

A comparison of acoustic measures among the microphone types for smartphone recordings in normal adults*

Jeong In Park¹ · Seung Jin Lee^{2,**}

¹Department of Speech Pathology & Audiology, Graduate School of Hallym University, Chuncheon, Korea

²Division of Speech Pathology and Audiology, Research Institute of Audiology and Speech Pathology, College of Natural Sciences, Hallym University, Chuncheon, Korea

Abstract

This study aimed to compare the acoustic measurements of speech samples recorded from individuals with normal voices using various devices: the Computerized Speech Lab (CSL), a unidirectional wired pin-microphone (WIRED) suitable for smartphones, the built-in omnidirectional microphone (SMART) of smartphones, and Bluetooth-connected wireless earphones, specifically the Galaxy Buds2 Pro (WIRELESS). This study included 40 normal adults (12 males and 28 females) who had not visited an otolaryngologist for respiratory diseases within the past three months. Participants performed sustained vowel /a/ phonation for four seconds and reading tasks with sentences (“Walk”) and paragraphs (“Autumn”) in a sound-treated booth. Recordings were simultaneously conducted using the four different devices and synchronized based on the CSL-recorded samples for analysis using the MDVP, ADSV, and VOXplot programs. Compared with CSL, the Cepstral Spectral Index of Dysphonia (CSID_v, CSID_s) and Acoustic Voice Quality Index (AVQI) values were lower in the WIRED and higher in the SMART. The opposite trend was observed for the L/H spectral ratios (SR_v and SR_s), and the WIRELESS demonstrated task-specific discrepancies. Furthermore, both the fundamental frequency (F0) and the cepstral peak prominence of the vowel samples (CPP_v) had intraclass correlation coefficient (ICC) values above 0.9, indicating high reliability. These variables, F0 and CPP_v were considered highly reliable for voice recordings across different microphone types. However, caution should be exercised when analyzing and interpreting variables such as the SR, CSID, and AVQI, which may be influenced by the type of microphone used.

Keywords: smartphone, microphone, acoustic analysis, cepstral analysis, reliability

1. 서론

이비인후과 음성클리닉의 임상현장에서 음성장애를 진단 및

평가하기 위한 검사 기기 가운데 CSL(Computerized Speech Lab, Model 4150B; KayPENTAX, Lincoln Park, NJ, USA)은 음성샘플을 신뢰성 있게 표집하고 재생 및 분석하기 위한 목적으로 가장

* This research was supported by Hallym University Research Fund, 2024 (HRF-202401-018).

** sjl@hallym.ac.kr, Corresponding author

Received 29 April 2024; Revised 27 May 2024; Accepted 27 May 2024

© Copyright 2024 Korean Society of Speech Sciences. This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

널리 사용되어 왔다. CSL은 그 자체로 여러 모듈 프로그램의 집합을 지칭하는 말이기도 하며, 고성능 내장형 오디오카드에 연결된 외장형 오디오 인터페이스 기기를 일컫는 말이기도 하다. 그러나 CSL에 사용되는 내장형 오디오카드는 데스크탑 컴퓨터에만 설치가 가능하고, CSL 본체는 USB 기반의 연결을 지원하지 않는다. 따라서 기기의 이동이 제한적이고 다양한 상황에서 사용하기에 어려움이 있는 실정이며(Yun et al., 2015), 코로나19 경험과 맞물려 이러한 고가의 장비 대신 아날로그-디지털 컨버터와 프리앰프 기능을 갖춘 USB 기반의 오디오 인터페이스에 다양한 마이크를 연결하여 사용하거나, 스마트폰과 같은 다른 대체 장비에 대한 관심이 전세계적으로 증가하고 있다.

CSL에는 본체 외에 다양한 부속 기기들이 제공되는데, 그 중 수음을 위한 장비인 마이크는 녹음샘플의 품질을 좌우하는 핵심부분 중 하나이다. CSL 구매 시 디폴트로 제공되는 마이크로는 Shure사의 SM-48 혹은 SM-58 등 단일지향성을 가진 다이내믹 마이크(unidirectional dynamic microphone)가 주로 사용되어 왔다. 그러나 말의 분석을 위하여 전문가 수준의 녹음이 필요한 상황에서 사용하도록 권고되는 마이크는 다이내믹 마이크보다 주파수 응답성(frequency response) 측면에서 월등한 성능을 가진 컨덴서 마이크(condenser microphone)이다(Švec & Granqvist, 2010). 높은 성능수준에도 불구하고 컨덴서 마이크는 사용 시 환경 소음을 잘 통제해야 하고, 비교적 가격이 비싸며, 프리앰프나 팬텀 전원의 연결과 같이 고려해야 할 것이 많다는 제한점들(Awan et al., 2022)이 있어 임상에서는 다이내믹 마이크가 그대로 이용되는 경우가 대부분이다. 이러한 경향은 연구에서도 마찬가지로, 한 문헌연구에서 음향학적 분석에 관한 21개의 논문들을 분석한 결과, 평탄(flat)한 주파수 응답성, 컨덴서 타입의 트랜스듀서, 높은 신호대잡음비와 다이내믹 레인지 등의 권고사항(Švec & Granqvist, 2010)을 만족하는 마이크를 사용한 논문은 4개에 불과하였다(Castro-Tighe & Inostroza-Moreno, 2020).

최근에는 팬데믹 상황에서 편리하게 이용하였던 비대면 서비스에 대한 경험으로 인하여 비대면 음성치료에 대한 관심도 높아졌으며(Lee, 2022), 음성 녹음을 위한 대체 장비나 플랫폼 및 관련 요인에 대한 연구들이 많이 진행되고 있다. Sevitz et al.(2021)와 Weerathunge et al.(2021)의 연구에서는 비대면 플랫폼(telepractice platforms)으로 Zoom과 Microsoft Teams를 활용하여 음성 샘플을 녹음하는 것을 제안하였다. 그러나 비대면 화상회의의 플랫폼을 사용한 음성 녹음은 플랫폼의 종류와 인터넷 속도가 음향 측정치에 영향을 미치는 것으로 나타났다. Awan et al.(2022)의 연구에서는 환자들에게 저렴한 비용의 헤드셋 마이크를 제공하고 컴퓨터 원격 지원을 통하여 음성녹음과 분석을 하는 방법을 제안하였으나, 저가형 헤드셋 마이크를 통한 음성녹음은 해당 연구와 유사한 마이크 특성을 가진 헤드셋 마이크와 신호 획득 시스템을 사용해야 한다는 제한점이 있다고 하였다.

스마트폰은 가장 범용성이 높고 보급률이 높아 각광받고 있는 음성 녹음장비 중 하나이며, 실제로 국내외에서 이를 활용하여 많은 연구가 진행되고 있다. 음향분석을 위한 장비로 갤럭시

스마트폰 녹음기능에 대한 유용성을 살펴본 연구(Yun et al., 2015)에서는 갤럭시 스마트폰(model SM-G906S)과 CSL, 디지털 녹음기(model PCM-M10)를 사용한 녹음샘플로 음향지표를 분석 후 비교하였다. 그 결과 스마트폰으로 녹음한 녹음샘플의 음향지표(F0, jitter, shimmer, NHR, F1, F2, F3, F4) 측정값이 다른 장비로 녹음한 음향지표와 통계적으로 유의미한 차이가 없어 스마트폰이 다른 장비들과 비교했을 때 녹음기능이 큰 차이 없이 유사하다고 하였다. 또 다른 관련 연구들(Jannetts et al., 2019; Manfredi et al., 2017)에서도 F0 같은 특정 음향학적 변수에 있어서는 스마트폰 녹음의 높은 신뢰성이 보고되었다. 스마트폰 기반 음성 건강 지수 어플리케이션에 대한 연구에서도 스마트폰 5종에서 측정치 간 강한 상관관계가 있었으며, 진단 정확도도 우수하고, 스마트폰 녹음 샘플의 측정치와 스튜디오 마이크로 수집한 녹음에서 얻은 측정치 간 높은 상관관계를 보고하였다(Uloza et al., 2023b). 스마트폰의 임상적 유용성을 살펴본 다른 연구(Lee et al., 2018)에서도 스마트폰과 CSL을 통한 녹음샘플을 분석 및 비교하였을 때, 기기 간 측정치의 차이는 있었으나, 동일 변수 간 양의 상관관계를 보였고, 곡선 아래 면적의 차이가 없어 스마트폰이 선별검사 도구로서 유용성이 있다고 하였다.

그러나 스마트폰을 활용한 음성녹음은 각 마이크마다 주파수 응답성이 다르고(Awan et al., 2022, 2023), 스마트폰 기기의 종류에 따라 내장 마이크의 차이로 음향학적 측정치의 차이가 있다는 제한점이 있다(Lee et al., 2018; Yun et al., 2015). 이와 같은 제한점으로 인해 Lee(2022)는 스마트폰에 연결 가능한 단일지향성 유선마이크를 사용하는 방법을 제안하기도 하였는데, 아직 이러한 유선마이크와 같은 외부마이크의 사용이 스마트폰의 자체 내장 마이크를 사용하는 것에 비해 어떠한 이득이 있는지 실증적으로 검증된 바는 없다. 따라서 스마트폰뿐만 아니라 그에 연결하여 사용 가능한 다양한 외부마이크에 대해 유용성을 탐색해볼 필요가 있다. 여기에는 마이크 기능만 갖춘 기기뿐만 아니라, 마이크의 기능을 갖추고 있으면서 보다 범용성 있게 널리 사용되는 기기, 가령 블루투스로 연결되는 무선이어폰이 함께 포함될 수 있다고 여겨진다.

한국갤럽조사연구소(<https://www.gallup.co.kr/>)의 ‘2012-2023 스마트폰 사용률 & 브랜드, 스마트워치, 무선이어폰에 대한 조사’에 따르면 2023년 7월 11-13일 동안 국내에 거주하는 만 18세 이상 남녀 1,001명을 대상으로 무선이어폰 사용에 대하여 조사한 결과, 전체의 56%가 무선이어폰을 사용하였다. 특히 남녀 20-40대의 60% 이상, 50대도 40%-50% 이상이 무선이어폰을 사용하는 것으로 나타났다. 이러한 보급률과 범용성을 고려할 때 무선이어폰은 녹음기기로서의 활용 가능성을 탐색할 필요가 있다고 할 수 있다. 다만 무선이어폰의 마이크는 통화 시 이어셋 용도로 탑재된 만큼 무지향성의 특성을 가지고 있고, 일반 마이크에 비해 녹음의 품질이 비교적 낮다. 또한 성능이 제품에 따라 다르고 모두 투명하게 공개되어 있지는 않으며, 자체적인 소음제거기능이 작동할 수 있다는 한계가 있으므로 고려해야 한다.

따라서 본 연구에서는 CSL, 스마트폰에 적용 가능한 단일지

향성 유선 편마이크, 스마트폰의 자체 내장 무지향성 마이크, 그리고 블루투스로 연결되는 무선 이어폰을 사용하여 동시에 수집한 모음 및 문장 샘플에서 녹음장비 별 음향학적 측정치를 비교함으로써, 스마트폰 및 연결가능한 외부 마이크로 얻은 음성 녹음이 실제 음향학적 분석을 위한 샘플로서 가지는 유용성과 신뢰도를 알아보고자 하였다.

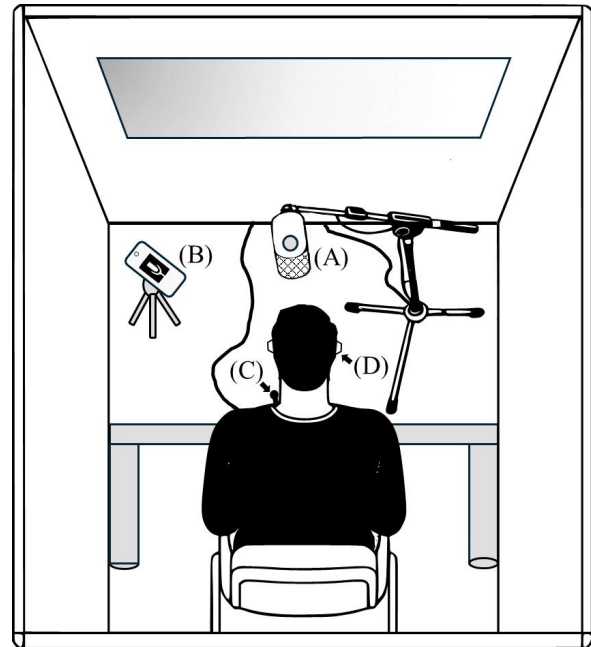
2. 연구방법

2.1. 연구 대상

본 연구의 대상자는 평균 23.0±2.4세의 정상 성인 남녀 총 40명(남 12명, 여 28명)으로 최근 3개월 이내 감기 등의 호흡기 또는 음성 증상으로 인하여 이비인후과에 내원한 적이 없으며, 연구 참여 시점에 호흡기 계통 및 청각적 문제가 없는 자로 한정하였다. 또한 음성 활동 및 참여 프로파일-한국판(Korean version of the voice activity and participation profile, K-VAPP)을 시행하여 총점이 14.5점 이상(Lee et al., 2016)인 대상자의 경우에는 잠재적인 음성 문제를 배제하기 위해 연구 대상에서 제외하였다.

2.2. 연구 절차

모든 대상자는 소음이 통제된 방음 부스에서 녹음을 진행하였다(그림 1). 녹음 장비로는 CSL과 AKG P420(AKG, Wien, Austria) 마이크(CSL), 갤럭시 노트 10+(SM-N976; Samsung, Suwon, Korea)의 자체 내장 마이크(SMART), 갤럭시 버즈2프로(SM-R510; Samsung, Suwon, Korea; WIRELESS) 그리고 USB-C 젠더를 통해 스마트폰에 연결 가능한 단일지향성 특징을 가진 마이크 EIM-009+(EDUTIGE, Seoul, Korea; WIRED)를 사용하였으며 CSL과 스마트폰 자체 마이크를 제외한 나머지 외부 마이크는 모두 삼성 갤럭시 탭7(SM-T870; Samsung, Suwon, Korea)에 연결하여 녹음을 진행하였다.



CSL, Computerized Speech Lab; SMART, built-in microphone of a smartphone; WIRED, unidirectional microphone wired to a smart device; WIRELESS, wireless earbuds connected to a smart device via Bluetooth.

그림 1. 음성녹음 시 (A) CSL, (B) SMART, (C) WIRED, (D) WIRELESS 마이크의 배치

Figure 1. Microphone placement during voice recording: (A) CSL, (B) SMART, (C) WIRED, (D) WIRELESS

사전에 적절한 무선 이어폰을 선정하기 위해서 갤럭시 버즈2프로와 애플 에어팟 프로를 이용하여 파일럿 녹음을 진행하였다. 그 결과, 에어팟 프로에서는 모음 연장 샘플 녹음 시 갑작스러운 전체 파형의 강도 저하가 지속적으로 관찰되는 등 적절하게 녹음이 이루어지지 않는 현상이 발생하였다. 이는 전화 통화 시 깨끗한 음성 전달을 중요시하는 헤드셋 마이크 특성상 잡음을 감쇄하기 위한 기능이 자동적으로 작동하여 발생한 현상으로 보이며, 설정 변경을 통해 이러한 기능을 off 시킬 수 없었다. 반면 갤럭시 버즈2프로에서는 그림 2에서 확인할 수 있듯이, 약 5-7 kHz의 주파수대역에서 불규칙적인 에너지 감쇄효과가 관찰되긴 하였으나, 전체 파형의 지속적인 감쇄효과는 관찰되지 않았다. 따라서 본 연구의 무선 이어폰 녹음 기기로 채택하였다.

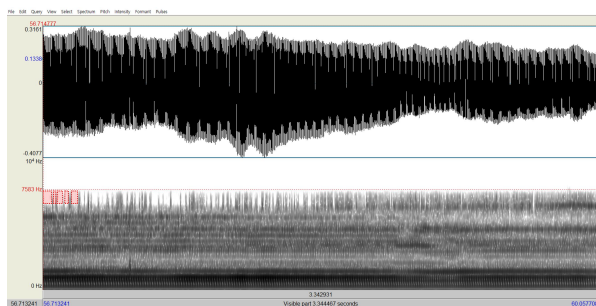


그림 2. 무선이어폰으로 녹음한 모음 발성 샘플에서 발생한 5 - 7 kHz 주파수 대역의 불규칙한 감쇄효과 예시
Figure 2. Example of irregular attenuation effects in the 5-7 kHz frequency range in a vowel sample recorded with wireless earphones

SMART와 WIRED의 경우, 안드로이드 어플리케이션인 스마트 레코더(Smart Recorder)를 사용하여 녹음을 진행하되, Gain factor를 3.0으로 동일하게 유지하였다. WIRELESS의 경우, 스마트 레코더 어플리케이션의 옵션에서 인풋으로 지정하는 것이 불가능하였으므로, 블루투스 기기를 인풋으로 지정 가능한 어플리케이션인 이어버즈 레코더 프로(Earbuds Voice Recorder pro)를 사용하여 녹음하였다. 모든 샘플은 표본추출률 44.1 kHz, 확장자는 WAV로 통일하였으며, 네 가지 기기로 동시에 녹음하였다. 마이크의 배치는 (A) CSL의 경우 대상자의 입으로부터 마이크가 약 10 cm 거리를 유지하도록 하였으며, (B) SMART의 경우 튼튼한 삼각대에 고정하여 입으로부터 45도 각도로 약 30 cm 거리를 유지하도록 하였다. 또한 (C) WIRED는 대중적으로 사용되는 핀마이크 위치인 대상자의 옷깃에 착용하였으며, (D) WIRELESS는 양쪽 귀에 착용하여 녹음을 진행하였다. 모음 연장 과제와 '가을' 모음의 경우 4초간 3회 산출하도록 하여 안정적으로 녹음된 샘플을 사용하여 분석하였으며, 연속 발화 과제로는 "가을" 문단(Kim, 2012) 전체와 "산책" 문단 중 일부(예: 넓게 펼쳐있는 바다를 바라보면 내 마음 역시 넓어지는 것 같다; Kim et al., 2018)를 읽도록 하였다.

2.3. 자료분석

수집된 샘플 중 /아/ 모음의 경우 CSL 이외에는 녹음 샘플이 자동적으로 분절되지 않으므로, 선행연구(Lee et al., 2018)에서 제시한 동기화 절차와 Sony Sound Forge 12.0(Sony Creative Software, Middleton, WI, USA)를 사용하여 CSL의 4초 녹음샘플을 기준으로 정확히 4초 길이로 트리밍하여 분석하였다. 연속 말 과제인 "가을" 문단의 경우 문단 중 앞의 두 번째 문장("무엇보다도 산에 오를 땀 더욱더 그 빼어난 아름다움이 느껴진다")을 별도로 트리밍하여 분석하였다.

위 절차를 통해 총 160개의 모음 녹음샘플(기기 4종×대상자 40명, vowel phonation, V)과 320개의 문장 녹음샘플(기기4종×대상자 40명×연속 발화 과제 2개, sentence production, S)이 마련되었다. 각 샘플들 중 모음샘플에 대해서는 MDVP(Multi-Dimensional Voice Program, model 5105, KayPENTAX)을 통해 F0, 지터퍼센트(jitter percent, Jitt), 씬머퍼센트(shimmer percent, Shim), 소음

대배음비(noise-to-harmonic ratio, NHR)를 분석하였다. 또한 ADSV(Analysis of Dysphonia in Speech and Voice, model 5109, KayPENTAX) 프로그램을 통해 모음샘플의 켈스트럼 피크 현저성(cestral peak prominence, CPP_V), L/H 스펙트럼 비율(low-to-high spectral ratio, SR_V), 켈스트럼-스펙트럼 발생장애 지수(Cepstral Spectral index of Dysphonia, CSID_V), 가을 문장샘플의 CPP_S, SR_S, CSID_S를 분석하였다. CPP와 SR의 경우 각 측정치의 표준편차(σ CPP_V, σ SR_V, σ CPP_S, σ SR_S) 또한 분석하였다. 모음 및 산책 문장에 대해 VOXplot 프로그램(<https://voxplot.lingphon.com/en/>)을 사용하여 음향학적 음성 질 지수(Acoustic Voice Quality Index, AVQI)를 분석하였으며, CSID의 경우 ADSV 프로그램으로 분석시 대상자의 성별에 맞게 추정치를 택하여 분석을 진행하였다.

2.4. 통계 방법

통계분석 프로그램으로는 IBM SPSS Statistics 26.0(IBM-SPSS, Armonk, NY, USA)를 사용하였다. 마이크 유형에 따른 측정치의 차이가 있는지 알아보기 위해 반복측정 분산분석(repeated measures of ANOVA)을 시행하였으며, 사후분석은 Bonferroni 방법을 이용하였다. CSL을 기준으로 하였을 때 다른 기기들과의 음질 관련 측정치들의 상관관계를 살펴보기 위해 스피어만 상관분석을 시행하였고, Python 3.9(Van Rossum & Drake, 2009)을 사용하여 히트맵을 작성하였다. 마지막으로 CSL을 기준으로 신뢰도를 알아보기 위해 급내상관계수(intraclass correlation coefficient, ICC)를 산정하였다. 유의수준은 .05로 설정하였다.

3. 연구결과

3.1. 마이크 유형에 따른 측정치 비교

마이크 유형에 따른 음향학적 측정치를 비교한 결과가 표 1에 제시되어 있다. 분석 결과, Jitt($F=10.172$, $p<.001$), Shim($F=7.414$, $p=.001$), CPP_V($F=21.464$, $p<.001$), σ CPP_V($F=11.735$, $p<.001$), SR_V($F=116.433$, $p<.001$), σ SR_V($F=38.511$, $p<.001$), CSID_V($F=59.155$, $p<.001$), CPP_S($F=13.924$, $p<.001$), σ CPP_S($F=23.487$, $p<.001$), SR_S($F=191.580$, $p<.001$), σ SR_S($F=190.075$, $p<.001$), CSID_S($F=203.876$, $p<.001$), AVQI($F=57.953$, $p<.001$)에서 기기 간 유의한 차이가 있었다. 반면에 F0($F=0.983$, $p=.328$), NHR($F=0.944$, $p=.409$)은 기기 간 유의한 차이가 없었다.

기기 간 유의한 차이를 보인 변수들에 대하여 사후검정을 실시한 결과, WIRED의 경우 CSL에 비해 CSID_V($p<.001$), σ SR_S($p<.001$), CSID_S($p<.001$), AVQI($p<.001$)가 낮았던 반면, CPP_V($p<.001$), SR_V($p<.001$), CPP_S($p=.001$), σ CPP_S($p<.001$), SR_S($p<.001$)는 더 높게 나타났다. SMART의 경우 CSL에 비해 CPP_V($p=.006$), SR_V($p<.001$), σ CPP_S($p<.001$), SR_S($p<.001$)가 낮았던 반면, Jitt($p<.001$), CSID_V($p<.001$), CSID_S($p<.001$), AVQI($p=.008$)는 높았다. WIRELESS의 경우 CSL에 비해 CSID_S($p<.001$), AVQI($p=.018$)가 낮았으며, Shim($p=.013$), σ CPP_V($p=.011$), SR_V($p=.001$), σ SR_V($p<.001$), CSID_V($p<.001$), CPP_S($p<.001$), SR_S($p<.001$), σ

SRs($p<.001$)가 높았다.

표 1. 음향학적 측정치의 마이크 유형 간 비교 결과

Table 1. Comparison of acoustic measurement across microphone types

Parameter	CSL	WIRED	SMART	WIRELESS	F	p-value
F0 (Hz)	183.55 (55.21)	186.43 (47.78)	186.50 (47.77)	186.48 (47.79)	0.983	.328
Jitt (%)	0.94 (0.71)	0.82 (0.56)	1.42 (1.05)	1.05 (0.87)	10.172	<.001***
Shim (%)	3.82 (1.94)	3.79 (1.66)	4.35 (2.00)	5.04 (2.27)	7.414	.001**
NHR	0.14 (0.02)	0.14 (0.02)	0.14 (0.02)	0.14 (0.02)	0.944	.409
CPP _V (dB)	11.13 (2.27)	11.71 (1.88)	10.83 (2.00)	11.13 (2.24)	21.464	<.001***
σCPP _V (dB)	0.67 (0.35)	0.67 (0.32)	0.67 (0.31)	0.81 (0.33)	11.735	<.001***
SR _V (dB)	34.66 (4.23)	37.32 (5.05)	23.05 (4.95)	39.39 (8.32)	116.433	<.001***
σSR _V (dB)	1.61 (0.51)	1.63 (0.49)	2.02 (1.17)	4.34 (2.53)	38.511	<.001***
CSID _V	8.10 (12.83)	2.87 (11.51)	23.64 (13.02)	24.71 (19.05)	59.155	<.001***
CPP _S (dB)	5.96 (0.79)	6.32 (0.80)	6.01 (0.72)	6.37 (0.77)	13.924	<.001***
σCPP _S (dB)	3.44 (0.52)	3.61 (0.46)	3.29 (0.44)	3.48 (0.57)	23.487	<.001***
SR _S (dB)	31.64 (2.85)	35.42 (2.86)	23.47 (2.95)	35.07 (3.59)	191.580	<.001***
σSR _S (dB)	8.96 (1.05)	8.17 (0.99)	8.98 (0.76)	12.96 (1.60)	190.075	<.001***
CSID _S	5.84 (12.2)	-0.68 (10.87)	17.40 (11.20)	-13.82 (10.34)	203.876	<.001***
AVQI	2.63 (1.02)	1.62 (1.06)	3.09 (1.13)	2.17 (0.96)	57.953	<.001***

Values are presented as mean (SD). ** $p<.01$, *** $p<.001$.

CSL, Computerized Speech Lab; WIRED, unidirectional microphone wired to a smart device; SMART, built-in microphone of a smartphone; WIRELESS, wireless earbuds connected to a smart device via Bluetooth; F0, fundamental frequency; Jitt, jitter percent; Shim, shimmer percent; NHR, noise-to-harmonic ratio; CPP, cepstral peak prominence; SR, low-to-high spectral ratio; σ, standard deviation; CSID, Cepstral Spectral index of Dysphonia; V, vowel phonation; S, sentence production; AVQI, Acoustic Voice Quality Index.

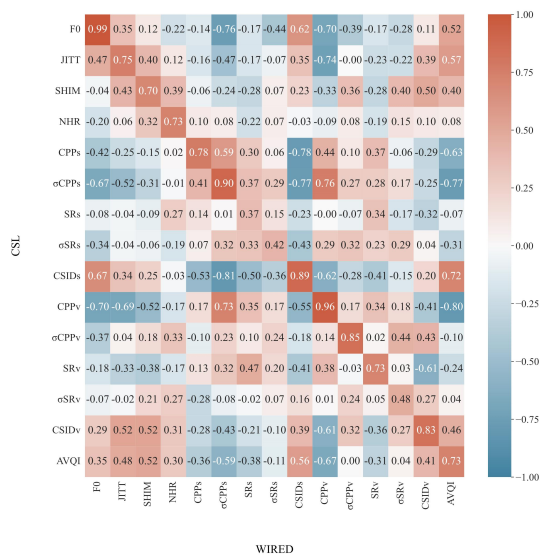
3.2. CSL과 마이크 유형 간 음향학적 측정치의 상관관계 및 신뢰도

CSL의 음향학적 측정치와 각 마이크 유형의 음향학적 측정치의 상관관계를 분석한 결과를 그림 3-5에 제시하였다. 전반적으로 중간 이상 수준의 양의 상관관계가 있었으며, 마이크 유형과 무관하게 F0와 CPP_V에서 강한 양의 상관관계가 있었다.

WIRED의 경우 F0($\rho=.995$, $p<.001$), CPP_V($\rho=.961$, $p<.001$), σCPP_S($\rho=.904$, $p<.001$)에서 매우 강한 양의 상관관계가 있었다. 또한 σSR_V($\rho=.475$, $p=.002$), SR_S($\rho=.373$, $p=.018$), σSR_S($\rho=.420$, $p=.007$)를 제외한 나머지는 강한 양의 상관관계를 보였다.

SMART의 경우에도 F0($\rho=.995$, $p<.001$), CPP_V($\rho=.962$,

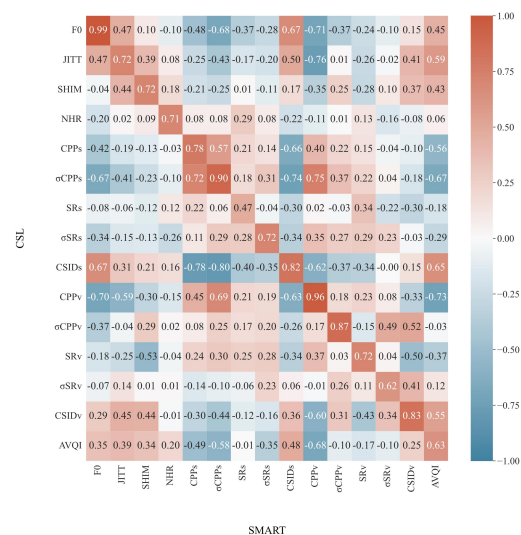
$p<.001$), σCPP_S($\rho=.901$, $p<.001$)에서 매우 강한 양의 상관관계가 있었다. 그 외 변수들은 SR_S($\rho=.474$, $p=.002$)를 제외하고는 모두 중간-강함 수준의 양의 상관관계를 보였다.



CSL, Computerized Speech Lab; F0, fundamental frequency; WIRED, unidirectional microphone wired to a smart device; Jitt, jitter percent; Shim, shimmer percent; NHR, noise-to-harmonic ratio; CPP, cepstral peak prominence; SR, low-to-high spectral ratio; σ, standard deviation; CSID, Cepstral Spectral index of Dysphonia; V, vowel phonation; S, sentence production; AVQI, Acoustic Voice Quality Index.

그림 3. CSL과 WIRED 간 음향학적 측정치의 상관관계 히트맵

Figure 3. Correlation heatmap of acoustic measurements between CSL and WIRED

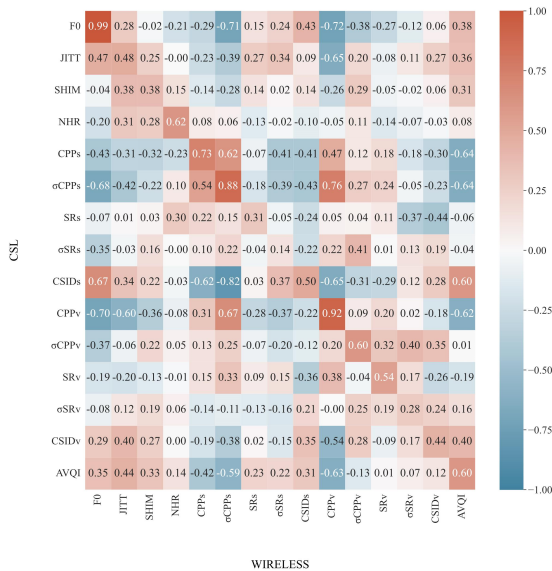


CSL, Computerized Speech Lab; SMART, built-in microphone of a smartphone; F0, fundamental frequency; Jitt, jitter percent; Shim, shimmer percent; NHR, noise-to-harmonic ratio; CPP, cepstral peak prominence; SR, low-to-high spectral ratio; σ, standard deviation; CSID, Cepstral Spectral index of Dysphonia; V, vowel phonation; S, sentence production; AVQI, Acoustic Voice Quality Index.

그림 4. CSL과 SMART 간 음향학적 측정치의 상관관계 히트맵

Figure 4. Correlation heatmap of acoustic measurements between CSL and SMART

WIRELESS의 경우에도 $F0(\rho=.994, p<.001)$, $CPP_V(\rho=.916, p<.001)$ 는 매우 강한 양의 상관관계를 보였다. 그러나 $Jitt(\rho=.476, p=.002)$, $Shim(\rho=.383, p=.015)$, $\sigma SR_V(\rho=.282, p=.077)$, $CSID_V(\rho=.444, p=.004)$, $SR_S(\rho=.306, p=.055)$, $\sigma SR_S(\rho=.140, p=.390)$, $CSID_S(\rho=.497, p=.001)$ 등의 변수는 SMART에 비해 다소 낮은 강도의 상관관계를 보였다. 이들을 제외한 나머지 변수들에서 중간 이상의 양의 상관관계가 관찰되었다.



CSL, Computerized Speech Lab; WIRELESS, wireless earbuds connected to a smart device via Bluetooth; F0, fundamental frequency; Jitt, jitter percent; Shim, shimmer percent; NHR, noise-to-harmonic ratio; CPP, central peak prominence; SR, low-to-high spectral ratio; σ , standard deviation; CSID, Cepstral Spectral index of Dysphonia; V, vowel phonation; S, sentence production; AVQI, Acoustic Voice Quality Index.

그림 5. CSL과 WIRELESS 간 음향학적 측정치의 상관관계 히트맵
Figure 5. Correlation heatmap of acoustic measurements between CSL and WIRELESS

표 2에서는 CSL과 여러 유형의 마이크의 음향학적 측정치들 간의 ICC를 살펴보았다. WIRED에서 가장 높은 ICC를 보인 변수는 $F0(.954)$ 였고, 가장 낮은 변수는 $SR_S(.048)$ 였다. SMART의 경우 가장 높은 ICC를 보인 변수는 $CPP_V(.960)$ 였으며, 가장 낮은 변수는 $SR_S(-.500)$ 였다. WIRELESS의 경우 가장 높은 ICC를 보인 측정치는 $F0(.954)$ 였으며, 가장 낮은 변수는 $\sigma SR_S(-.673)$ 였다. 세 기기에서 모두 .75 이상의 ICC를 보인 변수는 $F0$, CPP_V , σCPP_S 였으며, 그중 $F0$ 와 CPP_V 의 경우 .90 이상의 높은 신뢰도를 보였다(표 2).

표 2. CSL과 마이크 유형 간 음향학적 측정치의 신뢰도
Table 2. Reliability of acoustic measurements across CSL and microphone types

Parameter	ICC (95% CI)		
	WIRED	SMART	WIRELESS
F0 (Hz)	.935 (.881-.965)	.935 (.881-.965)	.935 (.881-.965)
Jitt (%)	.803 (.658-.890)	.618 (.385-.778)	.560 (.307-.740)
Shim (%)	.779 (.622-.877)	.594 (.353-.762)	.273 (-.035-.535)
NHR	.778 (.620-.876)	.763 (.597-.867)	.699 (.500-.828)
CPP _V (dB)	.913 (.843-.953)	.960 (.927-.979)	.922 (.858-.958)
σCPP_V (dB)	.954 (.916-.975)	.949 (.907-.973)	.635 (.409-.788)
SR _V (dB)	.629 (.401-.785)	-.349 (-.592-.104)	.231 (-.080-.502)
σSR_V (dB)	.650 (.429-.798)	.211 (-.101-.487)	-.297 (-.553-.010)
CSID _V	.813 (.675-.896)	.352 (.051-.594)	.123 (-.190-.414)
CPP _S (dB)	.705 (.509-.832)	.840 (.718-.912)	.623 (.392-.781)
σCPP_S (dB)	.842 (.722-.913)	.859 (.750-.923)	.882 (.790-.936)
SR _S (dB)	.048 (-.262-.349)	-.500 (-.699-.228)	-.009 (-.314-.299)
σSR_S (dB)	.277 (-.310-.538)	.663 (.448-.806)	-.673 (-.812-.461)
CSID _S	.749 (.576-.859)	.475 (.198-.682)	-.111 (-.403-.203)
AVQI	.451 (.168-.665)	.617 (.384-.777)	.496 (.224-.697)

CSL, Computerized Speech Lab; WIRED, unidirectional microphone wired to a smart device; SMART, built-in microphone of a smartphone; WIRELESS, wireless earbuds connected to a smart device via Bluetooth; F0, fundamental frequency; Jitt, jitter percent; Shim, shimmer percent; NHR, noise-to-harmonic ratio; CPP, central peak prominence; SR, low-to-high spectral ratio; σ , standard deviation; CSID, Cepstral Spectral index of Dysphonia; V, vowel phonation; S, sentence production; AVQI, Acoustic Voice Quality Index.

4. 논의 및 결론

본 연구에서는 정상음성 성인을 대상으로 CSL과 스마트폰에 적용 가능한 단일지향성 유선 핀마이크, 스마트폰 자체 내장 무지향성 마이크, 블루투스로 연결 가능한 무선 이어폰인 버즈2 프로로 녹음한 모음과 문장 샘플을 분석하여 음향학적 측정치에서의 차이를 확인하고, CSL과 각 마이크 유형들 간의 상관관계와 신뢰도를 알아보았다. 그 결과 CSL과 다른 마이크 유형에서 $F0$ 와 NHR 을 제외한 나머지 측정치들에서 유의한 차이가 나타났다. 이러한 차이에도 불구하고 CSL의 측정치들과 각 기기에서 측정한 동일 변수 간에는 대개 중간-강함 수준의 양의 상관관계가 나타났으며 몇몇 변수를 제외하고 상당 수준의 신뢰도가 관찰되었다. 이러한 결과로 볼 때, 각 마이크 유형에 따라 측정 변수를 선별적으로 사용함과 동시에 해석에도 주의를 기

울인다면 비교적 높은 신뢰성을 가지고 유용하게 이용할 수 있는 가능성을 일부 확인하였다고 할 수 있다.

먼저 마이크 유형 간 측정치 차이를 살펴보았을 때, F0와 NHR을 제외한 Jitt, Shim, CPP_V, σCPP_V, SR_V, σSR_V, CSID_V, CPP_S, σCPP_S, SR_S, σSR_S, CSID_S, AVQI에서 차이가 있었다. 특히 SR_V, CSID_V, SR_S, CSID_S, AVQI의 경우 모든 마이크 유형에서 유의한 차이가 나타났다. SR_V, CSID_V, SR_S, CSID_S의 경우 각 마이크 유형별 주파수 응답성이 SR과 같은 스펙트럼 비율 측정치나 스펙트럼과 캡스트럼 측정치를 결합한 측정치인 CSID와 같은 음향 지표에 영향을 미친다는 선행연구(Awan et al., 2022) 결과의 맥락을 같이하는 것으로 보인다.

다만 AVQI의 경우에는 스튜디오 마이크와 스마트폰 내장 마이크로 녹음된 샘플 간 차이가 나타나지 않았다는 선행연구(Uloza et al., 2023a)와는 다른 결과를 보였다. 이는 선행연구에서는 본 연구와는 달리 스튜디오 마이크와 스마트폰 마이크의 기술적 차이로 인한 영향을 줄이기 위해 별도의 필터링 과정을 적용했기 때문으로 보인다. AVQI는 Praat script를 이용하여 측정하거나 VOXplot에서도 쉽게 분석할 수 있으며, 음성평가를 위한 다중 매개변수로서 평활화된 CPP, 배음 대 소음비(harmonic-to-noise ratio, HNR), 강도 변동률(shimmer_{local}), 강도 변동률dB(shimmer_{local} dB), 장구간 평균 스펙트럼의 기울기(general slope of the spectrum, Slope), 장구간 평균 스펙트럼의 추세선 기울기(tilt of the regression line through the spectrum, Tilt)와 같은 다양한 변수들을 포함하여 계산된다(Kim et al., 2018, 2021; Latoszek et al., 2023). 따라서 AVQI도 SR과 CSID와 마찬가지로 마이크 유형에 따른 주파수 응답성의 차이로 인해 포함되는 변수들의 측정치가 달라지고, 이러한 영향을 받아 그 값이 달라진 것으로 판단된다. 즉 마이크의 주파수 응답성은 녹음된 음성 파형의 스펙트럼에 수정을 가하여 측정하고자 하는 변수의 식별 능력을 감소시킬 수 있다는 것이다(Parsa et al., 2001).

또한 본 연구에서 사용된 각 유형의 마이크들 가운데 특히 무지향성 마이크들은 상대적으로 음성문제의 중증도가 더 높은 방향으로 측정치들이 변화하는 양상이 나타났다. 이는 조용한 검사실에서 스마트폰 내장 마이크를 이용하여 진행한 선행연구 결과와는 일치하나, 본 연구가 방음 부스 내에서 시행된 점을 고려하면 다소 놀라운 결과이다. 이러한 결과는 한쪽 방향으로만 소리가 입력되는 단일 지향성 마이크와 달리 모든 방향에서 입력되는 무지향성 마이크의 특성상 방음 부스 내라도 대상자의 작은 움직임으로 인한 미세 소음이 영향을 미쳤을 것이라 생각된다. 특히 WIRELESS의 경우 감쇄효과도 이러한 결과에 영향이 있었을 것이라 판단된다. 또한 방음 부스 내가 아닌 일상 생활 환경에서 이러한 마이크를 활용할 경우 설사 조용한 방이라고 하더라도 마이크의 영향이 더욱 커질 가능성을 배제할 수 없다. 따라서 이와 같은 연구 결과를 통해 추후 외부 마이크를 사용한 음성분석과 해석 시 환경 소음의 변화에 따라 각 마이크가 측정치에 미치는 영향을 염두에 두고 해석하여야 할 필요가 있을 것으로 보인다.

CSL과 마이크 유형 별 측정치 간의 상관관계를 살펴본 결과,

모든 변수들에서 양의 상관관계가 나타났다. 이러한 결과는 음성장에 환자들을 대상으로 스마트폰의 선별검사로써 유용성을 살펴본 선행연구(Lee et al., 2018)의 결과와 동일하였다. 이는 CSL과 스마트폰뿐만 아니라 다른 외부 마이크를 사용하여도 음성에 따른 측정치 변화의 경향성이 유사하며, CSL과 다른 기기들 간 측정치에 차이가 있더라도 유의한 상관관계를 보이는 것은 높은 신뢰도를 가지고 측정이 가능하다는 선행연구들(Lee et al., 2018; Manfredi et al., 2017; Yun et al., 2015)과도 맥락을 같이 하는 것으로 보인다. 또한 본 연구에서 세 가지 기기 모두에서 가장 강한 상관관계를 보인 변수는 F0로 CSL과 스마트폰에서 F0가 높은 상관관계가 나타났다는 선행연구(Yun et al., 2015)와 일치하였다. 이는 F0가 다른 측정치들에 비해 녹음에 사용된 장비에 덜 영향을 받는다는 것을 시사하며, 측정의 메커니즘이 비교적 단순하고, 고주파수대의 환경 소음이나 기식 소음의 영향을 비교적 덜 받는 낮은 주파수대에서 측정되기 때문인 것으로 풀이된다. CPP_V 또한 세 가지 기기 모두에서 가장 강한 상관관계를 보인 변수들 중의 하나였으며, 이는 Lee et al.(2018)의 연구 결과와도 일치한다. CPP가 음성문제의 중증도나 기식성과도 깊은 관련이 있다는 점에서 이러한 결과는 의미가 있다. 반면 SR과 CSID 관련 변수와 AVQI의 경우 모두 .75이하의 상관관계 강도를 보였는데, 이는 위에서 설명한 것과 같이 마이크 유형에 따른 주파수 응답성에 크게 영향을 받은 것으로 판단된다.

WIRED와 SMART의 경우 모든 변수에서 중간-강함 정도의 상관관계를 보였으며, 몇몇 변수를 제외하고 WIRED에서 더 강한 상관관계를 보였다. 이러한 기기 간 차이는 스마트폰 어플리케이션에서 내장 마이크와 외부 마이크를 이용하여 소음을 측정하였을 때 외부마이크에서 정확성과 정밀도가 향상되었다는 선행연구(Kardous & Shaw, 2016)의 결과와 맥락을 같이한다. 반면 WIRELESS에서는 이 두 마이크보다 낮은 강도의 상관관계를 보이는 경향이 관찰되었다. 이는 전술하였듯이 5-7 kHz 주파수 대역의 음향 에너지가 불규칙하게 손실된 것과 무관하지 않은 결과로 보인다. 추후에는 충분한 성능을 갖춘 마이크가 장착되어 있으면서도 이와 같은 잡음 감쇄 메커니즘의 작동여부를 조작할 수 있는 무선 이어폰을 활용하여 후속연구를 할 수 있으리라 생각된다.

위의 상관관계 분석에서 나타난 경향은 CSL과 여러 유형의 마이크의 음향학적 측정치들 간의 ICC를 살펴보았을 때에도 유사하게 나타났다. 세 기기에서 모두 .75 이상의 ICC를 보인 변수는 F0, CPP_V, σCPP_S였으며, 그중 .90 이상의 ICC를 보인 변수는 F0, CPP_V였다. 이는 F0가 가장 높은 ICC를 보였다는 선행연구(Yun et al., 2015)와 CPP_V가 가장 높은 ICC를 보였다는 선행연구(Lee et al., 2018) 결과와 일치한다. 또한 McKenna et al.(2023)의 연구에 따르면 스마트폰으로 녹음 시 귀 위치, 독서할 때의 책 위치, 스피커폰 위치, 입술에서 8.5 cm 거리를 두고 지면과 평평하게 했을 때의 4가지 휴대폰 위치 상황에서 F0와 CPP는 휴대폰 위치에 영향을 받지 않는 반면 스펙트럼 측정값인 SR과 HNR은 영향을 받는다고 하였다. 이러한 결과는 같은 무지향성 마이크이면서도 다른 곳에 위치시켰던 SMART와 WIRELESS

에 있어서 F0와 CPP의 측정 신뢰도가 동일하게 높았던 결과를 설명해준다.

요약하자면 본 연구에서 사용한 스마트폰 자체 내장 마이크와 유선으로 연결가능한 USB-C 타입 마이크는 녹음 및 음향학적 분석을 위한 장비로써 유용성을 가진 것으로 판단되며, F0와 CPP와 같은 특정 변수를 측정을 위해서라면 무선 이어폰도 사용을 고려해 볼 만하다. 또한 스마트폰 녹음 시 스마트폰 기종별로 자체 내장 마이크가 달라 녹음 및 분석 시 차이가 발생할 수 있다는 점, 스마트폰의 내장마이크가 생활 소음에 상대적으로 취약한 무지향성 마이크라는 한계점은 단일지향성 유선 핀마이크를 사용하면 그 한계를 극복할 수 있을 것으로 판단된다.

본 연구의 가장 큰 한계점은 먼저 정상음성 사용자만을 대상으로 진행하였다는 점이다. 음성장애가 있는 환자에서도 스마트폰 내장 마이크 녹음이 선별검사 도구로서 유용성이 있으며 (Lee et al., 2018), 스마트폰 녹음 샘플에 대한 머신러닝 기반 분석을 통해 정상음성과 음성장애 식별이 가능하다는 결과(Di Cesare et al., 2024)는, 자택에서 환자가 스스로 스마트폰을 통해 녹음을 하고 이에 대한 원격의 분석을 진행함으로써 조기중재 서비스를 제공하는 음성 헬스케어의 가능성을 여전히 시사한다고 여겨진다. 또한 클리닉에 자주 내원하기 어려운 환자에 대한 반복적인 치료적 평가(therapeutic evaluation)에도 활용될 수 있다. 따라서 추후 후속연구에서는 기능적 음성장애와 같은 음성 장애에 문제가 있는 환자군을 대상으로 여러 외부 마이크를 사용한 음성 분석이 갖는 선별도구로서의 유용성을 확인할 수 있다면, 비대면 음성 헬스케어 분야에서 각 스마트폰 모델의 차이로 인한 한계를 극복하는 데 도움이 될 것으로 판단된다. 또한, 스마트폰에 적용 가능한 단일지향성 유선 핀마이크와 블루투스 로 연결 가능한 무선 이어폰의 경우 각각 한 가지의 기기만을 살펴보고, 기기의 가격이 상당하다는 한계가 있어, 추후 연구에서는 가격별로 다양한 장비들을 사용하여 연구를 진행해볼 필요가 있다. 마지막으로 본 연구에서는 각 마이크의 주파수 응답성과 녹음 어플리케이션이 통일되지 못하였다는 한계가 있다. 스마트폰을 통해 음성분석을 한 연구는 많이 진행되고 있으나, 아직 각 연구마다 마이크 설정이 상이하고, 스마트폰 녹음을 위한 적절한 프로토콜과 어플리케이션이 미비한 실정이므로 (Petrizzo & Popolo, 2021), 후속 연구를 통하여 가장 효율적이고 신뢰도 높은 프로토콜을 수립할 필요가 있을 것으로 판단된다.

References

Awan, S. N., Shaikh, M. A., Awan, J. A., Abdalla, I., Lim, K. O., & Misono, S. (2023). Smartphone recordings are comparable to “Gold Standard” recordings for acoustic measurements of voice. *Journal of Voice*. <https://doi.org/10.1016/j.jvoice.2023.01.031>

Awan, S. N., Shaikh, M. A., Desjardins, M., Feinstein, H., & Abbott, K. V. (2022). The effect of microphone frequency response on spectral and cepstral measures of voice: An examination of low-cost electret headset microphones. *American Journal of*

Speech-Language Pathology, 31(2), 959-973.

Castro-Tighe, S., & Inostroza-Moreno, G. (2020). Variability of microphones used for acoustic analysis of the voice in the last twenty years. *Revista de Investigación e Innovación en Ciencias de la Salud*, 2(2), 93-101.

Di Cesare, M. G., Perpetuini, D., Cardone, D., & Merla, A. (2024). Assessment of voice disorders using machine learning and vocal analysis of voice samples recorded through smartphones. *BioMedInformatics*, 4(1), 549-565.

Jannetts, S., Schaeffler, F., Beck, J., & Cowen, S. (2019). Assessing voice health using smartphones: Bias and random error of acoustic voice parameters captured by different smartphone types. *International Journal of Language & Communication Disorders*, 54(2), 292-305.

Kim, H. (2012). *Neurologic speech-language disorders*. Seoul, Korea: Sigma Press.

Kim, G. H., Lee, Y. Y., Bae, I. H., Park, H. J., & Kwon, S. B. (2018). Application of the new version of the acoustic voice quality index with Korean speakers. *Communication Sciences & Disorders*, 23(4), 1091-1101.

Kardous, C. A., & Shaw, P. B. (2016). Evaluation of smartphone sound measurement applications (apps) using external microphones: A follow-up study. *The Journal of the Acoustical Society of America*, 140(4), EL327-EL333.

Kim, G. H., von Latoszek, B. B., & Lee, Y. W. (2021). Validation of acoustic voice quality index version 3.01 and acoustic breathiness index in Korean population. *Journal of Voice*, 35(4), 660.E9-660.E18.

Latoszek, B. B. V., Mayer, J., Watts, C. R., & Lehnert, B. (2023). Advances in clinical voice quality analysis with VOXplot. *Journal of Clinical Medicine*, 12(14), 4644.

Lee, S. J. (2022). Current status and perspectives of telepractice in voice and speech therapy. *Journal of Korean Society of Laryngology, Phoniatics and Logopedics*, 33(3), 130-141.

Lee, S. J., Choi, H. S., Kim, H. 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(4), 695-708.

Lee, S. J., Lee, K. Y., & Choi, H. S. (2018). Clinical usefulness of voice recordings using a smartphone as a screening tool for voice disorders. *Communication Sciences & Disorders*, 23(4), 1065-1077.

Manfredi, C., Lebacqz, J., Cantarella, G., Schoentgen, J., Orlandi, S., & DeJonckere, P. H. (2017). Smartphones offer new opportunities in clinical voice research. *Journal of Voice*, 31(1), 111.E1-111.E7.

McKenna, V. S., Roberts, R. M., Friedman, A. D., Shanley, S. N., & Llico, A. F. (2023). Impact of naturalistic smartphone positioning on acoustic measures of voice. *The Journal of the Acoustical*

Society of America, 154(1), 323-333.

- Parsa, V., Jamieson, D. G., & Pretty, B. R. (2001). Effects of microphone type on acoustic measures of voice. *Journal of Voice*, 15(3), 331-343.
- Petrizzo, D., & Popolo, P. S. (2021). Smartphone use in clinical voice recording and acoustic analysis: A literature review. *Journal of Voice*, 35(3), 499.E23-499.E28.
- Švec, J. G., & Granqvist, S. (2010). Guidelines for selecting microphones for human voice production research. *American Journal of Speech-Language Pathology*, 19(4), 356-368.
- Sevitz, J. S., Kiefer, B. R., Huber, J. E., & Troche, M. S. (2021). Obtaining objective clinical measures during telehealth evaluations of dysarthria. *American Journal of Speech-Language Pathology*, 30(2), 503-516.
- Uloza, V., Ulozaitė-Stanienė, N., Petrauskas, T., Pribušis, K., Blažauskas, T., Damaševičius, R., & Maskeliūnas, R. (2023a). Reliability of universal-platform-based voice screen application in AVQI measurements captured with different smartphones. *Journal of Clinical Medicine*, 12(12), 4119.
- Uloza, V., Ulozaitė-Stanienė, N., Petrauskas, T., Pribušis, K., Ulozienė, I., Blažauskas, T., Damaševičius, R., ... Maskeliūnas, R. (2023b). Smartphone-based voice wellness index application for dysphonia screening and assessment: Development and reliability. *Journal of Voice*. <https://doi.org/10.1016/j.jvoice.2023.10.021>
- Van Rossum, G., & Drake, F. L. (2009). *Python 3 reference manual*. Scotts Valley, CA: CreateSpace.
- Weerathunge, H. R., Segina, R. K., Tracy, L., & Stepp, C. E. (2021). Accuracy of acoustic measures of voice via telepractice videoconferencing platforms. *Journal of Speech, Language, and Hearing Research*, 64(7), 2586-2599.
- 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(2), 104-111.

Email: sjl@hallym.ac.kr

관심분야: 음성장애, 음성언어의학, 신경말언어장애

• 박정인 (Jeong In Park)

한림대학교 일반대학원 언어병리청각학과 석사과정

강원특별자치도 춘천시 한림대학길1

Tel: 033-248-2223

Email: 1wjdmjs1@gmail.com

관심분야: 말장애, 음성장애, 음성언어의학

• 이승진 (Seung Jin Lee) 교신저자

한림대학교 자연과학대학 언어청각학부 및 청각언어연구소
조교수

강원특별자치도 춘천시 한림대학길1

Tel: 033-248-2223

정상 성인에서 스마트폰 녹음을 위한 마이크 유형 간 음향학적 측정치 비교*

박 정 인¹ · 이 승 진²

¹한림대학교 일반대학원 언어병리청각학과, ²한림대학교 자연과학대학 언어청각학부 및 청각언어연구소

국문초록

본 연구에서는 정상음성사용자를 대상으로 음성검사를 위한 고가의 음성 녹음 장비인 Computerized Speech Lab(CSL) 대신 스마트폰에 적용 가능한 단일지향성 유선 핀마이크(WIRED), 스마트폰의 자체 내장 무지향성 마이크(SMART), 블루투스 무선 이어폰인 갤럭시 버즈2 프로(WIRELESS)로 녹음된 음성샘플의 음향학적 측정치를 비교하고자 하였다. 연구대상은 최근 3개월 이내 호흡기 질환으로 이비인후과에 내원한 적이 없는 정상성인 40명(남 12명, 여 28명)이었으며, 소음이 통제된 방음 부스에서 모음 /아/ 연장 발성(4초) 과제와 ‘산책’ 문장, ‘가을’ 문단 읽기 과제를 네 가지의 기기로 동시에 녹음하였다. 4종의 샘플들에 대하여 CSL 녹음을 기준으로 동기화 작업을 진행하였으며, MDVP와 ADSV, VOXplot 프로그램을 이용하여 분석하였다. 연구 결과, F0, shimmer, noise-to-harmonic ratio를 제외한 다른 변수들에서 유의미한 차이가 있었다. 특히 SR_v, SR_s, CSID_v, CSID_s, AVQI의 경우 CSL에 비해 WIRED의 CSID_v, CSID_s, AVQI 중증도가 낮았던 반면, SMART에서는 높게 나타났다. SR_v, SR_s의 경우 반대의 경향이 나타났으며, WIRELESS는 과제에 따라 다른 경향이 있었다. CSL과 다른 마이크 유형들은 동일한 변수 간에는 모두 양의 상관관계를 보였으며, F0와 CPP_v가 모든 유형에서 공히 강한 양의 상관관계를 보였다. ICC 또한 F0와 CPP_v가 모두 0.9 이상으로 가장 높았다. 본 연구에서 사용된 마이크를 음향학적 분석을 위한 녹음 도구로 사용할 때, F0와 CPP_v의 경우 신뢰도 높은 분석 변수로 마이크 유형과 무관하게 포함할 수 있고, SR, CSID, AVQI의 경우 마이크 유형에 따라 분석 및 해석에 주의를 기울일 필요가 있을 것으로 판단된다.

핵심어: 스마트폰, 마이크, 음향학적 분석, 캡스트럼 분석, 신뢰도

참고문헌

- 김근호, 이연우, 배인호, 박희준, 권순복 (2018). 한국인 화자에서 새로운 버전의 Acoustic Voice Quality Index 적용. *언어청각장애연구*, 23(4), 1091-1101.
- 김향희 (2012). *신경언어장애*. 서울: 시그마프레스.
- 윤매화, 이재혁, 이상혁, 진성민 (2015). 음향분석에 사용할 녹음 장비로 갤럭시 스마트폰 녹음기능의 유용성. *대한후두음성언어의학회지*, 26(2), 104-111.
- 이승진, 이광용, 최홍식. (2018). 음성장애 선별검사 도구로서 스마트폰을 이용한 음성 녹음의 임상적 유용성. *언어청각장애연구*, 23(4), 1065-1077.
- 이승진, 최홍식, 김향희, 변형권, 임성은, 양민교 (2016). 음성 활동 및 참여 프로파일-한국판(K-VAPP): 타당성 검증 연구. *언어청각장애연구*, 21(4), 695-708.
- 이승진 (2022). 비대면 음성언어치료의 현황과 전망. *대한후두음성언어의학회지*, 33(3), 130-141.

* 이 논문은 2024년도 한림대학교 교비연구비(HRF-202401-018)에 의하여 연구되었음.