

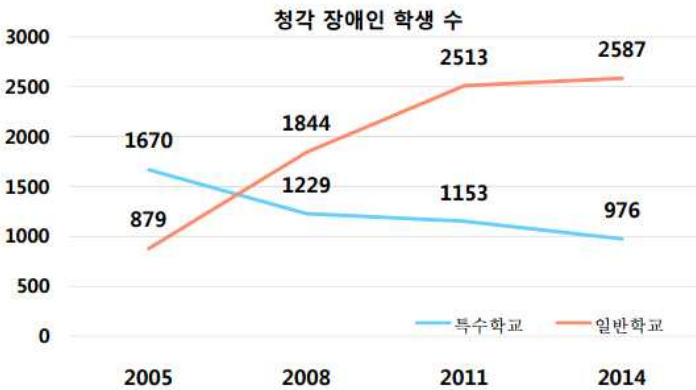
## 종합설계프로젝트 : 수행계획서

과제명	Talk to me : 음성인식을 이용한 실시간 인물 구분 타이핑 프로그램.				
과제목표	<ul style="list-style-type: none"> <li>청각장애인들이 혼자서도 쉽고 편리하게 속기기능을 사용할 수 있는 타이핑 프로그램을 개발한다.</li> <li>JAVA 언어에서 Naver 음성인식 API를 활용하여 마이크로 입력된 음성을 글자로 변환하여 스마트폰 화면에 출력한다.</li> <li>음성분석 프로그램인 PRAAT를 이용하여 들려온 음성을 목소리의 특징들을 참고하여 분석한다.</li> <li>머신러닝 알고리즘을 구현하여 분석된 음성을 학습하고 저장하며 인물을 특정 짓는다.</li> </ul>				
주요내용	마이크를 이용하여 목소리를 아날로그 음성으로 받는다. 진동수와 파의 모양, 진폭으로 음성을 분석한다. 머신러닝 알고리즘으로 개개인의 목소리를 학습하고 저장하고 저장된 음성과 들려온 목소리를 비교하고 구분하여 인물을 특정 분류한다. 음성인식 API를 이용하여 음성에서 글자를 인식한다. 마지막으로, 말한 인물과 그 내용을 실시간으로 화면에 출력한다.				
기대효과	청각장애인들이 혼자서도 수업을 듣고 비장애인과 대화를 하는 데에 있어서 편리함을 제공할 수 있다. 또한 비장애인에게 속기사의 활동이 필요할 때, 속기사 없이도 손쉽게 속기 타이핑 작업을 수행할 수 있다.				
팀대표	소속	IT 대학 컴퓨터 학부		성명	박소은
	연락처	010-3006-8495		Email	tdfqd3578@naver.com
과제 참여 학생	소속(학과)	학번	학년	성명	담당업무
	컴퓨터학부	2012105037	4	박소은	팀대표 (과제 총괄) 음성인식 API 분석/수정
	컴퓨터학부	2012105028	4	김한별	기자재 운반 음성패턴 수집/분석
	컴퓨터학부	2012105040	4	박지인	자료 수집/분석 출력 싱크 조율 & UI
	컴퓨터학부	2012105057	4	송은영	회계 및 문서관리 목소리 분류 알고리즘 설계

## 1. 과제 추진 배경

### 1.1. 청각장애 학생 수 가에 맞추지 못한 현행 학생 지원 서비스의 문제점

- 현재 청각장애인 학생으로서 일반학교에 재학 중인 학생의 수는 매해 증가 중이며, 2016년 기준 청각장애인 학생의 73%가 일반학교에 재학 중이다.



연도	특수학교	일반학교	일반학교 내 특수학급	일반학교 내 일반학급	총
2005	1670	879	494	385	2549
2008	1229	1844	795	1049	3073
2011	1153	2513	916	1597	3666
2014	976	2587	808	1779	3563

연도	특수학교	일반학교	일반학교 내 특수학급	일반학교 내 일반학급	총
2005	68%	34%	19%	15%	100%
2008	40%	60%	26%	34%	100%
2011	31%	69%	25%	44%	100%
2014	27%	73%	23%	50%	100%

- 하지만 교내 학습도우미 지원 및 속기사 지원 등의 서비스가 미흡하다.
- 청각 장애인의 평등한 학습권 보장 및 비장애인 학생과의 의사소통이 어려운 상황이다.

### 1.2. 유사 서비스 현황

- 유사 서비스인 '쉐어타이핑'의 경우, 속기된 문장을 자막으로 제공하는 것으로 속기사의 도움이 필요하다.
- 이 경우 속기사 파견을 신청하여 속기사와의 일정을 맞추거나 대기하여야 한다.



### 1.3. 진화된 시스템 개발의 필요성.

- 이러한 상황에서 청각장애인의 평등한 학습권 보장과 비장애인과의 의사소통을 도울 수 있도록 혼자서도 쉽고 편리하게 속기기능을 사용할 수 있는 프로그램을 개발하고자 한다.

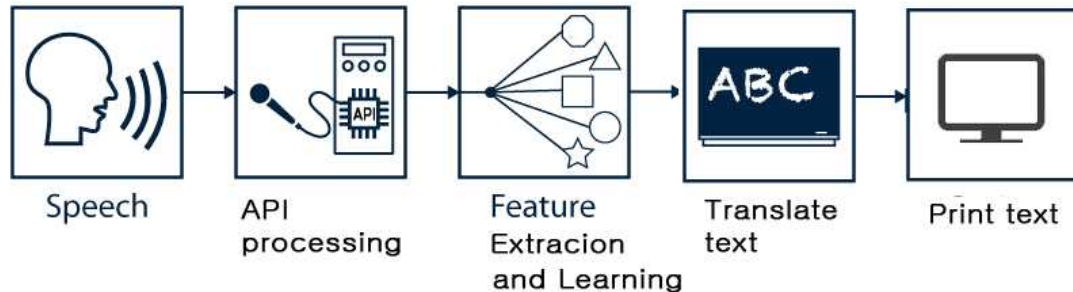
## 2. 과제 목표 및 수행 내용

### 2.1. 과제 개요

해당 프로젝트의 목표는 음성인식을 이용하여 실시간으로 대화나 강의를 하고 있는 사람들을 구분하고, 그들 개개인이 말하는 내용을 출력하는 타이핑 프로그램의 구현이다. 청각장애인들이 그들의 불편함을 해

소하기 위해 사용하게 되는 만큼, 정확하고 신속한 대화 내용 정보를 타이핑해주는 것을 주요한 목표로 하고 있다.

## 2.2. 과제 수행 목표



- 1) 마이크를 이용하여 사람의 목소리를 아날로그 음성으로 받아들인다. 이 때, 사람이 알아들을 수 있는 소리의 크기라면, 마이크도 최대한 그와 동일한 소리의 크기까지 받아들일 수 있도록 맞추어보는 것을 목표로 한다. 그리고 하나의 목소리에서 최대한 많고 효과적인 자료를 추출해내야 하는 만큼 마이크가 소리를 세세하게 받아들이는 것이 중요하다. 그러므로 이 과정에서는 마이크의 성능과 그 테스트가 중요하다. 또한 음성을 받아들이는 순간부터 스마트폰의 화면에 글자로 출력되는 순간까지의 시간이 짧아야 하는 만큼 음성전송 속도 등이 빠른 마이크를 선택하도록 한다.
- 2) 앞의 단계에서 받아온 아날로그 음성을 분석한다. 이 때, 소리의 세 가지 요소인 진동수와 파의모양, 진폭으로 아날로그 음성을 분석하도록 한다. 만약, 이에 기초하여 충분히 분석을 할 수 없을 시에는, 소리에 관련된 믿을 수 있는 자료를 참고하여 분석 요소들을 늘려가도록 한다. 그리고 아날로그 음성을 분석하기위해 사용할 프로그램은 음성 분석 프로그램인 Praat로 한다. 이 프로그램 또한 분석 상황과 효율에 따라 변동이 가능하다. 이 과정은 사람의 목소리 외의 기타 소리들을 학습 대상에서 제외시키고 효과적인 자료들만을 추출해내기 위한 음성 분석의 초기단계라고 할 수 있다.
- 3) 분석된 음성을 머신러닝 알고리즘을 이용하여 개개인의 목소리를 학습한다. 해당 프로젝트의 가장 주요 목표인 머신러닝 알고리즘 구현은 시간이 지날수록 더 정확한 인물을 특정하도록 설계하는 것을 목표로 한다. 이 때 실시되는 머신러닝 알고리즘은, 분석된 음성을 일정한 기준으로 학습하는 단계이다. 어떤 방식으로 그리고 어떤 정보를 중점으로 두어, 분석된 음성에서 자료를 추출해 낼 것인지를 잘 구성하여야 효율적으로 머신러닝 알고리즘을 사용할 수 있을 것이다. 그러므로 이 단계에서 충분한 구현에 관한 자료 수집과 테스트가 요구된다.
- 4) 앞서 추출한 음성에 대한 자료들을 바탕으로 이전에 분석된 음성 자료들과 비교하여 인물을 특정해낸다. 이 과정 또한 머신러닝 알고리즘의 일부라고 할 수 있다. 이 때는, 머신러닝 알고리즘을 실행한 초기에 비교할 자료가 많지 않은 상황에서 개개인의 목소리를 잘 구별해내도록 일정한 기준을 두어 비교 분석해내는 것에 신경을 써야할 것이다. 그리고 머신러닝 알고리즘에 자료가 모이면 모일수록, 가장 비교에 용이한 자료들을 선정하여 신속한 비교에 기여할 수 있도록 설계해야 한다. 예를 들면, 특정 목소리 분류를 대표하는 주요 자료들끼리 비교하여 인물이 특정지어진다면 바로 출력해내고 그렇지 않다면 더 세부적인 자료들을 비교해보는 방식이 있을 수 있을 것이다. 이렇게 효과적으로 비교하여 추출된 자료들이 분류되면, 인물 특정 작업을 완료하였다고 할 수 있다. 물론, 초기 상황에서 저장된 자료들이 부족하다면 인물을 특정해내지 못할 수 있다.
- 5) 다음 과정으로는, 추출한 자료들을 다음번에 인식되는 음성과 비교하기 쉽게 저장하는 것이다. 저장할 때에는, 얼마나 효과적인 저장 구조를 설계하느냐에 따라, 다음 음성 분류시기에 머신러닝 알고리즘의

실행시간에 좋은 영향을 줄 수 있을 것이다. 왜냐하면 이렇게 저장된 자료들은 다음번의 음성 분류시기에 사용되기 때문이다. 그리고 이 때, 모아진 자료들 중 비교에 용이해서 선정된 자료들을 저장하는 저장소가 따로 마련되어야 좀 더 빠른 실행시간을 실현시킬 수 있을 것이다.

- 6) 음성인식 API인 Naver 음성인식 API를 이용하여 아날로그 음성에서 글자를 인식한다. 머신러닝 알고리즘 실행과는 별개의 단계로, 머신러닝 알고리즘과 병렬식 수행도 가능하다. 이 때에는, 머신러닝 알고리즘이 인물을 특정해내는 속도와 사람이 말하는 속도에 아날로그 음성에서 글자를 인식하는 속도를 맞추어 더 느려지는 일이 없도록 하는 작업이 중요한 문제가 된다. 머신러닝 알고리즘의 인물 특정시간보다는 길게, 사람이 다음 순서의 말을 하는 속도보다는 빠르게 글자를 출력해주어야 사용자가 불편함을 느끼지 않을 수 있다. 그러므로 기존의 Naver 음성인식 API를 잘 분석하여 해당 프로젝트를 수행함에 있어서 필요한 부분만을 잘 추출해내어 글자 인식 속도를 짧게 만드는 것이 목표이다.
- 7) 마지막으로, 스마트폰의 화면에 말한 인물과 그 말의 내용을 실시간으로 화면에 출력한다. 해당 단계에서는 심미적으로 안정적인 출력을 목표로 한다. 앞의 모든 기능들이 잘 구현되어 속도가 빨라지더라도 사용자가 실시간으로 읽어보는 데에 불편함이 없도록, 부드러운 속도 및 표현의 출력과 읽기 좋은 출력 디자인을 선택하여 구성하도록 한다.

### 3. 추진 방법

#### 3.1. 프로젝트 기간.

- 2016. 03. 01 - 2016. 06. 08
- 15주 프로젝트

#### 3.2. 프로젝트 팀 구성

구성원	박소은	김한별	박지인	송은영
역할	활동총괄 음성인식 API 분석/수정	기자재 운반 음성패턴 수집/분석	자료 수집/분석 출력 싱크 조율 UI	회계 및 문서관리 목소리 분류 알고리즘 설계

#### 3.3. 프로젝트 비인적 자원

항목
개발&출력용 스마트폰
음성 입력용 마이크
Praat 프로그램
Naver 음성인식 API

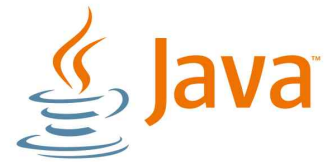
#### 3.4. 프로젝트 개발 환경

운영 체제 : Android

개발 환경 : Android Studio

개발 언어 : Java 언어

라이브러리 : Naver 음성인식 API



#### 4. 추진일정

시기	추진 내용	결과물
2주차(3/9)	주제 발표	주제제안서
3주차(3/16)	수행 계획서 발표	수행계획서
4주차(3/23)	시스템 설계 완료	시스템 설계 사양서
5주차(3/24)	구현 시작	
8주차(4/27)	기본기능 구현 완료	프로토타입
10주차(5/11)	정확도 개선	프로토타입2
11주차(5/12)	테스팅 시작	
13주차(6/1)	테스팅 완료	완성본
14주차(6/8)	결과 보고서 발표	결과보고서

#### 5. 기대효과 및 활용방안

##### 5.1. 기대효과

해당 프로젝트를 잘 수행하여 만들어진 'Talk to me' 프로그램은 청각장애인들의 귀가 잘 들리지 않아서

겪게 되는 불편함을 해소시킬 수 있을 것이다. 먼저, 청각장애인들이 수업에 참여할 때, 속기사 도우미를 포함한 다른 도우미에게 도움을 받지 않고도 교수님이 하시는 말씀이나 자료물의 음성을 쉽고 빠르게 이해하고 인지할 수 있게 될 것이다. 또한 청각장애인들이 비장애인과 함께 대화를 나눌 때, 비장애인의 입모양을 관찰하며 누가 어떤 말을 하는지 파악하려고 애쓰지 않아도, 해당 프로그램을 이용하면 누가 어떤 말을 하는지 편리하게 알 수 있을 것이다.

## 5.2. 활용방안

스마트워치를 활용하여 해당 프로그램을 좀 더 간편하게 사용할 수 있도록 한다면 사용성이 더 좋아질 것이다. 그리고 다양한 소리를 감지해서 분별할 수 있는 기술로 아기울음소리, 화재경보, 자동차경적, 초인종소리 등을 시각적으로 전달한다면, 청각장애인분들이 긴급 상황에 빨리 대처할 수 있는 데에 도움을 줄 수 있을 것이다.

또한 청각장애인들이 아니더라도 해당 프로그램을 여러 방면으로 사용할 수 있도록 향상시킨다면, 비장애인과 장애인 모두 편리하게 사용할 수 있는 유니버설 디자인 프로그램으로 거듭날 수 있을 것이다. 예를 들면, 법정에서 속기사들이 말하는 내용을 받아 적는 등과같이 기록이나 회의록의 작성이 필요한 경우, 향상된 해당 프로그램을 사용하여 전문 속기사 없이도 말하는 내용을 쉽게 타이핑하는 것이 가능할 것이다. 그리고 콜센터와 같이 고객의 음성을 저장해놓아야 할 경우, 실시간으로 말하는 내용을 텍스트로 바꾸고 고객의 음성 데이터베이스를 저장하여, 빅데이터 분석에 효과적으로 사용될 수 있을 것이다.