

마비말 장애인을 위한 발음 진단 애플리케이션

김시영⁰, 김민제, 장현아, 차요셉
한국산업기술대학교 컴퓨터공학과

{krbvpower⁰, alswp0210, jang9528, joseph704}@kpu.ac.kr

Pronunciation diagnosis application for the dysarthria

Si-Young Kim⁰, Min-Je Kim, Hyun-Ah Jang, Joseph Cha
Dept. of Computer Engineering, Korea Polytechnic University

요 약

마비말 장애인은 중추 혹은 말초 신경계의 이상으로 근육 부에 문제가 생겨 발화에 어려움을 겪는 말운 동장애로서 현재 마비말 장애인을 위한 발음 진단 애플리케이션은 존재하지 않는다. 따라서 본 연구에서 는 마비말 장애인의 발음 교정을 돕는 애플리케이션을 개발하고자 하였으며, 마비말 장애인이 언어 치료 를 받는 과정에서 언어 치료사의 보조도구로서 역할을 하고 시간, 공간적 제약이 없는 언어 치료 학습이 가능케 하여 언어 치료 효과를 향상시킨다. 마비말 장애인의 발음 진단 학습을 위해 STT(Speech To Text) 기술을 활용한다. 음성을 텍스트로 변환 후 단어를 음소별로 분할하는 알고리즘을 적용하여 마비말 장애인의 음성과 정확한 발음을 비교하고 정확도를 산출한다. 발음 정확도는 시각적인 도표로 제공되어 개선정도를 보기 쉽게 전달한다. 본 연구를 통해 현직 언어 치료사에게 새로운 언어 치료 프로세스를 제 시하고 발음 정확도 개선효과와 마비말 장애인의 언어능력 향상, 사회적응을 도우며 재활의지를 북돋아주 기를 기대한다.

1. 서 론

마비말 장애는 신경계 이상으로 조음 능력의 결함이 생 겨 발화 문제가 발생하는 언어 장애의 일종이다. 낮은 발화 명료도로 인해 의사소통에 어려움과 사회 적응 문제를 겪 는다. 언어 치료의 현 실정상 대부분의 마비말 장애인들이 언어 치료의 혜택을 받기 어렵고 시간, 공간적 제약을 받는다.

본 연구에서는 애플리케이션을 통해 언어치료를 받 을 보조하여 마비말 장애인을 치료하는 시간 이외에 시간, 공 간적 제약 없이 반복 학습을 가능하게 하여 언어 치료의 효율을 증가시킨다. 이로써 언어치료사와 다시 만나기 전까지 사용자의 자기 주도적인 언어 반복 학습이 가능 하게 된다.

마비말 장애인에게 맞춤 데이터를 제공하기 위하여 현 직 언어치료사에게 조언을 구해 언어 치료 프로세스를 설계하였다. 프로세스는 각 4단계로 구성, 모음의 정확도 를 확인하기 위한 무의미 음절, 이중모음이 들어간 무의 미 음절, 의미 있는 2음절 단어, 최소 대립쌍 단어로 구 성하였다. 현재 100개 이상의 단어들을 사용하여 언어 능력의 향상을 기대하고자 하였다. 최소 대립쌍 단어란 최소한의 변화를 주어 단어의 뜻과 말소리를 구별하기 위해 사용되는 단어로 발음 교정에 중요한 역할을 한다.

본 연구는 마비말 장애인을 위한 발음 진단 애플리케 이션을 설계한다. swift 언어를 기반으로 한 iOS 개발환경에서 개발하였고 발음 전문가의 발음 화면을 제공하며 스트리밍 화면과 마비말 장애인의 발음 연습 화면을 제 공, 자신의 입 모양과 전문가의 입 모양을 비교하며 발 음 교정을 가능하게 하였다. 실시간 스트리밍과 동시에,

마비말 장애인의 음성은 음성을 글자로 바꾸는 STT 기 술을 활용하여 텍스트로 변환시킨다. 텍스트로 변환된 마비말 장애인의 음성은 실시간으로 서버로 전달된다. 서버로 전달된 데이터는 서버 내부에서 단어를 음소별로 분할하는 알고리즘을 거쳐 정확한 단어의 비교를 통해 정확도를 산출하게 된다.

이러한 데이터들은 데이터베이스에 축적되어 애플리케 이션의 사용자에게 히스토리를 한눈에 보기 쉽게 제공된 다.

본 논문은 마비말 장애인을 위한 발음 진단 애플리케 이션의 특징에 관해 기술한다. 2장에서는 애플리케이션 내 언어 치료 프로세스 관련 연구를 소개하며 3장에서는 제안하는 프레임워크의 특징에 대한 구체적인 설계 및 구현 방안을 설명하고 4장에서 결론을 맺는다.

2. 관련 연구

2.1 언어 치료 프로세스

2.1.1 무의미 음절

아	이	우	어
오	야	으	예
우	아	이	아

<표 1> 무의미 음절 단어 프로세스

무의미 음절은 의미보다는 발음으로 정확도를 파악하 기 위해 사용하며 임의로 조음을 조작하여 구강 움직임 의 정확성을 보기 위한 의미로 사용한다. 따라서 구강의

움직임에 중점을 두어 혀의 전후 위치를 기준으로 서로 상대적 거리가 먼 모음을 짝을 지어 순차적으로 읽게 한다.

2.1.2 이중모음이 들어간 무의미 음절

예	애	여	유후
요	야	위	웨
워	와	의	
예	워	유후	야

<표 2> 이중모음 무의미 음절 단어 프로세스

이중모음이 들어간 무의미 음절은 혀의 전후 위치를 기준으로 상대적 거리가 먼 모음을 활용하여 좀 더 발음이 어려운 이중 모음을 중점적으로 교육하기 위해 활용한다. 이중 모음은 두 모음을 한 음절로 발음해야 하는데 마비말 장애인들은 미세한 근육 사용에 어려움을 겪으므로 이중 모음을 활용한다.

2.1.3 의미 있는 2음절 단어

사과	오이	우유
의자	소설	문학
음악	신호	그릇
그릇	숙제	부엌

<표 3> 의미 있는 2음절 단어 프로세스

의미 있는 2음절 단어는 발음이 어려운 동시에 흥미도가 낮은 마비말 장애인들에게 효과적인 치료 방식으로서 자신의 생활과 관련된 단어를 주로 활용한다. 따라 실생활에 자주 쓰이는 쉬운 단어들을 활용한다.

2.1.4 최소 대립쌍 단어

짐	김	물	불
술	슬	발	날
밖	밥	밤	담
논	돈	쌈	햄
시다	치다	바다	파다
거울	너울	정지	정리
포도	보도	가루	자루
구조	수조	마루	머루

쉬리	소리	사람	사랑
구실	구슬	장미	장비
보수	고수	상자	상가
배구	배추	찬물	단물
보리	꼬리	짜다	사다
가다	타다	배구	배추
머리	거리	포수	조수

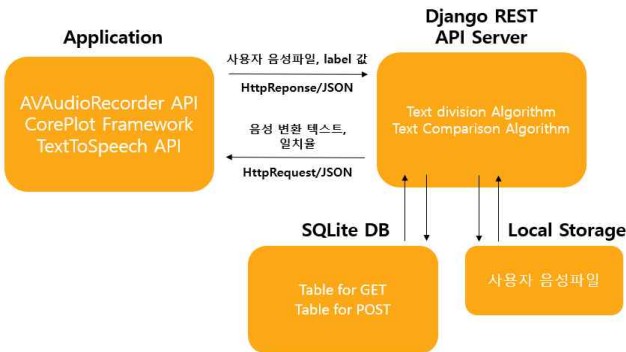
<표 4> 최소 대립쌍 단어 프로세스

최소 대립쌍 단어는 최소한의 변화를 주어 단어의 뜻과 발소리를 구별하기 위하여 언어 치료가 발음 연습에 자주 사용되는 방식이다. 모음의 변화를 주거나 파열음, 마찰음 등의 기준으로 제시되며 단어 두 개를 동시에 발음하여 근육의 활용도를 높이기 위해 활용한다.

3. 발음 진단 애플리케이션 설계

3.1 세부 설계 및 구현

3.1.1 시스템 구성



<그림 1> 시스템 구성도

애플리케이션은 전문가 화면 스트리밍, STT, 사용자 화면 실시간 출력 기능으로 이루어지며 각 정보는 서버로 전달, DB에 저장되며 서버에서 분석 및 결과물 내용을 도표를 통해 다시 애플리케이션으로 전달, 화면으로 전달시킨다.

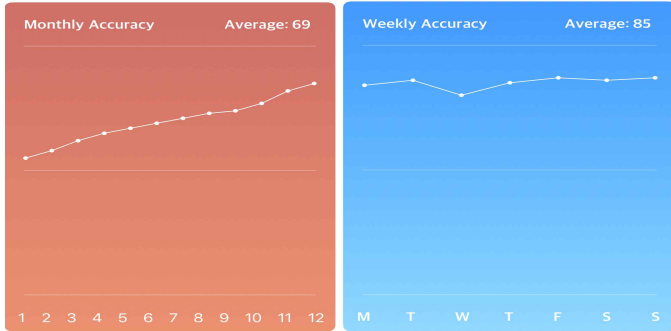
3.1.2 발음교정 화면



<그림 2> 애플리케이션 발음 교정 화면

좌측 전문가의 입 모양을 클릭하면 플래시 형태로 동영상 재생된다. 동영상과 우측 사용자 입 모양을 비교하면서 발음을 연습하고 녹음 버튼을 통해 음성을 녹음한다. 이후 재생 버튼을 통해 녹음된 음성을 확인할 수 있다.

3.1.3 히스토리 화면



<그림 3> 월간 정확도 내역(좌) 주간 정확도 내역(우)

3.2 모듈별 핵심 사항

3.2.1 화면 스트리밍 모듈

저장된 전문가 음성 발음 화면을 재생 시키면 각 화면을 재생, 사용자에게 전달, 자신의 입모양도 함께 출력시켜 주어 사용자는 자신의 입모양 및 전문가 입모양을 보며 발음 연습이 가능하도록 하였다.

3.2.2 STT(Speech-To-Text) 모듈

iOS에서 제공하는 Speech API 중 실시간 인지모듈을 활용했다. 사용자의 음성을 입력하면 음성데이터가 Speech API에 의해 텍스트 데이터로 변환된다.

3.2.3 녹음 및 재생 모듈

AVAudioRecorder API를 통해 사용자의 음성을 임시로 애플리케이션 로컬에 저장하고 AVPlayer를 통해 저장된 음성 파일을 재생한다.

3.2.4 데이터 전송 모듈

애플리케이션에서 STT 모듈을 통해 변환된 사용자의 텍스트 데이터는 HTTP 프로토콜을 통해 서버로 전송된다. 서버에서 처리된 텍스트 데이터도 HTTP 프로토콜을 통해 애플리케이션에 전달받는다.

3.2.5 서버

서버 내에서 애플리케이션의 데이터들을 저장하고 전송하기 위한 데이터베이스가 구축한다. 단어를 음소별로 분할하는 단어 분할 알고리즘과 사용자의 음성을 텍스트로 변환시킨 단어와 올바른 단어 데이터를 비교하는 비

교 알고리즘이 있다.

3.2.6 단어 분할 및 분석 모듈

애플리케이션으로부터 입력받은 label 단어 데이터와 STT 모듈로부터 반환받은 단어 데이터를 처리하는 모듈이다. 두 단어 데이터를 각각 초성, 중성, 종성으로 분할하여 비교 가능한 값으로 바꾸는 기능을 제공한다. 분할된 데이터를 비교하여 두 데이터의 일치율과 다른 음소 데이터에 대한 정보를 제공한다. 모듈로부터 반환된 데이터는 애플리케이션에서 일치율은 그래프, 도표를 통해, 다른 음소 데이터 정보는 컬러를 다르게 하여 시각적으로 나타내어질 수 있다.

3.2.7 Database

get_ddobaki		
FIELDTYPE	NAME	OPTION
IntegerField	<i>id</i>	<i>Auto-Created</i>
CharField	<i>user</i>	
CharField	<i>label</i>	
CharField	<i>Stt</i>	
DateFieldTime	<i>current_time</i>	
CharField	<i>divided_stt</i>	
CharField	<i>color</i>	<i>default=TRUE</i>

<표 5> Database Table

Database는 Django에서 기본으로 제공하는 SQLite3를 사용하였고 테이블을 구축하여 데이터들을 효율적으로 저장하도록 하였다.

4. 결 론

본 논문에서 제안한 발음 진단 애플리케이션은 사용자의 발음을 텍스트화 하여 음소 단위로 분할하고 올바른 발음 텍스트와 비교하여 정확도를 도출한다. 또한 N 정확도의 개선 정도를 도표를 통해 제시하여 언어 치료사의 새로운 발음 개선 프로세스를 제안한다. 현재는 마비말 장애인을 위한 애플리케이션이 부족한 실정이다. 본 연구의 애플리케이션을 통해 마비말 장애인의 시간, 공간적 제약 없는 언어 치료를 도우며 시각적으로 확인 가능한 서비스를 제공하여 언어 치료 시스템 구축 및 마비말 장애인의 재활에 큰 도움이 될 것으로 기대한다.

참고 문헌

- [1] 황유미, 김선희, 정민화 (2011). 마비 말 장애의 발화 명료도와 음성 인식을 비교. 한국HCI학회 학술대회, 355-358.
- [2] 어승우, 김영봉 (2006). 언어 장애인을 위한 발음 학습 시스템 설계. 한국멀티미디어학회 학술발표 논문집, 685-688.
- [3] 최여진, 심현섭 (2013) “뇌성마비로 인한 마비말 장애 성인의 최대 혀 및 입술 강도와 자음정확도 및 말명료도의 관계”. 말소리와 음성과학 제5권 제2호. 11-22
- [4] 이영미, 성지은, 심현섭 (2013). 뇌성마비로 인한 마비말장애 성인의 자음 오류 분석. 말소리와 음성과학, 5(1), 47-54.
- [5] 이옥분, 한지연, 박상희 (2010). 마비말장애 심각도에 따른 음절단위 말명료도와 모음공간. 말소리와 음성과학, 2(2), 85-92.