

Clinical Usefulness of Voice Recordings using a Smartphone as a Screening Tool for Voice Disorders

Seung Jin Lee^a, Kwang Yong Lee^b, Hong-Shik Choi^a

^aDepartment of Otorhinolaryngology and the Institute of Logopedics and Phoniatrics, Yonsei University College of Medicine, Seoul, Korea

^bDivision of Theatre, College of the Arts, Dongguk University, Seoul, Korea

Correspondence: Hong-Shik Choi, MD, PhD
Department of Otorhinolaryngology and the Institute of Logopedics and Phoniatrics, Yonsei University College of Medicine, 211 Eonju-ro, Gangnam-gu, Seoul 06273, Korea
Tel: +82-2-2019-3461
Fax: +82-2-3463-4750
E-mail: hschoi@yuhs.ac

Received: October 8, 2018

Revised: November 10, 2018

Accepted: November 10, 2018

Objectives: The purpose of the present study was to determine whether selected acoustic measures and estimates differ according to the recording devices, the Computerized Speech Lab (CSL) and a smartphone, and groups. In addition, correlation and reliability between the devices were explored. Lastly, the diagnostic ability of the two devices were compared.

Methods: A total of 180 participants (90 patients and 90 controls) participated in the study. Vowel samples of the participants were recorded via the CSL and an android smartphone. For each sample, acoustic parameters including jitter percent (Jitt), shimmer percent (Shim), noise-to-harmonic ratio (NHR), cepstral peak prominence (CPP), L/H ratio, Cepstral Spectral Index of Dysphonia (CSID) were measured. Two-way repeated measures of ANOVA, Pearson correlation analysis, and ROC curve analysis were performed. **Results:** Interaction effect for Jitt, Shim, NHR, and σCPP between groups and devices were found. F0 measures of the smartphone were higher than those of the CSL. CPP and SR of the CSL and patients were higher than those of the smartphone and controls, respectively. In contrast, CSID and σSR were higher for the smartphone. All the parameters showed positive correlation between devices. There was no difference in the area under curve between the devices, although substantially different cut-off scores were obtained. **Conclusion:** The current data showed a significant correlation of acoustic measures and no difference in the diagnostic ability between the devices, although differences in several measures and higher cut-off scores of the smartphone were noted. In conclusion, smartphones can be used as a screening tool for voice disorders.

Keywords: Voice disorders, Smartphone, Screening tool, Acoustic analysis, Cepstral analysis, Cepstral Spectral Index of Dysphonia (CSID)

음성에 대한 다면적 평가 방법들 가운데 하나인 음향학적 분석은 임상 및 연구에서 음성장애의 진단 및 평가에 널리 이용되어 왔다. 음향학적 분석을 위한 다양한 분석 도구들이 있는데, 대표적인 것 중의 하나로 ‘Computerized Speech Lab (CSL)’ (Model 4150B; KayPENTAX, Lincoln Park, NJ, USA)을 꼽을 수 있다. CSL은 자체적으로 신뢰성 있는 음성 샘플 표집 및 재생에 충분한 고성능의 하드웨어를 포함한다. 이 하드웨어가 CSL의 주 프로그램(main program) 외에 여러 다른 모듈 프로그램들과 호환된다는 점에서 볼 때, 대개 규격화된 하드웨어 없이 소프트웨어만으로 구성되는 다른 도구들과 차이가 있다. 또한 프로그램 내에서 국내의 자료는 아니나

자체적인 규준 자료와 다양한 프로토콜들을 제공하고, 시각화가 비교적 용이하다는 장점으로 인하여 임상에서 널리 이용되어 왔다.

이러한 호환 프로그램들 중에서 특히 ‘Multi-Dimensional Voice Program (MDVP)’ (Model 5105, KayPENTAX)이 널리 이용되어 왔으나, 최근 캡스트럼 기반의 분석이 임상 및 연구에서 활발히 이루어지면서 ‘Analysis of Dysphonia in Speech and Voice (ADSV)’ (Model 5109, KayPENTAX) 또한 활발히 이용된다(Lee, Lim, & Choi, 2017). MDVP를 통해서는 모음 샘플에 대하여 전통적으로 임상 및 연구에서 널리 분석에 활용되어 온, 시간 기반 분석을 토대로 한 변수들을 측정할 수 있으며, ‘기본주파수(fundamental fre-

quency, F0), ‘주파수변동율(jitter percent, Jitt)’, ‘진폭변동율(shimmer percent, Shim)’, ‘소음대배음비(noise-to-harmonic ratio, NHR)’ 등이 대표적인 예이다. 이들 변수들 중 특히 Jitt, Shim, NHR은 각각 1.04, 3.81, .19 등의 역치가 널리 알려져 임상에서 유용하게 이용되어 왔다(KayPENTAX, 2008). 반면, ADSV 프로그램에서 측정 가능한 대표적인 캡스트럼 측정 변수로는 캡스트럼 피크 현저성(cepstral peak prominence, CPP), L/H 비율(low-to-high spectral ratio, SR)과 각 표준편차(σ CPP, σ SR)도 측정할 수 있으며, 모음뿐만 아니라 연결발화 샘플에 대한 분석이 가능하다는 장점이 있다(KayPENTAX, 2011). 한국에서도 이미 여러 환자군에 적용된 연구들이 속속 보고되고 있으며(Kang & Koo, 2015; Kim et al., 2016; Park, Mun, Lee, & Jin, 2013; Seo et al., 2016; Shim et al., 2016), 최근 이러한 변수들에 대하여 한국어 정상 화자의 모음 및 연결발화 규준 자료에 대한 연구들이 이루어지기도 하였다(Choi & Choi, 2016; Lee, Pyo, & Choi, 2018).

이러한 캡스트럼 스펙트럼 분석을 기반으로 한 ADSV의 측정치들은 주기성을 전제로 한 MDVP의 전통적 측정치들에 비해 특히 중증도가 높은 장애 음성에서 유용하게 활용될 수 있는 것으로 알려져 있다(Lee, Pyo, et al., 2018). 나아가 이러한 측정치들을 기반으로 전문가의 청지각적 평정치, 보다 구체적으로는 Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V; Kempster, Gerratt, Abbott, Barkmeier-Kraemer, & Hillman, 2009)의 전반적 중증도 항목(overall severity)를 예측할 수 있는 유용한 지수인 캡스트럼 스펙트럼 발성장애 지수(Cepstral Spectral Index of Dysphonia, CSID)가 개발되었다(Awan, Roy, Zhang, & Cohen, 2016). CSID는 어떠한 말 샘플을 분석하였느냐에 따라 상이한 모델이 적용되며, ADSV 프로그램은 각 성별에 따라 모음에 해당하는 CSID, 그리고 성별에 무관하게 연결발화에 해당되는 CSID 수치를 자동으로 계산하여 제시해주는 유용한 기능을 지니고 있다(KayPENTAX, 2011). 각 변수별로 가중치를 달리 하여 만들어진 CSID의 회귀식은 비록 영어 말 자료를 이용하여 만들어진 모델이나, ‘산책’ 문단의 일부 문장 등 한국어 말 자료를 이용하여 유도한 말 샘플에 적용하는 것도 적절하다는 점이 선행연구들을 통해 증명되어 왔다. 예컨대, 내전형 연축성 발성장애(adductor spasmodic dysphonia)와 근긴장성 발성장애(muscle tension dysphonia) 환자의 과제에 따른 중증도 차이 측면에서 상이한 양상을 보이는 것이 CSID의 양상에서도 동일하게 드러남으로써, 이러한 적용 가능성성이 확인된다(Shim et al., 2016).

그러나 CSL은 프로그램과 자체 하드웨어를 포함하여 상용화되어 있어 비교적 고가이므로, 모든 임상 및 연구 세팅에서 구비할 수

는 없는 실정이다. 또한 자체 하드웨어 및 모니터링 스피커 등을 데스크탑 컴퓨터에 인스톨하여야 하므로, 휴대성이 떨어져 임상 및 연구에서 이용에 제한이 있다는 단점이 있다. 이러한 맥락에서 CSL을 대체할 수 있을 만한 장비들을 이용하여, 음향학적 분석의 신뢰도에 대한 연구들이 최근 비교적 활발하게 이루어져 왔으며(Van Leer, Pfister, & Zhou, 2017), 그 가운데 대표적인 것으로 스마트폰을 꼽을 수 있다. 스마트폰은 한편으로는 최근 성능 측면에서 상향 평준화가 이루어져 왔으며, 휴대성 및 이용 편의성 측면에서 큰 장점을 지니고 있는 기기이다. 또한 한국방송광고진흥공사(<https://kobaco.co.kr>)의 ‘2016년 시청점유율 결과보고서’에 따르면, 전체 응답자의 약 84.1%가 스마트폰을 사용하고 있다고 응답할 정도로 국내에 널리 보급되어 있다는 장점도 있다. 그러나 다른 한편으로는 기기나 제조사에 따라 자체 내장된 마이크나 스피커의 성능이 다르므로 녹음 및 재생 성능의 일관성이 담보되기 힘들며, 운영체제에 따라 녹음을 위한 애플리케이션이 다르므로, 이로 인하여 녹음의 세부적인 설정이 달라지고 이에 기인한 음성샘플의 차이가 발생할 가능성 또한 배제하기 어렵다는 단점도 있다.

이에 따라 스마트폰을 통해 녹음한 음성샘플을 분석하였을 때, 그러한 음향학적 측정치들이 과연 신뢰성이 있을 것인가에 대해서는 다소 논란의 여지가 있을 수 있다. 예컨대 스마트폰을 이용한 측정치가 기존 장비를 이용한 결과와 상당한 양의 차이를 보이거나(Vogel, Rosen, Morgan, & Reilly, 2014), 소음에 많은 영향을 받는다는 결과(Lebacq et al., 2017)들은 스마트폰의 활용도에 의문을 제기할 수 있는 연구결과들이다. 특히 CSL은 아니지만 고사양의 녹음기기와 아이폰, 전화통화 품질의 녹음을 비교한 선행연구에서는 고사양 녹음기기와 아이폰 간 비교 결과, F0와 NHR의 경우 비교적 높은 상관관계가 있었던 반면, Jitt, Shim, 그리고 특히 CPP에서 나타난 상관관계는 유의하지 않았으며, 기기 간 신뢰성 있는 측정이 가능하다고 볼 근거가 희박하다고 하였다(Vogel et al., 2014). 반면에 적어도 일부 측정치들은 기존의 기기들을 이용한 측정치들과 대동소이하다는 연구결과들(Manfredi et al., 2017; Maryn, Ysenbaert, Zarowski, & Vanspauwen, 2017; Oliveira, Fava, Baglione, & Pimpinella, 2017; Yun, Lee, Lee, & Jin, 2015) 또한 존재하고 있다. 예컨대 Yun 등(2015)은 삼성 갤럭시 S5 스마트폰과 디지털 녹음기, MDVP의 측정치를 비교한 결과 F0, Jitt, Shim, F1, F2, F3, F4의 기기 간 차이가 없고, 기기별 측정치 간 강한 상관관계가 있다고 보고하였다. 단순히 기기 간의 측정치 차이나 상관관계를 살펴보는 것을 넘어서서 스마트폰을 이용한 음향학적 측정치가 음성장애에 대하여 갖는 선별검사 도구로서의 유용성을 살펴본 연구는 드물지만, 적어도 선별 도구로 유용하게 이용될 수 있는 가능성성이 일부 제

시되었다(Uloza et al., 2015). 동일한 분석 프로그램을 사용하더라도 서로 다른 녹음기기들을 사용한다면, 기기별로 음향학적인 측정치가 서로 차이를 나타내는 것은 어찌 보면 당연한 일일 수도 있다. 이러한 차이에도 불구하고, 선별검사 도구로서의 이용 가능성이 확인될 경우 임상 및 연구에서의 활용도가 매우 클 것으로 보인다. 예컨대 스마트폰에서 적용할 수 있는 별도의 절단점이 마련될 수 있다면, 이를 토대로 선별검사를 간단하게 시행하여 결과를 살펴보고, 필요한 경우 보다 정밀하고 다면적인 음성검사와 의학적 진단을 위해 음성을 리닉에 내원하도록 권유함으로써 적절한 조기 중재를 시행할 가능성을 확대할 수 있을 것으로 보인다.

다만 스마트폰을 이용한 음향학적 분석 연구를 진행한다면, 적절한 사양의 스마트폰을 선정하는 일이 매우 중요할 것이라고 여겨진다. 전술한 바와 같이 스마트폰 기종에 따라 자체 내장된 마이크나 스피커의 성능뿐만 아니라 스마트폰 자체의 사양 또한 매우 다양할 것이기 때문이다. 그러나 선행연구들에서는 대부분 어떠한 스마트폰을 선택하였는지에 대한 근거를 제시하지는 않았다. 이러한 점들을 고려하면, ‘현재 가장 널리 보편적으로 이용될 만한 사양의 스마트폰 사양’을 추측하여 이를 적용함으로써 활용도를 높이는 방향이 바람직해 보인다. 먼저 출시된 다음 너무 오랜 시간이 경과한 기종이나, 혹은 반대로 가장 최신의 기종을 택할 경우 이와 유사한 사양의 스마트폰을 사용하는 사람의 수가 적어 연구결과의 실질적인 활용도가 다소 떨어질 우려가 있을 것으로 판단된다. 이러한 요인을 적절하게 고려하기 위해서는 근래의 우리나라 스마트폰 교체 주기를 살펴볼 필요성이 있는데, 한국인터넷진흥원(<https://www.kisa.or.kr/>)의 ‘2016년 인터넷이용실태조사 최종보고서’에 따르면 스마트폰 교체 주기는 평균 2년 7개월로 점차 길어지는 추세를 보인다고 한다. 이는 스마트폰의 전반적인 고사양화와 무관하지 않을 것으로 보인다. 이러한 점을 고려하면 출시된 후 대략 2-3년이 경과한 스마트폰을 이용하는 것이 적절할 것으로 여겨진다. 또한 앞서 언급한 ‘2016년 시청점유율 결과보고서’에서 안드로이드 운영체제 스마트폰 이용자의 비율이 78.2%로 나타난 점을 고려할 때, iOS 기반의 스마트폰보다는 안드로이드 운영체제의 스마트폰을 선택하여 연구에 이용하는 것이 활용도를 제고할 수 있는 방안인 것으로 판단된다. 아울러 국내의 한 선행연구에서 갤럭시 S5를 선택한 점을 고려하여 동일 제조사나 유사한 출시 시기(2014년)의 스마트폰을 선택하는 것도 고려해 봄 직하다.

따라서 이 연구에서는 CSL 하드웨어와 비교적 널리 보급된 정도의 사양을 지닌 기종의 안드로이드 스마트폰을 이용하여 녹음한 모음 발성 샘플에서 집단과 기기에 따른 음향학적 측정치의 차이가 있는지 알아보고, 각 기기별 측정치 간의 상관관계와 신뢰도를

알아보고자 하였다. 또한 각 기기별로 주요 음향학적 측정치들뿐만 아니라, 이들을 이용하여 얻은 전반적 중증도 추정치인 CSID가 음성장애에 대해 가질 수 있는 선별검사 도구로서의 유용성을 확인하고, 나아가 기기 간에 서로 비교하고자 하였다.

연구방법

연구대상

G*Power (Faul, Erdfelder, Buchner, & Lang, 2009)를 이용하여 집단(2개)과 기기(2개)에 따른 비교(two-way repeated measures of ANOVA)를 위한 최소의 표본수를 산정한 결과 총 대상자 수가 176명이었다(effect size $f=.15$, $\alpha=.05$, Power = .95). 따라서 대상자 수는 환자군과 대조군 각 90명으로 하였다.

먼저 환자군은 서울의 한 대학병원 이비인후과 음성클리닉에 내원하여 이비인후과 전문의에 의해 여러 병인의 음성장애로 진단된 19-60세의 음성장애 환자(남 25명, 여 65명, 평균연령 37.79 ± 10.92 세)로 하였다. 환자군 중 음성장애와 관련된 다른 신경학적 병력, 수술력이 있거나, 대조군 중 최근 1개월 내 음성장애나 흡연 이력이 보고된 자는 제외하였다. 정상 대조군 90명은 저자들의 지인들 중에서 서울 지역에 거주 중이며 한국어 모국어 화자인 정상 성인과 더불어, 갑상샘 수술 전 음성검사를 위해 내원한 대상자 가운데 환자군과 연령 및 성별이 정확하게 일치하도록 선정하였다. 잠재적 음성문제를 배제하기 위하여 면담, 연결 발화에 대한 청지각 평가, 심리측정적 평가, 후두 스트로보스코피를 시행하였다(Awan, Roy, & Cohen, 2014; Lee, Choi, & Kim, 2018). 이를 통해 후두의 기질적인 이상 소견을 배제함과 더불어 흡연력이 없으며, 최근 3개월 이내 감기 등의 증상으로 이비인후과에 내원한 적이 없으며, 심리측정적, 청지각적 평가에서 상당한 양의 음성장애가 보고되지 않은 자로 한정하였다. 심리측정적 기준으로는 음성 활동 및 참여 프로파일-한국판(K-VAPP)을 시행한 결과 총점이 14.5점 이상일 경우(Lee et al., 2016), 그리고 청지각적 기준으로는 각 대상자로 하여금 ‘가을’(Kim, 2012) 문단을 읽도록 하여 제1저자가 GRBAS 척도(Hirano, 1981)로 청지각적 평정을 시행한 결과, Grade (G) 척도에서 1 이상의 평정치를 보인 경우 연구대상에서 제외하였다. 각 집단에 따른 인구학적 데이터와 환자군의 진단명이 Table 1에 제시되어 있다.

연구절차

녹음 장비 및 소프트웨어로는 CSL 하드웨어, SM48 단일지향성 다이나믹 마이크(SHURE, Niles, IL, USA), MDVP advanced 프로그램과 2015년 출시된 안드로이드 스마트폰인 삼성 갤럭시 노트5

Table 1. Demographic data of the participants

Parameter	Patient group (N=90)	Control group (N=90)	p-value
Age (yr)	37.79±10.92 (19-60)	37.79±10.92 (19-60)	1.000
Gender			1.000
Male	25 (27.8)	25 (27.8)	
Female	65 (72.2)	65 (72.2)	
Professional voice user	44 (48.89)	39 (43.33)	.456
G of the GRBAS scale	2.18±.62	.00±.00	<.001***
K-VAPP			
Total score ^a (280)	103.15 (86.06)	1.55 (3.53)	<.001***
ALS (100)	34.74 (30.85)	.63 (1.70)	<.001***
PRS (100)	33.45 (31.23)	.38 (1.34)	<.001***
Diagnosis of the patients			
Vocal polyp, unilateral	19 (21.11)		
Vocal nodules, bilateral	15 (16.67)		
Vocal cord palsy, unilateral	14 (15.56)		
Spasmodic dysphonia, adductor type	12 (13.33)		
Intracordal cyst, unilateral	6 (6.67)		
Laryngopharyngeal reflux disease	6 (6.67)		
Sulcus vocalis, bilateral	6 (6.67)		
Leukoplakia of vocal cords	4 (4.44)		
Reinke's edema	2 (2.22)		
Contact granuloma	2 (2.22)		
Primary muscle tension dysphonia	1 (1.11)		
Vocal polyps, bilateral	1 (1.11)		
Mutational falsetto	1 (1.11)		
Glottic cancer	1 (1.11)		

Values are presented as mean±SD (range) or number (%).

K-VAPP=Korean version of the Voice Activity & Participation Profile; ALS=Activity Limitation Score; PRS=Participation Restriction Score.

^aCut-off score is 14.5.

***p<.001.

(Samsung Galaxy Note 5, SM-N920)의 자체 내장 마이크(built-in microphone)와 스마트 레코더(Smart Recorder)라고 하는 앤드로이드 애플리케이션을 이용하였다. MDVP 프로그램과 스마트 레코더 앱 모두 표본추출률(sampling rate)은 44.1 kHz로 설정하였고, 확장자는 wav로 통일하였으며, 스마트 레코더 앱에서 자동 게인 조절 기능(automatic gain control)은 꺼짐(off)으로 설정하였다. 해당 기능은 녹음 시 잡음을 방지할 수 있는 장점이 있는 반면, 음성 신호의 입력 수준에 대한 인위적인 조정이 가해짐으로써 음향학적인 측정치에 큰 영향을 줄 우려가 있기 때문이었다.

환자의 음성 샘플이 동시에 두 마이크에 안정적으로 녹음되도록 하기 위하여 듀얼 마이크 홀더(Toto Stand MA002)를 이용하였다. 먼저 T자형 스탠드에 듀얼 마이크 홀더를 결합하여 동일한 모델의 마이크 두 개를 평행을 이루도록 안정적으로 거치한 후, 두 마이크

위에 스마트폰을 안정적으로 거치하였다. 이때 스마트폰의 마이크 부분이 대상자 쪽을 향하도록 하였으며, 스마트폰이 두 마이크 위에 더욱 안정적으로 고정되도록 하기 위해 스마트폰 후면부에 스마트폰 링을 부착하여 이 스마트폰 링이 두 마이크 사이에 단단하게 고정되도록 설치하였다. 아울러 마이크에서 스마트폰 링이 닿는 부분에는 마스킹 테이프를 감아 미끄러짐을 방지하였다. 마지막으로 실험 전 검사자의 모음 발성을 사전에 녹음하여 적절한 수준으로 녹음되고 있는지, 잡음이나 클리핑, 하울링 현상 등이 발생하는지의 여부를 확인하였다. Figure 1은 두 개의 마이크와 스마트폰을 설치한 모습을 옆쪽과 위쪽에서 바라본 모습이다.

마이크들을 이와 같이 설치한 상태에서 각 대상자로 하여금 조용한 검사실에서 의사에 편안하게 앉되, 등을 떼고 허리를 편 바른 자세로 앉도록 한 후 녹음을 진행하였다. CSL의 마이크가 정면으로 대상자의 입으로부터 약 10 cm 거리를 유지되도록 하였다. 주변 환경 소음은 소음측정기(Voltcraft 322 datalogger; Conrad Electronic, Hirschau, Germany)를 이용하여 50 dB 이하로 유지되도록 지속적으로 모니터링하였다(Lee et al., 2015). 과제로는 각 대상자로 하여금 편안한 음도와 크기로 /아/ 모음을 4초 이상 산출하도록 하였다. 4초 동안 발성 유지가 힘든 환자는 분석 과정에서 안정구간 2초 샘플을 별도로 트리밍하였다.

자료분석

스마트폰에서 녹음한 샘플은 CSL의 녹음 샘플과 동시에 녹음된 부분이 정확하게 4초 길이로 따로 분리되어 있지 않으므로, 이 전체 샘플 가운데 모음을 산출한 구간을 MDVP에서 녹음한 샘플과 정확하게 동기화(synchronization)시키는 것이 매우 중요할 것으로 판단되었다. 따라서 이를 위해 별도의 태블릿 노트북에 설치된 Sony Sound Forge 12.0 (Sony Creative Software Inc., Middleton, WI, USA)을 이용하여 시각적, 청각적으로 확인해가며, 동일하고 정확하게 4초 길이로 트리밍하였다. Sound Forge 프로그램은 유료 프로그램으로 음성 샘플에 대한 매우 강력한 조정 및 편집 기능을 지니고 있어 특정 지점으로부터 양방향으로 지정한 지속시간만큼의 샘플을 정확하게 트리밍할 수 있다는 장점이 있다. 이같은 동기화의 구체적인 절차는 아래와 같다.

- (1) MDVP advanced 프로그램에서 CSL로 녹음한 파일(nsp)을 로드하고, Sound Forge에서는 스마트폰으로 녹음한 파일(wav)을 로드한다.
- (2) CSL로 녹음한 모음 샘플을 듣고, 파형(waveform)에서 진폭의 변화 추이를 주의 깊게 살펴 시각적으로 확인한다.
- (3) 스마트폰의 녹음 파일의 파형을 보고 해당 모음연장 발성으

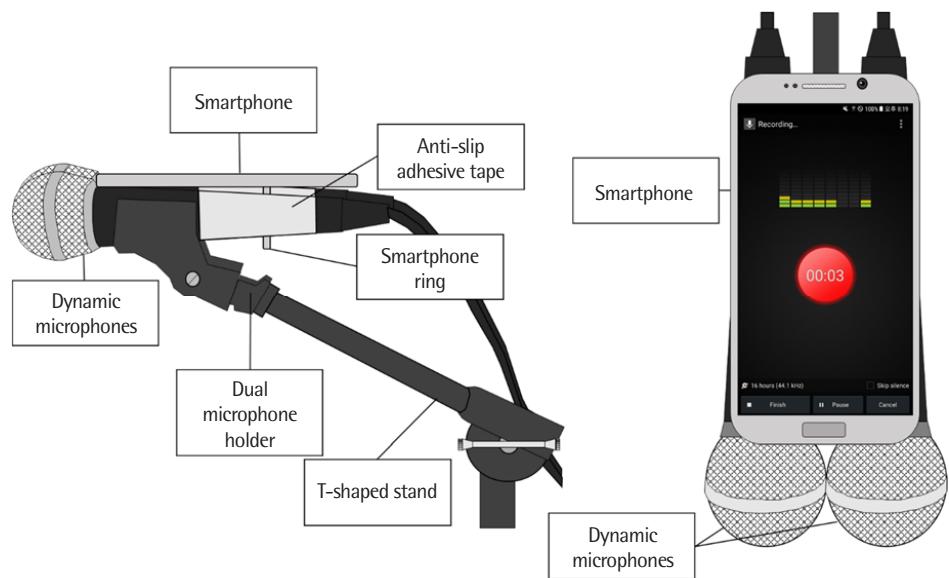


Figure 1. Schematic illustration of stably positioning the microphones and a smartphone using a dual microphone holder and a smartphone ring.

로 여겨지는 부분을 재생하여 동일 부분 여부를 청각적으로 확인한다.

(4) MDVP에 로드된 음성샘플의 파형에서 진폭이 뚜렷하게 증가하거나 감소한 부분을 확인하여 클릭하고, 4초 구간 내에서의 해당 지점의 시간(t)을 확인한다(예: 2.50602 seconds, 소수점 이하 다섯째 자리까지 확인 가능).

(5) Sound Forge의 스마트폰 모음 샘플의 파형에서 동일한 모양으로 진폭이 증가 혹은 감소한 부분을 확인한다. 파형의 상대적인 진폭 높이가 MDVP와 달라 확인이 어려운 경우, Ctrl+휠키 조작을 이용하여 진폭 높이를 조정하여 유사하게 한 뒤 정확하게 확인한다.

(6) 해당 지점에서부터 좌측으로 마우스를 드래그하여 t초 만큼 이전 지점을 지정한 뒤 클릭하여 해당 지점을 선택한 뒤, 해당 지점으로부터 정확히 4초 구간을 선택한다(Ctrl+Shift+D 누른 후 Length를 4초로 정확하게 지정 가능).

(7) CSL과 스마트폰의 두 샘플 구간을 동시에 재생(CSL 단축키 F4, Sound Forge 단축키 Spacebar)하여 청각적으로 두 구간의 샘플이 일치하면서 동시에 재생이 끝나는지를 확인함으로써 동기화 여부를 확인한다. 스마트폰 모음의 재생이 먼저 끝나면 스마트폰 샘플에서 현재 시작 지점보다 약간 앞의 지점을, 나중에 끝나면 뒤의 지점을 시작점으로 하여 다시 4초 구간을 선택하여 동시에 재생하며 동기화 여부를 재확인하는 작업을 반복한다.

(8) 두 샘플의 일치가 확인되면 스마트폰의 4초 구간을 별도의

wav 파일로 저장한다.

이와 같은 절차를 거쳐 총 360개의 모음 샘플 파일(기기 2개 × 집단 2개 × 집단별 90명)이 마련되었다. 각각의 파일을 대상으로 MDVP 프로그램을 이용하여 F0, Jitt, Shim, NHR을, ADSV 프로그램을 이용하여 CPP, σCPP, SR, σSR, CSID를 산정하였다. CSID는 ADSV 결과 창에서 대상자의 성별에 해당하는 추정치를 택하였다.

통계분석

통계 분석 방법으로는 먼저 집단(2개)과 기기(2개)에 따라 측정치의 차이가 있는지 알아보기 위해 이원 반복측정 분산분석(two-way repeated measures of ANOVA)과 ROC 곡선 분석을 시행하였다. 이원 반복측정 분산분석을 통해 윌크의 람다(Wilk's λ)를 계산하였다. 각 기기별 음향학적 측정치 간의 상관관계를 알아보기 위해서 피어슨 상관분석(Pearson correlation analysis)을 시행하고, 급내상관계수(intraclass correlation coefficient, ICC; 단일측도)를 산정하였다. 마지막으로 각 기기별 Jitt, Shim, NHR, CPP, SR, CSID 가 음성장애에 대하여 가지는 선별검사 도구로서의 유용성과 기기별 절단점을 확인하고, 기기 간에 비교하기 위해 수신자 조작 특성(receiver operating characteristic, ROC) 곡선 분석을 시행하여 곡선 아래 영역(area under curve, AUC)을 산정하고, 기기별 AUC 간 비교를 수행하였다. 이원 반복측정 분산분석과 피어슨 상관분석, ROC 곡선분석에는 SPSS 23.0 (IBM-SPSS Inc., Armonk, NY, USA) 프로그램을, ROC 곡선 간 비교에는 Med-Calc 통계 소프트웨어 버전 17.9.7 (MedCalc Software, Ostend, Belgium; <http://www.medcalc.org>)

Table 2. Comparison of the parameters of vowel /a/ samples recorded by the CSL and a smartphone

Parameter	Group	CSL	Smartphone	p-value		
				Interaction	Within-subject	Between-subject
F0 (Hz)	Patient	167.993±48.353	179.863±66.269	.053	.007**	.884
	Control	171.902±44.088	173.849±42.521			
Jitt (%)	Patient	3.049±2.705	4.034±3.594	.012*	<.001***	<.001***
	Control	.942±.601	1.376±1.063			
Shim (%)	Patient	7.601±5.841	12.270±7.346	.002**	<.001***	<.001***
	Control	3.469±1.124	6.040±3.712			
NHR	Patient	.198±.135	.288±.234	<.001***	<.001***	<.001***
	Control	.129±.019	.156±.041			
CPP (dB)	Patient	8.965±3.367	7.977±3.308	.107	<.001***	<.001***
	Control	12.598±1.756	11.397±1.807			
σCPP (dB)	Patient	1.315±.784	1.225±.671	<.001***	.228	<.001***
	Control	.600±.210	.651±.213			
SR (dB)	Patient	28.419±6.304	18.061±8.182	.102	<.001***	<.001***
	Control	32.315±4.804	23.223±5.744			
σSR (dB)	Patient	1.786±.856	1.954±.952	.133	.013*	.092
	Control	1.672±.569	1.713±.575			
CSID	Patient	22.920±21.397	39.279±23.551	.362	<.001***	<.001***
	Control	.814±7.997	16.047±9.113			

Values are presented as mean±SD.

CSL=Computerized Speech Lab; F0=fundamental frequency; Jitt=jitter; Shim=shimmer; NHR=noise-to-harmonic ratio; CPP=cepstral peak prominence; SR=low/high spectral ratio; CSID=Cepstral Spectral Index of Dysphonia derived from vowel samples.

* $p<.05$, ** $p<.01$, *** $p<.001$.**Table 3.** Correlation matrix between the acoustic measures of the vowel /a/ samples recorded by the CSL and a smartphone

CSL	Smartphone									
	F0	Jitt	Shim	NHR	CPP	σCPP	SR	σSR	CSID	
F0	<i>r</i>	.786	.046	.021	.034	-.254	-.210	-.080	.059	.036
	<i>p</i>	<.001***	.538	.777	.654	.001**	.005**	.288	.432	.634
Jitt	<i>r</i>	.237	.874	.664	.758	-.756	.463	-.406	.232	.720
	<i>p</i>	.001**	<.001***	<.001***	<.001***	<.001***	<.001***	<.001***	.002**	<.001***
Shim	<i>r</i>	.315	.755	.708	.847	-.709	.497	-.395	.310	.706
	<i>p</i>	<.001***	<.001***	<.001***	<.001***	<.001***	<.001***	<.001***	<.001***	<.001***
NHR	<i>r</i>	.346	.720	.637	.852	-.601	.404	-.269	.289	.592
	<i>p</i>	<.001***	<.001***	<.001***	<.001***	<.001***	<.001***	<.001***	<.001***	<.001***
CPP	<i>r</i>	-.386	-.806	-.697	-.696	.962	-.487	.590	-.197	-.860
	<i>p</i>	<.001***	<.001***	<.001***	<.001***	<.001***	<.001***	<.001***	.008**	<.001***
σCPP	<i>r</i>	.000	.572	.557*	.483	-.551	.946	-.376	.278	.631
	<i>p</i>	.998	<.001***	<.001***	<.001***	<.001***	<.001***	<.001***	<.001***	<.001***
SR	<i>r</i>	-.304	-.423	-.362	-.381	.453	-.210	.726	-.039	-.505
	<i>p</i>	<.001***	<.001***	<.001***	<.001***	<.001***	.005**	<.001***	.605	<.001***
σSR	<i>r</i>	.066	.222	.176	.316	-.188	.345	-.052	.727	.378
	<i>p</i>	.378	.003**	.018*	<.001***	.012*	<.001***	.489	<.001***	<.001***
CSID	<i>r</i>	.257	.810	.689	.739	-.861	.584	-.642	.381	.921
	<i>p</i>	<.001***	<.001***	<.001***	<.001***	<.001***	<.001***	<.001***	<.001***	<.001***

CSL=Computerized Speech Lab; F0=fundamental frequency; Jitt=jitter; Shim=shimmer; NHR=noise-to-harmonic ratio; CPP=cepstral peak prominence; SR=low/high spectral ratio; CSID=Cepstral Spectral Index of Dysphonia derived from vowel samples.

* $p<.05$, ** $p<.01$, *** $p<.001$.

calc.org; 2017)을 이용하였다.

연구결과

집단 및 기기에 따른 음향학적 측정치 비교

집단 및 기기에 따라 음향학적 측정치를 비교한 결과가 Table 2에 제시되어 있다. 먼저 MDVP 분석 변수 중에서는 Jitt (Wilks's $\lambda = .965$), Shim (Wilks's $\lambda = .949$), NHR (Wilks's $\lambda = .915$)에서 집단과 기기에 따른 상호작용이 있었으며, F0는 CSL보다 스마트폰에서 유의하게 높게 측정되었다(Wilk's $\lambda = .960$). ADSV 측정치 중 CPP (Wilks's $\lambda = .391$), SR (Wilks's $\lambda = .218$)은 CSL이 스마트폰보다, 대조군이 환자보다 높았다. 반면 CSID (Wilks's $\lambda = .213$)는 CSL이 스마트폰보다, 대조군이 환자보다 낮았다. σ CPP (Wilks's $\lambda = .903$)에서는 기기와 집단 간 상호작용이 있었고, σ SR (Wilks's $\lambda = .966$)은 스마트폰이 CSL보다 높은 측정치를 보였다.

Table 4. Intraclass correlation coefficients of the acoustic measures between the CSL and a smartphone

Parameter	ICC (95% CI)		
	Total	Patient group	Control group
F0	.773 (.707-.826)	.681 (.552-.778)	.950 (.926-.967)
Jitt	.839 (.790-.877)	.825 (.746-.882)	.495 (.322-.636)
Shim	.668 (.578-.742)	.671 (.539-.770)	.177 (-.030-.370)
NHR	.732 (.657-.794)	.730 (.617-.814)	.314 (.115-.488)
CPP	.962 (.949-.971)	.950 (.925-.967)	.930 (.895-.953)
σ CPP	.934 (.912-.950)	.920 (.882-.947)	.892 (.840-.928)
SR	.706 (.624-.772)	.640 (.499-.747)	.735 (.624-.817)
σ SR	.725 (.647-.787)	.747 (.639-.826)	.666 (.533-.767)
CSID	.918 (.892-.938)	.909 (.865-.939)	.696 (.571-.789)

CSL=Computerized Speech Lab; ICC=intraclass correlation coefficient; F0=fundamental frequency; Jitt=jitter; Shim=shimmer; NHR=noise-to-harmonic ratio; CPP=cepstral peak prominence; SR=low/high spectral ratio; CSID=Cepstral Spectral Index of Dysphonia derived from vowel samples.

각 기기의 측정치 간의 상관관계 및 신뢰도 분석

CSL과 스마트폰의 샘플에서 측정한 음향학적 측정치 간의 상관관계를 살펴본 결과가 Table 3에 제시되어 있다. 두 기기의 동일한 음향학적 측정치 간에는 모두 통계적으로 유의한 양의 상관관계가 나타났다. 각 변수에 따라 피어슨 상관계수를 살펴보면, F0 (.786), Jitt (.874), Shim (.708), NHR (.852), CPP (.962), σ CPP (.946), SR (.726), σ SR (.727), CSID (.921) 모두 .7 이상이었다.

두 기기를 이용하여 녹음한 모음 샘플의 측정치들 사이의 ICC를 산정한 결과가 집단별로 Table 4에 제시되어 있다. 먼저 전체 대상자에서 가장 높은 ICC를 나타낸 변수는 CPP (.962)였고, 가장 낮은 변수는 Shim (.668)이었다. .75 이상의 ICC를 보인 변수는 F0, Jitt, CPP, σ CPP, CSID였다. 환자군의 경우에도 가장 높은 ICC는 CPP

Table 5. Comparison of the diagnostic ability of selected acoustic measures and CSID between the devices

	AUC	95% CI	z	p-value	Cut-off score	Sensitivity	Specificity
Jitt			6.620	.105			
CSL	.845	.788-.901			1.233	.744	.744
Smartphone	.808	.746-.870			1.603	.722	.722
Shim			.320	.749			
CSL	.827	.767-.887			3.943	.722	.722
Smartphone	.818	.755-.880			6.884	.767	.767
NHR			.002	.998			
CSL	.771	.698-.843			.139	.722	.711
Smartphone	.771	.702-.839			.159	.678	.678
CPP			.987	.324			
CSL	.828	.768-.889			11.647	.756	.756
Smartphone	.815	.751-.879			10.394	.767	.767
SR			.208	.835			
CSL	.699	.622-.777			30.146	.633	.633
Smartphone	.693	.616-.769			21.092	.633	.633
CSID			.513	.608			
CSL	.830	.770-.891			6.138	.733	.733
Smartphone	.819	.757-.882			21.866	.711	.711

CSID=Cepstral Spectral Index of Dysphonia derived from vowel samples; AUC=area under curve; CI=confidence interval; CSL=Computerized Speech Lab; F0=fundamental frequency; Jitt=jitter; Shim=shimmer; NHR=noise-to-harmonic ratio; CPP=cepstral peak prominence; SR=low/high spectral ratio.

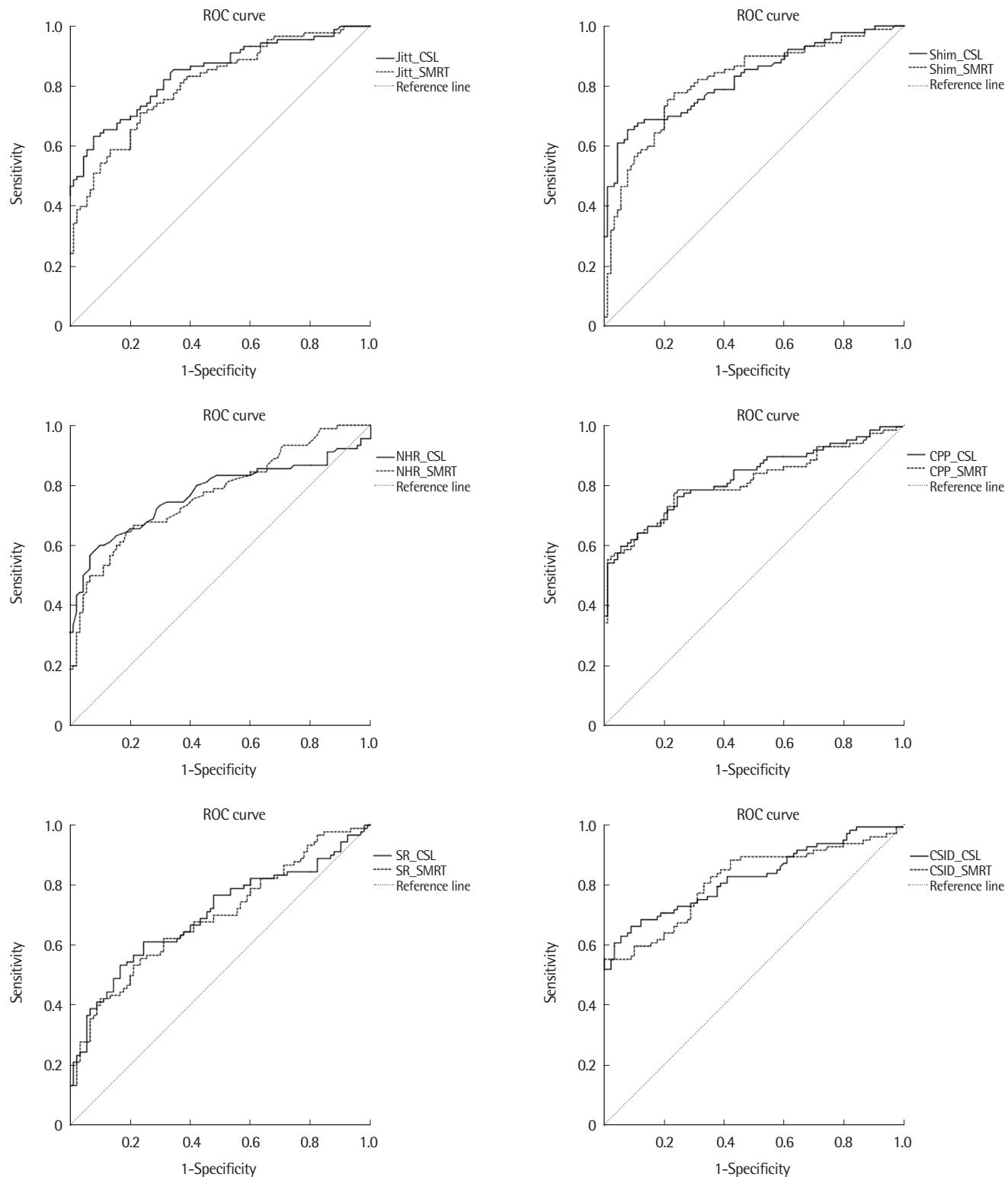


Figure 2. Receiver operating characteristic (ROC) curve of the acoustic measures and estimates derived from vowel samples recorded by the CSL and a smartphone. CSL = Computerized Speech Lab; SMRT = smartphone; Jitt = jitter; Shim = shimmer; NHR = noise-to-harmonic ratio; CPP = cepstral peak prominence; SR = low/high spectral ratio; CSID = Cepstral Spectral Index of Dysphonia derived from vowel samples.

(.950)였으나, 가장 낮은 변수는 SR로 서로 달랐다. 대조군에서는 ICC가 가장 높은 변수는 F0 (.950)였고 가장 낮은 변수는 Shim (.177)이었다. 세 가지 모든 경우에서 .75 이상의 ICC를 보인 측정 변수는 CPP와 σCPP였다.

측정 기기에 따른 선별검사 도구로서의 유용성 비교

CSL과 스마트폰의 각 측정치별 AUC와 비교 결과와 ROC 곡선이 각각 Table 5와 Figure 2에 제시되어 있다. CSL의 Jitt, Shim, NHR의 AUC는 각각 .845, .827, .771이었고, 스마트폰의 경우 각각 .808, .818, .771이었다. 이 세 시간 기반 변수들 모두에서 기기 간 AUC의 유의한 차이는 존재하지 않았다. CSID의 AUC는 각각 .830 (95% CI, .770-.891)과 .819 (95% CI, .757-.882)였으며, 두 측정기기의 AUC 간에는 유의한 차이가 없었다 ($z = .513, p = .608$). Jitt의 절단점은 CSL이 1.233 (민감도와 특이도 각 .744), 스마트폰이 1.603 (민감도와 특이도 각 .722)이었고, Shim의 절단점은 CSL이 3.943 (민감도와 특이도 각 .722), 스마트폰이 6.884 (민감도와 특이도 각 .767)였으며, NHR의 절단점은 CSL이 .139 (민감도와 특이도 각 .721과 .711), 스마트폰이 .159 (민감도와 특이도 각 .678)로 모두 스마트폰이 CSL보다 높은 절단점을 갖는 것으로 나타났다. CSID의 절단점 또한 CSL에서 6.138 (민감도와 특이도 각 .733), 스마트폰에서 21.866 (민감도와 특이도 각 .711)이었다. 반면 CPP의 절단점은 CSL에서 11.647 (민감도와 특이도 각 .756), 스마트폰에서 10.394 (민감도와 특이도 각 .767)로 스마트폰의 절단점이 더 낮았으며, SR 또한 CSL에서 30.146 (민감도와 특이도 각 .633), 스마트폰에서 21.092 (민감도와 특이도 각 .633)였다.

논의 및 결론

본 연구에서는 CSL 하드웨어와 스마트폰으로 녹음한 모음 샘플에서 집단과 기기에 따른 음향학적 측정치의 차이가 있는지 알아보고, 각 기기별 측정치 간의 상관관계와 신뢰도를 알아보았다. 아울러 기기별로 얻은 CSID가 음성장애 여부에 대해 갖는 선별검사 도구로서의 유용성을 확인하고, 기기 간의 비교를 시도하였다. 그 결과, 스마트폰을 이용한 음향학적 측정치는 CSL에서 녹음한 측정치와 일부 유의한 차이가 있고, 이러한 차이는 장애 음성에서 더욱 두드러지는 경향이 관찰되었다. 그럼에도 불구하고, 각 기기에서 측정한 동일 변수 간에는 유의한 양의 상관관계와 상당 수준의 신뢰도가 관찰되었을 뿐만 아니라, 스마트폰을 이용한 측정치가 갖는 선별 능력은 CSL과 차이를 보이지 않았다. 결론적으로, 스마트폰은 음성장애에 대한 선별검사 도구로서의 유용성을 가진 것으로 보인다.

먼저 F0, CPP, SR, CSID 등 핵심 변수들에서 기기 간 측정치의 유의한 차이가 있었다. 특히 스마트폰의 CPP가 낮고 CSID가 높은 경향이 관찰되었다. 기기 간 측정치가 유의한 차이를 보였다는 점은 같은 변수를 측정하더라도 그에 대한 임상적 해석에 매우 주의가 필요하다는 점을 시사한다. 예컨대 캡스트럼 측정치인 CPP와 SR이 낮게 측정된 것, 전통적 측정치가 높게 측정되었다면 이는 같은 음성이라도 더욱 음성의 질이 안 좋은 방향으로 측정이 되었음을 의미할 수 있다. 부분적으로 이는 스마트폰의 자체 마이크가 무지향성 마이크인 것에 반해 CSL에 단일지향성 다이나믹 마이크가 장착되는 것과 무관치 않아 보인다. 아울러 이는 소음에 많은 영향을 받는다는 선행연구 결과(Lebacq et al., 2017)와도 일맥상통한다. 주변 환경 소음을 통제했다 하더라도 방음실에서 실험을 진행하지 못했으므로, 스마트폰이 방향에 관계 없이 소리를 받아들이므로 더욱 많은 환경소음을 받아들였을 것이기 때문이다. 허나 그렇다고 하여 스마트 레코더 애플리케이션에서 소음을 줄일 목적으로 자동 개인 조정 기능을 켜 상태에서 녹음을 진행한다면, 소음의 영향은 배제할 수 있다 하더라도, 더욱 중요한 말 신호 자체의 진폭에 적지 않은 영향을 주어 음향학적 측정치에도 영향을 끼칠 것이므로 권하기 어려울 것으로 생각된다. 따라서 후속 연구에서는 음성 신호의 디지털 변환과 저장은 스마트폰 애플리케이션에서 진행하더라도, CSL과 마찬가지로 스마트폰에 잘 적용될 수 있는 모델의 단일지향성 다이나믹 마이크를 이용하면 주변 환경 소음의 영향을 배제한, 더욱 의미 있는 결과를 도출해 낼 수 있을 것으로 판단된다.

몇몇 변수에서는 집단과 기기 간 상호작용이 나타났다. 세부적인 차이들을 보면 특히 환자군에서 기기 간의 차이가 더욱 두드러지게 나타난 것으로 여겨진다. 그런데 전반적인 경향을 보면 상호작용을 보인 변수 중에는 캡스트럼 분석과 관련된 변수보다는 MDVP를 통해 측정한 전통적인 시간 기반 분석 변수들이 많았다. 구체적으로는 MDVP 분석 변수 가운데에는 Jitt, Shim, NHR 모두 상호작용을 보였으나, ADSV 분석 변수 가운데에는 상호작용을 보인 변수로는 σCPP가 유일하였다. 그뿐만 아니라 상호작용의 경향을 보다 구체적으로 살펴보면 집단 간의 차이는 Jitt, Shim, NHR의 경우 CSL보다 스마트폰이 더욱 컸으나, σCPP의 경우 오히려 스마트폰이 CSL보다 더 적은 집단 간 차이를 보였다. 이러한 결과로 미루어 볼 때, 장애 음성에 대한 측정 결과를 해석할 때 Jitt, Shim, NHR 등의 시간 기반 분석이 이루어지는 전통적인 측정 변수들의 경우에는 집단 간 차이가 더 크게 나타날 수 있다는 점을 염두에 두고 주의 깊게 해석할 필요가 있을 것으로 여겨진다.

기기별로 동일 변수의 측정치 간의 상관관계를 살펴본 결과, 모든 변수들에서 유의한 양의 상관관계가 나타났다. 이를 통하여 비

록 기기별 측정치 간에는 일부 유의미한 차이가 나타날지라도 주어진 음성에 따른 측정치 변화의 경향성은 일치한다는 점이 시사된다. 이는 여러 선행연구들에서 스마트폰을 통해 적어도 일부 변수들이 다른 녹음 기기들을 통해 얻은 수치들과 유의한 상관관계를 보임으로써 신뢰성 있게 측정될 수 있다는 결과가 보고된 것과 맥락을 같이 하는 것으로 보인다(Yun et al., 2015; Manfredi et al., 2017; Maryn et al., 2017; Oliveira et al., 2017). 그러나 구체적으로는 선행 연구와 차이가 있는 부분도 있었다. 예컨대 한 선행연구(Yun et al., 2015)에서는 CSL과 스마트폰이 가장 강한 상관관계를 보인 변수가 F0인 것으로 보고되었는데($r=1.000, p<.01$), 본 연구에서는 가장 높은 피어슨 상관계수를 보인 변수는 CPP였다($r=1.000, p<.001$). 이는 무엇보다도 해당 선행연구는 정상 성인을 대상으로 한 반면, 본 연구에서 많은 수의 음성장애 환자가 포함되었으며, 장애음성에서 켅스트럼 분석을 통해 얻는 대표적인 측정치가 CPP라는 점과 무관하지 않은 것으로 보인다. 따라서 추후 임상적 관점에서 장애음성을 포함한 스마트폰 녹음을 시행하여 임상 및 연구에 활용할 경우 CPP를 분석 변수에 포함할 필요성이 시사된다고 할 수 있을 것이다. 이와 더불어 해당 선행연구에서 스마트폰의 모음 샘플을 트리밍할 때 음성 신호의 시작(onset)에서부터 1.5초 구간을 자르는 등 세부사항에 있어서도 가급적 안정구간을 4초로 선택한 본 연구와 차이를 보였던 점도 영향이 있었을 것으로 여겨진다.

이러한 전반적인 경향은 급내상관계수를 이용한 집단별 분석에서도 유사하였다. 전체 대상자, 환자군, 대조군의 세 가지 모든 경우에서 .75 이상(Fleiss, 1986)의 ICC를 보인 측정 변수는 CPP와 σ CPP였으며, 특히 CPP는 세 경우 모두 .9 이상이었다. 이는 CPP가 특히 기기와 집단과 무관하게 가장 높은 측정의 신뢰도를 보인다는 점을 시사한다. 또한 대조군에서 가장 높은 ICC를 보인 변수가 F0였던 것은 선행연구와 일치한다(Yun et al., 2015). 다만 정상대조군에서 Jitt, Shim, NHR의 ICC가 현저히 낮았던 점은 위의 선행연구와는 다르며 마이크나 샘플 트리밍 방식의 차이를 고려한다 하더라도 다소 놀라운 결과이다. 이와 같은 결과는 정상군에서 Jitt, Shim, NHR의 기기 간 절대적 수치의 차이가 다른 변수들에 비해 미미하여 발생하였을 가능성이 있다. 그러나 전체 샘플을 살펴보았을 때 F0, Jitt, CPP, σ CPP, CSID 등 주요 변수에서 .75 이상의 ICC 가 산정된 점을 함께 고려하면, 정상군의 일부 변수가 ICC가 낮았다고 하여 스마트폰 측정의 신뢰도 전체를 문제로 있는 것으로 보기에는 어렵다고 여겨진다. 따라서 스마트폰을 이용하여 휴대성과 편의성을 유지하면서도 본 연구의 세팅보다 주변 환경 소음의 영향을 덜 받는 세팅으로서 앞서 언급하였던 추가적인 스마트폰용 다이나믹 마이크의 사용 등을 고려해 볼 필요가 있을 것으로 판단된다.

CSL과 스마트폰의 선별검사 도구로서의 유용성에 있어 유의한 차이가 없었던 것으로 보아 음성장애에 대한 선별검사 도구로 이용될 수 있는 가능성이 일부 확인된 것으로 보인다. 다만 역치값으로 활용될 수 있는 절단점들을 살펴보면, 스마트폰의 절단점이 대부분의 변수들에서 CSL보다 높았던 반면, CPP와 SR의 경우 CSL의 절단점이 더 높았다. 이는 앞서 언급하였듯이 자체 내장 마이크가 무지향성 마이크인 것과 무관하지 않은 것으로 보인다. 아울러 이러한 기기 간 차이는 별도의 절단점을 활용해야 한다는 시사점을 넘어, 스마트폰을 이용한 측정 결과를 해석할 때 주의가 필요하다는 점을 시사한다고 본다. 즉 스마트폰에서 측정한 음향학적 측정치나 CSID가 절단점을 넘어서라고 하여 선별리 음성장애로 진단하는 일은 다소 위험할 수 있다는 것이다. 다른 한편으로 스마트폰을 이용한 측정치가 정상 역치를 넘지 않는다고 하여 문제가 없는 것으로 판단하는 일은 이보다 더욱 위험하다고 판단된다. 선부른 정상 진단으로 인해 음성장애에 대한 조기 중재의 기회를 놓칠 우려가 크기 때문이다. 오히려 음성클리닉에 내원하여 숙련된 전문가의 정밀하고 다면적인 음성평가를 받을 필요성을 시사하는 것에 초점을 맞출 필요가 있을 것이다.

이러한 점에서 본 연구는 스마트폰이 음성장애의 예방을 위한 단순 선별검사 도구를 넘어 원격 임상서비스 제공의 핵심적인 수단이 될 수 있다는 선행연구결과(Grillo, 2017)에 이론적 토대를 더하는 예비적 연구였다고 할 수 있다. 스마트폰을 포함한 원격 의사 소통을 수단을 이용한 헬스케어는 편의성, 질, 안전성, 비용 대비 효율성을 증진시킬 수 있는 잠재력을 가지고 있다는 측면에서 최근 국제적으로도 널리 관심의 대상이 되고 있다(Yulzari et al., 2018). 이러한 적용은 특히 직업적 퍼포먼스가 음성의 지대한 영향을 받는 전문적 음성사용자(professional voice user)에 있어 그 임상적 유용성이 크다 할 것이며, 다양한 전문적 음성 사용의 수준에 따른 차이를 알아보는 일 또한 필요할 것으로 보인다. 또 다른 한 연구에서는 스마트폰이 휴대가 용이하다는 점에서 편측 성대마비 환자의 성대 내전을 위한 갑상성형술 수술 시 수술실에서 스마트폰을 이용한 음향학적 분석의 유용성을 밝혀내어 수술실에서의 또 다른 협업 가능성을 제시하기도 하였다(Guzman, Coleman, Rubin, Belannger, & Jackson-Menaldi, 2012). 본 연구에서 분석한 보다 다양한 음향학적 변수를 도입하거나, 수술실의 경우와 유사한 다른 적용 케이스를 탐색하는 일 또한 흥미로운 후속 연구가 될 것으로 보인다.

요약하자면 스마트폰은 음성장애의 선별검사를 위한 녹음장비로서 사용될 수 있는 가능성이 있다고 여겨진다. 다만 본 연구에서는 연결발화 샘플을 분석하지는 않았다는 점, 한 가지의 스마트폰 모델만을 선택하였다는 점, 스마트폰 이외의 다양한 기기나 세팅

간 비교를 수행하지는 못하였다는 점, 청지각적인 음성장애 중증도를 함께 살펴보지는 못하였다는 한계가 명확하다. 따라서 이를 고려한 후속 연구가 이루어짐으로써 스마트폰을 이용한 음성 녹음의 임상적 유용성을 제고할 수 있을 것으로 판단된다.

REFERENCES

- 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., 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.
- 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 Research*, 12, 157-163.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149-1160.
- Fleiss, J. L. (1986). *Design and analysis of clinical experiments*. New York, NY: Wiley.
- Grillo, E. U. (2017). An online telepractice model for the prevention of voice disorders. *Perspectives of the ASHA Special Interest Groups*, 2, 63-78.
- Guzman, M., Coleman, C., Rubin, A. D., Belanger, J., & Jackson-Menaldi, C. (2012). The use of the mobile voice laboratory in the operating room during type I thyroplasty with Gore-Tex. *Acta Otorrinolaringologica (English Edition)*, 63, 364-369.
- Hirano, M. (1981). *Clinical examination of voice*. New York, NY: Springer.
- Kang, Y., & Koo, B. (2015). Acoustic analysis of voice change according to extent of thyroidectomy. *Phonetica and Speech Sciences*, 7, 77-83.
- KayPENTAX. (2008). *Multi-Dimensional Voice Program (MDVP) Model 5105: software instruction manual*. Montvale, NJ: Author.
- KayPENTAX. (2011). *Analysis of Dysphonia in Speech and Voice Model 5109: software instruction manual*. Montvale, NJ: Author.
- Kempster, G. B., Gerratt, B. R., Abbott, K. V., Barkmeier-Kraemer, J., & Hillman, R. E. (2009). Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol. *American Journal of Speech-Language Pathology*, 18, 124-132.
- Kim, G. H., Lee, Y. W., Bae, I. H., Park, H. J., Lee, J. S., Wang, S. G., & Kwon, S. B. (2016). A cepstral analysis of voices with glottic cancer and laryngeal leukoplakia: sustained vowels and continuous speech. *Journal of Speech-Language & Hearing Disorders*, 26, 99-111.
- Kim, H. (2012). *Neurologic speech-language disorders*. Seoul: Sigmapress.
- Lebacq, J., Schoentgen, J., Cantarella, G., Bruss, F. T., Manfredi, C., & DeJongkere, P. (2017). Maximal ambient noise levels and type of voice material required for valid use of smartphones in clinical voice research. *Journal of Voice*, 31, 550-556.
- Lee, S. J., Cho, Y., Song, J. Y., Lee, D., Kim, Y., & Kim, H. (2015). Aging effect on Korean female voice: acoustic and perceptual examinations of breathiness. *Folia Phoniatrica et Logopaedica*, 67, 300-307.
- Lee, S. J., Choi, H. S., & Kim, H. (2018). A comparison of voice activities and participation profiles among etiological groups. *Journal of Voice*. Advanced online publication. DOI: 10.1016/j.jvoice.2018.04.016.
- Lee, S. J., Choi, H. S., Kim, H., Byeon, H. K., Lim, S. E., & Yang, M. K. (2016). Korean version of the Voice Activity and Participation Profile (K-VAPP): a validation study. *Communication Sciences & Disorders*, 21, 695-708.
- Lee, S. J., Lim, S. E., & Choi, H. S. (2017). A comparison of cepstral and spectral measures according to measurement position in a reading passage. *Communication Sciences & Disorders*, 22, 818-826.
- 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.
- Manfredi, C., Lebacq, J., Cantarella, G., Schoentgen, J., Orlandi, S., & DeJongkere, P. (2017). Smartphones offer new opportunities in clinical voice research. *Journal of Voice*, 31, 111.e1-111.e7.
- Maryn, Y., Ysenbaert, F., Zarowski, A., & Vanspauwen, R. (2017). Mobile communication devices, ambient noise, and acoustic voice measures. *Journal of Voice*, 31, 248.e11-248.e23.
- Oliveira, G., Fava, G., Baglione, M., & Pimpinella, M. (2017). Mobile digital recording: adequacy of the iRig and iOS device for acoustic and perceptual analysis of normal voice. *Journal of Voice*, 31, 236-242.
- Park, M. C., Mun, M. K., Lee, S. H., & Jin, S. M. (2013). Clinical usefulness of cepstral analysis in dysphonia evaluation. *Korean Journal of Otorhinolaryngology - Head and Neck Surgery*, 56, 574-578.
- Seo, I. H., Jung, D., Han, H. J., Moon, J. H., Chung, P. S., & Lee, S. J. (2016). Analysis of acoustic parameters to objectively reflect the change of voice quality before and after surgery in benign vocal fold mucosal disorders. *Korean Journal of Otorhinolaryngology-Head and Neck Surgery*, 59, 775-779.

- 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 spasmadic dysphonia. *Phonetics and Speech Sciences*, 8, 73-80.
- Uloza, V., Padervinskis, E., Vegiene, A., Pribuisiene, R., Saferis, V., Vaiciukynas, E., ... & Verika, A. (2015). Exploring the feasibility of smart phone microphone for measurement of acoustic voice parameters and voice pathology screening. *European Archives of Oto-rhino-laryngology*, 272, 3391-3399.
- Van Leer, E., Pfister, R. C., & Zhou, X. (2017). An iOS-based cepstral peak prominence application: feasibility for patient practice of resonant voice. *Journal of Voice*, 31, 131.e9-131.e16.
- Vogel, A. P., Rosen, K. M., Morgan, A. T., & Reilly, S. (2014). Comparability of modern recording devices for speech analysis: smartphone, landline, laptop, and hard disc recorder. *Folia Phoniatrica et Logopaedica*, 66, 244-250.
- Yulzari, R., Bretler, S., Avraham, Y., Sharabi-Nov, A., Even-Tov, E., & Gilbey, P. (2018). Mobile technology-based real-time teleotolaryngology care facilitated by a nonotolaryngologist physician in an adult population. *Annals of Otology, Rhinology & Laryngology*, 127, 46-50.
- Yun, M. H., Lee, J. H., Lee, S. H., & Jin, S. M. (2015). Feasibility of galaxy smartphone recording as portable recorder for acoustic analysis of voice. *Journal of the Korean Society of Laryngology, Phoniatrics and Logopedics*, 26, 104-111.

국문초록

음성장애 선별검사 도구로서 스마트폰을 이용한 음성 녹음의 임상적 유용성

이승진¹(조교수, 연구원, 제1저자) · 이광용²(교수) · 최홍식¹(교수, 연구소장, 교신저자)

¹연세대학교 의과대학 이비인후과학교실 및 후두음성언어의학연구소, ²동국대학교 예술대학 연극학부

배경 및 목적: 본 연구에서는 Computerized Speech Lab (CSL)과 스마트폰으로 녹음한 모음 샘플에서 음성장애 여부와 기기에 따른 음향학적 측정치의 차이를 확인하고, 기기 간 상관관계와 신뢰도를 알아보았으며, 스마트폰과 CSL이 음성장애에 대해 갖는 선별검사 도구로서의 유용성의 차이를 알아보았다. **방법:** 연구대상은 음성장애 환자와 대조군 각 90명이었으며, CSL과 스마트폰으로 녹음한 모음 /a/ 샘플의 주파수변동률(Jitt), 진폭변동률(shim), 소음대배음비(NHR), 캡스트럼 피크 현저성(CPP)과 표준편차, L/H 비율(SR)과 표준편차, 캡스트럼-스펙트럼 발성장애 지수(CSID)를 집단과 기기에 따라 비교하였다. 기기별 동일 측정치 간의 상관관계와 금내상관계수(ICC)를 산정하였다. ROC 곡선분석을 시행하여 AUC를 기기 간에 비교하였다. **결과:** Jitt, Shim, NHR, σCPP에서 집단과 기기에 따른 상호작용이 있었으며, 스마트폰의 F0는 CSL보다 높았다. CPP, SR은 CSL이 스마트폰보다, 대조군이 환자보다 높았다. 반면 CSID, σSR은 CSL이 스마트폰보다 낮았고, CSID는 대조군이 환자군보다 낮았다. 두 기기의 동일한 변수 간에는 모두 양의 상관관계가 있었다. 가장 높은 ICC를 보인 변수는 CPP (.962)였고, 가장 낮은 변수는 Shim (.668)이었다. 주요 변수의 AUC는 CSL은 .699-.845였고, 스마트폰은 .693-.819였으며, 기기 간 AUC의 차이는 없었다. **논의 및 결론:** 기기에 따른 측정치 차이가 있었으나, 기기 간 유의한 상관관계가 있었고, 선별검사 도구로서의 유용성에 차이가 없어 스마트폰이 선별검사를 위한 별도의 절단점을 가진 녹음 장비이자 선별검사 도구로서 유용성이 있는 것으로 판단되었다.

핵심어: 음성장애, 스마트폰, 선별검사, 음향학적 분석, 캡스트럼 분석, 캡스트럼-스펙트럼 발성장애 지수

참고문헌

- 강영애, 구본석(2015). 갑상선 수술범위에 따른 음성의 음향적 분석. *말소리와 음성과학*, 7, 77-83.
- 김근효, 이연우, 배인호, 박희준, 이재석, 왕수건, 권순복(2016). 성문암과 후두백반증 음성의 캡스트럼 분석 및 청지각적 평가: 연장모음과 연속발화. *언어치료연구*, 25, 135-145.
- 김향희(2012). *신경언어장애*. 서울: 시그마프레스.
- 박민철, 문명기, 이상혁, 진성민(2013). 음성장애 평가에서 캡스트럼 분석의 임상적 유용성. *대한이비인후과학회지-두경부외과학*, 56, 574-578.
- 서인호, 정동식, 한현주, 문정환, 정필상, 이상준(2016). 성대 양성점막 질환에서 수술 전후 음질변화를 객관적으로 반영하는 음향지표 분석. *대한이비인후과학회지-두경부외과학*, 59, 775-779.
- 심희정, 정훈, Lee, S. A., 최병흔, 허정화, 고도홍(2016). 내전형연축성 발성장애 음성에 대한 캡스트럼과 스펙트럼 분석. *말소리와 음성과학*, 8, 73-80.
- 윤매화, 이재혁, 이상혁, 진성민(2015). 음향분석에 사용할 녹음장비로 갤럭시 스마트폰 녹음기능의 유용성. *대한후두음성언어의학회지*, 26, 104-111.
- 이승진, 임성은, 최홍식(2017). 문단 내 위치에 따른 캡스트럼 및 스펙트럼 측정치 비교. *Communication Sciences & Disorders*, 22, 818-826.
- 이승진, 최홍식, 김향희, 변형권, 임성은, 양민교(2016). 음성 활동 및 참여 프로파일-한국판(K-VAPP): 타당성 검증 연구. *Communication Sciences & Disorders*, 21, 695-708.
- 이승진, 표화영, 최홍식(2018). 모음 발성 및 문단 읽기 과제 시 한국 성인의 캡스트럼 및 스펙트럼 측정치 정상 규준 연구. *Communication Sciences & Disorders*, 23, 208-217.
- 최성희, 최철희(2016). 한국 정상 화자의 캡스트럼과 스펙트럼 측정치의 성별과 발화 과제 효과. *한국청각언어재활학회지*, 12, 157-163.