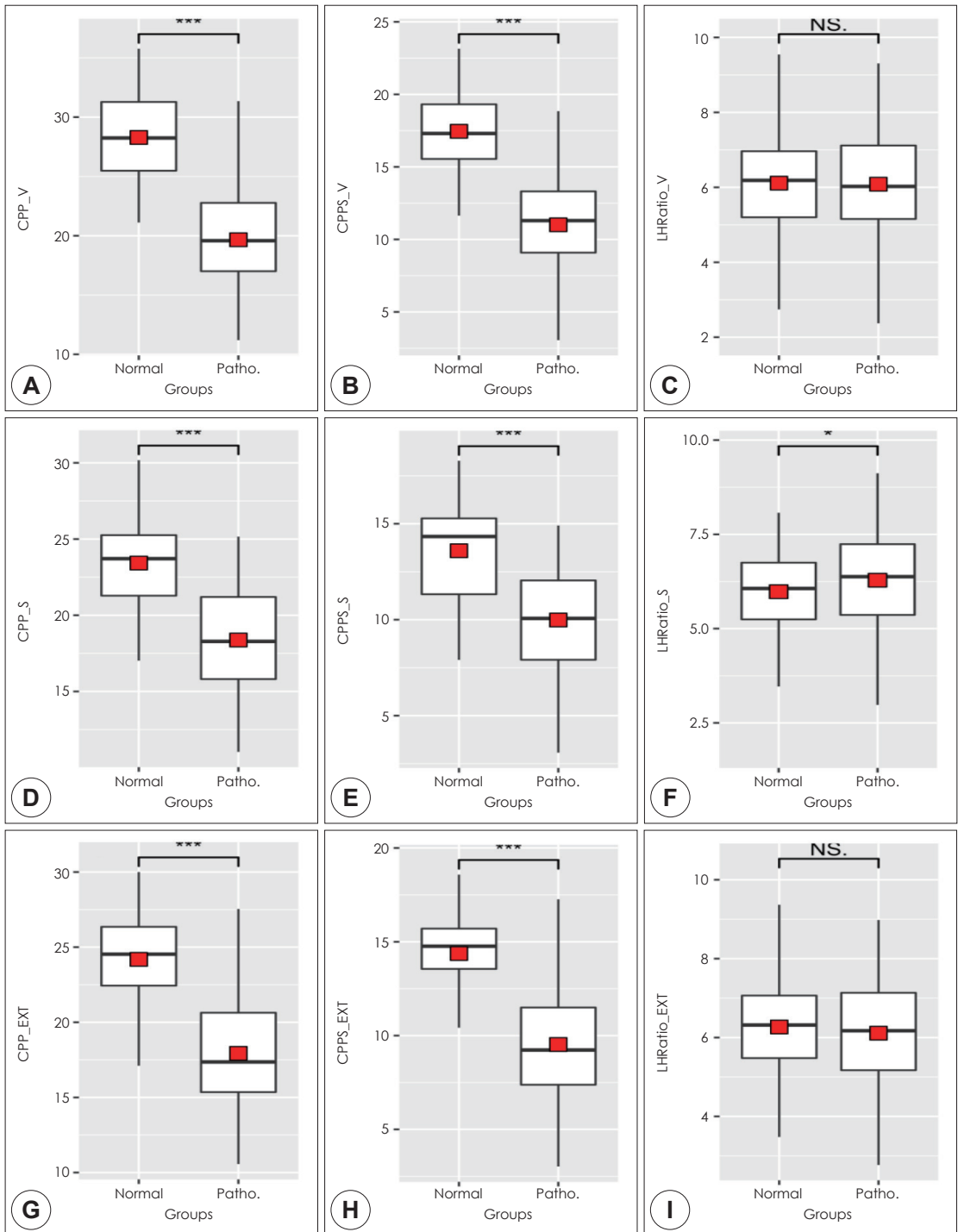


Fig. 1. Comparison of measured variables between normal and presbyphonia group. A : CPP_V. B : CPPS_V. C : LHRatio_V. D : CPP_S. E : CPPS_S. F : LHRatio_S. G : CPP_EXT. H : CPPS_EXT. I : LHRatio_EXT.

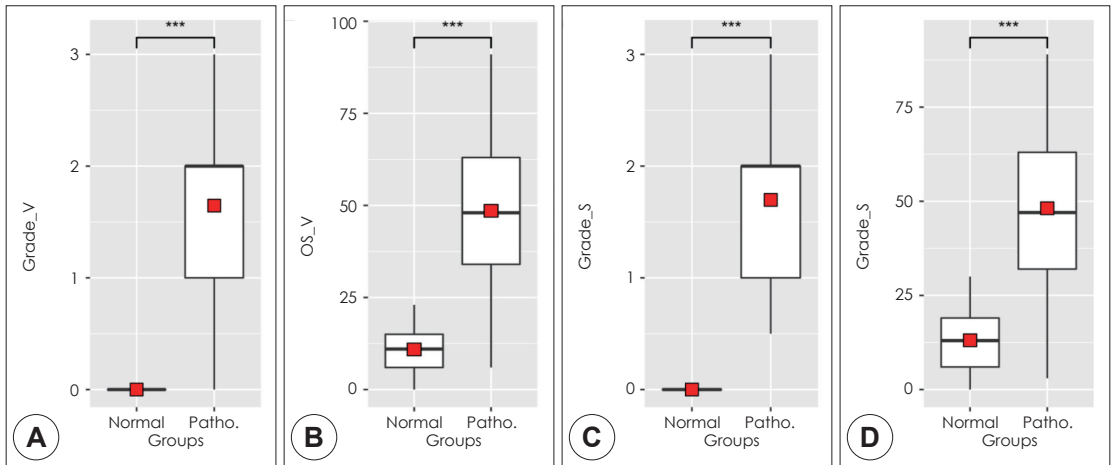


Fig. 2. Comparison of AP ratings between normal and presbyphonia group. A : Grade_V. B : OS_V. C : Grade_S. D : OS_S.

한다. 캡스트럼 분석은 주기성이 탐지되지 않는 기성음성에 대한 정량화가 뛰어난 것으로 알려져 있다. 노인성 음성장애는 심한 거칠고 쉼 목소리로 인해 섭동분석으로는 정확하게 정량화되지 않는다. 또한 노인들은 호흡의 길이가 매우 다양하기 때문에 그 자체만으로도 캡스트럼 측정치에 영향을 줄 수 있다. 이러한 호흡적 요소를 제외하고 발화 내의 유성음 구간만을 추출하여 분석한다면 음질에 대한 분석이 좀 더 정확하게 이루어질 수 있을 것이다.

본 연구의 결과, 노인성 음성장애 집단은 정상집단에 비하여 높은 CPP, CPPS, 낮은 청지각적 평정치(G, OS)를 나타냈고, LHRatio는 CS에서만 통계적으로 유의미한 차이를 보였다. 말과제에 따른 차이를 확인한 결과, 모든 측정변수들에서 CS와 EXT간에는 유의미한 차이가 나타나지 않았다. 그리고 CPP, CPPS는 청지각적 평가와 높은 상관관계가 있는 것으로 확인되었다. ROC 곡선 분석에서, CS와 EXT 모두 0.845 이상의 AUC로 높은 감별 예측력을 보였다. 연구 결과를 통해 다음과 같은 의의를 얻을 수 있었다.

첫째, 음향학적 변수들에 대한 두 집단의 결과를 비교하여 두 집단 간에 통계적으로 유의미한 차이가 있다는 것을 확인하였다. Praat을 이용한 CPP, CPPS 분석을 위해서 선행 연구에서 제시한 방법을 사용하였다. CPP, CPPS는 관련된 변수들의 값을 조정하여 각각 산출되었다. 기존 연구들도 SV의 측정치가 CS보다는 높게 나타

Table 1. Comparison of acoustic measurements and auditory-perceptual assessments according to groups

Parameters	Normal	Presby	p-value
Age	58.1 ± 12.1	56.1 ± 15.0	0.383
CPP_V	28.3 ± 3.6	19.7 ± 3.8	<0.01 †
CPPS_V	17.5 ± 2.5	11.0 ± 3.0	<0.01 †
LHRatio_V	6.1 ± 1.3	6.1 ± 1.4	0.851
CPP_S	23.4 ± 2.6	18.4 ± 3.3	<0.01 †
CPPS_S	13.6 ± 2.3	10.0 ± 2.6	<0.01 †
LHRatio_S	6.0 ± 1.2	6.3 ± 1.2	0.022*
CPP_EXT	24.2 ± 3.2	17.9 ± 3.5	<0.01 †
CPPS_EXT	14.4 ± 2.3	9.5 ± 2.8	<0.01 †
LHRatio_EXT	6.3 ± 1.2	6.1 ± 1.3	0.256
Grade_V	0.0 ± 0.0	1.6 ± 0.7	<0.01 †
OS_V	10.9 ± 5.8	48.5 ± 19.0	<0.01 †
Grade_S	0.0 ± 0.0	1.7 ± 0.7	<0.01 †
OS_S	13.1 ± 7.0	48.1 ± 18.7	<0.01 †

* : p < 0.05, † : p < 0.01. Presby : presbyphonia, CPP_V : cepstral peak prominence of sustained vowel, CPPS_V : smoothed cepstral peak prominence of sustained vowel, LHRatio_V : L/H spectral ratio of sustained vowel, CPP_S : cepstral peak prominence of continuous speech, CPPS_S : smoothed cepstral peak prominence of continuous speech, LHRatio_S : L/H spectral ratio of continuous speech, CPP_EXT : cepstral peak prominence of extracted continuous speech, CPPS_EXT : smoothed cepstral peak prominence of extracted continuous speech, LHRatio_EXT : L/H spectral ratio of extracted continuous speech, Grade_V : degree of Grade on sustained vowel, OS_V : degree of overall severity on sustained vowel, Grade_S : degree of Grade on continuous speech, OS_S : degree of overall severity on continuous speech

Table 2. Descriptive Statistics stratified by 'Task' and 'Groups'

Task	SV			CS			EXT		
	Normal	Presby	p	Normal	Presby	p	Normal	Presby	p
CPP	28.3±3.6	19.7±3.8	<0.01*	23.4±2.6	18.4±3.3	<0.01*	24.2±3.2	17.9±3.5	<0.01*
CPPS	17.5±2.5	11.0±3.0	<0.01*	13.6±2.3	10.0±2.6	<0.01*	14.4±2.3	9.5±2.8	<0.01*
LHRatio	6.1±1.3	6.1±1.4	0.851	6.0±1.2	6.3±1.2	0.022	6.3±1.2	6.1±1.3	0.256

* : p<0.01. Presby : presbyphonia, CPP : cepstral peak prominence, CPPS : smoothed cepstral peak prominence, LHRatio : L/H spectral ratio, SV : sustained vowel, CS : continuous speech, EXT : extracted continuous speech

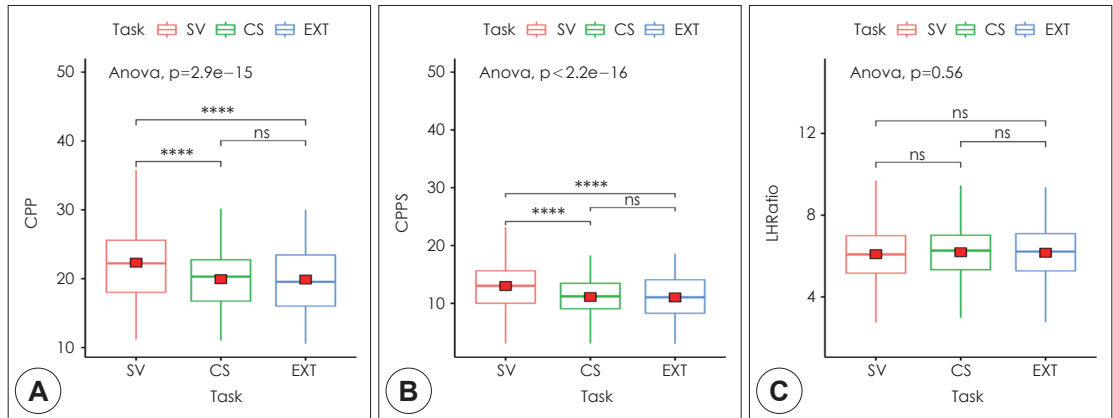


Fig. 3. Comparison of acoustic measurements according to voice task. A : CPP. B : CPPS. C : LHRatio.

났으며, 또한 CPP가 CPPS보다 큰 값으로 측정되었다고 보고되어 본 연구의 결과와 일치하였다.^{14,18)} LHRatio는 원래의 값들이 0~400 사이의 값들로 분포되어 다루기 힘들기 때문에 작게 축소하기 위해서 자연 로그(ln)를 적용하였다. LHRatio에서 유의미한 차이가 나타나지 않은 것은 자연 로그를 취해서 측정값들이 작아졌기 때문 인 것으로 생각할 수 있다. 추후 연구에서는 자연 로그를 취하지 않고 집단 간 비교를 하여 차이를 확인해야 할 것이다. 선행연구에서도 LHRatio는 음성장애의 중증도를 평가하는데 상대적으로 일관성이 떨어진다고 보고되었다.¹⁶⁾ 청치각적 평가에서도 두 집단 간에는 유의미한 차이가 나타났고 선행 연구의 결과들을 뒷받침하였다.^{19,20)}

둘째, 말과제에 따른 음향학적 측정치를 비교하였다. SV는 CS보다 큰 CPP, CPPS를 나타냈고, EXT는 CS보다 큰 CPP, CPPS를 산출하였다. SV가 CS보다 큰 CPP, CPPS를 갖는다는 것은 선행 연구들에서 많이 보고되었지만 유성음 구간만을 추출하여 분석한 연구는 많지 않다. 기존의 연구들은 본 연구와 같이 Praat 스크립트로 일관적으로 무성음 구간을 제거한 것이 아니라 연구자

Table 3. Correlation analyses between acoustic variables and AP ratings

Parameters	Grade_V	OS_V	Grade_S	OS_S
CPP_V	-0.86*	-0.85*		
CPPS_V	-0.86*	-0.86*		
LHRatio_V	-0.16	-0.20		
CPP_S			-0.83*	-0.86*
CPPS_S			-0.80*	-0.84*
LHRatio_S			-0.08	-0.14
CPP_EXT			-0.69*	-0.66*
CPPS_EXT			-0.68*	-0.66*
LHRatio_EXT			-0.16	-0.15

* : p<0.01

가 주관적으로 제거하여 분석하였다.^{21,22)} 이러한 문제로 인해 무성음 구간을 제거하기 보다는 그대로 포함시켜 캡스트럼 분석을 하도록 권고하였다.¹⁶⁾ 본 연구의 결과를 보면, 무성음 구간을 제거한 이후의 캡스트럼 측정치는 제거 전보다는 더 높은 CPP, CPPS 값을 나타내었다.

무성음 구간을 제거하는 것이 음성장애 감별 예측도를 상승시킬 수 있을지 확인하기 위해서 ROC 곡선 분석

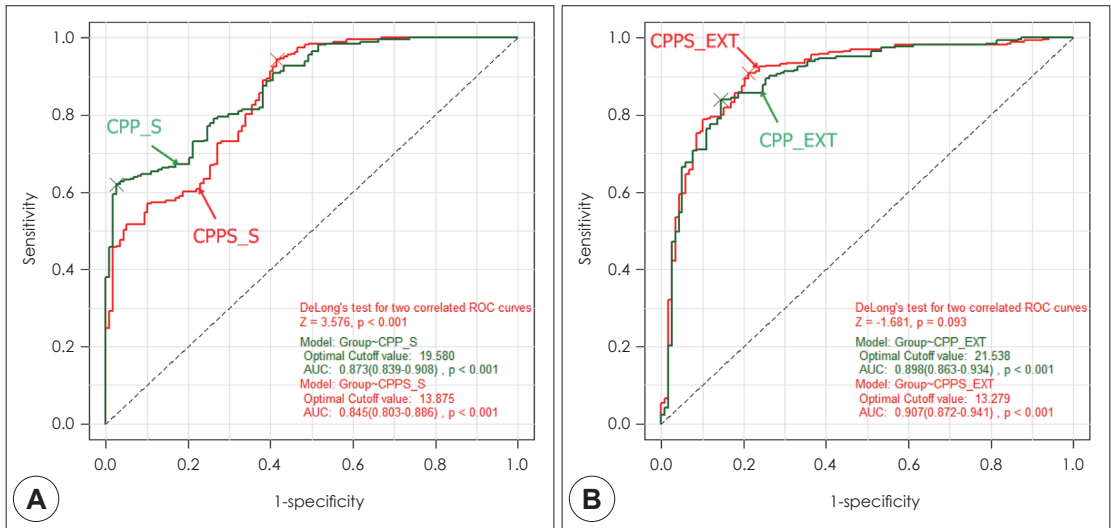


Fig. 4. ROC analysis. A : continuous speech. B : extracted continuous speech.

도 실시하였다. AUC를 확인하여 무성음 구간을 제거하는 것이 더욱 효과적인지 확인하였다. CS, EXT 과제에서 CPP, CPPS는 0.845 이상의 AUC를 보였고, DeLong's test 결과 CS에서 CPP, CPPS 간의 AUC는 유의미한 차이로 밝혀졌다. EXT에서는 CPP, CPPS간의 AUC는 유의미한 차이가 확인되지 않았다. 무성음 구간 제거 이후 CPP, CPPS의 감별 예측도는 모두 상승하였고 두 측정 변수들도 비슷한 감별 예측도로 확인되었다. 이 결과로 보아 무성음 구간 제거 이후에는 CPP, CPPS가 비슷한 수준의 감별 예측도를 나타내어 일정한 수행능력을 보였음을 알 수 있다.

셋째, 음향학적 측정과 청지각적 평가 간에는 높은 상관관계가 확인되었다. Grade_V에서 CPP, CPPS($r = -0.86$), OS_V에서는 CPP($r = -0.85$), CPPS($r = -0.86$)로 나타났다. 또한 Grade_S에서 CPP($r = -0.83$), CPPS($r = -0.80$), OS_V에서는 CPP($r = -0.86$), CPPS($r = -0.84$)로 유의미한 상관관계가 있는 것으로 파악되었다. 반면에 LHRatio는 청지각적 평가인 Grade, OS, 말과제인 SV, CS 모두에서 상관관계가 낮은 것으로 확인되었다. 이러한 결과들도 선행연구의 결과를 뒷받침한다.¹⁶⁾

연속발화 샘플의 무성음 추출에 대한 의견은 예전부터 논의의 대상이 되어왔다. 무성음 구간이 캡스트럼 분석의 측정치에 영향을 주기 때문에 모음 샘플과의 비교

에 제한이 있었다. 선행 연구에서는 연구자가 주관적으로 파형을 보면서 무성음 구간을 제거하여 분석하였다. 하지만 이러한 방식은 일관된 기준에 의해서 무성음이 제거되는 것이 아니기 때문에 신뢰성이 낮았다. 객관적으로 제거할 수 있는 방식이 제시되지 않아서 무성음 구간을 포함하여 분석하도록 권고하였다.^{21,22)}

본 연구에서 캡스트럼 분석에 영향을 주는 무성음 분절을 Praat 스크립트를 통해서 추출하고 분석하였다. 본 연구의 결과를 통해서 이러한 방식이 정상음성과 병리적 음성을 감별하는데 도움을 주고 CPP, CPPS의 감별 예측도 또한 높은 수준을 유지할 수 있다는 것을 확인하였다. 노인 성대의 병태생리학적 변화로 인해 성문폐쇄가 나타나고 병리적 음성의 산출을 유발하였다. 이러한 특성이 캡스트럼 측정치에 반영이 되었고 본 연구를 통해서 정상음성과의 차이를 확인할 수 있었다. 성문폐쇄를 유발하는 음성장애 질환은 다양하며 추후 연구에서는 노인성 음성장애를 포함한 대부분의 음성장애군을 포함하여 이러한 무성음 추출 방식이 음성장애의 감별 예측도를 향상시키는지 확인해야 할 것이다.

결론

본 연구에서는 노인성 음성장애 집단을 대상으로 연

속발화 내의 유성음 구간만을 추출하여 캡스트럼 분석을 실시하였고 음성장애를 감별하는데 도움이 된다는 것을 확인하였다. 유성음 구간을 추출한 음성샘플을 분석하여 CPP, CPPS의 감별 예측력이 향상되고 두 변수의 AUC가 비슷한 수준으로 증가하였다. 임상현장에서 유성음 추출 후의 캡스트럼 분석이 노인성 음성장애의 진단 및 평가에 도움이 될 것으로 생각된다.

중심 단어 : 음성장애 · 음향학 · 노화 · Praat.

REFERENCES

- 1) Martins RH, Benito Pessin AB, Nassib DJ, Branco A, Rodrigues SA, Matheus SM. Aging voice and the laryngeal muscle atrophy. *Laryngoscope* 2015;125(11):2518-21.
- 2) Bradley JP, Hapner E, Johns MM 3rd. What is the optimal treatment for presbyphonia? *Laryngoscope* 2014;124(11):2439-40.
- 3) Jin SM. Voice Evaluation. *J Clinical Otolaryngol* 2015;26(1):14-22.
- 4) Ravi SK, Shabnam S, George KS, Saraswathi T. Acoustic and aerodynamic characteristics of choral singers. *J Voice* 2018;S0892-1997(17)30587-8.
- 5) Dejonckere PH, Remacle M, Fresnel-Elbaz E, Woisard V, Crevier-Buchman L, Millet B. Differentiated perceptual evaluation of pathological voice quality: reliability and correlations with acoustic measurements. *Rev Laryngol Otol Rhinol (Bord)* 1996;117(3):219-24.
- 6) Hillenbrand J, Houde RA. Acoustic correlates of breathy vocal quality: dysphonic voices and continuous speech. *J Speech Lang Hear Res* 1996;39(2):311-21.
- 7) Patel RR, Awan SN, Barkmeier-Kraemer J, Courey M, Delyski D, Eadie T, et al. Recommended protocols for instrumental assessment of voice: american speech- language-hearing association expert panel to develop a protocol for instrumental assessment of vocal function. *Am J Speech Lang Pathol* 2018;27(3):887-905.
- 8) Kim GH, Lee YW, Bae IH, Park HJ, Wang SG, Kwon SB. Efficacy of the Acoustic Voice Quality Index As a Measure of Dysphonia Severity in Age-Related Dysphonia. *J Clinical Otolaryngol* 2018;29(1):63-71.
- 9) Awan SN, Roy N, Dromey C. Estimating dysphonia severity in continuous speech: application of a multi-parameter spectral/cepstral model. *Clin Linguist Phon* 2009;23(11): 825-41.
- 10) Crawley BK, Dehom S, Thiel C, Yang J, Cragoe A, Mouselli I, et al. Assessment of clinical and social characteristics that distinguish presbylaryngis from pathologic presbyphonia in elderly individuals. *JAMA Otolaryngol Head Neck Surg* 2018;144(7):566-71.
- 11) Boersma P. Praat, a system for doing phonetics by computer. *Glott International* 2001;5(9):341-7.
- 12) 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.e1-7.
- 13) 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.
- 14) Phadke KV, Laukkanen AM, Ilomaki I, Kankare E, Geneid A, Svec JG. Cepstral and perceptual investigations in female teachers with functionally healthy voice. *J Voice* 2018; S0892-1997(18)30310-2.
- 15) Delong ER, DeLong DM, Clarkepearson DI. Comparing the areas under 2 or more correlated receiver operating characteristic curves-a nonparametric approach. *Biometrics* 1988; 44(3):837-45.
- 16) 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.
- 17) Lowell SY, Colton RH, Kelley RT, Hahn YC. Spectral- and cepstral-based measures during continuous speech: capacity to distinguish dysphonia and consistency within a speaker. *J Voice* 2011;25(5):e223-32.
- 18) 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 Otorrinolaringol Esp* 2018;69(3):134-40.
- 19) 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.
- 20) Karnell MP, Melton SD, Childes JM, Coleman TC, Dailey SA, Hoffman HT. Reliability of clinician-based (GRBAS and CAPE-V) and patient-based (V-RQOL and IPVI) documentation of voice disorders. *J Voice* 2007;21(5):576-90.
- 21) Awan SN, Roy N. Acoustic prediction of voice type in women with functional dysphonia. *J Voice* 2005;19(2):268-82.
- 22) Awan SN, Roy N, Jetté ME, Meltzner GS, Hillman RE. Quantifying dysphonia severity using a spectral/cepstral-based acoustic index: comparisons with auditory-perceptual judgments from the CAPE-V. *Clin Linguist Phon* 2010;24(9):742-58.