JSLHD 2025. Vol. 34, No. 4. 093-102
ISSN 1226-587X / eISSN 2671-7158
Journal of Speech-Language & Hearing Disorders 언어치료연구

ORIGINAL ARTICLE
https://doi.org/10.15724/jslhd.2025.34.4.093
http://jslhd.org/

Standardization of a Dysarthria Assessment System in Healthy Adults

Ji-Yeong Kim[1], Ki-Su Park[2,3], Ji-Wan Ha[4*]

[1] *Dept. of Speech-Language Pathology, Daegu University, Research Professor*
[2] *Neopons Inc., Daegu, Korea, CEO*
[3] *Dept. of Neurosurgery, School of Medicine, Kyungpook National University, Professor*
[4] *Dept. of Speech-Language Pathology, Daegu University, Professor*

Purpose: This study aimed to validate the reliability and validity of the Dysarthria Assessment System (DAS), a web-based auditory-perceptual tool designed for the evaluation of dysarthria, and to establish normative data in healthy older adults.
Methods: The Dysarthria Assessment System comprises 80 items across three domains—speech mechanism, speech, and speech motor function—scored on a 3-point scale (2=normal, 1=mild, 0=severe). Data were collected from 328 adults aged 50 years and older. Inter-rater reliability, internal consistency, construct validity, criterion validity, content validity, and system usability were assessed. One-way ANOVA was performed to establish age-based norms.
Results: Inter-rater agreement was 99.04%, and internal consistency was acceptable (Cronbach's α=.808). The Dysarthria Assessment System total score showed significant correlations with subdomain scores and corresponding measures from established tools, confirming construct and criterion validity. Content validity was high (S-CVI=.97), and system usability scores (SUS) indicated excellent user acceptability. Norms for speech mechanism and speech were stratified into two age groups (50~79 and ≥80 years), while speech motor norms were established across four age groups (50s, 60s, 70s, ≥80s) based on percentile ranks.
Conclusions: The Dysarthria Assessment System is a standardized, reliable, and valid web-based instrument for assessing dysarthria. Its automated scoring and reporting system offers clinical and research utility for the diagnosis, severity classification, and monitoring of speech motor impairments in adult populations.

Correspondence: Ji-Wan Ha, PhD
E-mail: jw-ha@daegu.ac.kr

Received: August 23, 2025
Revision revised: September 23, 2025
Accepted: October 31, 2025

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2022S1A3A2A03089435).

ORCID
Ji-Yeong Kim
https://orcid.org/0000-0002-6971-6731
Ki-Su Park
https://orcid.org/0000-0002-4829-6299
Ji-Wan Ha
https://orcid.org/0000-0002-1191-791X

Keywords: Dysarthria, assessment system, standardization, reliability, validity

Ⅰ. 서론

마비말장애(구음장애, dysarthria)는 신경계 손상으로 발생하는 신경인성(neurogenic) 말장애로, 신경의사소통장애 환자 중 약 48%의 출현율을 나타낸다고 보고되고 있다(Duffy, 2019). 마비말장애 환자는 말운동 근육에 마비, 약화, 불협응 등이 초래되어 발화 시 호흡, 발성, 공명, 조음 및 운율의 한 영역 또는 전반에 결함을 보인다(Freed, 2020).

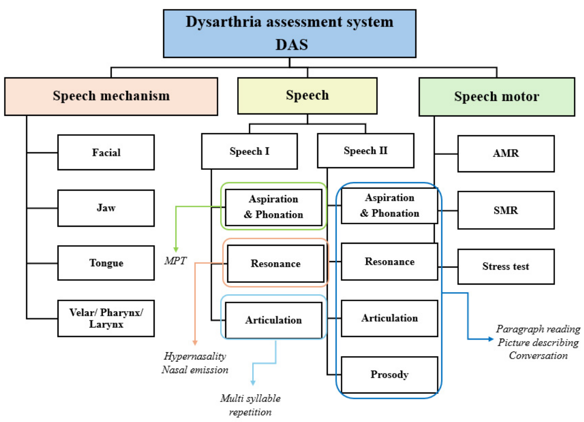
대부분의 연구자들이 마비말장애 감별진단을 위한 핵심 평가과제로 첫째, 모음 연장 과제, 둘째, 교대운동속도(alernating motion rate: AMR)와 일련운동속도(sequential motion rate: SMR), 셋째, 자발화나 읽기 등의 연결발화 과제를 보고하였다(Enerby & Palmer, 2008; Freed, 2020). 세 과제를 통해 얻은 발화자료에서 환자들의 말 특성을 언어치료사가 청지각적으로 평가하는 것이 가장 보편적인 마비말장애 진단 절차이며, 대부분의 국외 마비말장애 검사도구들(Frenchay Dysarthria Assessment 2, Enerby & Palmer, 2008; Intelligibiliy Rating Scale, Duffy, 2005; Assessment of Intelligibility of Dysarthric Speech, Yorkston & Beukelman; 1981)도 이 같은 청지각적 평가를 기반으로 하고 있다. 이 가운데 Frenchay Dysarthria Assessment 2(FDA-2)는 전 세계적으로 가장 널리 쓰이는 표준화 검사도구로 말산출에 관여하는 다양한 말운동 메커니즘과 기능 영역을 청지각적으로 평가하도록 설계되어 있다. 해당 도구는 구조적 기능과 실제 발화를 모두 포괄하고 있으며 여러 신경학적 질환 환자에게 모두 사용 가능하다는 장점을 지니고 있다. 하지만 종이 기반의 수기 채점 방식에 의존한다는 점에서 데이터 관리의 효율성에는 한계가 있다.

국내의 경우 마비말장애 환자를 위한 표준화 검사도구는 현재 부재한 실정이다. 임상 현장에서는 자체 프로토콜을 이용한 비표준화 간이검사를 실시하거나 조음기관 구조·기능 선별검사(Speech Mechanism Screening Test: SMST, Shin et al., 2010), 우리말 조음·음운검사 2(Urimal Test of Articulation &

Phonology 2: UTAP2, Kim et al., 2020)와 같은 관련 표준화 검사 결과를 참고하여 마비말장애 유무를 판단하고 있다(Kim et al., 2024). 그러나 이 같은 간이검사나 관련 표준화 검사만으로는 검사의 객관성과 정확성 측면에 제한이 있다. 또한, 검사자마다 평가 방식이 상이하여 결과의 신뢰성을 확보하는 데 어려움이 있으며, 무엇보다 검사의 타당성 측면에 문제가 있다. 마비말장애 평가에 타당성과 신뢰성을 갖춘 표준화 검사를 사용할 경우, 말산출 체계와 기능에 대한 객관적 평가가 가능하기 때문에 마비말장애 하위 유형 분류와 진단에 실질적인 도움이 될 수 있다(Hong et al., 2018).

더 나아가, 마비말장애 평가를 위한 웹 기반 표준화 검사도구는 국내뿐 아니라 국외에서도 찾아보기 어려운데, 검사도구가 웹 기반으로 제작된다면 기존 검사방식의 단점들을 보완할 수 있다. 기존 종이 기반 검사는 채점 및 결과 산출 과정에서 검사자 간 변이가 발생하기 쉽고, 검사시간과 보고서 작성에 많은 시간이 소요되며, 평가 결과의 장기적 보관 및 추적 비교가 어렵다. 반면, 웹 기반의 디지털 검사는 검사 절차의 표준화와 데이터의 자동 저장·분석을 가능하게 하여 평가의 신뢰성을 높이고, 시간과 공간의 제약을 최소화할 수 있다. 또한 온라인 평가방식은 대규모 데이터 축적을 통한 규준의 지속적 업데이트에도 유리하다는 장점이 있다. 이러한 측면에서, 국내 환경에 적합한 웹 기반의 디지털 표준화 마비말장애 평가도구의 개발과 보급은 임상 및 연구 현장에서 매우 필요한 과업이라 할 수 있다.

이에, 본 연구자들은 웹 기반의 디지털 마비말장애 평가시스템인 Dysarthria Assessment System(DAS)를 개발하였다(Kim et al., 2024). DAS는 대상자들의 말산출 증상과 특성에 기반하여 마비말장애 여부 및 정도, 그리고 하위유형을 진단할 수 있도록 개발된 검사도구로, 말기제(speech mechanism), 말(speech), 말운동(speech motor)의 세 가지 하위평가로 구성되어 있다(Figure 1). DAS의 전체 구성도는 Figure 1과 같으며, DAS의 개발 과정은 Kim 등(2024)의 연구에 자세히 기술되어 있다.



Note. AMR=alternating motion rate; SMR=sequential motion rate; MPT=maximum phonation time.

Figure 1. Components for each subtask in the DAS

아울러 DAS의 임상 활용성을 확보하기 위해서는 연령군별 규준 설정이 선행되어야 한다. 이는 환자의 점수를 상대적으로 해석할 수 있는 준거를 제공함으로써 중재 계획 수립에 타당성을 높여줄 수 있다. 더불어 검사도구에 대한 신뢰도와 타당도를 확보하는 작업 또한 표준화 과정에서 핵심적 과제이다. 신뢰도는 평가 결과의 일관성을, 타당도는 측정 도구가 의도한 개념을 얼마나 정확하게 반영하는지를 의미한다(Cohen & Swerdlik, 2018). 이같은 필요성에 따라 연구 1에서는 DAS의 규준 정립을, 연구 2에서는 DAS의 신뢰도 및 타당도 검증을 목적으로 하였다. 요약하면, 본 연구는 일반 성인을 대상으로 연령군별 규준을 설정하고 신뢰도와 타당도를 검증함으로써, 마비말장애 환자를 위한 표준화 검사로서 DAS의 타당성을 입증하고자 한다.

II. 연구 방법

1. 연구 1: DAS의 규준 정립

1) 연구 대상

DAS의 규준 작업을 위해 복지관, 노인정, 평생교육원 등을 통해 50세 이상 일반 성인 남녀를 모집하였다. 일반 성인의 기준은 뇌손상, 정신질환, 신경학적 질환 병력이 없고 한국판 간이정신상태검사(K-MMSE)에서 연령대비 2%ile 이상인 자로 설정하였다. 1차에서 모집된 참여자는 50대 57명, 60대 54명, 70대 62명, 80대 이상 51명의 총 224명이었다. 이들의 성별 분포는 남성 66명, 여성 156명으로, 연령대별 성별에는 유의한 차이가 없었다(χ^2=6.127, p>.05). 참여자들의 평균 교육년수는 10.8년(SD=4.6)이었으며, 평균 K-MMSE 점수는 26.5점(SD=3.4)이었다. 각 연령대별 대상자의 교육년수 및 K-MMSE 점수에 대한 기술통계는 Table 1과 같다. 수집된 데이터는 DAS의 말기제 평가와 말 평가의 규준 분석을 위해 활용되었다.

Table 1. Demographic characteristics of the normal group for speech mechanisim and speech task by age (N=224)

Age group	Years of education[a] M (SD)	K-MMSE[b] M (SD)
50 years	14.00 (2.11)	28.67 (1.35)
60 years	12.80 (3.01)	27.43 (2.37)
70 years	9.68 (4.14)	26.26 (3.05)
≥80 years	6.35 (4.44)	23.39 (4.05)

Note. K-MMSE=Korean Mini Mental State Examination.
[a] Values are presented as year. [b] Values are presented as raw score.

AMR과 SMR을 포함하는 말운동 평가에 대해서는 보다 정밀한 기준을 마련하고 백분위 기반의 상대적 비교를 가능하게 할 필요가 있다. 이에 2차 모집을 통해 말운동 평가에 134명의 참여자를 추가 모집하였다. 따라서 말운동 평가 참여자는 1차와 2차 모집을 합하여 총 358명(50대 58명, 60대 79명, 70대 156명, 80대 이상

65명)이었다. 이들의 성별 분포는 남성 113명, 여성 245명으로, 연령대별 성별에는 유의한 차이가 없었다(χ^2=3.279, p>.05). 참여자들의 평균 교육년수는 10.0년(SD=4.5)이었으며, 평균 K-MMSE 점수는 26.2점(SD=3.1)이었다. 각 연령대별 대상자의 교육년수 및 K-MMSE 점수에 대한 기술통계는 Table 2와 같다. 해당 데이터는 DAS의 말운동 평가의 규준 분석을 위해 활용되었다.

Table 2. Demographic characteristics of the normal group for speech motor task by age (N=358)

Age group	Years of education[a] M (SD)	K-MMSE[b] M (SD)
50 years	13.28 (2.07)	28.67 (1.29)
60 years	10.49 (3.35)	26.80 (2.27)
70 years	8.76 (4.39)	25.99 (2.66)
≥ 80 years	7.45 (5.25)	23.78 (3.96)

Note. K-MMSE=Korean Mini Mental State Examination.
[a] Values are presented as year. [b] Values are presented as raw score.

2) 검사 도구 및 절차

본 연구에서 사용한 DAS는 말산출 능력을 정량적으로 평가할 수 있도록 고안된 웹 기반 디지털 마비말장애 평가도구이다. DAS는 크게 말기제 평가, 말 평가, 말운동 평가라는 세 가지 소검사로 구성되어 있다. 총 80개의 문항으로 이루어져 있으며, 각 항목 당 3점 척도(정상=2점, 경도=1점, 심도 이상=0점)로 채점하여 모두 정상일 경우 총점은 160점이 된다. 말기제 평가는 안면, 턱, 혀, 연구개 등 말산출 기제의 구조 및 기능을 평가하며, 말 평가에서는 MPT를 측정하고, 독립음, 구문, 자발화 수준에서 과제를 실시하여 호흡, 발성, 공명, 조음 및 운율 특성에 대해 평가한다. 말 평가 영역은 마비말장애 평가에서 핵심 평가로 영역별과 종합적 말 평가로 구분하여 다양한 수준에서의 말산출 특성을 파악할 수 있도록 하였다. 말운동 평가에서는 AMR과 SMR을 측정하여 조음기관의 최대 운동성을 확인한다. MPT, AMR 및 SMR 과제의 경우 녹음파일을 재생하면 파형이 함께 제시되어 이를 통해 결괏값을 카운트할 수 있다. 특히 AMR과 SMR은 재생 속도를 조절할 수 있으며, 7초간 발화가 녹음되고 5초 구간이 기본 설정되어 있어 검사자가 발화의 시작지점을 설정할 수 있다. 해당 구간의 반복횟수를 입력하면 1초 당 반복횟수는 자동으로 계산되어 평가화면에 기록된다. 검사자는 대상자를 관찰하거나 과제의 수행을 청지각적으로 분석하여 평가하며, 점수가 낮을수록 말장애의 중증도가 높은 것으로 해석할 수 있다.

DAS 검사의 시행은 소정의 교육을 이수한 언어재활사 3인이 4주에 걸쳐 수행하였다. 검사자들은 주 2회, 1년 이상 신경언어장애 환자와 일반 성인의 언어 및 인지 평가를 시행한 경험이 있으며, 300건 이상의 성인 발화를 전사한 경험을 가진 석사 및 박사과정 중인 언어재활사로 선정하였다. 모든 검사자는 검사 실시에 앞서 매뉴얼을 숙지하였다. 또한 20대 성인 3인을 대상으로 예비검사를 시행하여 검사결과를 모니터링 하였으며, 검사자 간 불일치한 항목에 대해서는 합의를 통해 구체적인 채점기준을 설정하는 과정을 거쳤다. 검사 시행 전 모든 참여자에게 녹음 및 데이터 활용에 대한 동의를 얻은 후 평가를 진행하였다. 평가는 독립된 공간에서 검사자와 피검자가 마주 앉아 노트북 또는 태블릿 기기를 사용하여 실시하였다. 음성 녹음 및 채점은 시스템 내에서 처리되었으며 결과 데이터는 웹 기반의 서버에 저장되도록 하였다.

3) 연령군 구분 및 표준점수 기준 설정

우선, DAS의 하위평가(말기제 평가, 말 평가, 말운동 평가) 각각에서 연령대별 평균 점수 변화의 추이를 살펴보았다. 이후 표준점수 산출 시 적용할 연령군 구분을 위해, 50대, 60대, 70대, 80대 이상 집단 간 하위평가 점수에 유의한 차이가 있는지를 일원배치 분산분석(one-way ANOVA)을 이용해 알아보았고, 유의한 차이를 보인 경우 Bonferroni 사후검정을 적용하였다. 통계분석은 SPSS 29.0을 사용하였으며, 유의수준은 .05로 설정하였다.

2. 연구 2: DAS의 신뢰도와 타당도 검증

1) 데이터 수집

신뢰도와 타당도 검증은 규준 작업을 위해 1차 모집되었던 대상자들의 자료에서 각 연구의 목적에 맞게 일부 또는 전체 자료를 추출해 실시하였다. 내용 타당도와 시스템 사용성 평가를 위해 정해진 기준을 충족하는 전문가들을 별도로 모집하여 설문을 진행하였다.

2) 신뢰도 분석

신뢰도 검증을 위해 검사자 간 신뢰도와 내적 일치도를 분석하였다. 우선, 검사자 간 신뢰도는 DAS 채점의 일관성을 검증하기 위해 두 검사자가 30명의 대상자를 독립적으로 채점하여 전체 2,400문항에 대한 검사자 간 일치율(percentage agreement)을 산출하였다. 검사자 간 일치율은 두 검사자가 동일하게 판단한 항목 수/전체 평가 항목 수 × 100으로 산출하였다.

내적 일치도는 검사 문항들이 동일한 개념을 얼마나 일관되게 측정하고 있는지를 나타내는 지표로 DAS의 전체 80개 문항 간 일관성을 평가하기 위해 50세 이상 일반 성인 224명의 검사문항에 대한 Cronbach's α 계수를 산출하였다.

3) 타당도 분석

타당도 검증에는 구인 타당도, 공인 타당도 및 내용 타당도를 분석하였다. 먼저, DAS의 구인 타당도를 확인하기 위해 일반 성인 224명의 DAS 수행자료를 활용하여 DAS 총점과 각 하위평가 항목 점수 간 상관분석을 실시하였다.

공인 타당도 검증을 위해 50대 이상 일반 성인 30명을 대상으로 DAS 하위검사와 기존 신경말장애 평가 프로토콜(Neurological Speech Assessment Protocol: N-SAP, Kim, 2024)의 하위검사를 각각 실시하여 두 검사 간 수행 점수의 Pearson 상관분석을 실시하였다.

내용 타당도 검증은 신경의사소통장애 평가 및 치료 경험이 있는 언어재활 전문가 35인에게 설문방식으로 진행하였다. 전문가들

은 DAS의 전체 문항에 대해 각 문항별로 관련성, 명확성, 적절성, 대표성의 4개 항목을 5점 척도(1=전혀 관련 없음, 2=관련성 낮음, 3=어느 정도 관련 있음, 4=관련 있음, 5=매우 관련 있음)로 평가하였다. 각 문항별로 문항 수준 내용 타당도 지수인 I-CVI(item-level content validity)와 척도 수준 내용 타당도 지수인 S-CVI(scale-level content validity index)를 산출하였다. I-CVI는 전체 해당 문항에 대해 4점 이상을 준 전문가 수/전체 전문가 수로 산출하였으며, S-CVI는 전체 문항의 I-CVI 평균값/전체 문항 수로 산출하였다. 내용 타당도에 참여한 전문가 집단 정보는 Appendix 1에 수록하였다.

4) 시스템 사용성 평가

DAS의 온라인 기반 실용성 확인을 위해 SUS(system usability scale)를 활용하여 6명의 사용자에게 총 10개 문항에 대한 평가를 실시하였다. SUS는 시스템, 제품, 서비스의 사용성을 신속하게 평가하기 위한 표준화된 설문 도구로 사용자 경험(user experience: UX) 평가 분야에서 가장 널리 사용되는 평가 척도이다(Brooke, 1996). 사용성 평가는 해당 검사도구를 10회 이상 사용한 경험이 있는 언어재활사 6명이 담당하였다. 사용성 평가의 각 문항은 5점 척도(1=전혀 그렇지 않다, 2=그렇지 않다, 3=보통이다, 4=그렇다, 5=매우 그렇다)로 평가하며, 10개 문항 중 5개의 문항은 긍정 문항이며, 나머지 5문항은 부정 문항이다. 부정 문항은 역코딩하여 평균과 100점 만점의 변환 평균을 산출하였다. SUS 항목별 세부 문항은 Appendix 2에 제시하였다.

Ⅲ. 연구 결과

1. 연구 1: DAS의 규준 정립

1) 말기제 평가 및 말 평가 규준

전체 대상자들의 평균 말기제 평가 점수는 59.64점(SD=0.94), 말 평가 점수는 99.00점(SD=2.13), DAS 총점은 158.65점(SD=2.63)이었다. 말기제 평가, 말 평가, DAS 총점에 대한 연령대별 평균 점수는 Figure 2, 3, 4와 같다. 전반적으로 연령이 증가함에 따라 점수가 떨어지는 양상을 보여주었다.

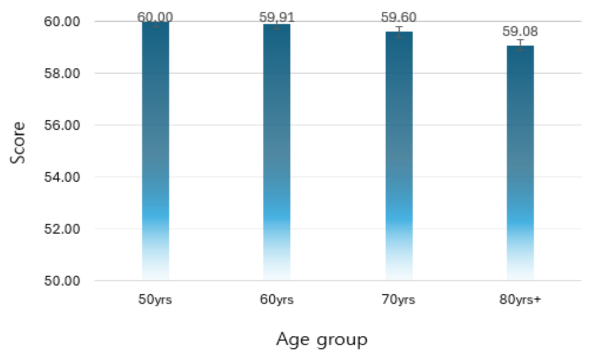


Figure 2. Speech mechanism scores by age group

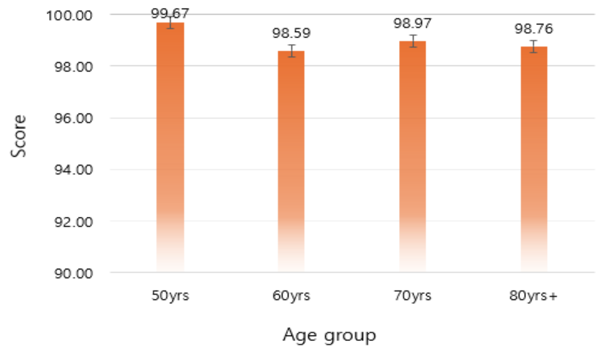


Figure 3. Speech scores by age group

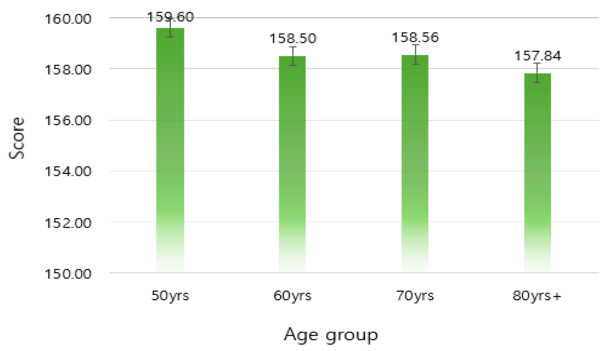


Figure 4. Total DAS score by age group

표준점수 산출 시 적용할 연령군 구분을 위해, 말기제 평가 점수, 말 평가 점수, DAS 총점 각각에 대해 네 연령집단 간 차이를 ANOVA 분석을 통해 알아보았다. 분석 결과, 말기제 평가 점수($F_{(3, 220)}$=10.541, p<.001, partial η^2=.126), 말 평가 점수($F_{(3, 220)}$=2.779, p<.05, partial η^2=.037), DAS 총점($F_{(3, 220)}$=4.330, p<.01, partial η^2=.056) 모두에서 집단 간 차이가 유의하였다. 사후검정 결과, 말기제 평가와 말 평가 점수에서는 80대 이상 집단이 50대(p<.001), 60대(p<.001), 70대 집단(p<.01)에 비해 유의하게 낮았으며, DAS 총점에서는 50대와 80대 이상 집단 간 차이가 유의하였다(p<.01).

Table 3. Demographic characteristics of the normal group for speech mechanisim and speech task by age (N=224)

Age group	Gender		Years of education[a]			K-MMSE[b]
	Male	Female	0~5	6~15	≥16	M (SD)
50~79 years	54	119	4	121	48	27.42 (2.57)
≥ 80 years	12	39	11	37	3	23.39 (4.05)

Note. K-MMSE=Korean Mini Mental State Examination.
[a] Values are presented as year. [b] Values are presente as raw score.

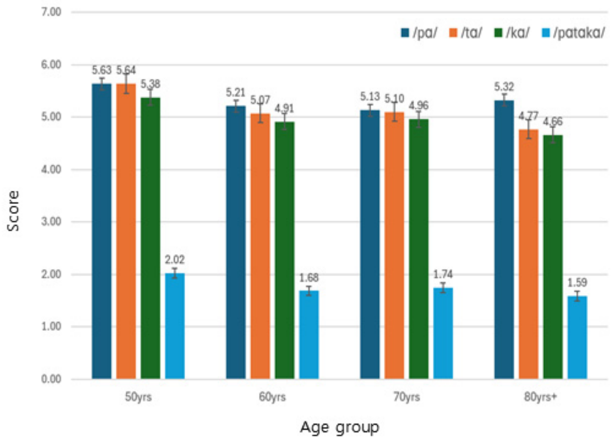
이에 본 검사는 기존 네 집단을 50~79세(n=173) 집단과 80세 이상(n=51) 집단으로 이원화하여 규준을 산출하였다. 해당 규준에 따라 재분류된 정상군 표본의 기본 정보는 Table 3과 같다. 이후 두 집단에 대해 10분위 단위로 백분위 값을 산출하여,

각 집단의 10%ile 이하는 ‘약한 기능 저하’로, 2%ile 이하는 ‘심한 기능 저하’로 해석하도록 규준을 설정하였다. 이 같은 해석 기준은 DAS 매뉴얼에 상세히 기술되어 있으며, 온라인 상에서 DAS를 실시한 후 본 규준에 기반하여 얻을 수 있는 결과 화면의 예시를 Appendix 3에 수록하였다.

2) 말운동 평가 규준

말운동 평가 결과, 전체 AMR 평균은 /파/, /타/, /카/ 각각 5.19회(SD=0.74), 5.12회(SD=0.78), 4.96회(SD=0.76)였으며, SMR /파타카/ 평균은 1.75회(SD=0.32)였다. /파/, /타/, /카/, /파타카/ 항목에 대한 연령대별 AMR과 SMR 평균은 Figure 5와 같다. 전반적으로 연령이 증가함에 따라 AMR과 SMR 수치가 떨어지는 양상을 보여주었다.

통계 분석 결과, /파/, /타/, /카/와 /파타카/ 초 당 반복횟수에서 모두 연령대 간 유의한 차이를 보였다(/파/: $F_{(3, 354)}$=11.498, p<.001, partial η^2=.089, /타/: $F_{(3, 354)}$=14.638, p<.001, partial η^2=.110, /카/: $F_{(3, 354)}$=9.985, p<.001, partial η^2=.078, /파타카/: $F_{(3, 354)}$=24.780, p<.001, partial η^2=.174). 사후검정 결과, /파/에서 50대가 60대(p<.01), 70대(p<.001), 80대 이상(p<.001) 집단과 각각 유의한 차이를 보였으며, 60대가 80대 이상 집단에 비해 유의하게 높았다(p<.001). /타/에서도 50대는 다른 세 집단에 비해 유의하게 높았으며(p<.001), 70대도 80대 이상 집단에 비해 유의하게 높은 반복횟수를 보였다(p<.05). /카/의 반복횟수에서도 50대는 60대(p<.01), 70대(p<.01) 및 80대 이상(p<.001) 집단에 비해 유의하게 높았으며, 70대와 80대 이상 집단 간 차이도 유의하였다(p<.05). /파타카/에서는 50대가 60대(p<.001), 70대(p<.001), 80대 이상(p<.001) 집단에 비해 유의하게 높았으며, 70대와 80대 이상 집단 간 차이도 유의하였다(p<.01).



Note. AMR=alternating motion rate; SMR=sequential motion rate.
Figure 5. AMR and SMR by age group

이에 말운동 평가 결과는 50대, 60대, 70대, 80대 이상의 총 네 개의 연령군으로 구분하여 각 연령군별 평균 초당 반복횟수 및 10단위 백분위를 산출하여 대상자의 수행 결과를 각 연령대별로 상대적 비교가 가능하도록 기준을 마련하였다. 개별 연령군별 cut-off 점수 및 백분위 점수는 온라인 결과보고서를 통해 자동으로 제공되도록 하였으며, 온라인상 출력되는 말운동 평가 결과 화면 일부를 Appendix 4에 제시하였다.

2. 연구 2: DAS의 신뢰도와 타당도 검증

1) 검사자 간 신뢰도

DAS 채점의 일관성 검증을 위해 검사자 간 일치율을 분석하였다. 그 결과, 전체 2,400문항의 채점 사례에 대한 평균 검사자 간 일치율은 99.04%로 매우 높은 검사자 간 신뢰도를 보였다. 평가 영역별로 살펴보면, 말기제 평가에서는 99.00%, 영역별 말 평가에서는 98.13%, 종합적 말 평가의 일치율은 100%였다.

2) 내적 일치도

각 평가 영역의 문항 간 일관성을 검증하기 위해 224명 참여자의 전체 문항(80×226)에 대한 Cronbach’s α 계수를 산출하였다. 그 결과, 전체 문항에 대한 신뢰도 계수는 α=.808로 우수한 수준(≥.80)의 내적 일치도를 나타냈다. 하위 검사별로 살펴보면 말기제 평가에서 α=.606으로 ‘보통’ 수준이었으며, 말 평가에서 α=.788로 ‘양호한’ 수준을 보였다.

3) 구인 타당도

DAS의 총점과 각 하위평가 항목 간 상관분석을 실시한 결과 DAS 총점은 말기제 평가 점수(r=.669, p<.001), 말 평가 점수(r=.925, p<.001)와 각각 유의한 정적 상관을 보였다. 말운동 지표에서는 /타/의 초당 반복횟수와 DAS 총점 간 유의한 정적 상관을 나타냈다(r=.216, p<.01).

4) 공인 타당도

DAS와 기존 신경말장애 평가 프로토콜인 N-SAP검사의 하위검사 간 Pearson 상관분석을 실시한 결과, DAS의 총점과 N-SAP의 모음검사 점수(r=.416, p<.05), 자음검사 점수(r=.599, p<.01)와 각각 유의미한 정적 상관을 보였다. 단, DAS의 말기제 점수와 영역별 말 평가 점수, N-SAP의 비구어 점수와 명료도 점수의 경우 대상자들의 수행이 척도의 상한에 도달하여 천장효과(ceiling effect)를 보였기 때문에 분석에서 제외하였다.

5) 내용 타당도

문항별 I-CVI를 산출한 결과, 전체 35명의 전문가 중 4점 또는 5점으로 평가한 비율은 .88~1.00범위로 대부분의 문항에서 높은 타당도를 나타냈다. S-CVI는 말기제 평가에서 .97, 말 평가에서 .97로 나타났으며 DAS 문항 전체 S-CVI 평균은 .97로 나타났다.

6) 시스템 사용성 평가

시스템 사용성 평가(SUS) 결과, 전체 평균 SUS점수는 32.17점(SD=0.55)이었으며, 변환 평균은 80.42점(SD=1.38)으로 ‘우수(excellent)’ 수준으로 나타났다. 유용성 및 전문성(기술지원) 항목에서 가장 높은 평균 점수를 획득하였으며, 전문성(사전학습) 영역에서 가장 낮은 점수를 보였다. 이러한 결과는 DAS 평가 시스템

이 사용자에게 쉽게 접근 가능하고 직관적인 검사도구로 평가받고 있음을 나타낸다. SUS의 각 항목별 평균과 변환 평균은 Table 4와 같다.

Table 4. Mean and converted mean scores for SUS items

No.	Item	*M* (*SD*)	Convert *M* (*SD*)
1	Usefulness	3.67 (0.52)	9.17 (1.29)
2	Complexity	3.00 (0.63)	7.50 (1.58)
3	Simplicity	3.50 (0.55)	8.75 (1.37)
4	Professionalism (technical support)	3.67 (0.82)	9.17 (2.04)
5	Intergrability	3.17 (0.41)	7.92 (1.02)
6	Consistency	3.33 (0.52)	8.33 (1.29)
7	Learnability	3.00 (0.00)	7.50 (0.00)
8	Convenience	3.33 (0.52)	8.33 (1.29)
9	Satisfaction	3.17 (0.75)	7.92 (1.88)
10	Professionalism (prior learning)	2.33 (0.82)	5.83 (2.04)
	Total	32.17 (0.55)	80.42 (1.38)

Ⅳ. 논의 및 결론

본 연구에서는 마비말장애 평가시스템(DAS)의 신뢰도 및 타당도를 분석하고 규준을 마련하고자 하였다. DAS는 마비말장애 환자 평가를 위한 웹 기반의 청지각적 평가시스템으로 말기제 평가, 말 평가, 말운동 평가라는 세 가지 주요 평가 영역으로 구성된 검사도구이다.

연구 결과, DAS의 세부 평가과제인 말기제 평가, 말 평가의 경우에는 50~79세 집단과 80세 이상 집단 간 점수 차이가 유의하여, 80세 미만과 80에 이상으로 집단 구분을 이원화하여 규준값을 제공하였다. 이에 반해, 말운동 평가의 경우에는 50대, 60대, 70대, 80대 이상의 네 집단 간 수행 차이가 유의하여, 네 집단별로 규준값을 제공하였다. 신뢰도 검증 결과, 평균 검사자 간 채점의 일치율은 99.04%로 높게 나타났으며, 문항 간 일관성 검증에서도 우수한 수준의 내적 일치도를 보였다. 이러한 결과는 DAS 검사도구가 채점자의 주관적 편차가 적고, 동일한 채점 기준과 절차가 유지되었음을 의미하며 검사의 하위 문항들이 동질적이고 높은 신뢰도를 보인다는 것을 시사해 준다. 또한 타당도 결과는 DAS는 하위 항목과 총점 간 높은 상관을 통해 도구의 내적 구조가 일관됨을 보여주었다. 기존 타당화된 평가도구와의 상관을 통해 임상적 신뢰성을 확보했으며, 전문가 검증 결과 매우 높은 내용 타당도를 보여 평가 영역을 포괄적으로 반영하는 도구임을 입증하였다.

DAS의 평가 항목은 표면적으로 드러나는 증상에 기반하여 결함 유무를 체크하는 방식으로 해당 결함이 장애로 인한 필연적 결함인지, 장애를 위한 보상적 전략인지 구분하는 것은 쉽지 않다. 특히, 마비말장애는 하위 유형 간 겹치는 증상이 많기 때문에 증상에만 기반하여 하위 유형을 판단하는 데는 어려움이 있다. 따라서 해당 증상이 나타나면 그 정도에 따라 ‘경도’ 또는 ‘심도’에 체크한 후, 검사자의 축적된 경험과 전문적 지식을 활용하여 피검자의 핵심적 문제를 파악하여 판단해야 할 것이다.

DAS는 마비말장애 환자를 위한 표준화 검사도구로 장애등급 판정 기준 마련, 수술 및 치료를 통한 말영역 기능 회복 정도 평가, 뇌졸중 및 뇌손상 후 말영역 결함 여부 진단과 중증도 판단, 노화과정 중 정상 범주를 벗어나 유의미한 말기제 또는 말산출 기능저하를 보이는 대상자 선별에 사용가능하다. 그 외에도 뇌성마비 아동이나 말운동 기능저하가 의심되는 말소리장애 아동 등에게도 평가를 시행할 수 있다. 다만 이 경우에는 규준 자료가 제공되지 않으므로 아동의 각 영역별 강약점을 파악하는 정도로 활용할 수 있을 것이다. DAS는 웹 기반의 평가도구로 검사시간의 단축, 평가 및 결과 산출의 자동화, 보고서 작성의 용이성으로 병원 현장뿐 아니라 마비말장애 연구 현장에서도 폭넓게 활용될 수 있을 것으로 기대한다.

다만 본 연구에서는 정상군만을 대상으로 데이터를 수집하였기 때문에 다음과 같은 제한점이 있다. 첫째, 본 연구에서는 DAS의 민감도와 특이도를 직접적으로 평가하지 못하였다. 둘째, DAS가 주로 구어운동 능력을 평가한다는 특성상, 정상 성인 집단에서는 여러 과제에서 천장효과가 나타났다. 그 결과 공인 타당도를 검증하는 과정에서 과제 변별력이 제한되었다. 셋째, 검사자 간 신뢰도 분석에서 대부분의 평가항목이 ‘정상’으로 채점되어 분포의 불균형이 발생하였으며 이로 인해 단순 일치율만을 산출하였다는 제한점이 있다. 마지막으로 본 연구의 표본은 주로 복지관이나 노인정 등 특정 지역 기반 시설을 통해 모집되었기 때문에 일반 성인을 충분히 대표한다고 보기 어렵다. 표집 과정에서 사회·경제적 배경이나 지역 분포 등의 요인이 통제되지 않았으므로, 연구 결과를 전체 성인 집단으로 일반화하는 데는 한계가 있다. 이러한 제한점을 보완하기 위해 추후 연구에서는 다양한 병리 집단(예: 뇌졸중, 파킨슨병, 외상성 뇌손상 등)을 포함한 대상군을 확보하여 규준을 확장함으로써 절단점(cut-off score)을 산출하고 DAS의 진단 정확도를 검증해야 할 것이다.

Reference

Brooke, J. (1996). SUS: A quick and dirty usability scale. In P. W. Jordan, B. Thomas, I. L. McClelland, & B. Weerdmeester (Eds.), *Usability evaluation in industry* (pp. 189-194). Taylor & Francis.

Cohen, R. J., & Swerdlik, M. E. (2018). *Psychological testing and assessment: An introduction to tests and measurement* (9th ed.). New York: McGraw-Hill.

Duffy, J. R. (2005). *Motor speech disorders: Substrates, differential diagnosis, and management* (2nd ed.). MO: Elsevier Mosby.

Duffy, J. R. (2019). *Motor speech disorders: Substrates, differential* *diagnosis, and management* (4th ed.). MO: Elsevier Mosby.

Enerby, P. M., & Palmer, R. (2008). *FDA-2: Frenchay dysarthria assessment: Examiner's manual.* Austin, TX: Pro-Ed.

Freed, D. B. (2020). *Motor speech disorders* (3rd ed.). San Diego, CA: Plural Publishing.

Hong, S. M., Jeong, P., & Sim, H. S. (2018). Comparison of perceptual assessment for dysarthric speech: The detailed and general assessments. *Communication Sciences & Disorders, 23*(1), 242-253. doi:10.12963/csd.18457

Kim, H. H. (2024). *Neurologic Speech Assessment Protocol* (N-SAP). Seoul: Primemedi.

Kim, J.-Y., Jo, D. H., Yoon, J., Park, K.-S., & Ha, J.-W. (2024). Development of the auditory-perceptual assessment for dysarthria: The dysarthria assessment system (DAS). *AAC Research & Practice, 12*(2), 1-21. doi:10.14818/aac.2024.12.12.2.1

Kim, Y. T., Shin, M. J., Kim, S. J., & Ha, J. W. (2020). *Urimal Test of Articulation and Phonology 2* (UTAP2). Seoul: Hakjisa.

Shin, M. J., Kim, J. O., Lee, S. B., & Lee, S. Y. (2010). *Speech Mechanism Screening Test* (SMST). Seoul: Hakjisa

Yorkston, K. M., & Beukelman, D. R. (1981). *Assessment of the intelligibility of dysarthric speech (manual).* Oregon: CC Publications.

Appendix 1. Expert panel information for content validity evaluation (N=35)

기본 정보		비율
성별	남	22.9%
	여	77.1%
연령	20대	8.6%
	30대	68.6%
	40대	22.9%
최종 학력	석사 수료	8.6%
	석사 졸업	65.7%
	박사 수료	8.6%
	박사 졸업	8.6%
자격증	1급 언어재활사	94.3%
	2급 언어재활사	5.7%
경력	5년 이상~10년 미만	48.6%
	10년 이상~15년 미만	42.9%
	15년 이상~20년 미만	5.7%
	20년 이상	2.9%
소속기관	병원	54.3%
	사설언어재활기관	14.3%
	대학	11.4%
	복지관	5.7%
	연구소	2.9%
	기타	11.4%

Appendix 2. Items of the system usability scale (SUS)

번호	항목	내용
1	유용성	나는 이 시스템을 자주 사용하고 싶다.
2	복잡성	나는 이 시스템이 불필요하게 복잡하다고 생각했다.
3	단순성	나는 이 시스템이 사용하기 쉽다고 생각했다.
4	전문성(기술지원)	나는 이 시스템을 사용하려면 기술적 지원이 필요하다고 생각한다.
5	통합성	나는 이 시스템에 다양한 기능들이 잘 통합되어 있다고 생각했다.
6	일관성	나는 이 시스템에 일관성이 부족하다고 생각했다.
7	학습 용이성	대부분의 사람들이 이 시스템을 빠르게 배울 수 있을 것이라고 생각한다.
8	편리성	나는 이 시스템이 매우 번거롭다고 느꼈다.
9	만족감	나는 이 시스템을 사용하는 데 매우 자신감이 있었다.
10	전문성(사전학습)	나는 이 시스템을 시작하기 전에 많은 것을 배워야 했다.

Appendix 3. Sample screen of the speech mechanism and speech assessment results

SPEECH: 말 평가

SPEECH MECHANISM : 말기제 평가

Appendix 4. Sample screen of the speech motor assessment results

마비말장애 평가 시스템 Dysarthria Assessment System 이건강님

SPEECH MOTOR : 말운동 평가

AMR	초당 반복횟수	연령별 평균횟수	백분위
파	3.65	4.9(회)	2-10
타	3.89	4.77(회)	2-10
카	3.29	4.66(회)	2-10

SMR	초당 반복횟수	연령별 평균횟수	백분위
파타카	0.78	1.59(회)	2-10

언어치료연구 2025년 제34권 제4호 093-102
ISSN 1226-587X / eISSN 2671-7158
JSLHD (Journal of Speech-Language & Hearing Disorders)
ORIGINAL ARTICLE
https://doi.org/10.15724/jslhd.2025.34.4.093
http://jslhd.org/

마비말장애 평가시스템 개발을 위한 정상 성인 대상 표준화 연구

김지영[1], 박기수[2,3], 하지완[4*]

[1] 대구대학교 언어치료학과 연구교수
[2] 네오폰스(주) CEO
[3] 경북대학교 의과대학 신경외과 교수
[4] 대구대학교 언어치료학과 교수

목적: 본 연구는 마비말장애 평가를 위한 웹 기반의 청지각적 평가도구인 Dysarthria Assessment System(DAS)의 신뢰도 및 타당도를 검증하고, 일반 성인을 대상으로 한 규준을 마련하는 데 그 목적이 있다.

방법: DAS는 말기제 평가, 말 평가, 말운동 평가의 세 영역으로 총 80문항으로 구성되어 있다. 정상일 경우 0점, 경도일 경우 1점, 심도일 경우 0점으로 표기하는 3점 척도를 사용한다. 본 연구에서는 DAS를 사용하여 50세 이상 일반 성인 328명의 데이터를 분석하였으며, 평가자 간 신뢰도, 내적 일치도, 구인 타당도, 공인 타당도, 내용 타당도 및 시스템 사용성(SUS) 평가를 실시하였다. 말기제 평가, 말 평가 및 말운동 평가에 대한 연령군별 규준을 설정하기 위해 일원분산분석을 실시하였다.

결과: 신뢰도 검증 결과, 평가자 간 일치율은 99.04%로 높았고, Cronbach's α 계수는 .808로 양호한 내적 일치도를 보였다. DAS 총점은 하위 항목 및 기존 평가도구와 유의한 상관을 보여 타당도가 입증되었으며, 내용 타당도 지수(S-CVI)는 .97로 매우 높았다. 또한 시스템 사용성 평가 결과도 '우수' 수준으로 나타났다. 말기제 평가와 말 평가의 연령군별 규준은 '50~79세'와 '80세 이상'으로 구분하여 제시하였으며, 말운동 평가는 50대, 60대, 70대, 80대 이상으로 구분하여 연령대별 백분위 기준을 제공하였다.

결론: 본 연구는 마비말장애 환자를 위한 웹 기반의 표준화 평가 시스템인 DAS를 개발하여 일반 성인을 대상으로 표준화하였다. DAS는 검사시간의 단축, 평가 및 결과 산출의 자동화, 보고서 작성의 용이성으로 병원 현장뿐 아니라 마비말장애 연구 현장에서도 폭넓게 활용될 수 있을 것이다.

검색어: 마비말장애, 평가 시스템, 표준화, 신뢰도, 타당도

교신저자 : 하지완(대구대학교)
전자메일 : jw-ha@daegu.ac.kr

게재신청일 : 2025. 08. 23
수정제출일 : 2025. 09. 23
게재확정일 : 2025. 10. 31

이 연구는 2022년도 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2022S1A3A2A03089435).

ORCID
김지영
https://orcid.org/0000-0002-6971-6731
박기수
https://orcid.org/0000-0002-4829-6299
하지완
https://orcid.org/0000-0002-1191-791X

참고문헌

김영태, 신문자, 김수진, 하지완 (2020). **우리말 조음·음운 검사 2(UTAP2).** 서울: 학지사.

김지영, 조대형, 윤장혁, 박기수, 하지완 (2024). 마비말장애 환자를 위한 청지각적 평가 시스템(DAS) 개발. **보완대체의사소통연구, 12**(2), 1-21.

김향희 (2024). **신경말장애 평가 프로토콜(N-SAP).** 서울: 프라임메디.

신문자, 김재옥, 이수복, 이소연 (2010). **조음기관 구조·기능 선별검사(SMST).** 서울: 학지사.

홍새미, 정필연, 심현섭 (2018). 마비말장애 발화의 청지각적 평가방법 비교: 세부평가와 범주평가. Communication Sciences & Disorders, 23(1), 242-253.