



Comparisons of voice quality parameter values measured with MDVP, Praat, and TF32

Hye-Ju Ko^{1,*} · Mee-Ryung Woo² · Yaelin Choi³

¹Department of Psychological Rehabilitation, Myongji University, Seoul, Korea

²Rehabilitation Medical Center, National Health Insurance Corporation Ilsan Hospital, Goyang, Korea

³Department of Psychological Rehabilitation & Speech-Language Pathology, Myongji University, Seoul, Korea

Abstract

Measured values may differ between Multi-Dimensional Voice Program (MDVP), Praat, and Time-Frequency Analysis software (TF32), all of which are widely used in voice quality analysis, due to differences in the algorithms used in each analyzer. Therefore, this study aimed to compare the values of parameters of normal voice measured with each analyzer. After tokens of the vowel sound /a/ were collected from 35 normal adult subjects (19 male and 16 female), they were analyzed with MDVP, Praat, and TF32. The mean values obtained from Praat for jitter variables (J local, J abs, J rap, and J ppq), shimmer variables (S local, S dB, and S apq), and noise-to-harmonics ratio (NHR) were significantly lower than those from MDVP in both males and females ($p < .01$). The mean values of J local, J abs, and S local were significantly lower in the order MDVP, Praat, and TF32 in both genders. In conclusion, the measured values differed across voice analyzers due to the differences in the algorithms each analyzer uses. Therefore, it is important for clinicians to analyze pathologic voice after understanding the normal criteria used by each analyzer when they use a voice analyzer in clinical practice.

Keywords: MDVP, Praat, TF32, jitter, shimmer, noise-to-harmonics ratio (NHR)

1. 서론

음성장애는 누구나 경험할 수 있고 다양한 특성을 가지고 있으며 치료 접근법 또한 다양하다. 그러므로 환자 특성에 맞는 최적의 치료를 위해서는 가장 먼저 적절한 평가가 시행되어야 한다. 보통 임상현장에서 언어재활사들이 음성장애 평가를 위해 사용하는 방법으로는 목소리의 음향 신호를 다양한 도구로 분석하는

음향학적 검사, 환자의 음성을 귀로 듣고 평가하는 청지각적 평가, 환자 스스로 설문지를 작성하여 자신의 음성 상태에 대해 보고하는 자기보고식 평가가 있다(Pyo & Song, 2010). 이중에서도 음향학적 검사는 음성장애 진단 및 치료 효과를 알아보기 위해 개인의 음성 특성을 평가하는 방법으로 매우 중요하다. 개인의 음성은 여러 측면에서 다르게 실현될 수 있고 매개변수들도 매우 다양하여 측정하기 쉽지 않다(Ko, 2003; Yoo et al., 2003). 그러나

* juya2501@nate.com, Corresponding author

Received 5 June 2020; Revised 14 September 2020; Accepted 14 September 2020

© Copyright 2020 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.

음향학적 검사를 이용하면 음성과 관련된 다양한 매개변수들의 객관적 수치를 비침습적인 방법으로 빠르고 쉽게 얻을 수 있다(Choi et al., 2005).

음성장애 평가에 주로 사용되는 음향학적 검사로는 KayPENTAX사의 Computerized Speech Lab(이하 CSL)의 Multi-Dimensional Voice Program(이하 MDVP), Visi-Pitch, Multi-Speech, Voice Range Profile, Tiger Electronics사의 Dr. Speech, Nagashima사의 Phonatory Function Analyzer, Praat, PCQuirer, CSpeech 등이 있다(Pyo & Sim, 2007; Pyo & Song, 2010; Yoo et al., 2003). 이 중에서도 MDVP와 Praat이 음성 평가를 위한 대표적인 도구이며(Yoo et al., 2003) 최근에는 컴퓨터의 DOS 환경에 적합했던 CSpeech를 대체하여 개발된, 윈도우 환경에서 실행 가능한 Time-Frequency Analysis software(이하 TF32)가 국내외에서 점차 사용되고 있는 추세이다(Paul, 2018; TF32 and CSpeech, 2005). 이 세 가지 도구에 대해 구체적으로 살펴보면 다음과 같다.

MDVP는 음성 분석을 위한 주요 프로그램으로 전문 문헌에 많이 인용되어 전 세계에서 일반적으로 많이 사용되고 있다. MDVP로 음성 분석 시 33개의 음성 매개변수가 쉽고 빠르게 도출된다. 또한 임상가가 즉시 음성을 평가할 수 있도록 각 매개 변수의 평균, 표준편차, 역치를 제공한다. 뿐만 아니라 그 결과를 방사형 다이어그램으로 나타내어 임상가가 음성의 병리적 특성을 시각적으로 쉽게 분석할 수 있도록 도와준다(KayPENTAX, 2005; Oğuz et al., 2011).

Praat은 음성 분석 및 음성 변형 프로그램으로 Praat 홈페이지(<https://www.praat.org>)에서 무료로 다운 받을 수 있으며 현재까지 정기적으로 업그레이드 되고 있다. 뿐만 아니라 Praat의 스크립트를 이용하면 많은 양의 데이터를 한번에 빠르게 처리할 수 있어 전 세계의 많은 임상가와 과학자들이 사용하고 있다(Oğuz et al., 2011; Styler, 2017; Van Lieshout, 2017).

TF32는 말소리 또는 가청주파수 파형(audio-frequency) 등을 분석할 수 있는 32비트 윈도우 기반의 시간-주파수 분석 프로그램으로 최근 음성 과학자 및 음성 임상가들이 점차 많이 사용하고 있다(Natour & Saleem, 2009; Paul, 2018). TF32에는 데모 단계(demo level), 기본 단계(basic level), 연구 자동화 단계(lab automation level)의 총 3가지 단계가 있다. 데모 단계는 TF32 홈페이지(<http://userpages.chorus.net/cspeech>)에서 무료로 다운 받아 음성을 손쉽게 분석할 수 있다. 그러나 파일 재생 및 파형 가져오기만 지원하며, 녹음 또는 파형 저장은 지원하지 않는다는 단점이 있다. 음성 녹음 및 파형 저장을 하려면 기본 단계 또는 연구 자동화 단계가 필요한데 이 경우 비용을 지불하고 구입을 해야 한다(TF32 and CSpeech, 2005).

위와 같이 자동화 된 음향학적 검사는 음성에 대한 주관적인 지각적 판단에 대한 의존도를 잠재적으로 줄이며 음성 신호의 정량화 된 측정을 제공하므로 임상가의 실수 가능성을 줄일 수 있다는 장점이 있다(Natour & Saleem, 2009). 그러나 이러한 검사 도구로 음성을 분석 했더라도 음성 수집 시 마이크 유형, 소음 수준, 데이터 수집 시스템, 데이터 분석에 이용되는 소프트웨어 등에 따라 그 음성 분석 결과의 타당성과 신뢰성에 영향을 주게 된다(Amir et al., 2009; Deliyski et al., 2005; Deliyski et al., 2006). 그 중

에서도 특히 각 도구 자체의 분석 알고리즘 차이로 인해 같은 음성에 대해 다양한 음성 분석 도구 간의 데이터 값에 차이가 있을 수 있다(Choi et al., 2005; Yoo et al., 2003).

실제로 그동안 국내외에서는 이러한 문제를 인식하여 같은 음성을 다양한 음향학적 검사 도구로 분석하였을 때 나타날 수 있는 수치 차이에 대해 몇몇 연구가 이루어졌다.

Maryn et al.(2009)은 50명의 음성장애 환자(성대 결절, 성대 폴립, 성대 낭종, 레인케 부종 등)를 대상으로 /아/ 모음을 MDVP와 Praat으로 비교 분석하였다. 그 결과 주파수 변동율(이하 jitter), 진폭 변동율(이하 shimmer) 변수에서 모두 MDVP보다 Praat의 수치가 통계적으로 유의하게 낮다고 하였다.

Amir et al.(2009)은 58명의 여성 음성장애 환자(성대 결절, 성대 폴립, 성대 낭종, 기능적 발성장애)를 대상으로 /아/와 /이/ 모음을 MDVP와 Praat으로 비교 분석하였다. 그 결과 jitter, shimmer, 소음 대 배음비(Noise-to-Harmonics Ratio, 이하 NHR), 무성음 정도(Degree of Voiceless, DUV)의 변수에서 MDVP보다 Praat의 수치가 통계적으로 유의하게 낮다고 하였다. Oğuz et al.(2011) 또한 정상인 18명, 음성장애 환자 29명(일측성 성대마비, 성대 낭종, 중증 근무력증, 성대 결절 등), 총 47명의 남성 및 여성 음성을 대상으로 MDVP와 Praat을 비교 분석하였다. 그 결과 jitter 변수와 NHR에서 MDVP보다 Praat의 수치가 통계적으로 유의하게 낮다고 하였다. Burris et al.(2014)은 성인 남성 및 여성, 아동의 합성모음(synthesized vowel), 자연모음(natural vowels)을 대상으로 CSL, Praat, TF32, Wavesurfer의 4가지 음향학적 검사의 수치를 비교 분석하였다. 그 결과 제조사가 권장하는 기본 설정을 사용했을 때 성인 남성의 음성에서는 CSL보다 Praat, TF32, Wavesurfer에서 가장 적합한 수치를 얻을 수 있고, 성인 여성과 아동의 음성에서는 4가지 음향학적 검사 모두 정확도가 저하되어 있다고 하였다.

Bielamowicz et al.(1996)은 CSpeech, CSL, SoundScope로 경도에서 중도까지의 음성장애 환자 총 50명(남성 29명, 여성 21명)의 음성을 분석한 결과 각 도구 간의 jitter와 shimmer 수치가 낮은 상관관계를 보였다고 하였다. Shim et al.(2014)의 연구에서도 정상 성인 여성 10명, 성대 결절 환자 20명 총 30명을 대상으로 /아/ 모음을 MDVP와 Praat으로 비교 분석하였다. 그 결과 정상 성인의 음성은 jitter 변수, shimmer 변수, NHR에서, 성대 결절 환자의 음성은 jitter 변수와 NHR에서만 MDVP보다 Praat의 수치가 통계적으로 유의하게 낮다고 하였다. Yoo et al.(2003)의 연구에서는 정상 성인 총 45명(여성 30명, 남성 15명)을 대상으로 MDVP, Praat, Dr. Speech 간의 음향학적 측정치에 관한 상관을 살펴보았다. 그 결과 MDVP와 Praat의 기본주파수(Fundamental frequency, 이하 F_0)와 shimmer 수치에서, MDVP와 Dr. Speech의 F_0 , jitter, NHR에서 높은 상관관계가 있다고 하였다. Choi et al.(2005)은 경도에서 심도의 음성장애 환자(성대 결절, 성대 마비, 성대 폴립, 연축성 발성장애 등) 총 19명(남성 9명, 여성 10명)을 대상으로 MDVP, Praat, TF32와 저자들이 새로 개발한 다채널 분석 프로그램의 수치를 비교 분석하였다. 그 결과 각 도구에 따라 jitter와 shimmer 변수의 수치에 통계적으로 유의한 차이가 있다고 하였다.

이와 같이 도구에 따라 분석 결과에 차이가 나타나는 이유는

각 도구에서 사용되는 알고리즘의 차이로 인한 것이라 생각된다. 많은 음향학적 검사 도구 중 대표적으로 MDVP, Praat, TF32의 F_0 검출과 변동을 분석을 위한 알고리즘이 어떻게 다른지 구체적으로 살펴보면 다음과 같다. 먼저 F_0 를 검출하기 위한 방법은 여러 가지가 있는데 처리 영역에 따라 크게 병렬처리법, 자기상관함수법(Autocorrelation Function), 교차상관함수법(Cross Correlation), Average Magnitude Difference Function(AMDF), 영교차율(Zero Crossing Rate) 등이 포함된 시간 영역(time domain)법과 고조파분석법, Comp-filtering, Harmonic Product Spectrum(HPS) 등이 포함된 주파수 영역(frequency domain)법으로 나눌 수 있다(Back et al., 2005; Park, 2005; Rabiner & Schafer, 1978). 이 중에서 MDVP는 자기상관함수법을, TF32는 교차상관함수법을 사용한다. Praat의 경우, 억양 분석 시에는 자기상관함수법을, 음성 분석 시에는 교차상관함수법을 선택하도록 하였다. 여기서 자기상관함수법은 동일한 음성에서 임의의 신호와 그 신호를 지연시킨 후 자기 신호와의 유사성을 살펴보는 것이고, 교차상관함수법은 서로 다른 신호 간의 유사성을 살펴보는 방법이다(Choi et al., 2005; Paul, 2018; Paul & David, 2018; Shim et al., 2014). 변동을 분석을 위한 알고리즘을 각 도구별로 살펴보면 MDVP의 경우 파형이 최대인 시간 위치를 찾는 peak-picking 방식을, Praat의 경우 두개의 연속적인 파형이 최대로 비슷한 시간 거리를 찾는 방법인 waveform-matching 방식을, TF32의 경우 두 개의 연속적인 주기 구간에서의 차이를 계산하는 방식인 최소자승법(least mean square)을 사용한다(Boersma, 2009; Choi et al., 2005; Milenkovic, 1987; Paul, 2018; Paul & David, 2018; Shim et al., 2014).

이를 종합해보면, 자동화 된 음향학적 검사 도구에서는 F_0 검출과 변동을 분석 시 분석 도구에 따라 다양한 알고리즘을 사용할 수 있으며 사용된 알고리즘에 따라 그 결과는 다르게 검출 될 수 있다(Natour & Saleem, 2009). 이는 같은 음성 자료라 할지라도 다른 분석 도구를 사용할 경우, 분석 도구에 따라 데이터 값에 차이가 있을 수 있다는 것을 의미한다.

그럼에도 불구하고 몇몇 음성 분석 도구들은 각 도구에 맞는 역치 기준이 제시되어 있지 않다. 그래서 임상가들이 특정 음성 분석 도구에서 얻은 결과를 다른 검사 도구에서 얻은 결과와 비교하여 분석해야 하는 경우가 많다. 실제로 MDVP는 결과지에 역치 기준이 제시되어 있으나 Praat이나 TF32의 경우에는 각 도구에 맞는 역치 기준이 제시되어 있지 않다. 그래서 임상가들이 Praat이나 TF32에서 얻은 결과값을 MDVP의 역치 기준과 비교하여 병적 음성에 대해 판단하는 경우가 대부분이다. 이러한 경우 분석 도구에 따른 데이터 값에 어떠한 차이가 있는지 고려하지 않는다면, Praat이나 TF32에서 얻은 결과값에 대해 잘못된 분석을 하게 될 가능성이 높아진다. 그러므로 음성 분석 도구에 따라 데이터 값에 어떠한 차이가 있는지 알아볼 필요가 있다.

그러나 최근까지의 연구가 대부분 병적 음성 위주로 진행되어 정상 음성일 때 각 검사 도구에 따라 구체적으로 어떠한 차이가 있는지에 대한 연구가 부족하다. 그러므로 본 연구에서는 음성 분석에 유용한 MDVP, Praat, TF32 각각의 음향학적 검사 도구로 정상 음성 측정치를 비교 분석하여 분석 도구에 따른 음향학

적 검사 변수의 구체적인 차이를 살펴보고자 하였다. 이를 통해 임상가들이 임상현장에서 특정 음향학적 검사 도구를 사용하더라도 다른 분석 도구에서 얻은 데이터 값과 쉽게 비교할 수 있을 뿐만 아니라 병적 음성에 대해서도 손쉽게 파악할 수 있도록 도움을 주고자 하였다.

본 연구의 연구문제는 다음과 같다.

첫째, 동일 성별에서 MDVP, Praat의 공통 매개변수의 측정치에 차이가 있는가?

둘째, 동일 성별에서 MDVP, Praat, TF32의 공통 매개변수의 측정치에 차이가 있는가?

2. 본론

2.1. 연구 대상

본 연구는 현재 서울에 거주하며 연구 참여에 동의한 만 18-45세 이하의 정상 성인 남녀 각 40명, 총 80명을 대상으로 실시하였다. 대상자의 나이를 위와 같이 제한한 이유는 만 18세에 음성 상태가 변성기를 지나 안정기에 접어들어 정상 성인의 음성 수준에 도달하고(Ko et al., 2013), 만 45세부터는 여성의 폐경으로 인해 음성의 변화가 있을 수 있기 때문이다(Lindholm et al., 1997; Linville, 1996). 이들 중 다음과 같은 기준을 모두 충족하는 남성 19명, 여성 16명만을 최종 분석 대상으로 하였다(표 1). 첫째, 음성 상태에 영향을 줄 수 있는 음성장애, 호흡기 질환, 소화기 질환, 갑상선 질환에 대한 병력이 최근 1년 이내에 없었고, 둘째, 언어치료 경력 10년 이상의 1급 언어재활사 2인이 대상자의 음성을 GRBAS 척도(grade, roughness, breathy, asthenic, strained)(Hirano, 1981)로 평가한 결과 대상자 모두 G0(0: 정상, 1: 경도, 2: 중도, 3: 고도)을 받았다. GRBAS는 음향학적 분석을 위해 모음 /아/ 발성 중 중간의 안정적인 구간 1.5초로 편집된 음성 파일을 언어재활사 각자 조용한 환경에서 랜덤으로 재생 후 평가 하였다. 셋째, /아/ 발성 시 MDVP를 이용하여 음질 검사를 한 결과, jitter, shimmer, NHR의 수치 중 하나라도 MDVP에서 제공하는 역치 기준(jitter: 1.040%, shimmer: 3.810%, NHR: 0.190)을 초과한 경우 그 대상자는 제외하였다. 넷째, 월경 주기에 따라 음성 변화 가능성이 있으므로 월경 중인 여성도 대상자에서 제외 하였고(Chae et al., 2001), 다섯째, 현재 흡연 중인 경우도 대상자에서 제외하였다.

표 1. 연구 대상자 정보
Table 1. Information of subjects

성별	인원(명)	연령(세)
남성	19	21.0±3.2
여성	16	24.0±6.3

mean±SD.

2.2. 연구 방법

2.2.1. 자료 수집

음성 자료 수집은 소음이 통제된 음성검사실에서 실시하였다.

대상자는 평소 사용하는 편안한 음도와 강도로 ‘○○○입니다.’라는 발화에 뒤이어 모음 /아/를 연장발성 하도록 하였다. 이때 모음 /아/의 길이가 총 3초 이상이 되도록 길게 발성 하도록 하였고, 총 2회 반복하여 녹음하였다(Yun & Kwon, 1998). 음성 녹음 시 각각의 음향학적 검사 도구에서 받을 수 있는 영향을 배제하고자 portable digital recorder(TASCAM DR-07, Japan)를 사용하였고, 마이크는 TASCAM DR-07에 내장된 스테레오 마이크를 이용하였다. 샘플링 44,100 Hz, 양자화 16 bit, 스테레오였고, 녹음기는 대상자의 입에서 6cm 떨어진 위치에 15도 각도로 고정된 후 사용하였다(Baek et al., 2012).

2.2.2. 분석 절차

2.2.2.1. 음성 자료 편집

수집된 음성 자료 편집 또한 각각의 음향학적 검사 도구에서 받을 수 있는 영향을 배제하고자 오디오 편집기인 Cool Edit Pro version 2.1을 사용하였다. 수집된 3초의 음성 중 발성의 첫 부분과 마지막부분을 제외하고, 안정 구간 1.5초를 선택하였다(Pyo et al., 2000). 1.5초로 자른 음성을 스테레오에서 모노로 변환하기 위해 Cool Edit Pro의 편집 메뉴에서 샘플 유형 변환(convert sample type)을 선택한 후 샘플링 44,100 Hz, 양자화 16 bit를 선택하였다. 또한 채널에서 음성 녹음 시 좌우 채널의 특성을 모두 반영하기 위해 왼쪽 믹스(left Mix) 50%, 오른쪽 믹스(right Mix) 50% 입력 후 모노(mono)로 변환하였다. 이는 스테레오에서 모노로 변환할 경우 가장 일반적인 믹싱 방법이라는 Cool Edit Pro의 도움말에 근거한 것이다(Syntrillium, n.d.).

이후 1.5초로 편집되고 스테레오에서 모노로 변환된 모음 /아/를 CSL의 MDVP(Model 5105, KayPentax, Lincoln Park, NJ, USA), Praat(version 6.0.42, Paul Boersma and David Weenink), TF32(Lab Automation level, Milenkovic, 2001, Madison, WI, USA) 각 도구로 분석하였다. 총 2회씩 반복 측정된 모든 자료를 분석한 후 그 평균 값을 비교하였다.

2.2.2.2. 분석 도구의 설정

분석을 위해 세 가지 음성 도구 모두 각 도구의 표준 설정을 참고하였다.

MDVP 분석을 위해 MDVPvoice 옵션 메뉴의 기본주파수 분석 범위(F_0 Analysis Range) 탭에서 기본주파수 분석 범위를 정상 범위인 70–625 Hz로 선택하였다(KayPENTAX, 2005).

Praat은 MDVP와의 음도 범위를 맞추기 위해 음도 세팅(pitch setting)에서 음도 범위(pitch range)를 70–625 Hz로 설정하였다. 분석 방법(analysis method)은 음성 분석에 적합한 교차상관함수법(cross-correlation)으로 설정하였다(Paul & David, 2018).

TF32는 시간-주파수 세팅(time-frequency settings)에서 주파수

범위(frequency range) 5.060 kHz, Floor -75 dB, Dynamic range 48 dB로 맞추었다. 이후 jitter, shimmer, NHR 분석을 위해 열기(open) 메뉴의 jitter 탭에서 tokens 10을 입력한 후 계산되어 나온 값 중 Avg 값을 분석하였다(Paul, 2018).

2.2.2.3. 분석한 공통 매개변수

분석은 MDVP와 Praat의 공통 매개변수인 평균기본주파수(Mean fundamental frequency, 이하 MF_0), jitter의 4개 변수인 국소적 주파수 변동율(Jitter local, 이하 J local), 절대적 주파수 변동율(Jitter absolute, 이하 J abs)¹, 상대적 평균 변동율(Jitter relative average perturbation, 이하 J rap), 주기 변동율 지수(Jitter period perturbation quotient, 이하 J ppq), shimmer의 3개 변수인 국소적 진폭 변동율(Shimmer local, 이하 S local), 진폭 변동율 데시벨(Shimmer dB, 이하 S dB), 진폭 변동율 지수(Shimmer amplitude perturbation quotient, S apq), NHR의 총 9개 변수를 비교 분석하였다. 또한 MDVP, Praat, TF32의 공통 매개변수인 MF_0 , J local, J abs, S local의 총 4개 변수를 비교 분석하였다. 각 매개변수에 대한 구체적인 설명은 부록 1에 제시하였다(KayPENTAX, 2005; Oğuz et al., 2011; Paul, 2018; Paul & David, 2018; Shim et al., 2014).

2.3. 통계 분석

통계분석은 Statistical Product and Service Solution(SPSS, version 21.0) 통계 프로그램을 이용하였다.

동일 성별에서 두 가지 음향학적 검사 도구(MDVP, Praat)간의 공통 매개변수에 차이가 있는지를 살펴보기 위해 대응표본 t 검정(paired t -test)을 실시하였다. 또한 동일 성별에서 세 가지 음향학적 검사 도구(MDVP, Praat, TF32) 간의 공통 매개 변수에 차이가 있는지를 살펴보기 위해 반복측정 분산분석(repeated measures of ANOVA)을 실시하였다. 반복측정 분산분석 결과 본 연구에서는 S local에서만 구형성 가정을 만족하였다. 나머지 변수에서는 유의 수준이 0.05 미만으로 구형성 가정을 만족하지 않아 그린하우스-가이저(Greenhouse-Geisser)의 유의수준을 사용하였고 0.05 미만에서 검정하였다. 또한 세 가지 음향학적 검사 도구 간의 차이가 통계적으로 유의한 경우, 어떤 도구에서 차이가 있는지를 살펴보기 위해 Bonferroni 사후검정을 실시하였다.

3. 연구 결과

3.1. MDVP, Praat, TF32의 공통 매개변수 차이

MF_0 평균의 경우, MDVP, Praat, TF32 각각 남성은 109.377 Hz, 109.376 Hz, 109.379 Hz, 여성은 203.902 Hz, 203.906 Hz, 203.906 Hz로 각 도구에 따라 통계적으로 유의한 차이는 없었다(그림 1, 2). J local 평균의 경우, MDVP, Praat, TF32 각각 남성은 0.443%, 0.333%, 0.304%, 여성은 0.501%, 0.312%, 0.277%로 각 도구에 따

1 J abs 단위의 경우, MDVP와 Praat은 마이크로 세컨드(millisecond; msec)를, TF32는 밀리세컨드(millisecond;msec)를 사용하여 분석 시 TF32의 수치를 마이크로 세컨드로 변환하여 분석함(1 msec=1,000 msec).

라 통계적으로 유의한 차이가 있었다(표 2, 그림 3, 4). J abs 평균의 경우, MDVP, Praat, TF32 각각 남성은 41.144 μ s, 30.767 μ s, 28.026 μ s, 여성은 24.851 μ s, 15.551 μ s, 13.844 μ s로 각 도구에 따라 통계적으로 유의한 차이가 있었다(표 2). S local 평균의 경우, MDVP, Praat, TF32 각각 남성은 2.472%, 2.048%, 1.740%, 여성은 2.417%, 1.857%, 1.517%로 각 도구에 따라 통계적으로 유의한 차이가 있었다($p<.001$; 표 2, 그림 5, 6).

즉 jitter 변수(J local, J abs), shimmer 변수(S local) 평균의 경우 남성과 여성 모두 MDVP, Praat, TF32 순으로 그 수치가 통계적으로 유의하게 낮아졌다.

표 2. MDVP, Praat, TF32 공통매개변수 측정치

Table 2. Measurement values of common parameters in MDVP, Praat, TF32

변수 (단위)		성별 [†]	Mean±SD			p값
			MDVP	Praat	TF32	
MF ₀ (Hz)		M	109.377 ±2.903	109.376 ±2.903	109.379 ±2.902	.700
		F	203.902 ±16.366	203.906 ±16.366	203.906 ±16.359	.571
Jitter 변수	J local (%)	M	0.443 ±0.168	0.333 ¹⁾ ±0.081	0.304 ³⁾⁵⁾ ±0.075	<.001***
		F	0.501 ±0.212	0.312 ²⁾ ±0.120	0.277 ³⁾⁵⁾ ±0.108	<.001***
	J abs (μs)	M	41.144 ±17.554	30.767 ¹⁾ ±7.778	28.026 ³⁾⁵⁾ ±7.344	<.01**
		F	24.851 ±11.069	15.551 ¹⁾ ±6.423	13.844 ³⁾⁵⁾ ±5.833	<.001***
Shim mer 변수	S local (%)	M	2.472 ±0.648	2.048 ²⁾ ±0.806	1.740 ³⁾⁵⁾ ±0.793	<.001***
		F	2.417 ±0.751	1.857 ¹⁾ ±0.750	1.517 ⁴⁾⁵⁾ ±0.552	<.001***

[†]M, Male, F, Female, * $p<.05$, ** $p<.01$, *** $p<.001$.

¹⁾MDVP와 Praat의 차이 $p<.01$,

²⁾MDVP와 Praat의 차이 $p<.001$,

³⁾Praat과 TF32의 차이 $p<.001$,

⁴⁾Praat과 TF32의 차이 $p<.01$,

⁵⁾MDVP와 TF32의 차이 $p<.001$.

MDVP, Multi-Dimensional Voice Program; TF32, Time-Frequency Analysis software; MF₀, Mean fundamental frequency.

3.2. MDVP와 Praat의 공통 매개변수 차이

남성의 경우, MDVP에서 각 변수의 평균은 J rap 0.251%, J ppq 0.267%, S dB 0.216dB, S apq 2.100%, NHR 0.130이었다. Praat에서 각 변수의 평균은 J rap 0.181%, J ppq 0.203%, S dB 0.178dB, S apq 1.850%, NHR 0.011로 모든 변수에서 MDVP보다 Praat의 수치가 통계적으로 유의하게 낮았다($p<.01$; 표 3).

여성의 경우에도, MDVP에서 각 변수의 평균은 J rap 0.303%, J ppq 0.293%, S dB 0.210dB, S apq 1.692%, NHR 0.123이었다. Praat에서 각 변수의 평균은 J rap 0.185%, J ppq 0.184%, S dB 0.162dB, S apq 1.259%, NHR 0.006으로 모든 변수에서 MDVP보다 Praat의 수치가 통계적으로 유의하게 낮았다($p<.01$; 표 3).

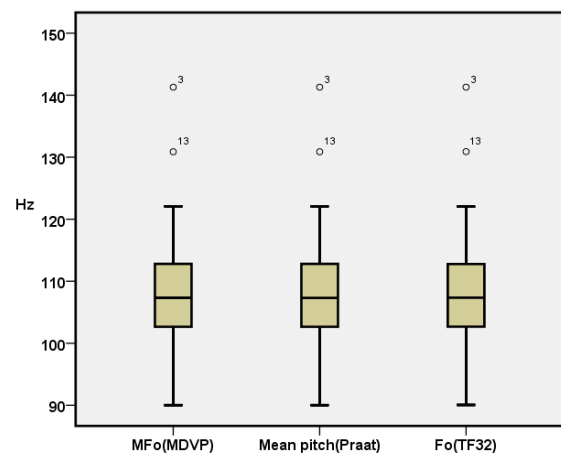
표 3. MDVP, Praat 공통매개변수 측정치

Table 3. Measurement values of common parameters in MDVP and Praat

변수 (단위)		성별 [†]	Mean±SD		p값
			MDVP	Praat	
Jitter 변수	J rap (%)	M	0.251±0.102	0.181±0.051	<.01**
		F	0.303±0.131	0.185±0.075	<.001***
	J ppq (%)	M	0.267±0.108	0.203±0.051	<.01**
		F	0.293±0.125	0.184±0.068	<.001***
Shim mer 변수	S dB (dB)	M	0.216±0.058	0.178±0.071	<.001***
		F	0.210±0.065	0.162±0.066	<.001***
	S apq (%)	M	2.100±0.717	1.850±0.891	<.01**
		F	1.692±0.543	1.259±0.509	<.01**
NHR		M	0.130±0.015	0.011±0.007	<.05*
		F	0.123±0.015	0.006±0.004	<.001***

[†]M, Male, F, Female, * $p<.05$, ** $p<.01$, *** $p<.001$.

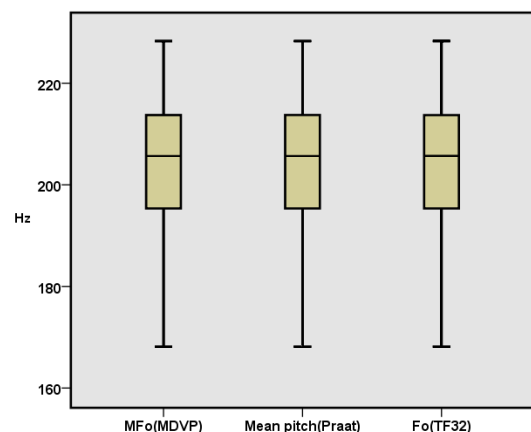
MDVP, Multi-Dimensional Voice Program; TF32, Time-Frequency Analysis software; NHR, noise-to-harmonics ratio.



MF₀, Mean fundamental frequency.

그림 1. 세 가지 음성분석 도구의 MF₀ 수치(남성)

Figure 1. MF₀ values in three voice analyzers(male)



MF₀, Mean fundamental frequency.

그림 2. 세 가지 음성분석 도구의 MF₀ 수치(여성)

Figure 2. MF₀ values in three voice analyzers(female)

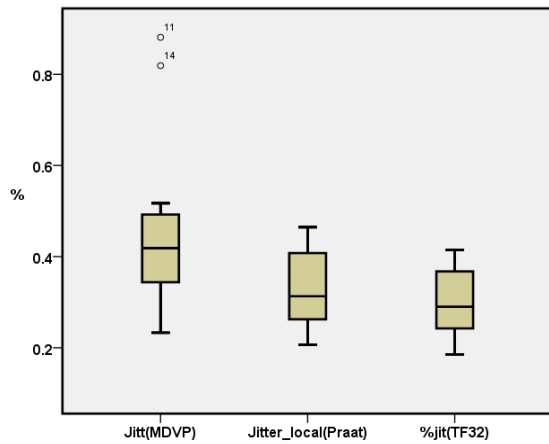


그림 3. 세 가지 음성분석 도구의 J local 수치(남성)
Figure 3. J local values in three voice analyzers(male)

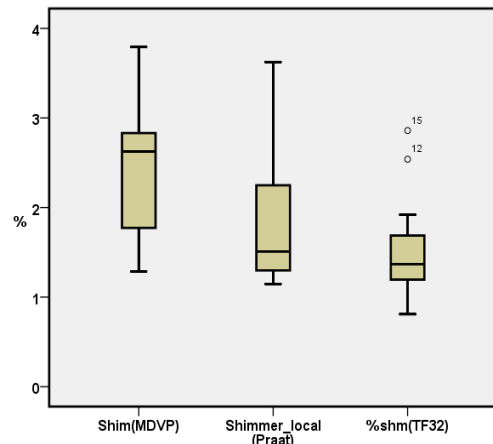


그림 6. 세 가지 음성분석 도구의 S local 수치 차이(여성)
Figure 6. S local values in three voice analyzers(female)

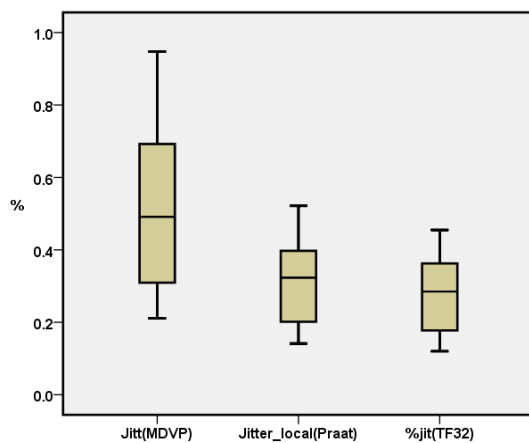


그림 4. 세 가지 음성분석 도구의 J local 수치 차이(여성)
Figure 4. J local values in three voice analyzers(female)

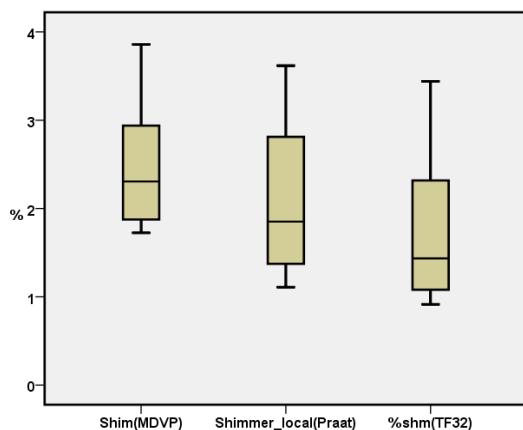


그림 5. 세 가지 음성분석 도구의 S local 수치 차이(남성)
Figure 5. S local values in three voice analyzers(male)

4. 논의 및 결론

음성장애를 진단하기 위해서는 음향학적 검사로 음성을 분석하는 것이 매우 중요하다. 그러나 도구 자체의 알고리즘 차이로 인해 음향학적 검사 도구 간의 수치에 차이가 있을 수 있다(Choi et al., 2005; Yoo et al., 2003). 그러므로 본 연구에서는 음성 분석에 유용한 MDVP, Praat, TF32로 동일한 음성을 비교 분석하여 각각의 도구에 따라 음향학적 변수에 차이가 있는지 살펴보고자 하였다.

MF₀ 평균의 경우, 남성은 MDVP, Praat, TF32 각각 109.377 (±2.903) Hz, 109.376(±2.903) Hz, 109.379(±2.902) Hz, 여성 203.902(±16.366) Hz, 203.906(±16.366) Hz, 203.906(±16.359) Hz로 MDVP, Praat, TF32 간의 통계적으로 유의한 차이는 없었다. 이는 다양한 음향학적 검사 도구 간 F₀에는 큰 차이가 없었다는 여러 선행연구 결과와 일치한다(Amir, 2009; Choi et al., 2005; Nam et al. 2005; Oğuz et al., 2011; Shim et al., 2014; Yoo et al., 2003). 본 연구에서 F₀ 검출 시 MDVP는 자기상관함수법을, Praat과 TF32는 교차상관함수법을 사용하였음에도(Choi et al., 2005; Paul, 2018; Paul & David, 2018; Shim et al., 2014), 세 도구 간의 F₀에는 유의한 차이가 없었다. 그 이유를 추정해보면 다음과 같다. 본 연구에서 사용된 모음 /아/는 억양이 포함되지 않은 단음도로 발생되었고, 특히 그 중에서도 음도가 안정적인 구간을 분석하였다. 이로 인해 대부분의 구간에서 F₀가 일정하여 동일한 음성 내에서 신호를 추출하는 두 가지 방법에 따라 F₀에 큰 차이가 없었을 것으로 추정된다. 또한 F₀ 검출 시 처리 과정상 주파수 영역으로의 변환 과정이 필요한 주파수 영역법과 달리 시간 영역법은 시간 영역에서 직접 처리 하여 다른 영역으로의 변환이 필요 없다. 이로 인해 프로세싱으로 인한 정보의 손실이 적어 도구 간 F₀의 절대적 수치가 다르더라도 그 패턴은 일정하게 유지된다(Park, 2005; Shin et al., 2000; Yoo et al., 2003). 그러므로 자기상관함수법과 교차상관함수법은 모두 시간 영역법에 해당되어 세 가지 음성분석도구 간의 MF₀에 차이가 없었을 것으로 판단된다. 이는 자기상관함수법을 사용한 MDVP와 교차상관함수법을 사용한 TF32의 F₀에 차이가 없었다는 Choi et al.(2005) 및 Nam et al.(2005)의 연구 결과가 이를

지지한다.

세 가지 음성분석도구 간 F_0 에는 차이가 없었으나, MDVP, Praat, TF32의 공통 매개변수인 jitter 변수(J local, J abs)와 shimmer 변수(S local)는 남녀 모두 MDVP, Praat, TF32 순으로 통계적으로 유의하게 수치가 낮아졌다($p < .01$). 또한 MDVP와 Praat의 공통 매개변수인 jitter 변수(J rap, J ppq)와 shimmer 변수(S dB, S apq), NHR에서도 남녀 모두 MDVP보다 Praat의 수치가 통계적으로 유의하게 낮았다($p < .05$). 이는 jitter, shimmer, NHR에서 MDVP보다 Praat의 수치가 낮았고(Amir, 2009; Maryn et al., 2009; Nam et al., 2005; Yoo et al., 2003), MDVP보다 TF32의 수치 또한 낮았다(Choi et al., 2005)는 선행연구 결과와 일치한다. 반면 Oğuz et al.(2011)의 연구에서는 jitter와 NHR의 수치가 MDVP보다 Praat이 낮았으나, shimmer 수치는 오히려 MDVP보다 Praat이 높아 본 연구와 다른 결과를 보였다. 또한 Choi et al.(2005)의 연구에서도 shimmer 수치는 MDVP, Praat, TF32 순으로 낮아졌으나, jitter 수치는 MDVP, TF32, Praat 순으로 낮아져 본 연구 결과와 일치하지 않았다. 이는 이 두 선행연구(Choi et al., 2005; Oğuz et al., 2011) 모두 병적 음성을 대상으로 연구를 진행하여 정상 음성을 대상으로 한 본 연구와 차이가 있었을 것으로 추측된다. 실제로 정상 음성과 병적 음성 모두를 대상으로 한 Shim et al.(2014)의 연구를 살펴보면 정상 음성은 shimmer 변수의 수치가 MDVP보다 Praat이 통계적으로 유의하게 낮은 반면, 병적 음성은 MDVP와 Praat 간 shimmer 변수에 통계적으로 유의한 차이가 없어 이를 뒷받침한다. 즉, 성대의 진동이 불규칙할 때의 진폭은 MDVP와 Praat이 비슷한 방식으로 처리하여(Shim et al., 2014) 정상 음성과 병적 음성의 shimmer 변수에 대한 도구 간 처리 방식에 차이가 있었을 것으로 판단된다.

이처럼 동일한 음성에 대해 음성분석도구에 따라 jitter 변수, shimmer 변수, NHR의 수치가 현저히 달라지는 것은 각 도구 간 알고리즘 차이로 인한 것이라 판단된다(Choi et al., 2005; Shim et al., 2014; Yoo et al., 2003). 실제로 Boersma(2009)는 MDVP가 사용하는 peak-picking과 Praat이 사용하는 waveform-matching에 차이가 있는지를 연구하였다. 그 결과 잡음이 포함되지 않는 0.001%-20% 사이의 기본 jitter 값을 갖는 소리의 경우 두 가지 방식에 따라 측정치에 차이가 없었다고 한다. 반면 동일한 소리에 잡음을 추가한 후 두 가지 방식에 따라 다시 측정한 결과, peak-picking 방식이 waveform-matching 방식보다 28배 더 민감하게 측정되었다고 하였다. 이는 소리에 잡음이 포함될 때 어떠한 방식을 사용하느냐에 따라 매우 다른 결과를 나타낸다는 것을 의미한다. 그 이유는 peak-picking은 파형이 최대인 시간 위치를 조사하여 두 개의 가장 높은 정점(peak) 사이의 시간차를 측정하므로 잡음에 크게 영향을 받지만, waveform-matching은 파형의 전체 모양을 고려하여 가장 일치하는 파형 모양으로 그 값을 결정하므로 잡음에 비교적 적게 영향을 받기 때문이다(Boersma, 2009). 결국 사람의 음성은 잡음이 포함될 수밖에 없어 이러한 결과가 나타났으리라 추정된다.

본 연구에서도 정상 음성을 대상으로 하였음에도 jitter 변수, shimmer 변수, NHR에서 MDVP의 수치가 Praat이나 TF32의 수치

보다 현저히 높아 MDVP가 다른 음향학적 검사 도구에 비해 음성의 변동율에 매우 민감하게 반응하는 것을 알 수 있었다. 특히 J local의 경우, 그림 3의 상자 도표에서 알 수 있듯이 Praat이나 TF32와는 달리 MDVP에서만 최대값을 넘어서는 이상치(outlier)가 발견되어 다른 음성분석도구에 비해 MDVP가 jitter 변화에 매우 민감하다는 것을 지지한다(Shim et al., 2014). 반면 TF32는 jitter 변수와 shimmer 변수에서 MDVP와 Praat에 비해 그 수치가 현저히 낮았다. TF32에서 변동율 분석을 위해 사용된 최소자승법은 잡음을 제거하고 시스템의 성능을 향상시키는 필터로(Ahn & Oh, 2012), 이로 인해 이러한 결과가 나타난 것으로 추정된다. 때문에 TF32의 경우 병적 음성을 비교적 좋은 음질의 음성으로 분석할 가능성이 있다.

이처럼 음질 분석에서 매우 중요한 jitter 변수, shimmer 변수, NHR의 수치가 어떤 음성분석도구를 사용했는지에 따라 현저한 차이가 나타나므로 임상현장에서 임상가들이 각각의 음향학적 검사 도구로 병적 음성을 파악할 때 주의가 필요하다.

결론적으로 MDVP, Praat, TF32에서 동일한 음성을 사용하여 음성분석을 한 결과, 음향학적 검사 도구 간에 차이가 있었다. 이는 각 도구에서 사용된 알고리즘 차이로 인한 것으로 임상가들이 임상현장에서 각각의 도구를 사용할 때 각 도구의 정상 수치에 대해 이해한 후 병적 음성을 분석하는 것이 중요하다는 것을 의미한다.

본 연구는 음질 분석에 매우 유용하고 현재 임상 현장에서 많이 쓰이고 있는 MDVP, Praat, TF32의 측정치에 차이가 있다는 것을 밝혔는데 그 의의가 있다. 반면, 본 연구의 제한점은 정상 음성을 대표하기에는 연구 대상자의 수가 적다는 데 있다. 추후 대상자를 더 추가하여 Praat과 TF32의 정상 기준치에 대한 연구가 더 필요할 것이다.

감사의 글

음성데이터 분석에 도움을 주신 이아름 선생님께 감사드립니다.

References

- Ahn, C. S., & Oh, S. Y. (2012). CHMM modeling using LMS algorithm for continuous speech recognition improvement. *Journal of Digital Convergence*, 10(11), 377-382.
- Amir, O., Wolf, M., & Amir, N. (2009). A clinical comparison between two acoustic analysis softwares: MDVP and Praat. *Biomedical Signal Processing and Control*, 4(3), 202-205.
- Back, S. E., Kim, J. Y., Na, S. Y., & Choi, S. H. (2005). Speaker separation based on directional filter and harmonic filter. *Phonetics and Speech Sciences*, 12(3), 125-136.
- Back, Y., Kim, S., Kim, E., & Choi, Y. (2012). Vocal acoustic characteristics of speakers with depression. *Phonetics and Speech Sciences*, 4(1), 91-98.

- Bielamowicz, S., Kreiman, J., Gerratt, B. R., Dauer, M. S., & Berke, G. S. (1996). Comparison of voice analysis systems for perturbation measurement. *Journal of Speech, Language, and Hearing Research*, 39(1), 126-134.
- Boersma, P. (2009). Should jitter be measured by peak picking or by waveform matching? *Folia Phoniatica et Logopaedica*, 61(5), 305-308.
- Burris, C., Vorperian, H. K., Fourakis, M., Kent, R. D., & Bolt, D. M. (2014). Quantitative and descriptive comparison of four acoustic analysis systems: Vowel measurements. *Journal of Speech, Language, and Hearing Research*, 57(1), 26-45.
- Chae, S. W., Choi, G., Kang, H. J., Choi, J. O., & Jin, S. M. (2001). Clinical analysis of voice change as a parameter of premenstrual syndrome. *Journal of Voice*, 15(2), 278-283.
- Choi, S. H., Nam, D. H., Lee, S. H., Jung, W. H., Kim, D. W., & Choi, H. S. (2005). Jitter and shimmer measurements of dysphonia among the different voice analysis programs. *Journal of The Korean Society of Laryngology, Phoniatrics and Logopedics*, 16(2), 140-145.
- Deliyski, D. D., Shaw, H. S., & Evans, M. K. (2005). Influence of sampling rate on accuracy and reliability of acoustic voice analysis. *Logopedics Phoniatrics Vocology*, 30(2), 55-62.
- Deliyski, D. D., Shaw, H. S., Evans, M. K., & Vesselinov, R. (2006). Regression tree approach to studying factors influencing acoustic voice analysis. *Folia Phoniatica et Logopaedica*, 58(4), 274-288.
- Hirano, M. (1981). "GRBAS" scale for evaluating the hoarse voice & frequency range of phonation. *Clinical Examination of Voice*, 5, 83-84.
- KayPENTAX(2005). *Multi-Dimensional Voice Program(MDVP) Model 5105. Instruction Manual*. A Division of PENTAX Medical Company 2 Bridgewater Lane Lincoln Park, NJ.
- Ko, D. H. (2003). A study of extracting acoustic parameters for individual speakers. *Phonetics and Speech Sciences*, 10(2), 129-143.
- Ko, H. J., Kang, M. J., Kwon, H. J., Choi, Y., Lee, M. G., & Choi, H. S. (2013). Acoustic characteristics on the adolescent period aged from 16 to 18 years. *Phonetics and Speech Sciences*, 5(1), 81-90.
- Lindholm, P., Vilkman, E., Raudaskoski, T., Suvanto-Luukkonen, E., & Kauppila, A. (1997). The effect of postmenopause and postmenopausal HRT on measured voice values and vocal symptoms. *Maturitas*, 28(1), 47-53.
- Linville, S. E. (1996). The sound of senescence. *Journal of voice*, 10(2), 190-200.
- Maryn, Y., Corthals, P., De Bodt, M., Van Cauwenberge, P., & Deliyski, D. (2009). Perturbation measures of voice: A comparative study between multi-dimensional voice program and praat. *Folia Phoniatica et Logopaedica*, 61(4), 217-226.
- Milenkovic, P. (1987). Least mean square measures of voice perturbation. *Journal of Speech, Language, and Hearing Research*, 30(4), 529-538.
- Nam, K. C., Lee, S. H., Choi, J. N., Choi, H. S., Nam, D. H., & Kim, D. W. (2005, May). Comparison of vowel pitch results among several commercial voice analysis programs. *Proceedings of the KIEE Conference* (pp. 54-56). Seoul, Korea.
- Natour, Y. S., & Saleem, A. F. (2009). The performance of the time-frequency analysis software (TF32) in the acoustic analysis of the synthesized pathological voice. *Journal of Voice*, 23(4), 414-424.
- Oğuz, H., Kiliç, M. A., & ŞAFAK, M. A. (2011). Comparison of results in two acoustic analysis programs: Praat and MDVP. *Turkish Journal of Medical Sciences*, 41(5), 835-841.
- Park, S. B. (2005). *A voice signal transformation scheme for voice-based music retrieval* (Master's thesis). Ajou University, Korea.
- Paul, B., & David, W. (2018). *Praat manual* (version 6.0.42). Amsterdam, the Netherlands: University of Amsterdam.
- Paul, H. M. (2018). *TF32 User's manual*. Madison, WI: University of Wisconsin-Madison. Retrieved from <http://userpages.chorus.net/cspeech/TF32.pdf>
- Pyo, H. Y., Sim, H. S., & Lim, S. E. (2000). The change of correlation between GRBAS scales and MDVP parameters according to the different length of voice samples for MDVP analysis. *Phonetics and Speech Sciences*, 7(2), 71-81.
- Pyo, H. Y., & Sim, H. S. (2007). A study for the development of Korean voice assessment model for the patients with voice disorders: A qualitative study. *Phonetics and Speech Sciences*, 14(2), 7-22.
- Pyo, H. Y., Sim, H. S., Song, Y. K., Yoon, Y. S., Lee, E. K., Lim, S. E., Hah, H. R., & Choi, H. S. (2002). The acoustic study on the voices of Korean normal adults. *Phonetics and Speech Sciences*, 9(2), 179-192.
- Pyo, H. Y., & Song, Y. (2010). Recent trends in evaluation and diagnosis of voice disorders: A literature review. *Communication Sciences and Disorders*, 15(4), 506-525.
- Rabiner, L. R., & Schafer, R. W. (1978). *Digital processing of speech signals*. Englewood Cliffs, NJ: Prentice Hall.
- Shim, S. Y., Kim, H. H., Kim, J. O., & Shin, J. C. (2014). Difference in voice parameters of MDVP and praat programs according to severity of voice disorders in vocal nodule. *Phonetics and Speech Sciences*, 6(2), 107-114.
- Shin, D. S., Kim, J., & Bae, M. J. (2000). *On a pitch detection using AME in Transition Region* (Research report). Jincheon, Korea: National IT Industry Agency.
- Styler, W. (2017). Using Praat for linguistic research. Version 1.8.1. Retrieved from <http://savethevowels.org/praat> on August 1, 2018.
- Syntrillium. (n.d.). Cool Edit Pro [Computer Software]. <http://www.syntrillium.com>

- TF32 and CSpeech. (2005). Retrieved from <http://userpages.chorus.net/cspeech/>
- Van Lieshout, P. (2017). Praat short tutorial (version 5.0). Toronto, ON: University of Toronto. Retrieved from https://www.researchgate.net/publication/270819326_PRAAT_-_Short_Tutorial_-_An_introduction
- Yoo, J. Y., Jeong, O. R., Jang, T. Y., & Ko, D. H. (2003). A Correlation study among acoustic parameters of MDVP, Praat, and Dr. Speech. *Phonetics and Speech Sciences*, 10(3), 29-36.
- Yun, S. Y., & Kwon, D. H. (1998). Acoustic characteristics of normal children's voice of 5 to 11 years old. *Journal of Speech-Language & Hearing Disorders*, 7(1), 67-78.

• **고혜주 (Hye-Ju Ko)**

명지대학교 일반대학원 심리재활학협동과정 박사 수료 & 대
구보건대학 언어치료과 강사
서울시 서대문구 거북골로 34
Tel: 010-2768-7048
Email: juya2501@nate.com
관심분야: 음성장애, 신경언어장애, 삼킴장애

• **우미령 (Mee-Ryung Woo)**

국민건강보험 일산병원 재활치료센터 언어재활사 & 명지대
학교 일반대학원 심리재활학협동과정 박사 수료
경기도 고양시 일산동구 백석동 1232
Tel: 010-8922-4542
Email: meeryung0622@hanmail.net
관심분야: 신경언어장애, 언어발달장애, 음성장애

• **최예린 (Yaelin Choi) 교신저자**

명지대학교 심리재활학협동과정 & 언어치료학과 교수
서울시 서대문구 거북골로 34
Tel: 02-300-0882
Email: ylchoi@mju.ac.kr/yaelinchoi@gmail.com
관심분야: 음성장애, 말운동장애, 삼킴장애, 연구방법론

부록 1. MDVP, Praat, TF32의 공통 매개변수 정의

공통 매개변수 (약어)/단위/	MDVP 매개변수	Praat 매개변수	TF32 매개변수	정의
Mean Fundamental Frequency (MF ₀) /Hz/	Mean Fundamental Frequency (MF ₀)	Mean pitch	F ₀	분석된 구간의 F ₀ 전체 평균
Jitter local (J local) /%/	Jitter Percent (Jitt)	Jitter (local)	%jit	음도의 매우 짧은 기간에서 주기간(cycle-to-cycle) 변동성에 대한 상대적 평가로 연속 주기 내에서 평균 절대값 차이를 평균 주기로 나눈 값
Jitter absolute (J abs) /μs/	Absolute Jitter (Jita)	Jitter (local, absolute)	jit (absolute jitter) /ms/	음도 주기의 주기간 변동성에 대한 평가로 연속된 주기 사이의 평균 절대값 차이. 단위는 마이크로세컨드(microseconds)임. 단, TF32의 단위는 밀리세컨드(millisecond)임
Jitter relative average perturbation (J rap) /%/	Relative average perturbation (RAP)	Jitter (rap)	-	음도의 3주기 내에서 주기간 변동성에 대한 상대적 평가로 한 주기와 그 주기와 인접한 두 주기의 평균사이의 평균 절대값 차이를 평균 주기로 나눈 값
Jitter period perturbation quotient (J ppq) /%/	Pitch perturbation quotient (PPQ)	Jitter (ppq5)	-	음도의 5주기 내에서 주기간 변동성에 대한 상대적 평가로 한 주기와 그 주기와 가장 인접한 4주기의 평균 사이의 평균 절대값 차이를 평균 주기로 나눈 값
Shimmer local (S local) /%/	Shimmer Percent (Shim)	Shimmer (local)	%shm	매우 짧은 기간에서 진폭간(peak-to-peak amplitude)의 주기간 변동성에 대한 상대적 평가로 연속된 주기의 진폭 사이의 평균 절대값 차이를 평균 진폭으로 나눈 값
Shimmer dB (S dB) /dB/	Shimmer in dB (ShdB)	Shimmer (local, dB)	-	매우 짧은 기간에서 진폭간주기간 변동성에 대한 dB 평가
Shimmer amplitude perturbation quotient (S apq) /%/	Amplitude perturbation quotient (APQ)	Shimmer (apq11)	-	11주기에서 진폭간의 주기간 변동성에 대한 상대적 평가로 한 주기에 대한 진폭과 그 주기와 가장 인접한 10 주기에 대한 진폭의 평균 사이의 평균 절대값 차이
Noise-to-Harmonics Ratio (NHR)	Noise to Harmonic Ratio (NHR)	Mean noise-to-harmonics ratio	-	70-4,200 Hz 주파수 범위 내에서 배음 스펙트럼 에너지에 대한 잡음 스펙트럼 에너지 비율의 평균

TF32, Time-Frequency Analysis software.

MDVP, Praat, TF32에 따른 음향학적 측정치에 대한 비교

고혜주¹ · 우미령² · 최예린³¹명지대학교 심리재활학과학과간협동과정, ²국민건강보험 일산병원 재활치료센터,³명지대학교 심리재활학과학과간협동과정 & 언어치료학과

국문초록

음질 분석에 매우 유용한 Multi-Dimensional Voice Program (MDVP), Praat, Time-Frequency Analysis software (TF32)는 각각의 음향학적 검사에 사용된 알고리즘 차이로 인해 그 측정치에 차이가 있을 수 있다. 그러므로 본 연구에서는 각각의 음향학적 검사 도구로 음성 측정치를 비교 분석하여 분석 도구에 따른 음향학적 검사 변수의 차이를 살펴보고자 하였다. 정상 성인 총 35명 (남성 19명, 여성 16명)을 대상으로 모음 /아/를 수집한 후, 동일한 음성을 MDVP, Praat, TF32 각각의 음향학적 검사 도구로 분석하였다. 그 결과 jitter 변수(J local, J abs, J rap, J ppq), shimmer 변수(S local, S dB, S apq), noise-to-harmonics ratio (NHR) 평균의 경우, 남성과 여성 모두 MDVP보다 Praat의 수치가 통계적으로 유의하게 낮았다($p < .01$). 또한 J local, J abs, S local 평균의 경우, 남성과 여성 모두 MDVP, Praat, TF32 순으로 통계적으로 유의하게 낮아졌다. 결론적으로 각 음향학적 검사 도구에 사용된 알고리즘 차이로 인해 도구 간의 측정치에 차이가 있었다. 그러므로 임상가들이 임상현장에서 각각의 음향학적 검사 도구를 사용할 때 각 도구의 알고리즘에 대해 이해한 후 병적 음성을 분석하는 것이 중요할 것이다.

핵심어: MDVP, Praat, TF32, 정상 음성, jitter, shimmer, noise-to-harmonics ratio (NHR)

참고문헌

- 고도홍 (2003). 개별화자의 음성파라미터 추출에 관한 연구: 음성 파라미터의 상관관계를 중심으로. *말소리와 음성과학*, 10(2), 129-143.
- 고혜주, 강민재, 권혁제, 최예린, 이미금, 최홍식 (2013). 16-18세 청소년기 음성의 음향음성학적 특성. *말소리와 음성과학*, 5(1), 81-90.
- 남기창, 이승훈, 최재남, 최홍식, 남도현, 김덕원 (2005). 각종 음성 분석 상용 프로그램의 모음 기본주기 분석 결과 비교. *정보 및 제어 심포지엄(ICS '05)*, 54-56.
- 박상보 (2005). *음성 기반의 음악 검색을 위한 음성 신호 변환 기법*. 아주대학교 석사학위논문.
- 백승은, 김진영, 나승유, 최승호 (2005). Directional filter와 harmonic filter 기반 화자 분리. *말소리와 음성과학*, 12(3), 125-136.
- 백연숙, 김세주, 김은연, 최예린 (2012). 우울증 화자 음성의 음향음성학적 특성. *말소리와 음성과학*, 4(1), 91-98.
- 신동성, 김준, 배명진 (2000). *AME를 이용한 전이구간에서의 피치 검색에 관한 연구* (연구보고서). 진천: 정보통신산업진흥원.
- 심상용, 김향희, 김재욱, 신지철 (2014). 성대결절 음성 중증도에 따른 MDVP와 Praat 프로그램 별 파라미터 차이. *말소리와 음성과학*, 6(2), 107-114.
- 안찬식, 오상엽 (2012). 연속 음성 인식 향상을 위해 LMS 알고리즘을 이용한 CHMM 모델링. *디지털융복합연구*, 10(11), 377-382.
- 유재연, 정옥란, 장태엽, 고도홍 (2003). MDVP와 Praat, Dr. Speech 간의 음향학적 측정치에 관한 상관연구. *말소리와 음성과학*, 10(3), 29-36.
- 윤선영, 권도하 (1998). 5-11세 아동 음성의 음향학적 특성. *언어치료연구*, 7(1), 67-78.
- 최성희, 남도현, 이승훈, 정원혁, 김덕원, 최홍식 (2005). 각종 음성 분석기에 따른 음성장애 환자의 주기간 주파수 및 진폭변동을 분석. *대한후두음성언어의학회지*, 16(2), 140-145.
- 표화영, 송윤경 (2010). 음성장애 진단 및 평가의 최근 연구 동향: 문헌적 고찰. *언어청각장애연구*, 15(4), 506-525.
- 표화영, 심현섭 (2007). 음성장애 진단 및 평가에 관한 질적 연구: 진단 및 평가 모형 정립을 위한 기초연구. *말소리와 음성과학*, 14(2), 7-22.
- 표화영, 심현섭, 송윤경, 윤영선, 이은경, 임성은, 하현령, 최홍식 (2002). 한국 성인의 정상 음성에 관한 기본 음성 측정치 연구. *말소리와 음성과학*, 9(12), 179-192.
- 표화영, 심현섭, 임성은 (2000). 음성 Sample의 길이 변화에 따른 MDVP 측정치와 GRBAS 척도 간의 상관관계 변화 비교. *말소리와 음성과학*, 7(2), 71-81.