

음향분석에 사용할 녹음장비로 갤럭시 스마트폰 녹음기능의 유용성

성균관대학교 의과대학 강북삼성병원 이비인후과학교실

윤매화 · 이재혁 · 이상혁 · 진성민

= Abstract =

Feasibility of Galaxy Smartphone Recording as Portable Recorder for Acoustic Analysis of Voice

Mae-Hwa Yun, Jae-Hyuk Lee, MD, Sang-Hyuk Lee, MD and Sung-Min Jin, MD

Department of Otolaryngology-Head and Neck Surgery, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, Korea

Background and Objectives : Acoustic analysis of voice could be influenced so much by the quality of voice files which were recorded by recording device. In clinical practice, voice files that were recorded by analysis program directly or portable digital recording device were analyzed mostly. This study examined the feasibility of using Galaxy smartphone recordings for acoustic analysis of voice. **Materials and Methods** : Acoustic measures were compared between voice signals recorded from 30 normal speakers (15 males and 15 females) through Galaxy smartphone, portable digital recording device and CSL. Fo, jitter, shimmer, NHR (Noise-Harmony ratio) and Formant frequencies were analyzed by MDVP. **Results** : Fo, Jitter, Shimmer, NHR and formant frequencies from 3 devices were no significantly difference. The intraclass correlation coefficient (ICC) was higher between each of the voice perturbation measures. **Conclusion** : The findings indicated that Galaxy smartphone recording system was useful device for acoustic analysis of voice. Furthermore, Galaxy smartphone can be applied widely in various way for acoustic analysis of voice.

KEY WORDS : Galaxy smartphone · Recording devices · Acoustic analysis.

서 론

녹음된 음성신호의 질은 음향분석(acoustic analysis)에 많은 영향을 미치기 때문에 녹음환경, 녹음장비(microphone and recording device), 전송 상태(digitization of samples) 등의 요소들을 고려해야 보다 정확하고 의미 있는 결과를 얻을 수 있다.¹⁾

음성의 음향분석에 사용되는 음성분석 프로그램들은 최근 컴퓨터의 발달과 함께 다양하게 개발되고 있고 국내에도 CSL (Computerized Speech Lab), Dr. speech, Praat 등이 대표적으로 소개되고 있다. 그 중 CSL은 음향학적 분석과 관련하여 다양한 옵션 소프트웨어 프로그램들을 포함하고 있고, 자체

적으로 대상자의 음성에 대한 입력과 저장 및 분석이 가능하기 때문에 가장 널리 이용되고 있다.²⁾ 그러나 기기의 이동이 제한되기 때문에 다양한 상황에서 음성을 수집하는 녹음장비로 사용하기에 어려움이 있을 수 있다. 따라서 임상에서는 대상자의 음성을 휴대용 녹음기에 녹음한 후 음성분석 프로그램을 통해 음향지표들을 분석하기도 한다.

임상에서 사용하는 휴대용 녹음기는 전기신호의 처리과정에서 따라 디지털 녹음기와 아날로그 녹음기로 분류할 수 있다. 디지털 녹음기는 원신호의 연속적인 파형을 불연속적인 숫자 정보로 변환하고(analog-to-digital conversion), 2진법에 기초한 숫자로 정보를 저장하게 되며, 녹음 시 원신호의 왜곡(distortion)이 적고, 진폭범위(dynamic range)가 넓기 때문에 음향분석을 위한 녹음장비로 널리 사용되고 있다.³⁾

스마트폰은 최근 디지털 기술의 점진적인 발전으로 무손실 오디오 포맷을 통한 음성신호의 저장이 가능하게 되었다. 특히 최신 모델들은 높은 표본 추출률(sampling rate)이 제공되면서 원신호에 가까운 정보를 저장할 수 있게 되었고 보급 또한 빠르게

논문접수일 : 2015년 5월 13일

심사완료일 : 2015년 6월 2일

책임저자 : 진성민, 03181 서울 중로구 새문안로 29

성균관대학교 의과대학 강북삼성병원 이비인후과학교실

전화: (02) 2001-2264 · 전송: (02) 2001-2273

E-mail: strobojin@hanmail.net

게 확산되면서 많은 사람들이 이용할 수 있게 되었다. 따라서 스마트폰은 디지털 녹음기와 같은 녹음장비의 하나로 기능을 확대할 수 있다.^{4,5)} 그러나 새로운 기술의 발전에도 불구하고 스마트폰의 녹음기능에 대한 유용성을 살펴본 연구가 많지 않다.

따라서 본 연구에서는 스마트폰의 녹음기능이 기존의 녹음장비로 사용되었던 디지털 녹음기의 기능을 대체할 수 있는지를 알아보려 하였다. 이를 위하여 갤럭시 스마트폰과 디지털 녹음기, CSL 음성분석 프로그램을 통해 대상자의 음성을 동시에 녹음하고, 수집된 각 녹음장비 별 음성의 음향지표를 분석 후 비교하였다.

연구대상 및 방법

1. 연구 대상

본 연구는 정상 성인 남녀 15명씩, 총 30명을 대상으로 하였다. 대상자의 연령은 21~40세(평균 28.7세)였다. 연구에 참여할 당시 대상자들은 호흡기 계통 및 청력에 문제를 보이지 않았으며, 후두 스트로보스코피 검사상 후두 질환이 없음을 확인하였다.

2. 녹음장비

본 연구에 사용된 세 개의 녹음장비는 갤럭시 스마트폰(model SM-G906S, Samsung, Korea), CSL(model 4500, Kay PENTAX, USA), 디지털 녹음기(model PCM-M10, Sony, Japan)였다. 대상자의 음성을 녹음하기 위해 갤럭시 스마트폰과 디지털 녹음기는 기기 자체에 삽입되어 있는 내장마이크로폰을 사용하였고 CSL은 기기와 연결되어 있는 다이내믹 마이크로폰(model Shure SM48)을 사용하였다.

세 개의 녹음장비에서 기기마다 다른 값으로 설정되어 있던 두 가지 구성요소를 모두 같은 값으로 통일하였다.

첫째, 세 개의 녹음장비를 통해 녹음된 음성신호의 저장 확장자를 모두 .WAV파일로 통일하였다. 녹음장비 자체에 기본적으로 설정되어 있던 저장 확장자를 살펴보면, 갤럭시 스마트폰은 .m4a파일이고 디지털 녹음기는 .WAV파일, CSL은 .NSP 파일로 기기마다 모두 달랐다. 따라서 갤럭시 스마트폰의 .m4a 파일을 .WAV파일로 변환하기 위해 Switch Audio File Converter Software(NCH software, USA)를 사용하였으며, CSL은 프로그램 자체에서 .WAV파일을 선택하고 저장할 수 있었다.

둘째, 세 개의 녹음장비를 통해 음성신호를 녹음할 때 표본 추출률을 모두 44.1 kHz로 통일하였다. 갤럭시 스마트폰은 기기 자체에 44.1 kHz로 값이 정해져 있었으나 CSL과 디지털 녹음기는 연구자가 직접 선택 가능하였기 때문에 표본 추출률을 세 기기 모두 같은 값으로 통일할 수 있었다. 갤럭시 스

마트폰의 양자화(Quantization)는 24비트(bit)로 설정되었고 CSL과 디지털 녹음기의 양자화는 16비트로 설정되었다.

3. 자료수집

대상자의 음성을 수집하기 위해 대상자 앞에 갤럭시 스마트폰, CSL과 연결되어 있는 다이내믹 마이크로폰, 디지털 녹음기를 같은 높이와 같은 거리에 고정시켰다. 대상자는 자신의 입에서 세 개의 녹음장비까지 약 10 cm 떨어진 곳에 위치한 후 모음/아/를 4초간 편안하게 발성하였다. 대상자가 발성한 음성은 갤럭시 스마트폰의 녹음 앱과 디지털 녹음기, CSL의 Computerized Speech Lab 프로그램에 동시에 녹음되었다.

4. 자료분석

갤럭시 스마트폰, CSL, 디지털 녹음기를 통해 수집된 대상자의 음성은 CSL의 MDVP 프로그램을 사용하여 음향지표들을 분석하였다. 갤럭시 스마트폰, CSL, 디지털 녹음기를 통해 수집한 세 개의 음성파일의 분석구간은 최대한 동일한 구간을 분석하기 위해서 음성이 시작되는 처음부터 1.5초 구간까지로 선정하였다.

녹음장비를 통해 수집된 세 개의 음성신호는 기본주파수(F₀), 음도 변이를 나타내는 jitter, 강도 변이를 나타내는 shimmer, 그리고 잡음 대 정상음 간의 비율을 나타내는 NHR의 음향지표와 F1, F2, F3, F4의 공명주파수(Formant frequency)의 값을 각각 분석하였다. 공명주파수는 선형 예측 부호화(Linear Predictive Coding, LPC)를 사용하여 분석하였다.

5. 통계분석

통계분석은 IBM SPSS 22.0(IBM, Armonk, New York, USA) 통계 프로그램을 사용하여 실시하였다. 갤럭시 스마트폰, CSL, 디지털 녹음기에 동시에 녹음한 음성신호의 음향지표들의 값은 각 녹음장비 별 평균과 표준편차로 나타냈다. 세 개의 녹음장비에서 음성신호의 음향지표들의 평균차이를 비교하기 위하여 Repeated Measures ANOVA를 사용하였고, 유의한 차이가 나타난 음향지표들의 값은 Bonferroni 검정을 실시하여 사후분석을 실시하였다. 유의수준은 0.05로서 양측 검정을 시행하였다. 갤럭시 스마트폰의 음향지표들의 값과 CSL, 디지털 녹음기의 음향지표들의 값의 상관관계는 Pearson correlation coefficient를 통해 살펴보았다.

6. 신뢰도 분석

갤럭시 스마트폰, CSL, 디지털 녹음기를 통해 동시에 녹음한 세 개의 음성신호의 음향지표들의 값 간 신뢰도는 Intraclass correlation coefficients(ICC)를 사용하여 분석하였다.

결 과

1. CSL, 디지털 녹음기, 갤럭시 스마트폰 간 음향지표 값의 차이

30명의 대상자에서 CSL, 디지털 녹음기, 갤럭시 스마트폰 간 음향지표를 분석한 결과, F₀, jitter, shimmer, NHR에서 통계적으로 유의미한 차이가 없었다. F1, F2, F3, F4의 공명주

파수에서도 통계적으로 유의미한 차이가 없었다(Table 1).

성별을 구분한 남녀집단에서도 세 개의 녹음장비 간 음향 지표의 값이 통계적으로 유의미한 차이가 없는지를 살펴보고자 하였다. 여성 집단에서 CSL, 디지털 녹음기, 갤럭시 스마트폰 간 음향지표를 분석한 결과, F₀, jitter, shimmer, NHR, F2, F3에서 통계적으로 유의미한 차이가 없었고, F1과 F4에서 통계적으로 유의미한 차이가 있었다(Table 2). Bonferroni 검정

Table 1. Differences in the parameters of CSL, portable recorder, galaxy according to the population (n=30)

	Mean±SD			p-value of interaction
	CSL	Portable recorder	Galaxy	
F ₀	152.678±45.955	152.667±45.909	152.647±45.925	.225
Jitter	1.070±0.673	0.856±0.405	0.954±0.575	.609
Shimmer	4.395±1.180	5.293±1.450	5.034±1.448	.690
NHR	0.138±0.022	0.155±0.023	0.144±0.022	.132
F1	1090.531±567.215	1155.135±443.817	701.750±328.103	.174
F2	4489.821±1198.472	4506.510±1113.569	4385.678±1332.488	.746
F3	8406.628±1252.052	8362.410±1361.061	8115.685±1236.221	.264
F4	12295.960±1928.962	11831.343±1765.617	12673.642±1703.751	.380

* : p<0.05

Table 2. Differences in the parameters of CSL, portable recorder, galaxy according to the group of female (n=15)

Female	Mean±SD			p-value
	CSL	Portable recorder	Galaxy	
F ₀	195.719±15.211	195.655±15.273	195.649±15.265	.382
Jitter	1.409±0.758	1.100±0.409	1.207±0.627	.259
Shimmer	4.929±1.111	5.636±1.261	5.523±1.395	.068
NHR	0.140±0.013	0.151±0.022	0.144±0.018	.068
F1	1143.825±308.432	1260.664±240.649	931.667±250.052	.001*
F2	4512.073±954.185	4571.371±887.848	4470.389±1065.242	.720
F3	7958.752±1045.271	8031.347±1339.814	7976.753±1306.239	.939
F4	12190.697±1478.736	11842.633±1326.429	12980.085±1428.692	.033*

* : p<0.05

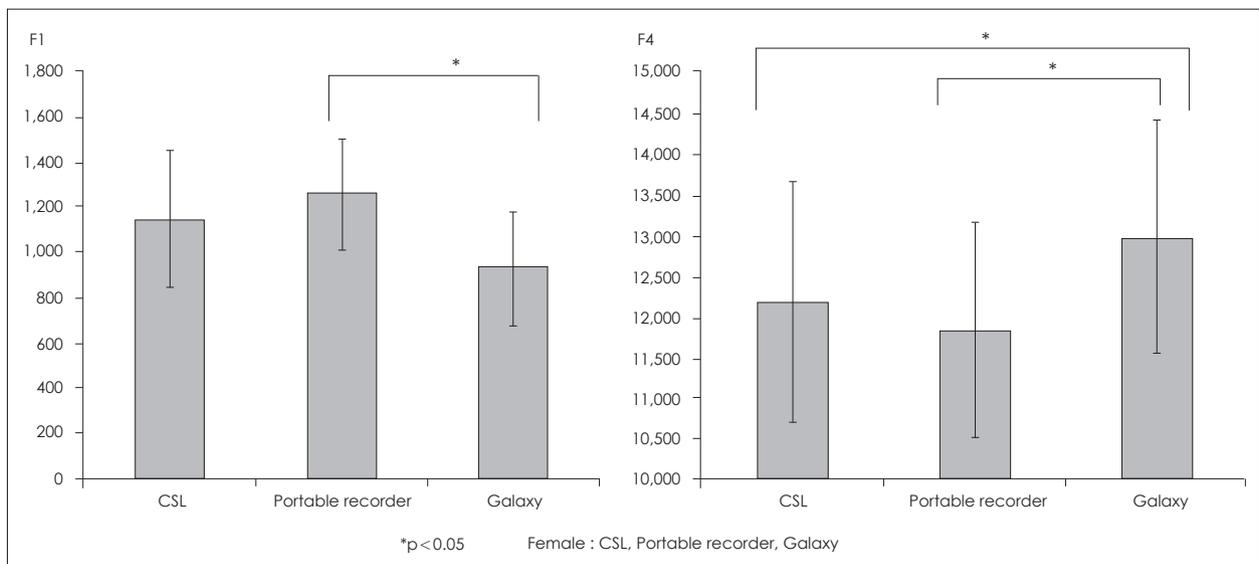


Fig. 1. post-hoc test (Bonferroni) of F1, F4 parameter according to the group of female.

을 이용하여 여성 집단에서 유의미한 차이를 나타낸 F1과 F4의 사후검정을 실시하였다. 그 결과, F1은 디지털 녹음기와 갤럭시 스마트폰 간의 통계적으로 유의미한 차이를 보였고, F4는 CSL과 갤럭시 스마트폰 간의 차이가 통계적으로 유의미한

차이를 보이는 것으로 나타났다(Fig. 1).

남성 집단에서 CSL, 디지털 녹음기, 갤럭시 스마트폰 간 음향지표를 분석한 결과, F₀, jitter, F2, F4에서 통계적으로 유의미한 차이가 없었고, shimmer, NHR, F1, F3에서 통계적으로

Table 3. Differences in the parameters of CSL, portable recorder, galaxy according to the group of male

Male	Mean ± SD			p-value
	CSL	Portable recorder	Galaxy	
F ₀	109.636 ± 13.170	109.678 ± 13.135	109.645 ± 13.167	.360
Jitter	0.730 ± 0.340	0.612 ± 0.212	0.700 ± 0.392	.402
Shimmer	3.862 ± 1.021	4.949 ± 1.586	4.546 ± 1.372	.011*
NHR	0.136 ± 0.029	0.160 ± 0.024	0.143 ± 0.025	.001*
F1	1037.238 ± 751.817	1049.605 ± 571.174	471.833 ± 217.237	.003*
F2	4467.569 ± 1436.571	4441.649 ± 1330.923	4300.967 ± 1589.891	.376
F3	8854.505 ± 1313.319	8693.472 ± 1344.370	8254.618 ± 1190.809	.001*
F4	12401.223 ± 2344.603	11820.052 ± 2167.439	12367.199 ± 1941.774	.339

* : p < 0.05

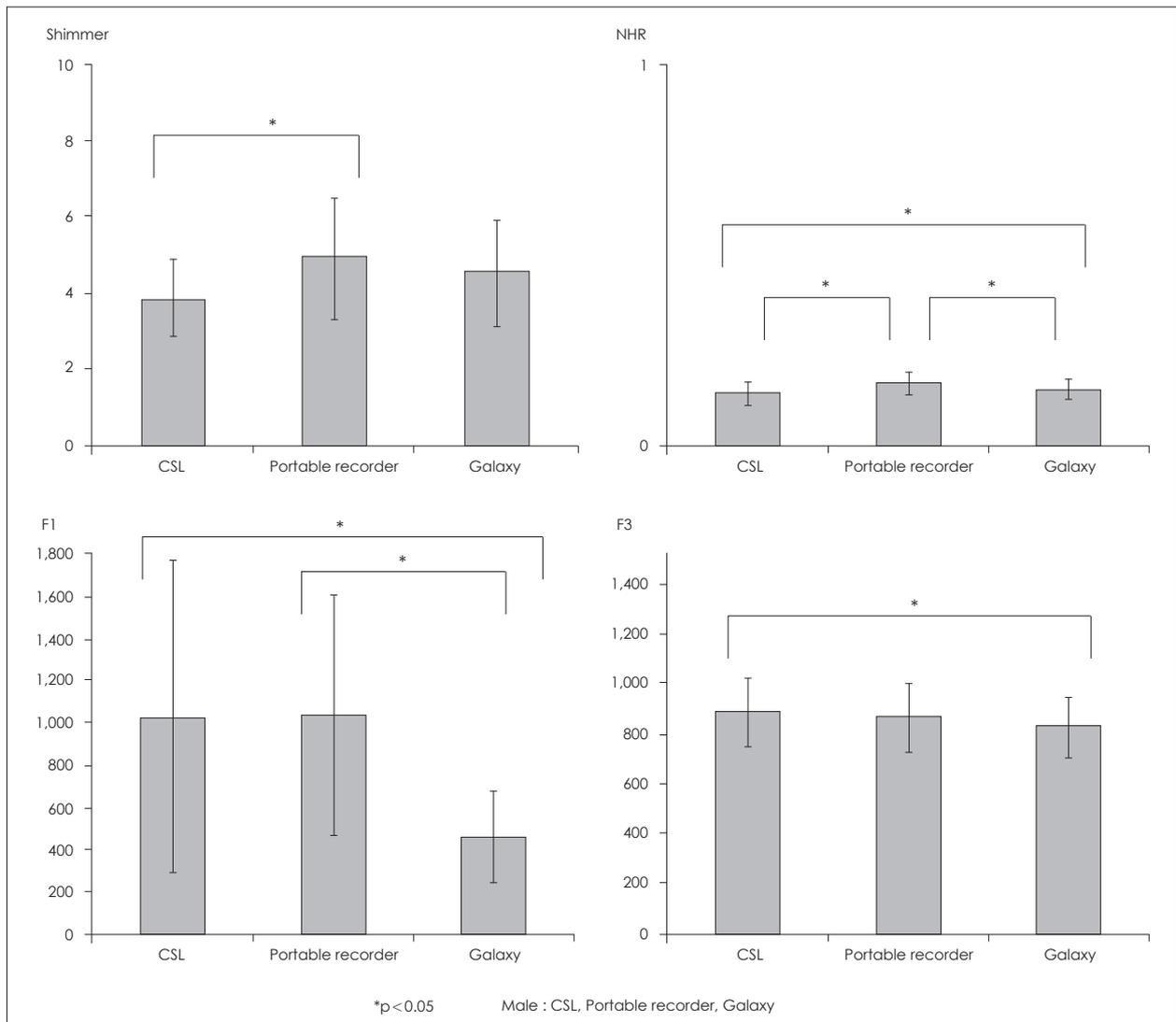


Fig. 2. post-hoc test (Bonferroni) of shimmer, NHR, F1, F3 parameter according to the group of male.

유의미한 차이가 있었다(Table 3). Bonferroni 검정을 이용하여 남성 집단에서 유의미한 차이를 나타낸 shimmer, NHR, F1, F3의 사후검정을 실시하였다. 그 결과, shimmer는 CSL과 디지털 녹음기 간 통계적으로 유의미한 차이를 보였고, NHR은 CSL과 디지털 녹음기 간, CSL과 갤럭시 스마트폰 간, 디지털 녹음기와 갤럭시 스마트폰 간 통계적으로 유의미한 차이를 보이는 것으로 나타났다. F1은 CSL과 갤럭시 스마트폰 간, 디지털 녹음기와 갤럭시 스마트폰 간 통계적으로 유의미한 차이를 보였고, F3는 CSL과 갤럭시 스마트폰 간 통계적으로 유의미한 차이를 보이는 것으로 나타났다(Fig. 2).

2. CSL, 디지털 녹음기, 갤럭시 스마트폰 간 음향지표 값의 상관관계

갤럭시 스마트폰의 음향지표 값과 CSL 및 디지털 녹음기와의 상관관계를 비교하여 위하여 Pearson correlation coefficient을 실시하였다.

30명의 대상자에서 갤럭시 스마트폰과 CSL 간 Fo, shimmer, NHR, F2, F3, F4의 Pearson's r은 1.000, 0.491, 0.829, 0.867, 0.606, 0.622로 강한 상관관계가 나타났다(p<0.01). 갤럭시 스마트폰과 CSL간 jitter의 Pearson's r은 0.438로 양의 상관관계가 있었고(p<0.05), F1의 Pearson's r은 0.185로 상관관계가 나타나지 않았다(Table 4).

30명의 대상자에서 갤럭시 스마트폰과 디지털 녹음기 간 Fo, jitter, shimmer, NHR, F2, F3, F4의 Pearson's r은 1.000, 0.596, 0.659, 0.731, 0.896, 0.528, 0.496으로 강한 상관관계가 나타났다(p<0.01). 갤럭시 스마트폰과 디지털 녹음기 간 F1의

Pearson's r은 0.165로 상관관계가 나타나지 않았다(Table 4).

성별을 구분한 남녀집단에서도 세 개의 녹음장비 간 음향지표의 값의 상관관계를 살펴보았다. 여성 집단에서 갤럭시 스마트폰과 CSL 및 디지털 녹음기 간 Fo의 Pearson's r은 1.000이고, 갤럭시 스마트폰과 CSL 및 디지털 녹음기 간 F2의 Pearson's r은 0.711, 0.771로 강한 상관관계가 나타났다(p<0.01). 갤럭시 스마트폰과 CSL 간 F4의 Pearson's r은 0.544로 강한 상관관계가 나타났으며(p<0.01), 갤럭시 스마트폰과 디지털 녹음기 간 shimmer, NHR의 Pearson's r은 0.839, 0.801로 강한 상관관계가 나타났다(p<0.01)(Table 5).

남성 집단에서 갤럭시 스마트폰과 CSL 및 디지털 녹음기 간 Fo의 Pearson's r은 1.000으로, NHR의 Pearson's r은 0.963, 0.729로 강한 상관관계가 나타났다(p<0.01). 갤럭시 스마트폰과 CSL 및 디지털 녹음기 간 F2의 Pearson's r은 0.937, 0.952로, F3의 Pearson's r은 0.907, 0.874로 강한 상관관계가 나타났다(p<0.01). 갤럭시 스마트폰과 CSL 간 F4의 Pearson's r은 0.692로 강한 상관관계가 나타났으며(p<0.01), 갤럭시 스마트폰과 디지털 녹음기 간 F4의 Pearson's r은 0.622로 양의 상관관계가 나타났다(p<0.05)(Table 6).

3. 신뢰도 분석

30명의 대상자에서 음향지표들의 신뢰도는 0.59(95% 신뢰구간, 0.25~0.79)에서 1.00(95% 신뢰구간, 1.00~1.00)의 범위에 있었다. F1의 신뢰도가 0.59(95% 신뢰구간, 0.25~0.79)로 음향지표들 중 가장 낮았고 F2는 0.96(95% 신뢰구간, 0.93~0.98), Fo는 1.00(95% 신뢰구간, 1.00~1.00)으로 신뢰도가 가장 높았다.

Table 4. Correlation between 8 parameters of portable recorder, CSL and galaxy according to the population (n=30)

	C_Fo	C_jitter	C_shim	C_NHR	C_F1	C_F2	C_F3	C_F4
G_Fo	1.000**	.419*	.412*	.086	.024	.046	-.232	-.017
G_jitter	.534**	.438*	.417*	.016	.338	.511**	.220	.332
G_shim	.305	.529**	.491**	.247	-.012	-.056	-.220	-.108
G_NHR	-.005	.220	.135	.829**	-.365*	-.182	-.174	-.250
G_F1	.566**	.412*	.344	.055	.185	-.052	-.405*	-.089
G_F2	.070	.044	-.063	-.158	.656**	.867**	.222	.594**
G_F3	-.101	-.225	-.238	-.004	.358	.549**	.606*	.484**
G_F4	.206	-.187	-.114	-.066	.493**	.597**	.318	.622**
	P_Fo	P_jitter	P_shim	P_NHR	P_F1	P_F2	P_F3	P_F4
G_Fo	1.000**	.614**	.160	-.228	.250	.096	-.167	-.032
G_jitter	.534**	.596**	.434*	.087	.635**	.550**	.243	.334
G_shim	.304	.391*	.659**	.450*	.046	.081	-.148	-.042
G_NHR	-.006	.103	.174	.731**	.011	-.112	-.116	-.107
G_F1	.566**	.339	.131	-.208	.165	-.048	-.214	.000
G_F2	.070	.095	.118	-.059	.507**	.896**	.227	.571**
G_F3	-.101	.061	-.067	-.108	.274	.510**	.528**	.349
G_F4	.206	.109	-.112	-.213	.490**	.601**	.354	.496**

Galaxy smartphone recording. * : p<0.05, ** : p<0.01. P : portable recorder, C : CSL, G : galaxy

Table 5. Correlation between 8 parameters of portable recorder, CSL and galaxy according to the group of female (n=15)

Female	C_Fo	C_jitter	C_shim	C_NHR	C_F1	C_F2	C_F3	C_F4
G_Fo	1.000**	-.282	.134	-.113	-.129	.181	.613*	.541*
G_jitter	.541*	.332	.205	-.016	.279	.678**	.727**	.505
G_shim	-.228	.517*	.404	.037	.118	.151	.223	-.173
G_NHR	-.369	.607*	.072	.476	-.271	-.010	-.294	-.407
G_F1	-.539*	.100	-.257	.230	.174	-.011	-.354	-.112
G_F2	-.029	.028	-.264	-.080	.596*	.711**	.274	.390
G_F3	-.035	-.097	-.172	.006	.383	.674**	.286	.294
G_F4	.345	-.420	-.093	-.118	.303	.707**	.346	.544*
Female	P_Fo	P_jitter	P_shim	P_NHR	P_F1	P_F2	P_F3	P_F4
G_Fo	1.000**	.307	-.123	-.292	.207	.121	.233	-.241
G_jitter	.535*	.485	.511	.154	.554*	.657**	.620*	.242
G_shim	-.234	.299	.839**	.494	.218	.257	.232	.207
G_NHR	-.377	.379	.307	.801**	-.189	-.066	-.065	-.021
G_F1	-.537*	-.320	-.405	-.052	.011	.033	.048	.308
G_F2	-.029	.106	.209	-.217	.556*	.771**	.235	.180
G_F3	-.036	.429	.189	-.146	.426	.586*	.196	.065
G_F4	.344	.124	-.035	-.154	.552*	.661**	.282	.245

Galaxy smartphone recording. * : p<0.05, ** : p<0.01. P : portable recorder, C : CSL, G : galaxy

Table 6. Correlation between 8 parameters of Portable recorder, CSL and galaxy according to the group of male (n=15)

Male	C_Fo	C_jitter	C_shim	C_NHR	C_F1	C_F2	C_F3	C_F4
G_Fo	1.000**	-.272	-.386	.000	-.304	.027	.221	-.182
G_jitter	.130	.066	.384	-.056	.493	.555*	.186	.382
G_shim	.099	.355	.395	.338	-.122	-.216	-.380	-.048
G_NHR	.121	-.228	.197	.963**	-.410	-.264	-.113	-.179
G_F1	-.513	.029	.381	-.159	.195	-.163	-.109	-.047
G_F2	.070	-.005	.007	-.195	.687**	.937**	.259	.689**
G_F3	.103	-.471	-.257	.010	.426	.497	.907**	.642**
G_F4	-.091	-.320	-.342	-.079	.560*	.554*	.464	.692**
Male	P_Fo	P_jitter	P_shim	P_NHR	P_F1	P_F2	P_F3	P_F4
G_Fo	1.000**	-.263	-.344	.007	.011	.149	.230	-.052
G_jitter	.132	.366	.257	.280	.870**	.630*	.102	.583*
G_shim	.102	.167	.474	.623*	-.157	-.063	-.373	-.207
G_NHR	.124	-.218	.098	.729**	.062	-.137	-.150	-.144
G_F1	-.515*	.127	.254	-.156	-.023	-.260	-.175	-.227
G_F2	.073	.046	.052	.056	.506	.952**	.269	.733**
G_F3	.103	-.368	-.242	-.121	.306	.502	.874**	.569*
G_F4	-.092	-.186	-.239	-.207	.458	.573*	.527*	.622*

Galaxy smartphone recording. * : p<0.05, ** : p<0.01. P : portable recorder, C : CSL, G : galaxy

Table 7. Intraclass correlation coefficients of groups

	Intraclass correlation coefficients (95% C.I.)		
	Total	Male	Female
Fo	1.00 (1.00-1.00)	1.00 (1.00-1.00)	1.00 (1.00-1.00)
Jitter	0.74 (0.52-0.87)	0.59 (0.02-0.85)	0.58 (0.01-0.85)
Shimmer	0.81 (0.65-0.90)	0.77 (0.46-0.92)	0.78 (0.48-0.92)
NHR	0.86 (0.75-0.93)	0.91 (0.79-0.97)	0.77 (0.46-0.92)
F1	0.59 (0.25-0.79)	0.55 (-0.08-0.84)	0.56 (-0.04-0.84)
F2	0.96 (0.93-0.98)	0.98 (0.95-0.99)	0.92 (0.81-0.97)
F3	0.86 (0.74-0.93)	0.97 (0.92-0.99)	0.66 (0.20-0.88)
F4	0.80 (0.64-0.90)	0.87 (0.70-0.95)	0.60 (0.06-0.86)

여성 집단에서 음향지표들의 신뢰도는 0.56(95% 신뢰구간, -0.04~0.84)에서 1.00(95% 신뢰구간, 1.00~1.00)의 범위에 있었다. 남성 집단에서 음향지표들의 신뢰도는 0.55(95% 신뢰구간, -0.08~0.84)에서 1.00(95% 신뢰구간, 1.00~1.00)의 범위에 있었다. 남녀 모두 F1의 신뢰구간에서 범위의 차이가 가장 컸으며(여 : -0.04~0.84, 남 : -0.08~0.84), F₀의 신뢰도가 가장 높았다(남, 여 모두 ICC=1.00)(Table 7).

고 찰

본 연구는 갤럭시 스마트폰과 CSL, 디지털 녹음기를 통해 동시에 녹음한 세 개의 음성신호의 음향지표 값들을 분석하고 세 개의 녹음장비 간 차이가 있는지를 비교하였다. 이를 통해 갤럭시 스마트폰의 녹음기능이 기존의 디지털 녹음기의 기능을 대체할 수 있는지를 알아보려고 하였다.

본 연구에서 30명의 대상자에서 갤럭시 스마트폰, CSL, 디지털 녹음기를 통해 동시에 녹음한 세 개의 음성신호 간 모든 음향지표들(F₀, jitter, shimmer, NHR, F1, F2, F3, F4)의 값에 통계적으로 유의미한 차이가 없었다. 또한 갤럭시 스마트폰의 F₀, shimmer, NHR, F2, F3, F4는 CSL 및 디지털 녹음기의 음향지표들과 강한 상관관계를 나타냈다(p<0.01). 이는 정상성인을 대상으로 아이폰과 노트북을 통해 동시에 녹음한 음성신호 간 F₀, jitter, F1, F2의 값이 통계적으로 유의미한 차이가 없었다는 Lin 등⁴⁾의 연구와 일치한다. 하드디스크 녹음기의 음향지표들의 값을 아이폰, 노트북, 전화(Landline)와 비교한 Vogel 등⁵⁾의 연구에서 다른 녹음장비에 비해 아이폰의 음향지표들의 값(F₀, jitter, shimmer, NHR, Voice turbulence index 등)이 하드디스크 녹음기와 평균적으로 가장 높은 상관관계가 있다는 결과와도 일치한다. 따라서 갤럭시 스마트폰의 녹음기능은 현재 음향분석의 녹음장비로 널리 사용되고 있는 CSL과 디지털 녹음기의 녹음기능과 큰 차이 없이 유사한 값을 갖는다는 것을 알 수 있었다.

본 연구에서 음향지표들을 구체적으로 살펴보면, 전체집단 및 남녀집단에서 갤럭시 스마트폰의 F₀는 CSL 및 디지털 녹음기의 F₀(r=1.000)와 가장 강한 상관관계를 나타냈다(p<0.01). 이러한 결과는 Lin 등⁴⁾, 유재연 등⁶⁾의 결과와 일치한다. F₀는 다른 음향지표들과는 달리 녹음장비의 녹음 기능에 덜 민감한 경향이 있으며, 음도 측정은 대부분 시간영역(time domain)의 분석이므로 주파수 영역(frequency domain)의 분석에 비해 녹음 및 분석 과정에서 정보의 손실이 적을 수 있기 때문에 강한 상관관계를 나타낸다고 볼 수 있다.^{4,6)}

전체집단에서 갤럭시 스마트폰의 shimmer와 jitter는 CSL 및 디지털 녹음기의 값과 상관관계가 있는 것으로 나타났다.

여성집단에서 갤럭시 스마트폰의 shimmer는 디지털 녹음기의 값과 강한 상관관계가 있는 것으로 나타났다. 이는 Lin 등⁴⁾의 연구에서 아이폰이 노트북에 비해 높은 shimmer 값을 나타낸다는 결과와 일치하지 않았다. 이것은 두 연구의 녹음환경, 녹음장비, 음성분석 프로그램에서 차이가 있기 때문인 것으로 볼 수 있다. 우선 두 연구에서 사용한 마이크로폰의 종류 및 위치, 녹음장비 등의 차이가 있었다. 음향분석을 위한 녹음환경에서 마이크로폰의 종류와 위치, 대상자의 입과 마이크로폰과의 거리 및 각도 등은 음향지표들의 민감성에 영향을 줄 수 있다.³⁾

다음으로 본 연구에서 사용한 음성분석 프로그램은 CSL의 MDVP 프로그램이었고, Lin 등의 연구에서 사용한 음성분석 프로그램은 CSpeech의 TF32(University of Wisconsin madison, USA) 프로그램이었다. 기존의 음성분석 프로그램들 간 음향지표의 값을 비교한 선행연구들은 음성분석 프로그램에 따라 측정된 음향지표들의 값에 차이가 있었다. 특히 F₀의 값은 매우 높은 상관관계를 보였으나 shimmer와 jitter에서는 연구마다 불일치한 결과를 도출했다.^{1,6,7)} 이는 음성분석 프로그램들이 서로 다른 알고리즘을 사용하기 때문에 차이가 나타난 것으로 볼 수 있다.⁵⁾ shimmer와 jitter 값은 F₀에 비해 기기의 소음 또는 주변 소음에 더 많은 영향을 받지만, 이들은 음성장에 대상자의 음질을 판단하는데 유용한 수단이며 CSL 또한 실험음성학 분야에서 현재까지 나와있는 음성분석 프로그램 중 가장 탁월한 성능을 인정받고 있다.^{4,8)} 이를 바탕으로 본 연구에서 갤럭시 스마트폰의 shimmer와 jitter 값이 CSL 및 디지털 녹음기와 유의미한 차이가 없고 강한 상관관계를 갖는다는 것은 갤럭시 스마트폰의 녹음기능이 음향분석에 사용할 녹음장비로서 의의가 있음을 알 수 있다.

전체집단에서 갤럭시 스마트폰의 F2, F3, F4는 CSL 및 디지털 녹음기의 F2, F3, F4와 강한 상관관계가 나타났다(p<0.01). 그러나 전체집단 및 남녀집단에서 갤럭시 스마트폰의 F1는 CSL 및 디지털 녹음기의 F1과 상관관계가 나타나지 않았다(p>0.05). 이는 Lin 등⁴⁾의 연구에서 아이폰과 노트북 간 F₀, F1, F2의 포먼트 값이 다른 음향지표의 값에 비해 가장 높은 상관관계를 보였다는 결과와 일치하지 않았다. 일반적으로 음향분석 시 음성의 가장 앞부분과 끝부분은 'edge phenomena' 효과가 발생할 수 있기 때문에 분석 구간에서 제외하며 전체 음성 길이 중 대상자의 음성을 가장 전형적으로 반영하는 부분을 선택하게 된다.^{9,10)} 하지만 본 연구에서는 세 개의 음성파일의 동일한 구간을 분석하고자 음성이 시작되는 처음부분을 분석구간으로 선택하였기 때문에 녹음장비 간 F1의 값에서 강한 상관관계가 나타나지 않은 것으로 볼 수 있다. Lin 등⁴⁾의 연구에서는 분석 구간에 대한 구체적인 언급이 없었기 때문

에 본 연구의 분석구간과 직접적으로 비교할 수 없었다.

본 연구에 사용된 녹음장비는 기기 자체에 정해져 있는 기본설정에 차이가 있었으며, 세 개의 녹음장비의 값을 통일하기 위해 기본설정 값의 변경이 가능한 경우 연구자가 직접 선택하기도 하였다. 우선 갤럭시 스마트폰과 디지털 녹음기는 기기 자체에 삽입되어 있는 내장마이크로폰을 사용하였고 CSL은 기기와 연결되어 있는 다이내믹 마이크로폰을 사용하였다. 둘째, 갤럭시 스마트폰의 양자화는 CSL, 디지털 녹음기의 양자화와 차이가 있었다. 양자화는 표본 추출률과는 달리 녹음 장비에 자체적으로 조절되어 나오기 때문에 연구자가 따로 설정을 할 수 없었다. 대부분의 말소리 연구에서는 12비트 이상으로 양자화 되는 것이 강도 범위를 넉넉히 확보할 수 있는 적절한 수준으로 보고 있으며 양자화가 높을수록 원신호에 더 가깝기 때문에 본 연구의 양자화 값은 기기마다 차이가 있지만 적절한 수준으로 설정되었다고 볼 수 있다.¹¹⁾ 셋째, 세 개의 녹음장비의 표본 추출률은 모두 44.1 kHz로 통일하였다. 사람의 목소리를 분석할 때 표본 추출률은 연구하고자 하는 말소리의 특성에 따라 10 kHz에서 20 kHz 사이에서 적절한 범위를 결정할 수 있으며 본 연구에서 사용했던 CSL의 MDVP 프로그램은 표본 추출률이 25 kHz 이상으로 설정되어 있어야 음향지표들의 분석이 가능하기 때문에 본 연구의 표본 추출률의 값은 적절한 수준으로 설정되었다고 볼 수 있다.¹²⁾ 넷째, 녹음장비마다 음성을 저장하는 저장 확장자의 차이가 있었다. 갤럭시 스마트폰의 저장 확장자인 .m4a파일은 애플에서 개발한 디지털 음악의 무손실 압축 오디오 코덱으로 휴대폰에서만 사용할 수 있다. CSL과 디지털 녹음기의 저장 확장자인 .WAV파일은 비압축 오디오 포맷으로 윈도우 시스템에서 쓰이는 기본 포맷이나 용도에 따라 여러 압축 코덱을 이용할 수 있다. 따라서 .m4a파일은 다른 오디오 시스템에서 사용하기 위해서 변환 과정을 거쳐야 하며 본 연구에서도 이 과정에서 음원의 소실 문제가 발생할 수 있었을 것이다. WAV파일은 용도에 따라 음원을 압축할 수 있기 때문에 CSL과 디지털 녹음기가 같은 저장 확장자를 사용하였더라도 음원의 압축형태에서 차이가 있을 수 있다.

결 론

본 연구는 갤럭시 스마트폰, CSL 및 디지털 녹음기 간에 서

로 다른 종류의 마이크로폰을 사용하고 있다는 점, 저장 확장자 및 양자화가 기기마다 자체적으로 다르게 설정되어 있다는 점, 그리고 녹음장비 간 저장 확장자를 통일하는 과정에서 예상되는 음원 소실이 있을 수 있다는 점 등의 기본적인 한계를 내포하고 있다. 그럼에도 불구하고 본 연구결과를 통해서 저자들은 갤럭시 스마트폰의 녹음기능이 음향분석을 위한 녹음 장비의 하나로 기존의 디지털 녹음기의 기능을 대체할 수 있는 가능성이 있음을 알 수 있었다.

중심 단어 : 갤럭시 스마트폰 · 녹음장비 · 음향분석.

REFERENCES

- 1) Oguz H, Kilic MA, Safak MA. *Comparison of results in two acoustic analysis programs: Praat and MDVP*. Turk J Med Sci 2011;41(5): 835-41.
- 2) 대한후두음성언어의학회. 음향분석. In:진성민, editors. 후두음성언어의학회II 후두음성질환의 이해와 치료. 1판. 일조각;2012. p.396-406.
- 3) Jiand J, Lin E, Hanson DG. *Effect of Tape Recording on Perturbation Measures*. JSLHR 1998;41:1031-41.
- 4) Lin E, Hornibrook J, Ormond T. *Evaluation iphone recordings for acoustic voice assessment*. Folia Phoniatr Logop 2012;64:122-30.
- 5) Vogel AP, Rosen KM, Morgan AT, Reilly S. *Comparability of modern recording devices for speech analysis: Smartphone, Landline, Laptop, and Hard Disc recorder*. Folia Phoniatr Logop 2014;66: 244-50.
- 6) Yoo JY, Jeong OR, Jang TY, Ko DH. *A correlation study among acoustic parameters of MDVP, Praat, and Dr. Speech*. Speech Sciences 2003;10(3):29-36.
- 7) Shim SY, Kim HH, Kim JO, Shin JC. *Difference in voice parameters of MDVP and Praat programs according to severity of voice disorders in vocal nodule*. Speech Sciences 2014;6(2):107-14.
- 8) 고도홍. *Computerized Speech Lab (CSL)의 사용법 및 응용*. In: 고도홍, 신희백, editors. 언어치료를 위한 음성언어의 측정, 분석 및 평가. 1판. 학지사;2015. p.12-31.
- 9) Burrell C, Vorperian HK, Fourakis M, Kent RD, Bolt DM. *Quantitative and descriptive comparison of four acoustic analysis systems: vowel measurements*. JSLHR 2014;57:26-45.
- 10) Pyo HY, Sim HS, Lim SE. *The change of the correlation between GRBAS scales and MDVP parameters according to the different length of voice samples for MDVP analysis*. Speech Sciences 2000;7(2):71-81.
- 11) Jin SM. *Introduction of acoustic analysis of voice*. Korean J Otolaryngol 2004;47:943-9.
- 12) 신지영. *음성학 음운론 연구의 기초를 위하여 말소리의 이해*. 2판. 한국문화사;2014. p.160-73.