

# STT(Speech To Text)기술을 이용한 청각장애인 대상의 강의지원 서비스

이병도, 장승원, 양다예, 김주희, 고석주\*, 김영옥\*\*

\*경북대학교 IT대학 컴퓨터학부

\*\*SK주식회사 C&C

e-mail : tij03105@gmail.com

## Lecture Support Service using STT(Speech To Text) Technology for a Hard of Hearing People

Byeong-Do Lee, Seung-Won Jang, Da-Ye Yang, Ju-Hee Kim,  
Seok-Ju Koh\*, Yeong-Ok Kim\*\*

\*School of Computer Science and Engineering, College of IT,  
Kyung-Pook National University

\*\*SK Holdings C&C

### 요 약

본 논문은 STT(Speech To Text) 기술을 중심으로 한 웹(Web) 기반의 강의 지원 서비스 프로그램의 설계 및 구현에 관한 내용을 다루고 있다. 점점 고학력 사회가 되어 가면서 전체적인 대학교 진학률이 높아지고 있지만, 일반학생과 특수교육 대상자의 대학교 진학률은 큰 차이가 있다. 이 원인 중 하나는 장애 학생이 비 장애 학생과 같은 수준의 교육 환경을 제공받지 못하기 때문이다. 따라서 본 프로그램은 STT 기술을 이용하여 교사의 음성을 실시간으로 텍스트로 변환하여 보여줌으로써 청각장애 학생들에게 비 청각장애 학생들과 동등한 수준의 교육 환경을 제공하기 위하여 기획 및 설계되었다. 본 프로그램의 목표는 STT 기능을 중심으로 하며 추가적으로 TTS(Text To Speech), 사전, 요약, 메모 등의 기능을 제공하여 더욱 효과적으로 청각장애 학생에게 향상된 교육 환경 제공 및 교육 참여의 기회를 제공하는 것이다.

## 1. 서론

개이지만 청각장애 특수학교 개수는 그 중, 14개에 불과하다.[2] 주변에 특수학교가 없는 일부 청각장애인들은 일반 학교에 진학하게 된다. 하지만 대부분의 일반 학교는 장애가 없는 일반 학생들을 기준으로 수업을 진행하기 때문에 청각장애인 학생이 교사의 말을 이해하는 것과 수업 중 질문을 하는 것이 어렵다. 이로 인해 청각장애인 학생들이 수업에 대한 참여도가 낮아지는 등 평등한 교육을 받을 수 없는 상황이다.

통계청에서 제시한 자료에 따르면 2017년 기준 ‘향후 지원이 필요한 특수교육 서비스’로 청각장애인 2,767명 중 550명(19.9%)이 “보조공학기기 지원”이라고 답했고, 548명(19.8%)이 “학습 보조기기 지원”이라고 답했다. 결과적으로 우리나라는 아직 청각장애인 교육에 대한 적절한 지원이 부족하다는 것이다.[3]

그림 1. 고등학교 졸업생의 대학교 진학률

국가지표체계에 따르면 2018년 기준 일반학생의 대학교 진학률은 69.7%에 이른다.[1] 이에 반해 2018년 기준 고등학교 특수교육대상자의 대학교 진학률은 47%이다. 20%가 넘는 진학률 차이의 원인 중 하나는 특수교육대상자에게 일반학생과 같은 수준의 교육 환경이 제공되지 않고 있기 때문이다. 2018년 기준 전국의 특수학교 개수는 175

## 2. 관련 서비스

### 2.1 실시간 속기 지원

청각장애인이 일반 학생들과 함께 수업을 듣는 환경에서 청각장애인의 학업을 도와주기 위해 여러 대학이 실시간으로 속기 서비스(Service)를 지원하고 있지만, 속기사

들의 수와 예산에 따른 어려움을 겪고 있다.[ 4 ]

## 2.2 음성인식 프로그램

청각장애인이 의사소통을 원활하게 할 수 있도록 사람의 음성을 텍스트(Text)로 변화해 주는 프로그램(Program)이 있다. 이러한 프로그램들은 교육 환경에서 청각장애인들이 다른 사람의 도움을 받지 않고도 교육에 참여할 수 있도록 도와준다. 또한, 이런 프로그램들은 모바일(Mobile) 기기의 보급으로 인해 사용자가 쉽게 접근하여 사용할 수 있는 장점이 있다.



그림 2. 음성인식 프로그램[5]

## 3. 요구사항 정의

본 프로그램은 STT(Speech To Text)를 이용하여 교사, 교수의 음성데이터를 텍스트 데이터로 변환하여 청각 장애인 학생들에게 텍스트로 강의를 제공한다. 추가로 프로그램의 사용자인 청각장애인들이 실제로 어떤 기능을 필요로 하는지를 알아보기 위해 “한국 농인 협회 대구광역시 본부” 및 “대구 청각, 언어 장애인 복지관”에서 인터뷰를 진행하였고 도출한 요구사항은 다음과 같았고 이를 바탕으로 [그림 3]을 작성하였다.

### 3.1 높은 음성 인식률의 필요성

기존의 STT API(Application Programming Interface)를 기반으로 한 프로그램들은 음성 인식률이 높지 않아서 음성데이터가 정확하게 텍스트 데이터로 변환되지 않았다. 이로 인해 사용자는 원하는 정확한 정보를 얻을 수 없었다. 그래서 본 프로그램은 블루투스(Bluetooth) 마이크를 이용해 화자와 마이크까지의 거리를 최소화하여 잡음의 간섭을 최소화하고 음성데이터의 정확도를 높였다.

### 3.2 한글에 익숙지 않은 사용자층 고려

청각장애인 중 일부는 미흡한 한글 교육 등의 이유로 어려운 한글 단어를 잘 모르는 경우가 있다. 이러한 사용자 계층을 고려하여 프로그램은 사전 API를 이용해 한글 단어에 대한 정보를 제공한다. 그리고 선택적으로 평상시 수어를 사용하는 사용자들을 위해 수어 동영상으로 단어 정보를 제공한다.

## 3.3 TTS(Text To Speech)를 이용한 의사소통 지원

청각장애인들은 언어에 대해 학습을 하기 어려운 환경 속에 있어서 일반인과의 의사소통 능력이 미숙한 경우가 있다. 이러한 사용자들을 위해 프로그램은 TTS 기능, 즉 텍스트 데이터를 음성데이터로 변환해 교사와의 의사소통을 지원함으로써 학생의 궁금증 해결을 돕는다. 그리고, 사용자는 자주 사용하는 텍스트 데이터를 데이터베이스(Database)에 저장하여 TTS 기능 실행 시 사용할 수 있다.

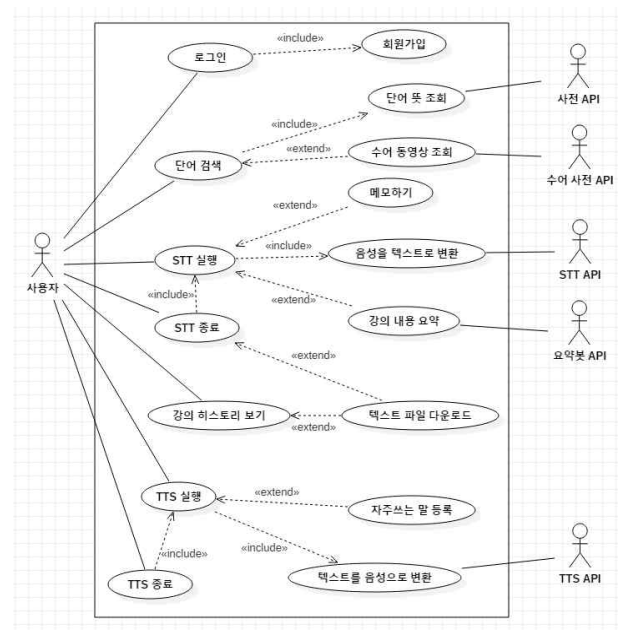


그림 3. Use case diagram

## 4. 프로그램 설계

전체 시스템 구조는 크게 음성데이터를 입력받아서 텍스트로 처리하는 부분과 사전을 이용한 단어 검색하는 부분, 텍스트 데이터를 음성데이터로 변환하는 부분, 음성데이터를 변환한 텍스트 데이터를 요약해주는 부분으로 구분된다.

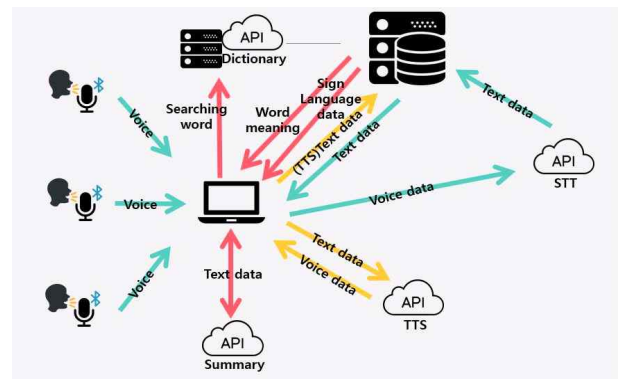


그림 4. 전체 시스템 구조

시스템 흐름을 보면, 블루투스 마이크를 이용한 화자가 음성데이터를 입력하면 STT API를 통해 음성데이터가 텍스트 데이터로 변환된다. 변환된 텍스트 데이터는 데이터베이스에 저장되고, 사용자에게 모니터(Monitor)로 출력이 된다. 사용자는 변환된 텍스트를 요약기능을 이용할 수 있고 내려받을 수 있다. 모니터에 출력되는 것과 별개의 화면으로 사용자가 수업 중 모르는 단어를 검색하면 사전 API를 통해 단어에 대한 정보를 사용자에게 출력된다. 이때 단어의 정보는 사용자의 옵션에 따라 텍스트 정보 또는 수어(Sign Language) 동영상을 이용하여 모니터에 출력한다. 마지막으로 사용자가 텍스트 데이터를 입력하면 TTS API를 통해 텍스트 데이터가 음성데이터로 변환된다. 변환된 음성데이터는 사용자의 질문을 돕는다.

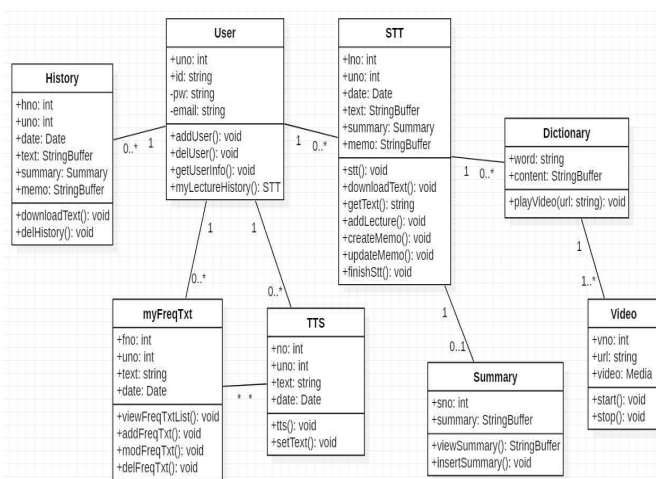


그림 5. Class diagram

[그림 5]는 위의 프로그램을 Class diagram으로 나타낸 것이다. 각각의 사용자들은 아이디와 비밀번호를 사용해 계정을 가진다. 프로그램을 사용하기 위해 로그인(Login) 기능을 이용하고, 계정을 이용하여 강의를 텍스트로 변환한 정보를 얻을 수 있다. 그리고 수업 중 질문을 위한 TTS 기능 이용 시, 자주 사용하는 단어들을 저장해두어 사용자에게 더욱 편리한 프로그램을 설계했다.

## 5. 프로그램 구현

### 5.1 개발 환경

프로그램을 구현하기 위한 개발 환경으로는 HTML(HyperText Markup Language)를 기반으로 사용하고 Front-End 부분의 웹 디자인을 위해 CSS를 이용한다. 그리고 동적인 Interface 및 UI/UX를 위해 React와 Javascript를 이용하여 구현한다. Back-End 부분의 기능을 구현하기 위해 Node.js를 이용하며 데이터베이스로 Firebase 및 NoSql를 사용한다.

## 5.2 구현

[그림 6]은 핵심 기능인 STT 기능의 UI를 구현한 것이다. ‘시작’ 버튼을 이용하여 음성인식을 시작할 수 있고, 마이크를 입력된 음성데이터는 텍스트로 변환되어 빈 화면에 출력된다. 사용자는 ‘종료’ 버튼을 이용하여 STT를 중단할 수 있다. 추가적인 기능인 사전을 이용한 텍스트 검색 및 텍스트 요약, 내려받기 그리고 TTS 기능은 차후에 구현할 예정이다.

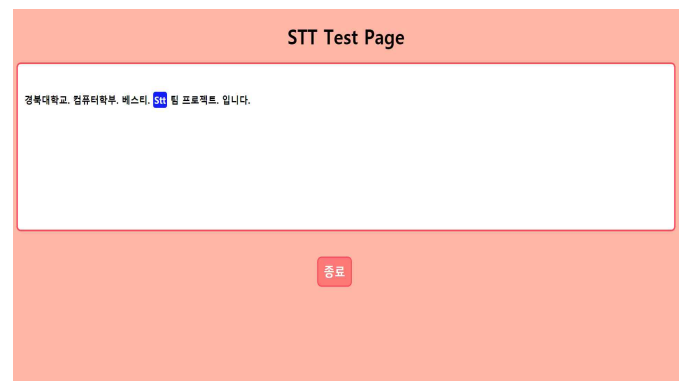


그림 6. STT 기능의 Prototype

## 6. 결론 및 기대효과

본 논문에서 설계한 시스템 구조를 이용해서 웹 프로그램을 구현할 수 있다. 프로그램을 이용하여 청각장애인 학생들이 일반학생과 같은 교육 환경에서도 동일한 수준의 교육을 받을 수 있을 것이다. 기존의 다른 프로그램들과 달리, 사전기능을 이용한 검색과 한글 텍스트에 어려움을 겪는 학생들을 위한 수어로 된 동영상 사전기능, 청각장애인 학생이 질문을 쉽게 할 수 있도록 TTS 기능을 추가하여 더 능동적이고 효과적으로 교육에 참여 할 수 있을 것이다.

## Acknowledgement

"본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음"(2015-0-00912), "This research was supported by the Korean MSIT (Ministry of Science and ICT), under the National Program for Excellence in SW supervised by the IITP(Institute of Information & communications Technology Planning & Evaluation)"(2015-0-00912)

## 참고문헌

- [1] 교육부. (2018). "교육통계분석자료집-유초중등교육통계편", 교육부, 한국교육개발원, 세종시, 진천, pp.59
- [2] 교육부. (2018). "2018년 특수교육통계". 교육부, 세종시, pp.5
- [3] 윤종옥, 권미은, 김라경, 서유진, 손창균, 이기성 (2017), "2017 특수교육 실태조사", 국립특수교육원, 아산, pp.237-238

[4] 박단비. (2016), "청각장애인을 위한 수강보조 모바일 소프트웨어 설계 및 구현", 석사학위논문, 성신여자대학교, 서울, 11 pages.

[5] 이기범. (2019). "구글, 청각장애인 위한 접근성 기술 선보여", <https://www.bloter.net/archives/333400> (2019-04-07 방문).