



## Development of the video-based smart utterance deep analyser (SUDA) application\*

Soo-Bok Lee<sup>1</sup> · Hyo-Jung Kwak<sup>2</sup> · Jae-Min Yun<sup>3</sup> · Dong-Chun Shin<sup>3</sup> · Hyun-Sub Sim<sup>2,\*\*</sup>

<sup>1</sup>Department of Speech Language Therapy and Aural Rehabilitation, Woosong University, Daejeon, Korea

<sup>2</sup>Department of Communication Disorders, Ewha Womans University, Seoul, Korea

<sup>3</sup>Bory Inc., Ansan, Korea

### Abstract

This study aims to develop a video-based smart utterance deep analyser (SUDA) application that analyzes semi-automatically the utterances that child and mother produce during interactions over time. SUDA runs on the platform of Android, iPhones, and tablet PCs, and allows video recording and uploading to server. In this device, user modes are divided into three modes: expert mode, general mode and manager mode. In the expert mode which is useful for speech and language evaluation, the subject's utterances are analyzed semi-automatically by measuring speech and language factors such as disfluency, morpheme, syllable, word, articulation rate and response time, etc. In the general mode, the outcome of utterance analysis is provided in a graph form, and the manger mode is accessed only to the administrator controlling the entire system, such as utterance analysis and video deletion. SUDA helps to reduce clinicians' and researchers' work burden by saving time for utterance analysis. It also helps parents to receive detailed information about speech and language development of their child easily. Further, this device will contribute to building a big longitudinal data enough to explore predictors of stuttering recovery and persistence.

**Keywords:** stuttering, longitudinal study, video-based utterance analyser, hybrid application, speech language therapy

### 1. 서론

말더듬은 언어 발달이 급속도로 이루어지는 시기인 2-5세에 주로 시작되는데(Guitar, 1998; Lee, 2005), 이들 중 75-85%는 발생 후 3-4년 이내에 자연적으로 회복되나 나머지 15-25%는 말더듬이 지속되어 만성적 말더듬이 될 수 있다(Yairi & Ambrose,

2005). 그러므로 말더듬을 조기 발견 및 조기 중재하는 것은 매우 중요하며, 말더듬의 시작 및 진전에 잠재적으로 작용하거나 회복 및 지속에 영향을 미치는 요인들을 찾아 증거 기반의 중재(evidence-based practice)를 하려는 노력이 필요하다.

말더듬의 회복 및 지속요인들에 관한 연구에 의하면 말더듬 발달에 다양한 요인들이 상호작용하며, 이러한 요인들 간의 상

\* This work was supported by the Ministry of Education of the Republic of Korea and National Research Foundation of Korea (NRF-2018S1 A5A2A03036976).

\*\* simhs@ewha.ac.kr, Corresponding author

Received 30 April 2020; Revised 25 May 2020; Accepted 25 May 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.

호 영향력은 시간이 지남에 따라 달라질 수 있다고 주장한다 (Smith & Kelly, 1997). 취학 전 아동은 신경학적, 생리학적, 정서적, 언어학적으로 성숙되어 가는 과정에 있기 때문에, 자연스럽게 비유창성이 나타날 수 있다. 이러한 비유창성과 초기 말더듬 증상을 구별하는 것은 쉽지 않기 때문에 정상적인 비유창성으로 판단하여 치료를 유보할 것인지, 말을 더듬을 가능성이 크고, 지속될 것으로 판단되어 중재를 시작해야 할지를 결정하는 것은 매우 중요하다(Byun et al., 2004). 또한 아동의 비유창성이 비정상적인 것으로 판단되면 이러한 패턴이 회복될 것인지를 정확하게 진단해야 하지만, 아동의 유창성 수준이 시점별로 다른 특성을 보일 수 있기 때문에 구어 유창성 능력이 제대로 반영된 데이터를 확보하는 것이 어렵다(Manning, 2001). 그러므로 특정 시점에서 수집된 발화보다는 장기간 지속적으로 발화를 수집하는 것이 필요하다. 연구자들은 이러한 점을 고려하여 시간의 흐름에 따라 데이터를 관찰하고 분석하는 종단연구 방법을 사용하여 연구를 진행한다. 그러나 종단연구는 횡단연구에 비해 많은 시간과 비용이 소요되며, 연구 도중에 연구대상자들이 이사를 가거나 연락이 끊기는 경우가 발생하여 자료를 수집하지 못하는 어려움이 발생할 수 있다(Kim et al., 2007).

또한 임상 현장에서 의사소통장애를 정확하게 진단하기 위해서는 선별검사, 표준화된 공식검사, 자발화 분석검사, 비공식 심화검사 등 다양한 검사가 실시되어야 한다. 특히 가정, 학교 및 일상적인 상황에서 아동의 자연스럽게 실제적인 말-언어능력을 평가하기 위해서 언어치료사들은 아동의 자발화를 수집하고 전사하고 분석해야 한다(Costanza-Smith, 2010). 자발화 분석 방법은 다른 검사들에 비해 더 많은 시간과 노력이 요구되고, 분석 도구 및 방식에 따라 그 결과가 달라지는 단점이 있다(Paul & Norbury, 2012). 이러한 단점을 보완하여 검사의 효율성을 높이기 위해서 컴퓨터 프로그램을 활용할 수 있다. 국외에서는 Systematic Analysis of Language Transcripts(SALT; Miller & Iglesias, 2015), Computerized Language Analysis(CLAN; MacWinney, 2000), Sampling Utterances and Grammatical Analysis Revised(SUGAR; Pavelko & Owens, 2017)가 사용되고 있다(Pezold et al., 2020). 미국의 경우 임상적인 측면에서 CLAN 프로그램의 활용도를 높이기 위한 방안으로 National Institute on Deafness and Other Communication Disorders(NIDCD)와 Eunice Kennedy Shriver National Institute of Child Health and Human Development(NICHHD)의 지원을 받아, Child Language Data Exchange System(CHILDES)의 데이터베이스를 사용하여 해당발화샘플에 대한 표준편차 및 여러 가지 측정치를 자동적으로 산출해줄 수 있는 KIDEVAL 분석기능이 개발 중이다(Garbarino et al., 2020). 국내에도 발화 분석의 효율성을 높이기 위해 웹 기반 한국어 발화 분석(Korean Language Analysis, KLA, Pae et al., 1998)이라는 소프트웨어가 개발된 적이 있으나 임상 현장에서 널리 사용되지 않고 있다(Ha et al., 2016). 국외에서 개발된 프로그램을 사용하더라도, 이 프로그램들이 다루고 있는 영역이 제한적이기 때문에 임상현장에서 효과적으로 사용되지 못하고 있는 실정이다. 나아가 언어치료사들은 프로그램 작동법 및 분석결과를 해석

하는 방법 등의 프로그램 사용방법을 익히는 과정에서 많은 시간이 소요되고, 효과적으로 사용하는데 어려움을 겪고 있다(Pezold et al., 2020).

이러한 실정을 고려하였을 때, 언어치료전문가들이 임상 및 연구 상황에서 시공간적인 한계에 얽매이지 않고 대상자의 발화를 좀 더 편리하고 수월하게 수집 및 분석할 수 있는 방안이 필요하다고 판단되어, 본 연구에서는 현재 빠르게 발전하고 있는 디지털 기술을 활용하는 방법을 제안하고자 한다.

디지털 기술이 계속해서 발전해감에 따라 누구나 일상생활에서 스마트 미디어(smart media)를 사용하는 모습을 볼 수 있는데, 이러한 스마트 미디어에는 스마트폰 및 태블릿 PC 등이 포함된다(Yim et al., 2014). 특히 스마트폰의 경우, 국내에서 특정 연령대와 계층에 국한되지 않고 보편적인 휴대용 통신기기로 사용되고 있으며, 82%의 사용자가 매일 스마트폰을 사용하는 것으로 밝혀져 스마트폰 사용량 또한 아시아·태평양 지역에서 가장 높은 수치를 보이는 국가이다(Lee et al., 2016; Yim et al., 2014). 이처럼 스마트폰은 우리 사회에서 필수품과 같은 존재가 되었으며, 어플리케이션은 가장 접근성이 좋은 디지털 기술이라고 할 수 있다(Lee et al., 2019). 이러한 어플리케이션은 현재 의료 및 재활 분야에서 스마트폰을 통해 질병의 진단을 돕고 치료 가이드라인을 제시하며, 가정 기반 재활(home based rehabilitation), 원격 재활(telerehabilitation) 목적으로 유용하게 사용되어지고 있다(Im et al., 2013). 또한 시공간의 제약을 극복할 수 있고, 개별적이고 다양한 자료를 활용한 교육이 가능하다는 점에서 일반 및 특수 교육 분야에서 스마트 러닝 목적으로 활용되고 있다(Lee et al., 2015; Lee & Lee, 2014). 언어치료 분야에서도 스마트폰을 활용하는 것이 의사소통장애 아동의 말·언어 발달에 유용할 것으로 판단되어 개발 및 배포하고자 하는 노력이 시도되고 있다. 하지만 임상 현장에서 발생하는 시간 및 공간적인 제약과 같은 현실적인 어려움이 해결되지 못하고 있으며, 대상자의 일상생활에 밀착하여 서비스를 제공하기 위한 목적으로 IT 기술을 활용한 연구가 많이 부족하다(Park, 2012; Yim et al., 2014). 또한 각 의사소통장애 군의 중재와 관련된 어플리케이션은 많이 개발되어지고 있으나(Go et al., 2017; Jeong, 2014; Lee & Lee, 2017; Lim & Park, 2012; Yang & Kang, 2011), 일상생활에서 부모와 아동 간의 상호작용 영상을 지속적으로 촬영함으로써 그들의 발화를 종단적으로 분석을 해주는 어플리케이션은 없는 실정이다. 그러므로 현재까지 발전된 IT 기술을 활용하여 데이터를 장기간에 걸쳐 축적하고, 축적된 데이터를 자동적으로 분석해주는 어플리케이션이 개발된다면, 임상적으로 뿐만 아니라 학문적인 측면에서 많은 도움이 된다. 예를 들면, 언어치료사들에게 의사소통장애의 진단 및 평가, 중재에 관한 중요한 임상적 정보를 제공할 수 있고, 또한 시간변화에 따른 의사소통장애 특성을 변화를 객관적으로 관찰함으로써 회복 및 지속에 관한 요인 파악도 가능하게 된다.

위와 같은 임상적 현실에 기초하여 본 연구는 일상생활에서 부모와 아동 상호작용 시의 유창한 발화뿐만 아니라, 비유창한 발화를 지속적으로 녹화하여 종단적으로 수집 및 분석함으로써

써, 시간의 흐름에 따라 아동의 말·언어발달 및 비유창성 변화 추이와 부모의 발화패턴을 동시에 살펴볼 수 있는 SUDA(Smart Utterance Deep Analyser)를 개발하고자 하였다. 본 연구를 통해 개발된 발화 분석기는 다른 의사소통장애의 진단 및 중재에 적용할 수 있지만, 특별히 한국형 말더듬 아동의 진단 및 중재에 적용할 수 있는 중단 데이터를 체계적으로 구축하는데 기여할 것이며(Lee et al., 2019), 구축된 데이터는 말더듬 회복 예측 요인들을 찾는 기초자료로 활용될 것이다.

## 2. 연구방법

### 2.1. SUDA 개발 절차

본 시스템은 다양한 이종 플랫폼에서 실행될 수 있도록 기존 오픈소스인 IONIC Framework를 통해 개발되었다. IONIC Framework는 HTML, CSS, JavaScript 같은 Web 기술을 이용하여 Progressive Web Apps(PWAs), High-performance mobile을 지원하기 때문에, iOS, ANDROID, WEB 등 다양한 이종 플랫폼에서 작동되는 하이브리드 앱을 개발할 수 있다(그림 1).

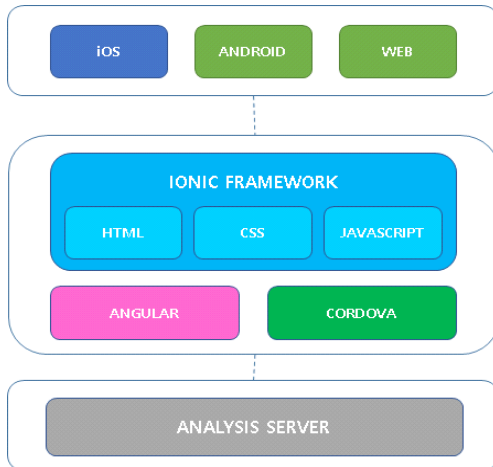


그림 1. SUDA 시스템의 하이브리드 앱 구조  
Figure 1. SUDA's hybrid application structure

### 2.2. SUDA 시스템 구성도

SUDA의 전체 시스템 구성도는 다음과 같다. 스마트폰이나 태블릿 PC에 설치된 앱에서 사용자가 동영상을 녹화하고 전송 버튼을 누르면, 녹화된 파일이 대용량 파일 전송 기능을 통해 클라우드로 전송된다. 본 시스템에서는 스마트폰, 태블릿 PC, 디지털카메라 등으로 녹화된 mp4 동영상 파일 포맷을 사용한다. mp4 확장자의 동영상 파일이 업로드 되면 동영상에서 오디오 데이터만 추출되는데, 추출된 오디오 데이터는 모노 채널 16,000샘플링레이트의 wav 확장자 파일 포맷으로 저장된다.

영상이 클라우드로 전송되면, 해당 영상에서 발화구간을 감지하여 기준에 따라 발화를 구분해준다. 발화(utterance)는 말끝을 내리거나 쉬거나, 혹은 다른 생각이 난 표시로 숨을 쉬는 것과 같은 행동에 의해서 서로 분리되는 문장이나 그보다 더 짧은

말을 의미하는데(Owens, 2016), 본 연구에서는 이를 바탕으로 끝이 분명하게 내려가거나 올라가는, 즉 억양 변화가 있는 경우에 발화를 구분하였다(그림 2).

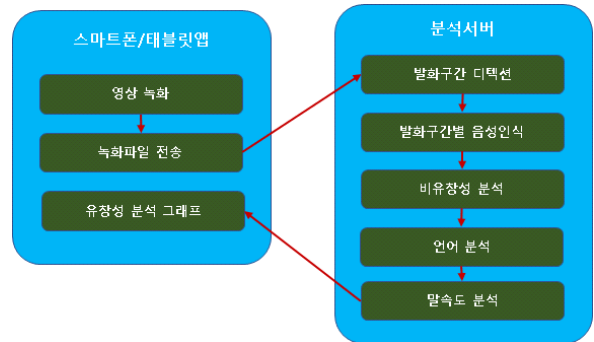


그림 2. SUDA 시스템 구성도  
Figure 2. SUDA's system configuration diagram

구분된 각 발화는 음성인식을 통해 전사되고, 비유창성, 언어, 말속도 면에서 자동적으로 분석된다. 비유창성 분석에서는 화자의 발화, 비유창성이 자동적으로 분석된다. 언어 분석에서는 각 발화의 형태소 수, 음절 수, 단어 수를, 말속도 분석의 경우, 각 발화의 말속도와 반응시간이 자동적으로 계산된다.

### 2.3. SUDA 앱 기능

SUDA는 ‘회원가입 및 로그인’, ‘동영상 녹화, 전송 및 업로드’, ‘발화 구분 및 전사’, ‘발화 비유창성 분석’, ‘발화 언어 분석’, ‘발화 말속도 분석’, ‘분석 결과 통계’ 기능을 제공한다. 또한 사용 목적에 따라 일반인용, 전문가용, 관리자용으로 구분된다. SUDA 앱의 세부적인 기능 설명은 다음과 같다.

#### 2.3.1. 회원가입 및 로그인 기능

SUDA는 회원가입 및 로그인 기능을 제공한다. 앱을 사용하기 위해서는 이메일 형식으로 된 아이디, 7자리 이상의 비밀번호, 사용자 이름으로 회원가입을 하고, 로그인을 해야 한다. 사용자가 비밀번호를 잊어버린 경우, 비밀번호 찾기 기능을 통해 찾을 수 있다. 그리고 스마트폰, 태블릿 PC뿐만 아니라 웹페이지를 통해서도 회원가입 및 로그인하여 앱을 사용할 수 있다(그림 3).



그림 3. 회원가입 및 로그인 화면  
Figure 3. Sign up and login screen

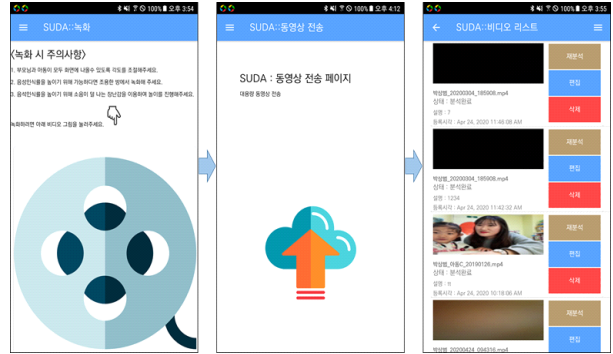


그림 5. 동영상 녹화, 전송, 업로드 화면  
Figure 5. Video recording, transmission, upload screen

### 2.3.2. 홈 화면 및 메뉴

홈 화면에서는 SUDA 사용설명서를 다운받아 앱 사용법에 대해 살펴볼 수 있다. 그리고 왼쪽 상단에 있는 버튼을 클릭하면 메뉴가 표시된다. 메뉴는 사용자에게 따라 차별화되어 구성된다.

일반인용에서는 동영상 녹화 및 업로드, 영상 삭제 요청, 분석된 결과를 그래프로 볼 수 있는 통계 기능이 제공된다. 전문가용에서는 일반인용 기능에 비유창성, 언어, 말속도 분석 기능이 추가되어 제공된다. 관리자용에서는 일반인용, 전문가용에서 업로드된 모든 영상을 볼 수 있고, 분석과 모든 영상 삭제 기능이 추가되어 제공된다(그림 4).



그림 4. 홈 화면 및 메뉴  
Figure 4. Home and menu screen

### 2.3.3. 동영상 녹화, 전송 및 업로드 기능

동영상 녹화 화면에서는 사용자들이 녹화 시 주의해야 할 사항에 대해 안내하고, 버튼을 클릭하여 영상을 녹화할 수 있도록 한다. 업로드 기능에서는 방금 녹화한 영상뿐만 아니라 이전에 녹화했었던 영상도 업로드하기를 원할 때 전송 및 업로드할 수 있다. 동영상이 업로드 되면 삭제 요청 기능이 활성화되며, 관리자가 이를 확인한 후, 해당 동영상을 삭제한다(그림 5).

### 2.3.4. 비유창성 분석 기능

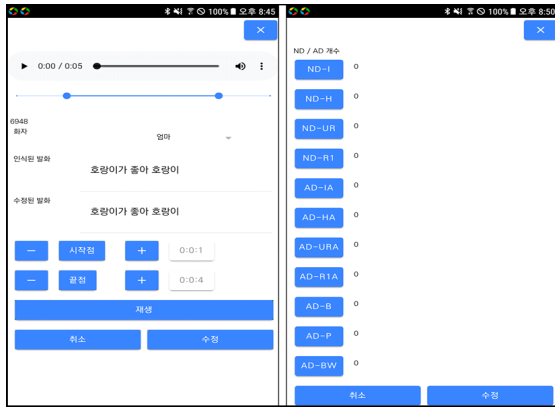
비유창성 분석 화면에서는 발화, 화자, 아동의 비유창성을 분석할 수 있고, 각 발화에 해당하는 영상을 시청할 수 있다(그림 6).



그림 6. 비유창성 분석 및 통계 화면  
Figure 6. Disfluency analysis and statistics screen

각 발화를 클릭하면, 발화, 화자, 영상 편집 화면이 나타난다. 앱이 자동으로 구분 및 전사해준 발화를 ‘인식된 발화’를 통해 확인할 수 있고, 전문가가 ‘수정된 발화’를 통해 직접 수정할 수 있다. 또한 화자와 영상도 전문가가 직접 수정할 수 있고, 비유창성 화면에서는 수정된 발화가 나타나게 된다.

각 발화의 ND와 AD를 클릭하면, 각 발화에서 나타난 ND와 AD의 유형별 수치를 입력할 수 있는 화면이 나타난다. Normal disfluency(ND)는 정상적 비유창성을, abnormal disfluency(AD)는 비정상적인 비유창성을 의미하며, 국내에서 개발된 유창성 장애 검사도구인 파라다이스 유창성 검사-II(Sim et al., 2010)의 비유창성 분류 방식에 따라 각 유형별 수치를 입력한다. 유형별 수치를 입력하면, 비유창성 화면에서 각 발화에서 나타난 ND와 AD의 총합이 자동적으로 계산되어진다(그림 7).



ND, normal disfluency; AD, abnormal disfluency.

그림 7. 발화, 화자, 영상 편집 화면 및 ND/AD 입력 화면

Figure 7. Utterance, speaker, video editing, and ND / AD input screen

ND = 각 발화의 ND 수 총합

AD = 각 발화의 AD 수 총합

통계 화면에서는 일간, 주간, 월간, 연간 기준에 따라 ND, AD 수치가 어떻게 변화하는지 그래프를 제시한다.

### 2.3.5. 언어분석 기능

언어분석 화면에서는 각 발화의 형태소 수, 음절 수, 단어 수를 분석할 수 있다. 세 가지 항목은 자동적으로 분석된다(그림 8).

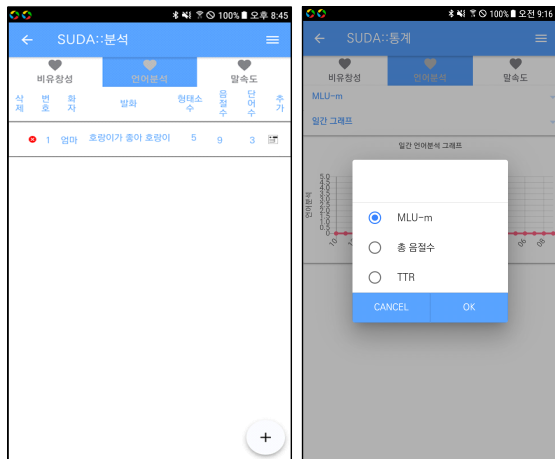


그림 8. 언어분석 및 통계 화면

Figure 8. Language analysis and statistics screen

각 발화의 '형태소 수'를 클릭하면 형태소 편집 화면이, '음절 수'를 클릭하면 음절 편집 화면이, '단어 수'를 클릭하면 단어 편집 화면이 나타난다. 각 언어단위는 '/' 기호를 통해 구분되어진다. 앱이 자동적으로 계산해준 수치를 확인하였을 때 잘못 계산된 경우, 전문가가 '수정된 형태소', '수정된 음절', '수정된 단어'를 통해 직접 수정할 수 있고, 언어분석 화면에서는 수정된 각 분석 수치가 나타나게 된다(그림 9).

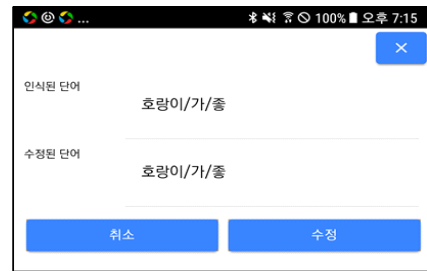


그림 9. 형태소, 음절, 단어 편집 화면

Figure 9. Morpheme, syllable, and word editing screen

각 발화에서 형태소, 음절, 단어분석이 완료되면 형태소의 경우, 평균 형태소 길이(Mean Length of Utterance in morphemes, 이하 MLU-m)로, 음절의 경우 총 음절 수로, 단어의 경우, 어휘 다양도(Type-Token Ratio, 이하 TTR)를 자동적으로 계산한다.

MLU-m은 각 발화의 형태소의 수를 총 발화의 수로 나누어 평균을 구한 것으로 지금까지 많이 사용되어 오고 있는 평균 발화 길이 척도이다(Kim, 2014). SUDA에서는 구체적인 대상이나 구체적인 상태를 나타내는 실질적 의미를 지닌 형태소인 실질 형태소(full morpheme)와 형식적 의미, 즉 문법적 의미를 지닌 형태소인 형식형태소(empty morpheme)를 모두 포함하였다(Lee, 2008).

$$MLU-m = \frac{\text{각 발화의 형태소수의 합}}{\text{총 발화수}} \quad (1)$$

TTR은 아동이 사용한 총 낱말 중에서 다른 낱말의 비율이 얼마나 되는지 산출하는 것으로 아동이 얼마만큼 다양한 낱말을 사용하는가를 측정하는 가장 널리 알려진 방법이다(Kim, 2014). SUDA에서는 어휘적 의미를 가지고 의미적 정보를 전달하는 내용어와 문법적 기능을 가지고 있는 기능어를 모두 포함하였다(Chon, 2010).

$$TTR = \frac{\text{아동의 총 발화에서 나타난 다른 낱말 유형의 수(전체 NDM)}}{\text{아동의 총 발화에서 나타난 전체 낱말의 수(전체 NTW)}} \quad (2)$$

통계 화면에서는 일간, 주간, 월간, 연간 기준에 따라 MLU-m, 총 음절 수, TTR 수치가 어떻게 변화하는지 그래프를 제시한다.

### 2.3.6. 말속도 분석 기능

말속도 화면에서는 각 발화로부터 말속도와 반응시간을 자동으로 계산해서 표시한다. 말속도는 각 발화의 음절수의 합을 전체 발화 소요시간으로 나눈 값으로 초당음절수(syllables per second, SPS)를 계산하였다(Kloth et al., 1998). 이때 2초 이상의 비정상적인 쉽의 지속시간은 제외하였다(Hall et al., 1999). 말속도는 값이 클수록 말속도가 빠르다는 것을 알 수 있다. 반응시간은 아동의 발화가 끝난 시간부터, 부모의 발화가 시작된 시간 간격이다. 반응 시간이 짧을수록 부모의 말에 아동이 빠르게 반응했음을 알 수 있다.

통계 화면에서는 일간, 주간, 월간, 연간 기준에 따라 말속도와 반응시간 수치가 어떻게 변화하는지 그래프를 제시한다(그림 10).

$$\text{말속도} = \frac{\text{각 발화음절수의 합}}{\text{총 발화소요시간}-2\text{초 이상의 비정상적 쉽(sec)}} \quad (3)$$

$$\text{반응시간} = \text{아동의 발화가 끝난 시간부터 부모의 발화가 시작된 간격의 시간길이 합(초)} \quad (4)$$



그림 10. 말속도 화면  
Figure 10. Articulation rate screen

## 3. 결론

본 연구에서는 일상생활에서 부모와 아동의 상호작용 시의 모습을 동영상으로 촬영하여 데이터를 추적하고, 이를 통해 아동의 발화와 부모의 발화를 자동적으로 분석하여 시간의 흐름에 따라 다양한 언어요소들의 변화를 파악할 수 있는 SUDA 하

이브리드 앱의 개발 절차를 제시하였다.

본 연구의 결과 및 의의를 요약하면 다음과 같다.

첫째, SUDA는 사용자가 동영상을 원하는 일시, 장소에서 촬영하여 업로드하면 자동으로 시스템 서버로 전달되기 때문에 시간 및 공간적인 제약, 비용 면에서 겪는 어려움을 보완하고, 대상자의 일상을 반영할 수 있기 때문에 실제적이고 신뢰로운 발화 데이터를 수집할 수 있다.

둘째, SUDA는 아동 및 부모의 발화를 자동으로 전사하고, 각 발화의 비유창성, MLU-m, 음절 수, TTR, 말속도를 반자동적으로 분석해주어 언어치료사와 연구자의 편의성을 높였다. SUDA가 자동적으로 분석 결과를 제공하면, 언어치료사가 이를 검토하고, 보완하여 의사소통장애 진단과 중재 효과 모니터링 시 활용할 수 있다.

셋째, 다양한 요인들(비유창성, 언어발달, 말속도 등)의 특성을 객관적으로 살펴볼 수 있으므로 유창성장애뿐만 아니라 언어치료사들이 다양한 의사소통장애 진단 및 중재 시 유용하게 사용할 수 있을 것이다.

넷째, SUDA는 사용자 모드(일반인, 전문가, 관리자)에 따라 다른 기능을 간편하고 쉽게 사용할 수 있다. 간편한 영상녹화 및 전송, 시각화된 분석 결과를 제시해준다. 이를 통해 일반인들이 발화에 대한 정보를 쉽고 다양하게 제공받을 수 있다.

다섯째, 다수의 사용자가 동시에 동영상을 전송해도 안정적으로 데이터를 수신할 수 있는 멀티스레드(multi thread) 기반의 서버가 구축되었으며, 동영상데이터 관리 서버와 시스템 운용 서버를 분리하여 데이터 관리에 안정성을 확보하였다.

마지막으로 SUDA 앱을 활용하여 말더듬아동의 중단데이터를 수집하고 말더듬 회복 예측 요인들을 찾는 한국형 아동말더듬의 빅데이터 구축에 기여할 수 있다.

본 연구의 결과를 바탕으로 추후에는 다음과 같은 문제점들을 보완하고자 한다.

첫째, Mahalanobis Distance, ZCR 등 기존에 많이 이용된 방식과 머신러닝 방법을 결합하여 디텍션 정확도를 높일 예정이다.

둘째, 하이브리드 앱 상에서 동영상 편집이 가능하도록 개발을 진행할 예정이다.

셋째, 데이터 축적을 통해 음성인식률을 향상시켜 발화 구간과 전사 내용의 정확도를 향상시킬 예정이다.

넷째, 인터넷 속도에 따른 동영상 업로드 속도를 확인하여 다중 사용자가 동시 이용 시에도 안정적으로 이용할 수 있도록 서버시스템을 보완할 예정이다.

## References

Byun, J. W., Lee, E. J., & Sim, H. S. (2004). A study on the characteristics of speech disfluencies in early childhood stuttering. *Communication Sciences & Disorders*, 9(1), 1-14.

Chon, H. (2010). Stuttering-like disfluencies in phonological words in preschool children who stutter. *Communication Sciences & Disorders*, 15(3), 422-432.

- Costanza-Smith, A. (2010). The clinical utility of language samples. *Perspectives on Language Learning and Education*, 17(1), 9-15.
- Garbarino, J., Ratner, N. B., & MacWhinney, B. (2020). Use of computerized language analysis to assess child language. *Language, Speech, and Hearing Services in Schools*, 51(2), 504-506.
- Guitar, B. (1988). *Stuttering: An integrated approach to its nature and treatment*. Philadelphia, PA: Lippincott Williams & Wilkins.
- Go, G., Kwon, Y., Noh, H., Song, K., & Choi, Y. (2017). The effects of speech-language therapy using a smartphone application to improve language ability in patients with aphasia. *Journal of Speech-Language & Hearing Disorders*, 26(4), 1-10.
- Ha, S., Seol, A., So, J., & Pae, S. (2016). Speech and language development patterns of Korean two-year-old children from analysis of spontaneous utterances. *Communication Sciences & Disorders*, 21(1), 47-59.
- Hall, K. D., Amir, O., & Yairi, E. (1999). A longitudinal investigation of speaking rate in preschool children who stutter. *Journal of Speech, Language, and Hearing Research*, 42(1), 1367-1377.
- Hugh-Jones, S., & Smith, P. (1999). Self-reports of short- and long-term effects of bullying on children who stammer. *British Journal of Educational Psychology*, 69(2), 141-158.
- Im, H., Song, J. Y., Cho, Y. K., Kim, Y. J., Kim, H. K., & Kang, Y. J. (2013). The use of smartphone applications in stroke rehabilitation in Korea. *Brain & NeuroRehabilitation*, 6(1), 33-40.
- Jang, Y. J., & Kim, C., W. (2010). The evolution of smartphone market & the effect by android. *Communications of the Korean Institute of Information Scientists and Engineers*, 28(5), 48-56.
- Jeong, P. Y. (2014). Speech-language therapy using apps for a student with cerebral palsy. *AAC Research & Practice*, 2(1), 119-126.
- Kloth, S. A. M., Janssen, P., Kraaimaat, F. W., & Brutton, G. J. (1998). Child and mother variables in the development of stuttering among high-risk children: A longitudinal study. *Journal of Fluency Disorders*, 24, 253-265.
- Kim, J., Lee, E., & Hwang, M. (2008). The relationships between stuttering severity, Perception of stuttering severity, and social anxiety of adults who stutter. *Communication Sciences & Disorders*, 13, 677-690.
- Kim, S. J., & Shin, J. Y. (2007). *Articulatory and phonological disorders*. Seoul, Korea: Sigma Press.
- Kim, Y. T. (2014). *Assessment and treatment of language disorders in children* (2nd ed.). Seoul: Hakjisa.
- Lee, B. W. (2015). *Korean grammar for speech therapists*. Seoul: Hakjisa.
- Lee, E. J. (2005). *The study of temperament of preschool children who stutter, temperament and parenting behavior of the mothers* (Doctoral dissertation). Ewha Womans University, Seoul, Korea.
- 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.
- Lee, H. R., & Lee, S. (2014). Parent survey on the effects of digital instrument on the language development in Korean children. *Communication Sciences & Disorders*, 19(2), 178-190.
- Lee, S. B., Park, H. Y., & Sim, H. S. (2019). Research trends on the longitudinal studies of childhood stuttering. *Communication Sciences & Disorders*, 24(2), 490-506.
- Lee, Y., Kim, Y. T., & Park, E. H. (2016). Analysis on current practice and needs based survey of parents and speech-language pathologists for reading intervention with smart device applications. *Communication Sciences & Disorders*, 21(3), 502-514.
- Lee, Y., & Lee, S. (2017). Effect of smartphone applications for aural habilitation in cochlear implanted children: Improvement for auditory memory skills. *Audiology and Speech Research*, 13(2), 176-188.
- Lee, Y., Lee, S., & Sung, M. (2015). Analysis of mobile application trends for speech and language therapy of children with disabilities in Korea. *Korean Society of Speech Sciences*, 7(3), 153-163.
- Lim, J. H., & Park, E. H. (2012). Development and research trends of application as the smart education media for ASD. *Journal of the Korean Association for Persons with Autism*, 12(1), 93-117.
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Erlbaum.
- Manning, W. H. (2001). *Clinical decision making in fluency disorders* (2nd ed.). San Diego, CA: Singular.
- Miller, J. F., & Iglesias, A. (2015). Systematic analysis of language transcripts [Computer software]. Middleton, WI: Salt Software.
- Owens, R. E. Jr. (2016). *Language development: An introduction* (9th ed.). Harlow, UK: Pearson.
- Pae, S., Kim, K. S., Sung, K. H., & Sung, J. A. (1998). Computer and language assessment: Korean computerized language analysis 1.0. *Communication Sciences & Disorders*, 3(1), 123-138.
- Paul, R., & Norbury, C. (2012). *Language disorders from infancy through adolescence* (4th ed.). St. Louis, MO: Mosby.
- Park, M. H. (2012). Development of a web-based fluency evaluation system for generalization programs. *Journal of Speech-Language & Hearing Disorders*, 21(3), 241-257.
- Pavelko, S. L., & Owens, R. E. (2017). Sampling utterance and grammatical analysis revised (SUGAR): New normative values for language sample analysis measures. *Language, Speech, and Hearing Services in Schools*, 48(3), 197-215.
- Peters, T. J., & Guitar, B. (1991). *Stuttering: An integrated approach to its nature and treatment*. Baltimore, MD: Williams & Wilkins.
- Pezold, M. J., Imgrund, C. M., & Storkel, H. L. (2020). Using

computer programs for language sample analysis. *Language Speech, and Hearing Services in Schools*, 51(1), 103-114.

Shin, M. S., & Park, M. H. (2014). A study of the web-based evaluation system model of adults who stutter. *Journal of Speech-Language & Hearing Disorders*, 23(3), 1-11.

Sim, H. S., Shin, M. J., & Lee, E. J. (2010). *Paradise-fluency assessment II*. Seoul, Korea: Paradise Welfare Foundation.

Smith, A., & Kelly, E. (1997). Stuttering: A dynamic, multifactorial model. In R. F. Curlee & G. M. Siegel (Eds.), *Nature and treatment of stuttering: New directions* (pp. 204-207). Boston, MA: Allyn and Bacon.

Van Riper, C. (1982). *The nature of stuttering* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Yairi, E., & Ambrose, N. G. (2005). *Early childhood stuttering for clinicians by clinicians*. Austin, TX: Pro-ED.

Yang, S. Y., & Kang, E. J. (2011). Development of serious mobile game improving social abilities for children with developmental disorders. *Journal of Children's Media & Education*, 10(2), 23-43.

Yim, D., Kim, S. Y., Park, W., Cheon, S., & Lee, Y. J. (2014). Analysis on needs based survey of parents and speech-language pathologists for smartphone programs. *Communication Sciences & Disorders*, 19(4), 486-500.

• 이수복 (Soo-Bok Lee)

우송대학교 언어치료청각재활학과 교수  
대전광역시 동구 동대전로 171  
Tel: 042-630-9222  
Email: sblee@wsu.ac.kr  
관심분야: 유창성장애, 중단연구, IT융합

• 곽효정 (Hyo-Jung Kwak)

이화여자대학교 언어병리학과 석사과정  
서울특별시 서대문구 이화여대길 52  
Tel: 02-3277-2120  
Email: yd61581@naver.com  
관심분야: 의사소통장애

• 윤재민 (Jae-Min Yun)

보리 주식회사 연구소장  
경기도 안산시 상록구 송호1길 17  
Tel: 031-409-0938  
Email: bory@bory.io  
관심분야: 음성대화, 언어공학, 지식공학

• 신동춘 (Dong-Chun Shin)

보리 주식회사 주임연구원

경기도 안산시 안산천동로1길 9

Tel: 031-409-0938

Email: dcshin@bory.io

관심분야: 음성인식, 자연어처리, 콘텐츠제작

• 심현섭 (Hyun-Sub Sim) 교신저자

이화여자대학교 언어병리학과 교수

서울특별시 서대문구 이화여대길 52

Tel: 02-3277-2120

Email: simhs@ewha.ac.kr

관심분야: 유창성장애, 음성장애



## 동영상 기반 자동 발화 심층 분석(SUDA) 어플리케이션 개발\*

이 수 북<sup>1</sup> · 곽 효 정<sup>2</sup> · 윤 재 민<sup>3</sup> · 신 동 춘<sup>3</sup> · 심 현 섭<sup>2</sup>

<sup>1</sup>우송대학교 언어치료청각재활학과, <sup>2</sup>이화여자대학교 언어병리학과, <sup>3</sup>보리(주)

### 국문초록

본 연구는 동영상을 기반으로 일상생활에서 녹화한 아동 및 성인의 발화를 자동으로 분석해주는 SUDA(smart utterance deep analyser) 하이브리드 앱 개발에 관한 것이다. 특히, 아동과 부모가 원하는 시간 및 장소에서 상호작용하는 장면을 촬영하여 업로드할 수 있고 시간의 흐름에 따라 데이터를 계속 축적하여 이를 관찰하고 분석할 수 있도록 도울 수 있다. SUDA는 안드로이드폰, 아이폰, 태블릿 PC 기반에서 구동되며, 대용량의 동영상을 녹화 및 업로드할 수 있고, 사용자의 목적(일반인, 전문가, 관리자)에 따라 차별화된 기능을 제공할 수 있다. 전문가 모드에서는 자동화된 시스템과 협업하여 대상자의 발화를 말 · 언어적인 측면(비유창성, 형태소수, 음절수, 단어수, 말속도, 반응시간 등)에서 세부적으로 분석할 수 있다. 즉, SUDA 시스템이 대상자의 발화를 반자동으로 전사 및 분석하면, 언어치료사가 이를 검토하고, 보완하여 의사소통장애 진단과 중재 시 활용할 수 있다. 일반인(부모)의 경우, 전문가가 분석한 결과를 그래프 형태로 제공 받아 모니터링 할 수 있고, 관리자는 발화 분석, 영상삭제 등 전체 시스템을 관리할 수 있다. 본 시스템은 발화 분석의 반자동화로 치료사와 연구자의 부담을 줄여주고, 부모가 자녀의 발화를 기반으로 하여 말 · 언어발달에 대한 정보를 쉽고 다양하게 제공 받을 수 있다는 점에서 임상적 의의가 있다. 또한, 한국형 말더듬아동 진단 및 중재에 적용할 수 있는 중단데이터를 구축하고, 말더듬 회복 예측 요인들을 찾는 기초 자료로 활용하고자 한다.

**핵심어:** 말더듬, 중단연구, 동영상 기반 발화 분석, 하이브리드 앱, 언어치료

### 참고문헌

고경백, 권유미, 노희연, 송기범, 최양규(2017). 어플리케이션을 활용한 언어중재가 실어증자의 언어능력 향상에 미치는 영향. *언어치료연구*, 26(4), 1-10.

김수진, 신지영(2007). *조음음운장애*. 서울: 시그마프레스.

김영태(2014). *아동언어장애의 진단 및 치료(제2판)*. 서울: 학지사.

김주연, 이은주, 황민아(2008). 말더듬 성인의 말더듬 중증도와 말더듬 지각 정도 및 사회불안 간의 상관연구. *언어청각장애연구*, 13(4), 677-690.

박만희(2012). 유창성 일반화 프로그램을 위한 웹기반 유창성 평가시스템 개발. *언어치료연구*, 21(3), 241-257.

배소영, 김광선, 성경훈, 성진아(1998). 컴퓨터와 언어능력 진단 평가. *언어청각장애연구*, 3(1), 123-138.

변재원, 이은주, 심현섭(2004). 초기 말더듬아동의 비유창성 특성 연구. *언어청각장애연구*, 9(1), 1-14.

신명선, 박만희(2014). Web 기반 말더듬 성인 평가 시스템 모형

연구. *언어치료연구*, 23(3), 1-11.

양심영, 강은진(2011). 발달장애 아동의 사회성 증진을 위한 기능성 모바일 게임 개발. *어린이미디어연구*, 10(2), 23-43.

이봉원(2015). *언어치료사를 위한 한국어 문법*. 서울: 학지사.

이수북, 박혜연, 심현섭(2019). 말더듬아동 중단연구에 관한 문헌 연구. *언어청각장애연구*, 24(2), 490-506.

이영미, 김영태, 박은혜(2016). 읽기 중재 어플리케이션 개발을 위한 장애 아동의 읽기 중재 실태 및 부모, 언어치료사의 요구 분석. *언어청각장애연구*, 21(3), 502-514.

이영미, 이수북, 성민경(2015). 국내 장애 아동을 위한 언어치료용 모바일 어플리케이션 현황 분석. *말소리와 음성과학*, 7(3), 153-163.

이영미, 이슬기(2017). 스마트폰 어플리케이션을 이용한 중재가 인공와우 아동의 청각기억에 미치는 영향. *한국청각언어재활학회*, 13(2), 176-188.

이은주(2005). *취학전 말더듬아동의 기질과 어머니의 기질 및 양육행동 연구*. 이화여자대학교 박사학위논문.

\* 이 논문은 2018년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2018S1A5A2A03036976).

- 이하나, 박준희, 유재연(2019). 스마트폰기반 음성치료 프로그램 개발연구. *말소리와 음성과학*, 11(1), 51-61.
- 이희란, 이승복(2014). 아동의 디지털 기기 사용과 언어발달에 관한 부모의 인식 조사. *언어청각장애연구*, 19(2), 178-190.
- 임동선, 김신영, 박원정, 천성혜, 이여진(2014). 스마트폰 프로그램 개발을 위한 의사소통장애 아동 부모 및 전문가 요구 분석. *언어청각장애연구*, 19(4), 486-500.
- 임장현, 박은혜(2012). ASD인을 위한 스마트 교육 미디어로서의 앱 개발 및 연구현황 분석. *자폐성장장애연구*, 12(1), 93-117.
- 임형준, 송제영, 조윤경, 김연준, 김현정, 강윤주(2013). 국내 뇌졸중 재활 분야에서 스마트폰 어플리케이션의 적용 및 이용 현황. *대한뇌신경재활의학회지*, 6(1), 33-40.
- 장윤정, 김철우(2010). 스마트폰 시장의 진화와 안드로이드의 영향. *정보과학회지*, 28(5), 48-56.
- 정필연(2014). 학령기 뇌성마비아동의 언어치료를 위한 앱 적용 사례. *보완대체의사소통연구*, 2(1), 119-126.
- 전희정(2010). 음운적 단어에 나타난 학령전기 말더듬 아동의 비유창성 발생 위치. *언어청각장애연구*, 15(3), 422-432.
- 하승희, 설아영, 소정민, 배소영(2016). 자발화 분석을 통한 만 2세 한국아동의 말·언어발달 특성. *언어청각장애연구*, 21(1), 47-59.