

Utility of a digital motor speech measurement program using an AI speech recognition module: A pilot study*

Sora Han · Do Hyung Kim · So Young Han · Jaewon Kim · Dae-Hyun Jang**

*Department of Rehabilitation Medicine, Incheon St. Mary's Hospital, College of Medicine,
The Catholic University of Korea, Seoul, Korea*

Abstract

This study evaluated the agreement between an AI-based speech recognition module and the acoustic analysis software Praat, focusing on the reliability and interchangeability of the two methods. Maximum phonation time (MPT) and diadochokinetic (DDK) tasks were used to analyze mean differences and limits of agreement. For MPT, the limits of agreement were 0.72 and -0.81, with 93% of the results falling within these limits and a mean difference of -0.04, indicating strong agreement and high accuracy. For the DDK tasks, 91% of the /pʌ/, /tʌ/, and /kʌ/ results fell within the limits, with mean differences of 0.23, 0.17, and 0.23, respectively, demonstrating reliable accuracy. The /pʌtʌkʌ/ task showed 93% of results within the limits, with a mean difference of 0.03, confirming very high agreement and accuracy. These findings suggest that AI-based automatic speech recognition technology has strong potential for clinical acoustic analysis, offering reliability and possibility of complementing or replacing traditional tools. This study also provides a foundation for developing speech recognition-based learning and therapeutic tools for children with and without speech sound disorders.

Keywords: maximum phonation time (MPT), diadochokinetic (DDK), automatic speech recognition, Praat

1. 서론

말 산출 과정에는 호흡, 발성, 조음 및 공명에 관여하는 기관 (조음기관) 간의 협응이 필요하다. 말소리 산출과 관련된 기관의 구조나 기능적인 원인으로 야기되는 의사소통의 어려움에는 조음장애, 음성장애, 유창성장애 등이 포함된다(Kim & Kim., 2016). 이러한 특성을 적절하게 치료하기 위해서는 말장애를 유

발할 수 있는 조음기관의 구조와 기능의 이상 유무를 먼저 살펴보는 것이 중요하다. 이를 통해 의사소통 문제의 치료계획을 구체적이고 체계적으로 수립할 수 있고, 진단 여부를 판단하는 근거로 활용할 수 있기 때문이다.

특히, 아동의 조음기관은 연령에 따라 지속적으로 발달 과정에 있어, 구개의 특성이나 혀의 움직임 범위 등이 성인과 차이를 보인다. 따라서 아동의 조음능력에 영향을 미치는 조음운동

* This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2022-NR071928), and by the Bio Industrial Technology Development Program (20017960) funded By the Ministry of Trade, Industry & Energy (MOTIE, Korea).

** dhjangmd@naver.com, Corresponding author

Received 29 October 2024; Revised 19 November 2024; Accepted 20 November 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.

성을 평가하기 위해서는 발달 과정을 고려한 검사 방법과 적합한 과제를 제시할 필요가 있다(Kim et al., 2018). 또한 호흡, 발성, 조음 및 공명기관의 구조와 기능을 수량화하는 평가가 필요하다(Kim & Kim., 2016).

국내에서 아동의 조음기관의 구조와 기능을 평가하는 대표적인 검사도구로 조음기관 구조·기능 선별검사(Speech Mechanism Screening Test for children, SMST-C; Shin et al., 2018)가 있다. 『SMST』는 조음장애나 언어장애가 없는 만 3-12세의 국내 정상 아동을 대상으로 표준화하여 조음기관의 구조와 기능의 이상 유무를 평가할 수 있는 기준치를 제시하고 있다.

『SMST-C』는 세 개의 하위 영역으로 구성되어 있는데, ‘조음 기관 구조와 기능’에서는 얼굴, 입술, 혀, 턱과 치아, 경구개 및 연구개, 인두 및 호흡기관 등 조음기관의 구조나 기능의 이상 여부를 확인한다. ‘발성·음성·조음 선별’에서 최대발성지속 시간(maximum phonation time, MPT)은 /아/ 모음을 발성할 때의 발성을 지속하는 시간을 측정하여, 조음 산출 시 기능이나 신경계 협응과 폐활량과 폐 내 압력 및 성대의 접촉 양상 그리고 성대를 통과하는 기류의 양과 압력 등을 간접적으로 평가할 수 있다(Kent et al., 1987). 조음기관을 구성하는 근육들의 최대 운동속도 및 규칙성과 정확성 평가를 위해 조음교대운동(diadochokinesis, DDK)을 실시한다. DDK 과제는 무의미 음절 /피/, /터/, /커/를 반복하여 산출하는 교대운동속도(alternating motion rate, AMR)와 3음절 /피터커/를 일련으로 산출하는 일련운동속도(sequential motion rate, SMR)로 구분되며, 일반적으로 주어진 시간 동안 반복한 음절 수를 세는 방식(count-by-time)을 통해 측정한다. DDK 평균 속도와 편차를 이용하여 아동의 조음능력을 간접적으로 알 수 있어 임상적으로 유용하게 활용할 수 있다(Yoo, 2018). Granocchio et al. (2021)는 Robbins & Klee 프로토콜을 확립하여 아동의 구강 운동성과 조음 운동성(DDK)을 평가하고, 발달 지연 여부를 파악하여 아동의 조음 및 구강 운동 발달 상태를 정량적으로 평가하는 데 신뢰성을 확보할 수 있음을 확인하였다.

일반적으로 임상현장에서 환자의 말소리를 평가하는 가장 보편적인 방법은 언어재활사가 귀로 듣고 평가하는 청지각적 평정이다. 청지각적 평정은 특별한 도구를 필요로 하지 않아, 임상현장에서 즉각적이고 편리하게 사용할 수 있는 장점이 있다. 그러나 청지각적 평정의 경우 측정 일관성과 반복 측정에 통일된 기준을 합의하기에 어려운 측면이 있다(Roy et al., 2013). 따라서 임상가는 음향학적 분석을 추가로 시행하여 음성 에 대한 정량화된 측정치를 확인하고, 청지각적 평정의 한계를 보완한다(Natour & Saleem, 2009). 음향학적 분석은 개인의 음성 특성을 파악하고, 이를 통해 진단과 진단도를 객관적으로 확인할 수 있는 중요한 평가 방법이다. 또한 음성과 관련된 다양한 매개변수들의 객관적 수치를 비침습적인 방법으로 빠르고 쉽게 얻을 수 있다는 장점이 있다(Choi et al., 2005). Praat은 음성 분석 및 음성 변형 프로그램으로 홈페이지(<http://www.praat.org>)에서 무상으로 다운 받을 수 있다. 또한 데이터를 효율적으로 처리할 수 있어 많은 임상가와 과학자들이 사용하고 있다(Oğuz

et al., 2011). 그러나 분석을 위해 음성 녹음이 선행되어야 하고, 평가 진행 이후 임상가가 프로그램을 이용하여 분석하기 때문에 즉각적인 결과 확인이 어렵다는 제한점이 있다. Jeong, et al. (2011)는 미취학 아동을 대상으로 DDK 과제에 대한 연령에 따른 반응태도를 분석하였다. 연구 결과 지시사항의 이해가 어려운 5세 미만 어린 연령 아동과 조음음운 능력이 낮은 아동은 말운동 능력 측정이 까다로워, 민감하게 반응할 수 있는 검사 방법이 필요하다는 점을 언급하였다. 이는 아동의 말운동 상태를 측정은 조음발달 수준과 이상 여부 등 중요한 정보를 확인할 수 있는 과제임에도 현실적인 문제로 인해 측정의 어려움이 있음을 시사한다. 따라서 아동의 말운동 능력 측정을 위해서는 검사자와 대상자의 편의성을 고려한 검사 방법의 도입이 필요하다.

최근 인공지능(artificial intelligence, AI) 음성인식 기술이 고도화되어 음성 신호 특징의 자동 인식과 분석 및 결과를 즉시 확인할 수 있는 수준에 도달하였다. 이러한 AI 음성인식 기술을 언어병리학에 도입한 학제간 융합 연구가 활발하게 진행되고 있다(Kang et al., 2022). 국외 연구 동향에 따르면 현재 언어병리학 분야에서의 AI 기술 적용은 임상에서 효과성 및 효용성을 확인하기 시작하였으며(Song et al., 2020), 국내에서도 진단의 정확성과 치료의 효율성을 확보하기 위한 연구가 진행 중에 있다(Yun & Im, 2022).

이에 따라 본 연구에서는 말운동 능력을 평가할 수 있는 과제(MPT, DDK)에서 AI 음성 인식 모듈을 이용한 자동 측정 결과와 전통적 분석 방법인 Praat의 결과를 비교하고자 한다. 전통적 측정 방법과의 일치도가 높을수록 결과의 신뢰성을 검증할 수 있으며, 이는 향후 임상에서의 활용 가능성을 탐색하는 데 중요한 기반이 될 것이다. 본 연구의 구체적인 목적은 다음과 같다. 1) 두 가지 음향 기기적 측정 방법(AI 음성인식 모듈, Praat) 간의 일치도를 평가하고, 2) AI 음성인식 모듈의 신뢰성과 대체 가능성을 검증하며, 3) 말소리장애 아동의 말운동 능력을 평가하는 데 기초자료로 제공하고자 한다.

2. 연구방법

2.1. 연구대상

본 연구는 가톨릭대학교 인천성모병원 임상연구심사위원회의 승인(IRB: OC24ONSI0086)을 받았으며, 가톨릭대학교 인천성모병원에 내원한 아동 중 본인 또는 보호자가 참여 의사를 밝힌 만 3-9세의 정상 발달 아동 30명을 대상으로 하였다. 1차적으로 보호자 설문문을 통해 아동의 발달사 및 구개파열, 뇌성마비 등의 조음기관 구조와 기능의 이상을 초래할 가능성 여부를 확인하였다. 이차적으로 1) 수용·표현어휘력검사(Receptive & Expressive Vocabulary Test, REVT; Kim et al., 2009)의 표현 어휘력 과제를 실시하여 -1SD 이하, 2) 자발화상에서 언어나 조음의 이상이 있다고 의심되는 경우, 3) 검사에 비협조적이거나 검사를 완료하지 못한 경우 대상에서 제외하였다(표 1).

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

집단	인원(명)	연령
미취학	11	4.27±1.10
취학	19	7.05±0.84
합계	30	6.03±1.65

Values are presented as mean (±SD).

2.2. 연구 방법

2.2.1. 음성 수집

본 연구는 독립되고 조용한 공간에서 실시되었으며, 각 연구자와 대상자는 1대 1로 검사를 시행하였다. 1급 언어재활사 1인과 2급 언어재활사 2인이 프로토콜에 따라 대상자의 MPT와 DDK를 측정하였다. 검사에는 평가자용 노트북, 아동 검사용 태블릿 PC(Galaxy Tab S7 FE, Samsung, Suwon, Korea), 구강경, 펜라이트, 설압자 및 초시계가 사용되었으며, 아동의 발화는 스마트폰(iphone pro max 12, Apple, Cupertino, CA, USA)을 사용하여 녹음하였다. Lee et al.(2019)의 연구에 따르면 스마트폰은 음성 선별검사를 위한 녹음 장비로써 유용성이 인정되어 본 검사에서도 활용하였다. 검사 시 아동과 태블릿 PC 그리고 스마트폰의 간격은 30 cm를 유지하였고, 스마트폰의 마이크가 아동을 향하도록 세팅하였다. 검사실에서 태블릿 PC 애플리케이션을 활용하여 아동의 MPT와 DDK를 자동으로 측정하였고, 검사가 끝난 후 녹음된 음성파일은 Praat을 이용해 분석을 실시하여 결과를 비교하였다.

2.2.2 사용 도구

자동측정에 사용된 애플리케이션(application, app)은 발화 시간을 측정하는 MPT 알고리즘을 적용한 ASR(automatic speech recognition) 엔진을 통합하고 있다. MPT 알고리즘은 발화 시작 지점부터 종료 지점까지의 시간을 측정하는 방식으로, 음성 감지 시 무음 구간을 오인하는 문제를 해결하는 데 중점을 두었다. 마이크로 입력되는 오디오의 음량이 일정하지 않을 경우, 일부 음성을 휴지 구간으로 인식하여 실제 발화 시간보다 짧게 측정되는 오류가 발생하였다. 이를 해결하기 위해 0.2초 이하의 짧은 휴지 구간은 무시하도록 설정하였으며, 발화가 끊기지 않고 정확하게 측정될 수 있도록 알고리즘을 개선하였다. 제시한 그림 1의 푸른 음영은 발성으로 인식한 부분으로 알고리즘 개선 전후의 결과를 비교하여 확인할 수 있다.

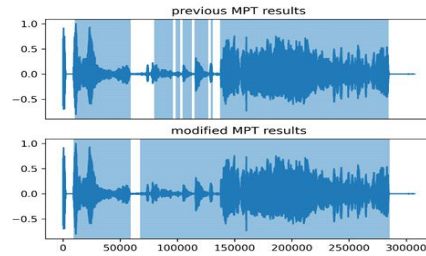


그림 1. MPT(Maximum phonation time) 알고리즘 개발 과정
Figure 1. MPT (Maximum phonation time) algorithm development process

DDK는 /퍼/, /터/, /키/ 또는 /퍼터커/와 같은 음절 반복 발화를 자동으로 측정하는 알고리즘을 적용하였다. 음성 데이터를 텍스트로 변환하는 복잡한 과정을 단순화해 주는 딥러닝 모델인 Wav2Vec 기반의 E2E STT(end-to-end speech-to-text) 엔진을 활용하여 음성 인식을 하였다. 파일럿 테스트에서 /kΛ:/와 같이 모음을 과도하게 강조하는 경우 두 번의 peak가 발생하여, 횡수 측정에 오류가 발생하였다. 이에 알고리즘의 peak 값을 0~1,000 Hz에서 2,000~3,000 Hz로 조정하여 민감도를 향상하였다. 제시된 그림 2에서 하단의 초록 마커는 기존 결과를 의미하며, 중간단의 파란 마커는 최종 개선된 사양을 의미한다.

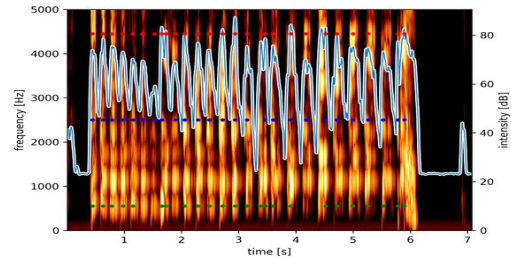


그림 2. DDK(Diachokinesis) 알고리즘 개발 과정
Figure 2. DDK (Diachokinesis) algorithm development process

2.2.2.1. MPT(Maximum phonation time) 측정

검사자는 아동에게 검사용 태블릿 화면을 보여주며 “숨을 크게 마시고, 화면에 마이크가 보이기 시작하면 /아/ 소리를 할 수 있는 만큼 오래 내세요”라고 지시하였다. 아동 연령에 따른 지시이해도를 고려하여 지시사항을 조절하여 제시하거나, 필요한 경우 검사자가 시범을 보였다. 아동이 1~2회 반복하여 연습한 후 검사가 가능할 것으로 판단되면 측정을 시행하였다. 검사자가 아동에게 지시한 후 평가자용 웹화면에서 시작버튼을 누르면, 태블릿 PC를 통해 아동의 발화가 시작되는 시점부터 MPT가 자동 측정되었다. 아동의 발화가 종료된 후 검사자가 웹화면의 종료버튼을 누르는 동시에 결과가 분석이 되며, 측정 결과 화면 예시는 그림 3에 제시하였다. 총 2회씩 MPT 측정을 시행하였으며, 가장 긴 결과치를 분석대상으로 채택하였다.



그림 3. MPT(Maximum phonation time) 결과 화면
Figure 3. MPT (Maximum phonation time) result interface

2.2.2.2. DDK(Diadochokinesis) 측정

DDK 검사는 MPT와 동일한 절차로 진행되며, /퍼/, /터/, /커/, /퍼터커/ 순서로 시행하였다. 검사자가 “화면에 마이크가 보이면 선생님이 그만~할 때까지 /퍼/를 최대한 빨리 말해보세요.”라고 지시한 후 아동의 발화가 시작되면 5초간의 음성이 자동 측정되어 분석되었다. 화면에 분석 결과가 제시되면 검사자는 청지각적 평정에 따라 규칙성과 정확성을 체크하였다. 과제별로 총 2회씩 시행하였으며, 가장 많이 시행한 결과치를 채택하였다(그림 4).



그림 4. Diadochokinesis(DDK) 결과 화면
Figure 4. Diadochokinesis (DDK) result interface

2.2.3. Praat 분석

대상자의 정확한 수행을 확인하기 위한 gold standard로 설정하기 위해 Praat의 수동 분석을 활용하였다. 평가 중 아이폰으로 수집된 M4P 녹음 파일은 과제별로 분리해 wav. 음성파일로 저장한 후 Praat을 이용해 분석하였다. MPT 길이는 연속된 음형과 형이 시작되는 지점부터 끝나는 지점 사이의 시간을 측정하였다. DDK 횟수는 음성파형의 시작 부분부터 5초간 산출된 음절(에너지 파형의 고점 발생 지점)의 수를 세어 측정하였다.

2.3. 자료분석

본 연구의 수집된 자료는 Statistical Product and Service Solution (SPSS, version 21.0) 통계 프로그램을 이용하여 분석하였다. 대상자의 인구학적 정보는 기술통계를 활용하여 평균 및 표준편차를 산출하였다. 두 측정치(app, Praat)의 일치도는 Bland-Altman plot을 활용하였다(Bland & Altman, 1986). Bland-Altman

plot은 두 측정 방법 간의 차이를 평균값과 비교하여 차이의 분포를 시각화하는 분석 기법으로 두 방법 간의 체계적 차이(평균 차이)를 확인하고, 95%의 데이터가 포함되는 일치한계(limits of agreement, LoA)를 설정하여 각 측정 방법이 일관된 값을 제공하는지 확인할 수 있다(Euser et al., 2008). 이를 통해 측정도구의 비교 평가에서 상관관계 분석의 취약점을 보완할 수 있으며(Giavarina, 2015), 두 측정 방법이 임상적으로 또는 연구적으로 대체 가능한지를 평가할 수 있다(Haghighyegh et al., 2020).

3. 연구결과

3.1. 측정 결과(표 2)

표 2. 측정 결과 기술통계
Table 2. Descriptive statistics of measurement tools

	app	Praat	평균 차이	상위 LoA	하위 LoA
MPT	8.61±3.31	8.62±3.43	-0.04	0.72	-0.81
/퍼/	21.48±4.79	21.37±5.07	0.23	1.21	-0.75
/터/	22.17±4.92	22.13±5.13	0.17	0.94	-0.59
/커/	21.2±4.46	21.03±4.61	0.23	0.23	-0.75
/퍼터커/	7.68±1.64	7.75±1.46	0.03	0.38	-0.32

Values are presented as mean (±SD).

MPT, maximum phonation time; app, application; LoA, limits of agreement.

3.2. MPT(Maximum Phonation Time) 결과 비교

MPT 측정 결과에서 상위 일치 한계는 0.72, 하위 일치 한계는 -0.81로 계산되었고, 93%에 달하는 결과가 일치 한계 내에 분포하여 높은 일치도가 확인되었다. 측정 평균 차이는 -0.04로 app이 Praat 보다 낮게 측정된 값이 있었다는 것을 의미한다(그림 5).

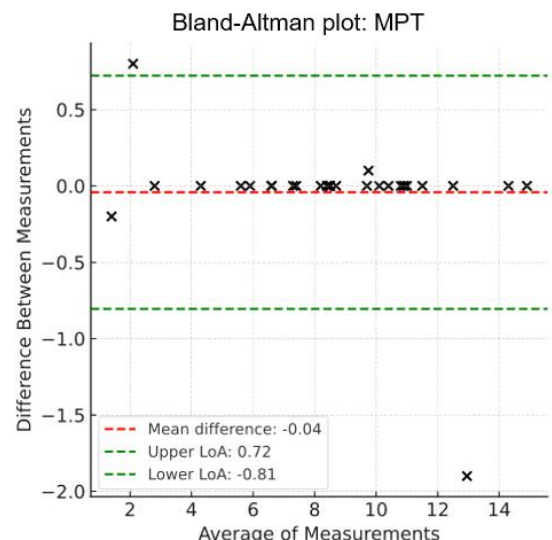


그림 5. App(Application)과 Praat의 maximum phonation time(MPT) 측정의 일치도

Figure 5. Agreement between maximum phonation time (MPT) measurements on application (app) and Praat

3.3. DDK(Diadochokinesis) 결과 비교

DDK 과제 중 /ㅍ/ 측정에서는 상위 일치 한계는 1.21, 하위 일치 한계는 -0.75로 계산되었으며, 91%의 결과가 일치 한계 내에 분포하였다. 또한 측정 평균 차이가 0.23으로 크지 않아, 두 측정 결과 간에 높은 일치도를 보였다(그림 6).

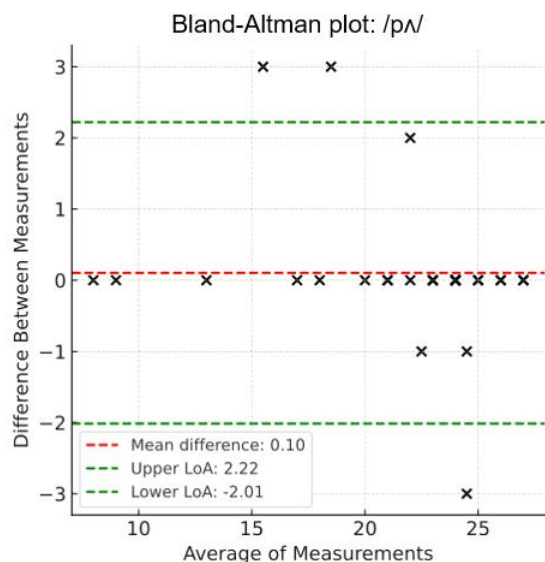


그림 6. App(Application)과 Praat 간 /ㅍ/ 측정의 일치도
Figure 6. Agreement between /p/ measurements on app (application) and Praat

/ㅌ/ 측정 결과에서 상위 일치 한계는 0.97, 하위 일치 한계는 -0.59로 계산되었으며, 91%의 결과가 일치 한계 내에 분포하였다. 또한 측정 평균 차이가 0.17으로 작아, 두 측정 결과 간에 높은 일치도가 확인되었다(그림 7).

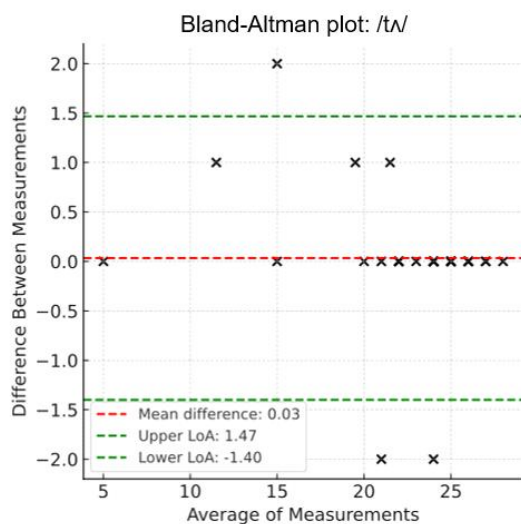


그림 7. App(Application)과 Praat 간 /ㅌ/ 측정의 일치도
Figure 7. Agreement between /t/ measurements on app (application) and Praat

/ㄱ/ 측정 결과에서 상위 일치 한계는 0.23 하위 일치 한계는 -0.75로 계산되었으며, 91%의 결과가 일치 한계 내에 분포하였다. 또한 측정 평균 차이가 0.23으로 크지 않아, 두 측정 결과 간에 높은 일치도를 보였다(그림 8).

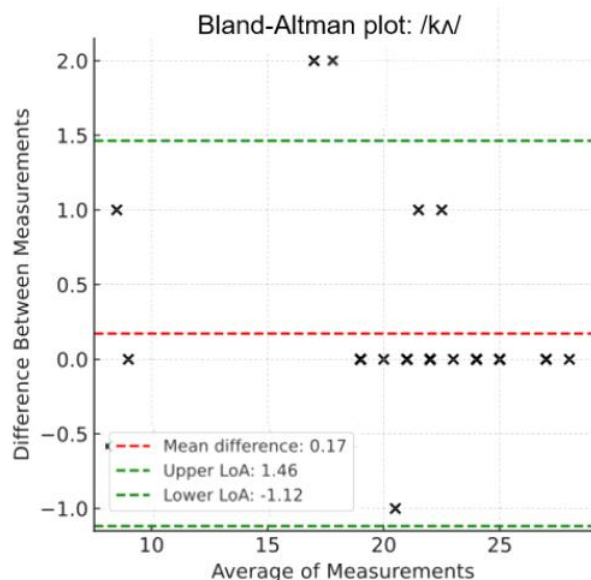


그림 8. App(Application)과 Praat 간 /ㄱ/ 측정의 일치도
Figure 8. Agreement between /k/ measurements on app (application) and Praat

/ㅍ터ㄱ/ 측정 결과에서 상위 일치 한계는 0.38 하위 일치 한계는 -0.32로 계산되었으며, 93%의 결과가 일치 한계 내에 분포하였다. 또한 측정 평균 차이가 0.03으로 작아, 두 측정 결과가 일치함을 시사하였다.

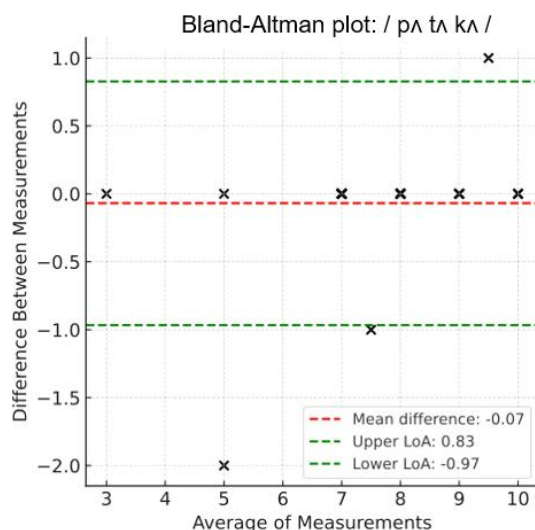


그림 9. App(Application)과 Praat 간 /ㅍ터ㄱ/ 측정의 일치도
Figure 9. Agreement between /pʌtʌkʌ/ measurements on app (application) and Praat

4. 논의 및 결론

본 연구는 AI 음성인식 모듈을 사용한 app과 Praat 간의 측정 결과의 일치도를 평가하고, app을 통한 측정 방법의 신뢰성과 대체 가능성을 검증하는 것을 목적으로 하였다. 이를 위해 MPT 및 DDK 과제를 통해 두 측정 결과 간의 평균 차이와 상위 및 하위 일치 한계를 분석하였다. 분석 결과, MPT, /파/, /터/, /커/, /퍼터커/의 app과 Praat 측정치 분포 비율이 91% 이상 일치 한계 내에 있었으며, 평균 차이는 최대 0.23으로 미미하여 두 결과 간 높은 일치도를 보였다. 이는 MPT와 DDK 과제에서 gold standard(Praat)와 본 연구를 통해 개발된 app의 일치도가 임상적 기준에 부합하여 평가 도구로 활용할 수 있음을 시사한다 (Borrie et al., 2019).

본 연구에서 사용한 Bland-Altman 분석은 두 측정 방법의 차이에 대한 평균과 표준편차를 사용하여 일치도를 평가하는 데 적합하다. 그러나 측정 방법 간 체계적 오차나 비선형적 상관관계가 존재할 가능성이 있다(Ribeiro et al., 2022). 이번 연구에서 한계선을 초과한 사례들은 한계선에 근접하여 일치도 분석에 영향을 미치지 않은 것으로 확인되었다(Ribeiro et al., 2022; Suh & Seo, 2022). 하지만 개별 아동의 음성 특성 차이, 발음 습관, 발성 방식의 차이 등을 고려하여 한계선 밖의 값들이 어떤 조건에서 발생하는지(Son et al., 2024)와 소음이 통제되지 않은 다양한 검사 환경에서의 측정한 결과를 추가로 분석할 필요가 있다. 이를 통해 강건한 알고리즘 성능을 구현하여 app 고도화에 적용할 수 있을 것이다.

따라서 향후 연구에서는 정상 음성을 대표하는 연구 대상자의 수를 추가적으로 확보하여 집단 내 결과에 영향을 미치는 요인의 유무를 확인할 것을 제안하는 바이다. 또한 병리적 음성을 분석할 때는 음성 분석 도구가 사용하는 알고리즘에 따라 민감도와 특이도가 영향을 받을 수 있다(shim et al., 2014). 이에 따라 향후 진행될 연구에서 말소리 장애 아동 음성 및 다양한 환자군에 따른 음성 측정 결과치를 확보해야 할 것이다.

본 연구 결과, AI 음성 인식 모듈을 활용한 앱과 전통적 측정 도구인 Praat을 사용한 음성 측정 결과 간에 높은 일치도가 확인되어, 두 측정 방식이 상호 대체 가능성을 입증하였다. 또한 MPT와 DDK 측정 결과를 실시간으로 확인할 수 있다는 점에서 연구와 임상 현장에서 편의성과 유용성을 향상시켰다는 의의를 가진다.

References

- Bland, J. M., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, 327(8476), 307-310.
- Borrie, S. A., Barrett, T. S., & Yoho, S. E. (2019). Autoscore: An open-source automated tool for scoring listener perception of speech. *The Journal of the Acoustical Society of America*, 145(1), 392-399.
- 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.
- Euser, A. M., Dekker, F. W., & le Cessie, S. (2008). A practical approach to Bland-Altman plots and variation coefficients for log transformed variables. *Journal of Clinical Epidemiology*, 61(10), 978-982.
- Giavarina, D. (2015). Understanding bland altman analysis. *Biochemia Medica*, 25(2), 141-151.
- Granocchio, E., Gazzola, S., Scopelliti, M. R., Criscuoli, L., Airaghi, G., Sarti, D., & Magazù, S. (2021). Evaluation of oro-phonatory development and articulatory diadochokinesis in a sample of Italian children using the protocol of Robbins & Klee. *Journal of Communication Disorders*, 91, 106101.
- Haghighyegh, S., Kang, H. A., Khoshnevis, S., Smolensky, M. H., & Diller, K. R. (2020). A comprehensive guideline for Bland-Altman and intra class correlation calculations to properly compare two methods of measurement and interpret findings. *Physiological Measurement*, 41(5), 055012.
- Jeong, H. J., Lee, O. B., & Sehr, K. H. (2011). Diadochokinetic skills in typically developing children aged 4–6 years: Pilot study. *Journal of the Korea Academia-Industrial Cooperation Society*, 12(7), 3149-3155.
- Kang, H. W., Kang, J. K., Lee, S. B., & Sim, H. S. (2022). Applications and performances of artificial intelligence in assessment and diagnosis of communication disorders: A systematic review of the literatures. *Communication Sciences and Disorders*, 27(3), 703-722.
- Kent, R. D., Kent, J. F., & Rosenbek, J. C. (1987). Maximum performance tests of speech production. *Journal of Speech and Hearing Disorders*, 52(4), 367-387.
- Kim, Y. S., & Kim J. (2016). A preliminary study to develop a speech mechanism screening test for preschool children. *Journal of Speech-Language & Hearing Disorders*, 25(3), 105-123.
- Kim, Y. T., Hong, G. H., Kim, K. H., Jang, H. S., & Lee, J. Y. (2009). Receptive & expressive vocabulary test (REVT). Seoul: Seoul Community Rehabilitation Center
- Kim, J., Shin, M. & Song, Y. K. (2018). Speech mechanism screening test for children: An evaluation of performance in 3- to 12-year-old normal developing children. *Communication Sciences and Disorders*, 23(1), 180-197.
- Lee, H. N., Park, J. H., & Yoo, J. Y. (2019). Development of smartphone-based voice therapy program. *Phonetics and Speech Sciences*, 11(1), 51-61.
- 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.

Ribeiro, E. O. S., Gosselink, R., de Moura, L. E. S., Correia, R. F., Leite, W. S., de Araújo, M. d. G. R., de Andrade, A. D., ... Campos, S. L. (2022). Agreement between two methods for assessment of maximal inspiratory pressure in patients weaning from mechanical ventilation. *Acute and Critical Care*, 37(4), 592-600.

Roy, N., Barkmeier-Kraemer, J., Eadie, T., Sivasankar, M. P., Mehta, D., Paul, D., & Hillman, R. (2013). Evidence-based clinical voice assessment: A systematic review. *American Journal of Speech-Language Pathology*, 22(2), 212-226.

Shim, SangYong, Kim, HyangHee, Kim, JaeOck, & Shin, JiCheol (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, M., Kim, J., Lee, S., & Lee, S. (2009). Speech Mechanism Screening Test (SMST) Seoul: Hakjisa.

Son, G., So, J., Ko, J., Lee, J. W., Lee, J. R., & Shin, W. S. (2024). Enhanced AI model to improve child speech recognition. *Journal of Digital Contents Society*, 25(2), 547-555.

Suh, M. H., & Seo, K. (2022). A comparative study on measurement of physical activity between smartphone app and self-reported questionnaire. *Journal of Muscle and Joint Health*, 29(2), 91-99.

Yoo. (2018). The Characteristics of Diadochokinesis in Older Preschooler. *Journal of speech-language & hearing disorders*, 27(3), 13-21.

Yun, E., & Im, I. (2022). Analysis of domestic research trends related to the development of digital therapeutics (DTx) in the field of communication disorders. *Phonetics and Speech Sciences*, 14(4), 57-66.

• 한소라 (Sora Han)

가톨릭대학교 인천성모병원 재활의학과 연구원
인천광역시 부평구 동수로 56
Tel: 032-280-5207
Email: wing21sr@gmail.com
관심분야: 신경언어장애, 언어발달장애

• 김도형 (Do Hyung Kim)

가톨릭대학교 인천성모병원 재활의학과 연구원
인천광역시 부평구 동수로 56
Tel: 032-280-5207
Email: ehgud1231@gmail.com
관심분야: 말소리장애, 언어발달장애

• 한소영 (So Young Han)

가톨릭대학교 인천성모병원 재활의학과 연구원
인천광역시 부평구 동수로 56
Tel: 032-280-5207
Email: hanqr379@gmail.com
관심분야: 말소리장애, 신경언어장애

• 김재원 (Jaewon Kim)

가톨릭대학교 인천성모병원 재활의학과 임상진료조교수
인천광역시 부평구 동수로 56
Tel: 032-280-6601
Email: jw2356@naver.com
관심분야: 음성인식, 소아재활

• 장대현 (Dae-Hyun Jang) 교신저자

가톨릭대학교 인천성모병원 재활의학과 교수
인천광역시 부평구 동수로 56
Tel: 032-280-5207
Email: dhjangmd@naver.com
관심분야: 음성인식, 소아재활

AI 음성인식 모듈을 통한 말운동 능력 측정 프로그램의 유용성: 예비연구*

한 소 라 · 김 도 형 · 한 소 영 · 김 재 원 · 장 대 현

가톨릭대학교 인천성모병원 재활의학과

국문초록

본 연구는 AI 음성인식 모듈을 적용한 앱과 음향 분석 소프트웨어 Praat 간의 측정 결과 일치도를 평가하고, 두 측정 방법의 신뢰성과 상호 대체 가능성을 검토하는 데 목적이 있다. 이를 위해 최대발성시간(MPT) 및 조음교대운동(DDK) 과제를 활용하여 두 측정 결과 간의 평균 차이와 일치 한계를 분석하였다. MPT 측정 결과, 상위 일치 한계는 0.72, 하위 일치 한계는 -0.81로 나타났으며, 93%의 결과가 일치 한계 내에 분포하였고, 측정 평균 차이는 -0.04로 두 측정 방법 간 높은 일치도가 확인되었다. DDK 과제에서는 /퍼/와 /터/에서 각각 91%의 결과가 일치 한계 내에 분포하였고, 측정 평균 차이는 각각 0.23과 0.17로 나타나, 임상적으로 두 측정 결과 간 유의미한 차이가 없었다. /키/ 과제의 경우에도 상위 일치 한계는 0.23, 하위 일치 한계는 -0.75로 나타났으며, 91%의 결과가 일치 한계 내에 분포하였다. /퍼터커/ 과제는 상위 일치 한계가 0.38, 하위 일치 한계가 -0.32로 나타났으며, 93%의 결과가 일치 한계 내에 분포하였고, 평균 차이는 0.03으로 두 측정 방법 간 매우 높은 일치율을 보였다. 이러한 결과는 AI 음성인식 기술이 음성 분석 도구로서 임상적 활용 가능성이 높음을 시사한다. 특히, 본 연구는 AI 기반 음성 평가가 기존 음성 분석 도구를 보완하거나 대체할 수 있는 신뢰성을 확인하였으며, 향후 정상 아동 및 말소리 장애 아동을 위한 음성 인식 기반 학습 및 치료 도구 개발의 기초 자료로 활용될 수 있을 것으로 기대된다.

핵심어: 최대발성지속시간, 조음교대운동, AI 음성 인식, Praat

참고문헌

- 강혜원, 강진경, 이수복, 심현섭. (2022). 의사소통장애의 평가 및 진단에서 인공지능 적용과 성과에 관한 체계적 문헌고찰. *한국언어척각임상학회*, 27(3), 703-722.
- 고혜주, 우미령, 최예린. (2020). MDVP, Praat, TF32에 따른 음향학적 측정치에 대한 비교. *말소리와 음성과학*, 12(3), 73-83.
- 김양선, 김재욱. (2016). 아동용 조음기관 구조 및 기능 선별검사 제작을 위한 학령전기 아동 대상 예비조사. *언어치료연구*, 25(3), 105-123.
- 김영태, 홍경훈, 김경희, 장혜성, 이주연(2009). 수용·표현어휘력 검사(REVT). 서울: 서울장애인종합복지관
- 김재욱, 신문자, 송윤경. (2018). 아동용 조음기관 구조 및 기능 선별검사: 3-12세 정상 발달 아동의 수행 특성 평가. *한국언어척각임상학회*, 23(1), 180-197.
- 박정인, 이승진. (2024). 정상 성인에서 스마트폰 녹음을 위한 마이크 유형 간 음향학적 측정치 비교. *말소리와 음성과학*, 16(2), 49-58.

- 박희준, 유재연. (2013). 공유소프트웨어의 언어치료 적용에 관한 고찰. *언어치료연구*, 22(3), 1-24.
- 서민희, 서경산. (2022). 스마트폰 앱과 자가보고식 설문지의 신체활동 측정 비교. *근관결건강학회지*, 29(2), 91-99.
- 손계원, 소준섭, 고주은, 이진우, 이정록, 신원선. (2024). 아동 음성 인식 향상을 위한 인공지능 모델 개선. *디지털콘텐츠학회논문지*, 25(2), 547-555.
- 신문자, 김재욱, 이수복, 이소연(2009). 조음기관 구조·기능 선별검사(SMST). 서울: 학지사.
- 심상용, 김향희, 김재욱, 신지철. (2014). 성대결절 음성 중증도에 따른 MDVP와 Praat 프로그램 별 파라미터 차이. *말소리와 음성과학*, 6(2), 107-114.
- 유재연. (2018). 취학 전 아동의 조음교대운동 특성. *언어치료연구*, 27(3), 13-21.
- 윤은미, 임익재. (2022). 의사소통장애 분야에서 디지털 치료제(DTx) 개발과 관련된 국내 연구동향 분석. *말소리와 음성과학*, 14(4), 57-66.
- 이하나, 박준희, 유재연. (2019). 스마트폰기반 음성치료 프로

* 이 연구는 정부(과학기술정보통신부)에서 지원하는 한국연구재단(NRF) (RS-2022-NR071928)과 산업통상자원부에서 지원하는 바이오산업기술개발사업(20017960)의 지원을 받아 수행되었습니다.

- 그램 개발연구. *말소리와 음성과학*, 11(1), 51-61.
- 정한진, 이옥분, 서경희. (2011). 학령전기 정상발달 아동의 자모음 교대운동특성: 예비연구. *한국산학기술학회 논문지*, 12(7), 3149-3155.
- 최성희, 남도현, 이승훈, 정원혁, 김덕원, 최홍식. (2005). 각종 음성분석기에 따른 음성장애 환자의 주기간 주파수 및 진폭변동을 분석. *대한후두음성언어의학회지*, 16(2), 140-145.