STT를 활용한 근본적 말하기 습관 분석 및 개선 연구

김다영*, 송민경*, 정수정*, 최서현*, 임성현**
*덕성여자대학교 컴퓨터학과
**우아한 형제들
e-mail:da1599@naver.com

A Study on Analysis and Improvement of Basic Speech Habits Using STT

Da-yeong Kim*, Min-gyung Song*, Su-Jeong Jeong* ,
Hyun-Choi Seo*,Sung-Hyun Lim**
*Dept of Computer Science, Duk-Sung University
**Woowa Brothers

요 약

지금까지 사람들의 말하기 습관을 분석하는 것은 전문가나 다른 사람들의 피드백을 통해 이루어졌다. 이는 평가하는 사람에 따라 다른 결과가 나타날 수 있기 때문에 다소 객관성이 떨어지며 직접 대면이 요구되기 때문에 훈련할 수 있는 시간과 공간에 제약이 있었다. 또 지금까지의 말하기 훈련이라 함은 면접 스피치나 발표 스피치 등 특정 목적을 가지고 진행되었기 때문에 말하기에 있어서의 통합적인 훈련을 받기 어려웠다. 본 연구에서는 가장 기초적인 말하기 습관에 초점을 맞추고 말하기 능력을 '어휘력', '계속성', '발음', '속도'라는 4개의 영역으로 나누어 각 영역을 분석하는 알고리즘을 제시한다. 이 알고리즘을 활용한다면 모두 같은 메커니즘을 통해 분석이 이루어지기 때문에 객관성을 보장할 수 있으며 어플리케이션을 활용해 시간과 공간의 제약을 받지 않고 말하기 습관 개선을 위한 훈련을 할 수 있을 것이다.

1. 서론

현대 사회에서는 사람들에게 점점 더 많은 능력을 요구하고 있다. 취업하기 위한 필수 조건인 면접은 물론 발표나 회의 등에서 자신의 의견을 조리 있게 말해 상대방을 설득하는 것이 중요해졌다. 뿐만 아니라 1인 미디어 시대를 살고 있는 우리는 강연자가 될 수도 있고 방송인이 될 수도 있다. 따라서 현대 사회에서 말하기 습관은 더욱 강조되고 있다.

20대 이상 성인남녀 89명을 대상으로 진행한 말하기 습관 관련 설문조사에 따르면 전체 응답자 중 약 33%가 말하기 문제를 겪고 있다고 답했으며 말하기 문제를 겪고 있지는 않지만 개선할 의지가 있는 응답자는 87%에 달하였다. 이는 많은 사람들이 말하기 습관을 하나의 능력으로 여기고 있으며, 관심을 보이는 부분이라고 할 수 있다.

기존 말하기 관련 어플리케이션은 면접 스피치 등 특정 목적에 맞춰 개발되었으며 이는 사용자를 제한하거나 사용자의 욕구를 만족시키지 못할 수도 있다. 본 연구에서는 말하기 요소를 내용과 기술로 구분하였다. 기술은 다시 어휘력, 계속성, 발음, 속도로 나누었고, 정확한 어휘력, 끊이지 않는 흐름, 정확한 발음, 신뢰성 있는 속도로 말을 한다면 말하고자 하는 내용이 미흡하더라도 말을 잘 하는 것처럼 느껴질 수 있다고 간주하였다. 따라서 본 연구는 말하기의 기술 측면에서 말하기 근본 습관에 대해 분석하며 개선하는 방법을 다룬다.

2. 사용된 기술

2.1 구글 Speech Recognition API

구글에서 배포한 Speech Recognition API는 사용자가의 음성을 녹음해 텍스트로 바꿔주는 기술이다. 딥 러닝 기술이 적용되었으며 현재 110가지 이상의 언어를 지원한다. 실시간으로 결과 출력이 가능하며 소음이 있는 환경에서도 효과적으로 잡음을 처리하고 사용자가 하는 말의 문맥을 인식해 결과를 내보낸다. 핸드폰, PC, 태블릿, IoT 디바이스 등 여러 기기에 적용시킬 수 있는 기술이다.[1)]

3. 말하기의 4개 영역

'어휘력', '계속성', '발음', '속도'를 말하기의 기본이 되는 4가지 영역으로 선정하였다. 이는 말하기 습관 관련 설문조사를 바탕으로 결정하였다.

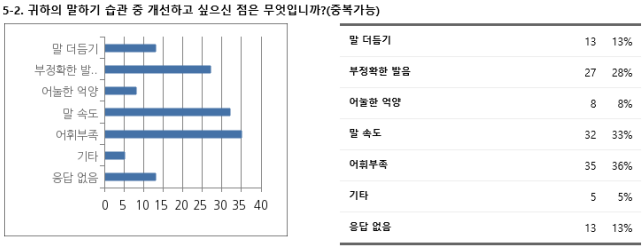

그림 1. 말하기 습관 관련 설문조사

위 설문조사에 따르면 말 속도, 어휘 부족, 부정확한 발음, 말 더듬기 순으로 높은 응답률을 보였다. 이를 말하기의 4개 영역으로 나누었으며 각 영역의 측정을 통해 객관적으로 말하기 습관을 분석할 수 있다.

3.1 어휘력 영역

어휘력이란 상황과 문맥에 따라 적절한 형태와 의미를 가진 단어를 사용에 관한 것이다. 어휘력은 총체적인 성격을 가지고 있기 때문에 그 범위가 매우 넓어 2가지 테스트로 구성해 측정한다.

첫 번째 테스트에서는 한국인이 자주 혼동하는 두 가지 단어를 비교하며 선택함으로써 어휘력을 향상시킬 수 있다. 가령 깨우치다/깨치다 와 같이 형태가 유사한 단어의 경우 앞뒤 문맥을 고려하여 사용해야만 한다. 이 테스트에서는 30초 동안 4문제가 주어지기 때문에 짧은 순간 올바른 어휘를 판단하여 답을 제시함으로써 학습효과를 높일 수 있다.

어휘력에 관한 두 번째 테스트는 관련성이 없어 보이는 5가지 단어를 조합하여 문장을 만드는 테스트이다. 사용자는 5가지 단어를 적절한 상황과 문맥에 맞춰 문장에 끼워 넣어야 하기 때문에 풍부한 어휘력이 요구되는 테스트이다.

어휘력 테스트에서는 구글 Speech API로 인식한 문장 내에 올바른 특정 어휘가 포함되어 있다면 점수를 부여한다.

3.1 계속성 영역

계속성은 말을 할 때 얼마나 끊기지 않고 적절한 흐름으로 말을 할 수 있는가에 관한 영역이다.

서론에서 언급했던 것과 같이 본 연구에서는 말할 때 내용보다는 기술에 집중하기 때문에 특정 시간 동안 언급한 어절의 개수로 계속성을 측정한다. 여기서는 30초 동안 적정 속도로 말을 할 때 이에 따른 적절한 어절의 수를 60~65개라고 두었다. 어절의 개수는 한 문장 내의 공백의 개수에 1을 더해 계산하였다. 계속성 영역의 문제는 대부분의 사람들이 일반적으로 쉽게 생각할 수 있지만 어느 정도 사고가 필요한 문제들로 구성했다. 가령 '인공지능 시대에 대비하는 우리의 자세에 대해 말해 보시오.'라는 질문에 적정 어절의 수를 초과하면 내용의 신뢰성이 떨어진다고 간주했으며 이상적인 어절의 수에 가까울수록 높은 점수를 얻게 된다.

계속성 훈련은 사용자가 말하기 전 머릿속으로 빠르게 해야 할 말을 정리하는 능력과 순발력을 기를 수 있도록 도와준다.

3.2 발음 영역

발음 영역은 사용자의 발음을 얼마나 구글 Speech API가 정확히 인식하는지에 관한 것이다.

발음 영역은 테스트 질문을 어떤 문장으로 선정할 것인지 고려하는 것이 중요하다. 테스트 문장을 너무 일반적인 문장으로 선정해 버리면 사용자가 잘못 발음해도 구글 Speech API가 문맥에 맞게 의미 있는 문장으로 바꿔서 결과를 출력한다. 따라서 '간장공장 공장장'과 같이 사전에 등록되어 있지 않은 문장을 사용하여야 한다. 발음 영역 점수는 사용자의 음성을 구글 Speech API가 인식해 출력한 문장과 테스트 문제가 요구하는 문장이 일치하는 정도에 따라 측정된다.

발음 영역의 문제는 4개의 문장으로 구성되어 있으며 처음 두 문제는 짧은 단어로 나머지 두 문제는 긴 문장으로 구성되어 있다. 이렇게 구성된 4문제가 30초 동안 발음하기 가장 적절한 분량이라고 판단하였다.

발음 영역에서는 비슷한 발음이 반복되는 단어나 문장을 반복적으로 연습할 수 있도록 도와준다.

3.1 속도 영역

속도 영역은 얼마나 적절한 속도로 말하는가에 관한 것이다.

여기서 말하는 적절한 속도[2][3]는 이미 기존에 논문으로 발표된 연구 자료를 참고하였다.

연구자	정상 발화
박경희(2008)	340음절/분당
정은이 · 정의철(2015)	300~380/분당
심홍임(2015)	306음절/분당

표 1. 정상 발화 속도

속도 영역의 문제는 우리가 쉽게 읽을 수 있는 문장으로 구성되어 있으며 질문은 기준보다 훨씬 많은 양의 글이 주어지기 때문에 사용자가 자연스럽게 읽는 속도를 측정할 수 있다.

4. 말하기 습관 개선 훈련

말하기 습관 개선 훈련은 사용자가 반복적인 과정을 통해 이루어질 수 있도록 한다.

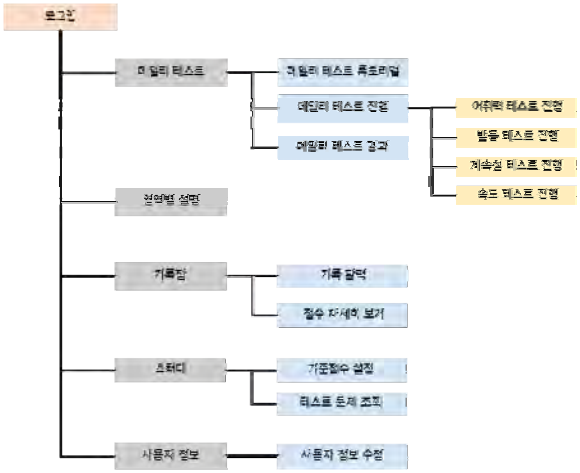

그림 2. 서비스 흐름도

그림 2에서 사용자는 데일리테스트, 스터디, 기록장 이렇게 3개의 과정을 순환하면서 훈련을 반복적으로 지속할 수 있다.

4.1 자가 피드백을 통한 말하기 습관 개선

말하기의 4개 영역의 테스트를 마쳤다면 자가 피드백을 통해 말하기 습관을 개선할 수 있다.

테스트를 마친 사용자는 스터디 메뉴를 통해 특정 점수 미만의 문제를 조회할 수 있다. 사용자는 영역별 스터디 메뉴에서 각 영역별 설명을 찾아보며 해당 영역의 점수가 어떻게 산출되었는지 확인한 후 사용자가 설정한 점수 미만의 테스트 문제를 통해 스스로 훈련할 수 있다.

4.2 기록장을 통한 학습의욕 고취

사용자는 기록장을 통해 지금까지 학습한 결과를 확인할 수 있다. 꺾은선 그래프를 통해 한 달 단위로 점수 편차를 한눈에 확인할 수 있으며 특정 날짜를 선택하여 영역별 상세 점수를 확인할 수 있다. 사용자의 점수를 확인하게 함으로써 학습의욕을 고취시키고 다음 단계인 데일리테스트를 진행하게 하도록 한다.

5. 결론

사전에 진행한 설문조사대상자인 89명의 모집단에서 10명의 표본집단을 추출하여 이들을 대상으로 5일 동안 말하기 테스트 진행하였다.

1	표본집단은 말하기 습관 개선에 대한 필요성을 인식하고 있는가?
2	표본집단은 본인의 말하기 습관 개선을 필요로 하는가?
3	표본집단의 연령은 모집단의 모든 연령을 포함하는가?
4	표본집단의 직업은 모집단의 응답률을 기준으로 하는 상위직군에 포함되는가?

표 2. 사후 설문조사 대상자 선정 기준

표 2. 는 사후 설문조사 대상자를 모집단에서 표본집단으로 선정한 기준을 정리한 표이다. 표본집단은 모집단의 사전 설문조사에서 말하기 습관 개선의 필요성을 인식하고 필요로 하는 응답자중 연령과 직업 분포를 기준으로 선정하였으며 5일 후 사용자들의 반응을 확인하기 위해 사후 설문조사를 진행하였다.

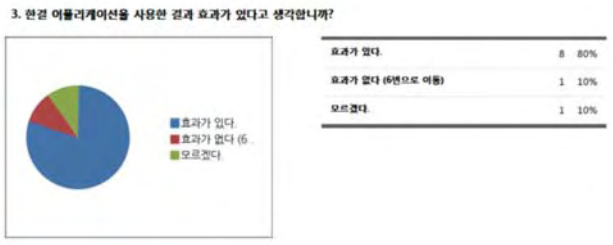

그림 3. 사후 설문조사 결과

그림 3. 은 어플리케이션을 이용해 말하기 테스트를 진행한 후 효과성에 대한 응답결과를 나타낸 그래프이다. 10명의 사용자 중 '효과가 있다' 항목의 응답률은 80%로 가장 높았다.

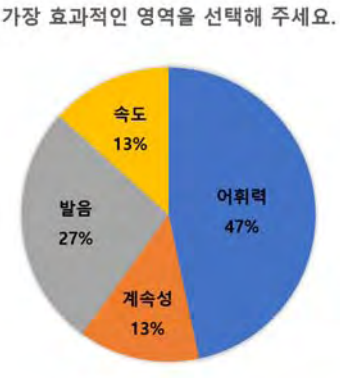

그림 4. 사후 설문조사 결과

그림 4. 은 사용자들이 가장 효과적으로 훈련했다고 생각하는 영역을 선택하는 질문에 대한 응답률을 나타낸 그래프이다. 가장 효과적이라고 생각하는 영역에서는 어휘력이 가장 높은 응답률을 보였다.

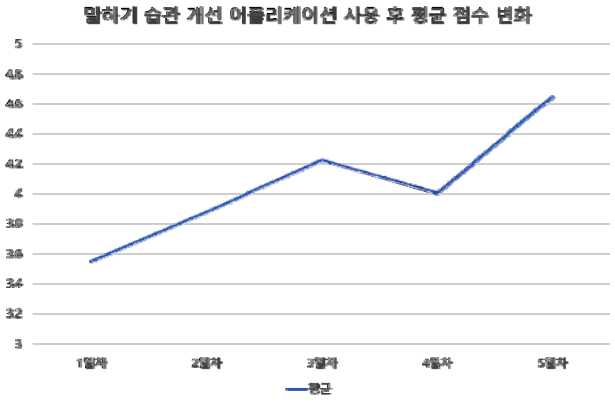

그림 5. 말하기 습관 개선 어플리케이션 사용 후 점수 변화 그래프

그림 4. 는 5일간 말하기 습관 개선 어플리케이션 사용 결과를 기록 후 점수의 변화를 나타낸 그래프이다. 4일차를 제외하고 점수 추이가 상승세를 보이는 것을 알 수 있다.

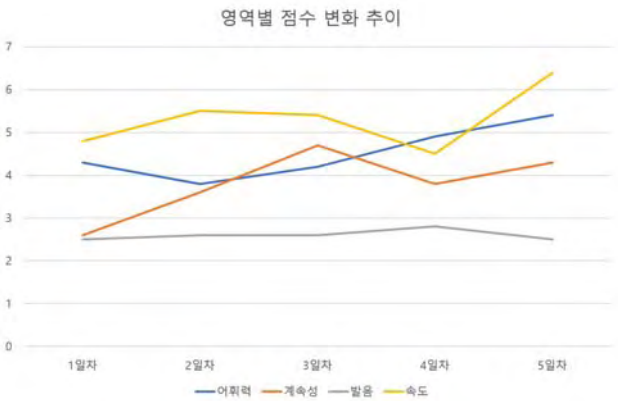

그림 6. 말하기 습관 개선 어플리케이션 사용 후 영역별 점수 변화 추이 그래프

그림 5. 는 영역별 점수 변화 추이를 나타낸 그래프이다. 1일차와 비교했을 때 모든 영역이 상승세를 보인다. 수치상으로 살펴보면 1일차와 5일차의 점수가 계속성 영역에서 1.7점의 차이를 보이면서 가장 큰 변화 폭을 보였으며 속도, 어휘력, 발음이 그 뒤를 이었다.

현대에 들어 국내외 스피치 시장규모가 증가함에 따라 말하기의 중요성이 강조되고 있다. 이는 말하기와 관련된 기술 습득의 수요가 늘어난다는 것과 동시에 말하기 관련 직업이 더욱 전문화되고 있다는 것을 의미한다.

말하기는 사람들의 직장과 대인관계와 같은 일상생활의 모든 분야에 스며들어가 영향을 미친다. 직장에서 업무를 수행하는 모든 과정은 말하기와 직접적 혹은 간접적으로 관련되어 있다. 또 대인관계에서도 긍정적인 말하기 습관은 사람들에게 호감과 신뢰감, 안정감을 주며 정확한 의견과 주장을 전달할 수 있도록 한다.

국내의 한국어 관련 스피치 서비스를 제공하는 어플리케이션은 얼마 되지 않을 뿐 만 아니라 제공하는 서비스는 사용자가 말하는 녹화 영상만을 보여주는 등 비전문적이고 간단한 기능만 포함하고 있다. 하지만 본 연구에서 다룬 4가지 영역을 분석하고 수치화하는 알고리즘을 활용해 어플리케이션을 배포한다면 말하기 근본 습관에 대해 수치화가 가능하다는 점에서 전문성과 객관성을 가지게 된다. 또 사용자의 금전적 부담을 줄여 주며 시간과 장소에 제한을 받지 않고 서비스를 이용할 수 있기 때문에 가격과 시간 투자 면에서 시장성을 가지게 된다. 따라서 본 연구에서 제공한 알고리즘은 한국어 스피치 시장을 대상으로 현시점에서 가장 활용도가 좋다고 할 수 있다.

참고문헌

[1] GOOGLE,“Google Cloud Speech API”, https://cloud.google.com/speech/

[2] 이소현.“대학생 말하기 향상을 위한 준언어적 구성요소와 교육 방법에 대한 모색: 발표를 중심으로”, 언어과학연구 제 79집, 2016.12

[3] 심홍임, "정상 성인의 비유창성 특징과 말 속도에 관한 연구", 2005