

- 자연어처리팀 주제분석 -

# 한국어 발음 연습을 위한 뉴스 대본 생성 서비스

박상훈  
곽동길  
박윤아  
김수진  
신민서

# 1. 프로젝트 개요

## 아이디어 발상 동기 | 수업 자료 제공의 한계

### 〈W스피치 아나운서 원고〉

미국 영화상 골든글러브 시상식이 할리우드에 보이콧으로 파행된 가운데
외신들은 오징어게임의 오영수가 한국배우 최초로 남우조연상을 받은 것을
올해 시상식의 주요 장면으로 꼽았습니다.
로이터 통신은 할아버지 오영수가 세계적으로 유명한 경쟁자들을 물리치고
상을 차지했다고 전했고,
CNN 방송도 오징어게임 스타 오영수가 역사를 만들었다고 보도했습니다.
포브스는 독창적인 한국드라마 오징어게임은 순식간에 넷플릭스 역사상
가장 인기드라마라는 명예를 얻었고,
극중 오영수는 팬들이 가장 좋아하는 캐릭터였으며 골든 글러브 수상에 따라
극중 오영수는 팬들이 가장 좋아하는 캐릭터였으며 골든 글러브 수상에 따라
일흔여덟살 그의 연기 이력은
결코 예전과 같지 않을 것이라고 박수를 보냈습니다.

아나운서 학원은 준비생들에게 정확한 발음 습득을 위해  
'된소리 발음 대본', '숫자가 많은 뉴스 대본' 등의  
연습 자료를 제공하고 있음



# 1. 프로젝트 개요

## 아이디어 발상 동기 | 수업 자료 제공의 한계

### 〈W스피치 아나운서 원고〉

미국 영화상 골든글러브 시상식이 할리우드에 보이콧으로 파행된 가운데  
외신들은 오징어게임의 오영수가 한국배우 최초로 남우조연상을 받은 것을  
올해 시상식의 주요 장면으로 꼽았습니다.

로이터 통신은 할아버지 오영수가 세계적으로 유명한 경쟁자들을 물리치고  
상을 차지했다고 전했고,  
CNN 방송도 오징어게임 스타 오영수가 역사를 만들었다고 보도했습니다.

포브스는 독창적인 한국드라마 오징어게임은 순식간에 넷플릭스 역사상  
가장 인기드라마라는 명예를 얻었고,  
극중 오영수는 팬들이 가장 좋아하는 캐릭터였으며 골든 글러브 수상에 따라

극중 오영수는 팬들이 가장 좋아하는 캐릭터였으며 골든 글러브 수상에 따라  
일흔여덟살 그의 연기 이력은  
결코 예전과 같지 않을 것이라고 박수를 보냈습니다.



그러나 적합한 뉴스를 일일이 검토해  
발음 훈련 대본을 만드는 작업은 시간이 많이 소요되며,  
개별 학습자에게 맞춘 고품질 자료 제공이 어려움

# 1. 프로젝트 개요

## 아이디어 발상 동기 | 수업 자료 제공의 한계

### 〈W스피치 아나운서 원고〉

미국 영화상 골든글러브 시상식이 할리우드에 보이콧으로 파행된 가운데  
와신들은 오징어게임의 오영수가 한국배우 최초로 남우조연상을 받은 것을  
올해 시상식의 주요 장면으로 꼽았습니다.

로이터 통신은 할아버지 오영수가 세계 최초로 유한한 영웅자료를 열었다고  
상을 차지했다고 전했다.  
CNN 방송도 오징어게임 스타 오영수가 역사를 만들었다고 보도했습니다.

포브스는 독창적인 한국드라마 오징어게임은 순식간에 넷플릭스 역사상  
가장 인기드라마라는 명예를 얻었고,  
극중 오영수는 팬들이 가장 좋아하는 캐릭터였으며 골든 글러브 수상에 따라

극중 오영수는 팬들이 가장 좋아하는 캐릭터였으며 골든 글러브 수상에 따라  
일본여덟살 그의 연가 이력은  
결코 예전과 같지 않을 것이라고 박수를 보냈습니다.

딥러닝 기반 대본 생성 자동화로 학원의 수작업 부담을 덜고,

아나운서 준비생들에게는 지식적 어려움과 물리적 어려움을

동시에 고려한 체계적인 발음 훈련 자료 제공하고자 함

그러나 정한한 뉴스로 인원이 검토해  
받은 훈련 대본을 만드는 작업은 시간이 많이 소요되며,  
개별 학습자에게 맞춘 고품질 자료 제공이 어려움



## 1. 프로젝트 개요

### 서비스 기능 ① 뉴스 대본 생성

풀무원식품이 요리 서바이벌 프로그램 '흑백요리사'에 출연한 재미교포 요리사  
에드워드 리를 브랜드 모델로 발탁했습니다. 풀무원은 에드워드 리의 요리...

지난해에 이어 올해도 역사상 가장 더운 해가 될 것으로 보인다고 전망이  
나왔는데요, 기후 변화를 부정해 온 트럼프 미국 대통령 당선인은 취임 직후... (중략)...

최신 기사 내용이 반영될 수 있도록 실시간으로 다양한 뉴스 대본 생성

## 1. 프로젝트 개요

### 서비스 기능 ② 뉴스 대본 주제 및 난이도 선택

주제 : 정치 ☒ 경제 ☐ 지역 ☐ 국제 ☐ 사회 ☐ 문화 ☐ 스포츠 ☐ IT\_과학 ☐

난이도 : 초급 ☒ 중급 ☐ 고급 ☐

주제 및 연습하고 싶은 발음 난이도 설정 가능



이번 미국 대선의 또 다른 승자로, 트럼프 당선인을 전폭적으로 지원한 일론 머스크 테슬라 최고경영자를 꼽는데 이견이 없을 겁니다. ... (중략) ...

## 1. 프로젝트 개요

서비스 기능 ② 뉴스 대본 주제 및 난이도 선택

예) '정치' 섹션 내 대본들을 난이도 측정

현재 우리나라 법에서 정한 정년은 60세입니다. 그런데, 국민연금을 받기 시작하는 나이는 출생 연도에 따라서 조금씩 다릅니다. 지금 50 대 중반인 ...



난이도 1

일론 머스크 CEO가 트럼프 당선인 정부의 **신설 위원회**에서 **중추적 역할**을 할 것이란 관측이 나오며, 자율주행기술 규제 완화에 대한 **낙관론**이 급부상하고 있는데요, ...



난이도 5

사용자의 난이도 선택에 따라 제공되는 대본이 달라짐



## 1. 프로젝트 개요

### 서비스 기능 ③ 발음 주석 태깅

주차장에서 처음 본 남성의 금품을 빼앗고 살해유기한 40대 남성이

[주:차장에서 처음 본 남성에 금품을 빼앗고 살해유기한 사:십 대 남성이]

검거됐습니다. 증거인멸을 위해 피해자 차량에 불까지 질렀는데,

[검:거됐습니다. 증거인멸을 위해 피:해자 차량에 불까지 질렀는데,]

빼앗은 돈 12만 원으로 일부 복권을 산 것으로 드러났습니다.

[빼앗은 돈 시비마 뉘스로 일부 복권을 산 거스로 드러났습니다.]

제공되는 뉴스 대본의 각 문장 밑에 주석으로 정확한 발음을 태깅



01

2주차 피드백

# 1. 2주차 피드백 분석 및 구현 흐름

## 대본 생성 모델 개발

GPT 기반

프롬프트 러닝

+

파인 튜닝



대본 생성 모델 완성

## S2P 변환 모델 개발

Automatic speech  
recognition 활용

발음 정확도 산출

## G2P 태깅 모델 개발

한국어 기초사전  
API 활용

LLM application



발음 태깅 모델 완성

## 발음 연습 난이도 지표 제작

회귀 분석



지표 계산 공식 결정

# 1. 2주차 피드백 분석 및 구현 흐름

## 대본 생성 모델 개발

GPT 기반

프롬프트 러닝

+

파인 튜닝



대본 생성 모델 완성

## S2P 변환 모델 개발

Automatic speech  
recognition 활용

발음 정확도 산출

## G2P 태깅 모델 개발

한국어 기초사전

API 활용

LLM application



발음 태깅 모델 완성

## 발음 연습 난이도

지표 제작

회귀 분석

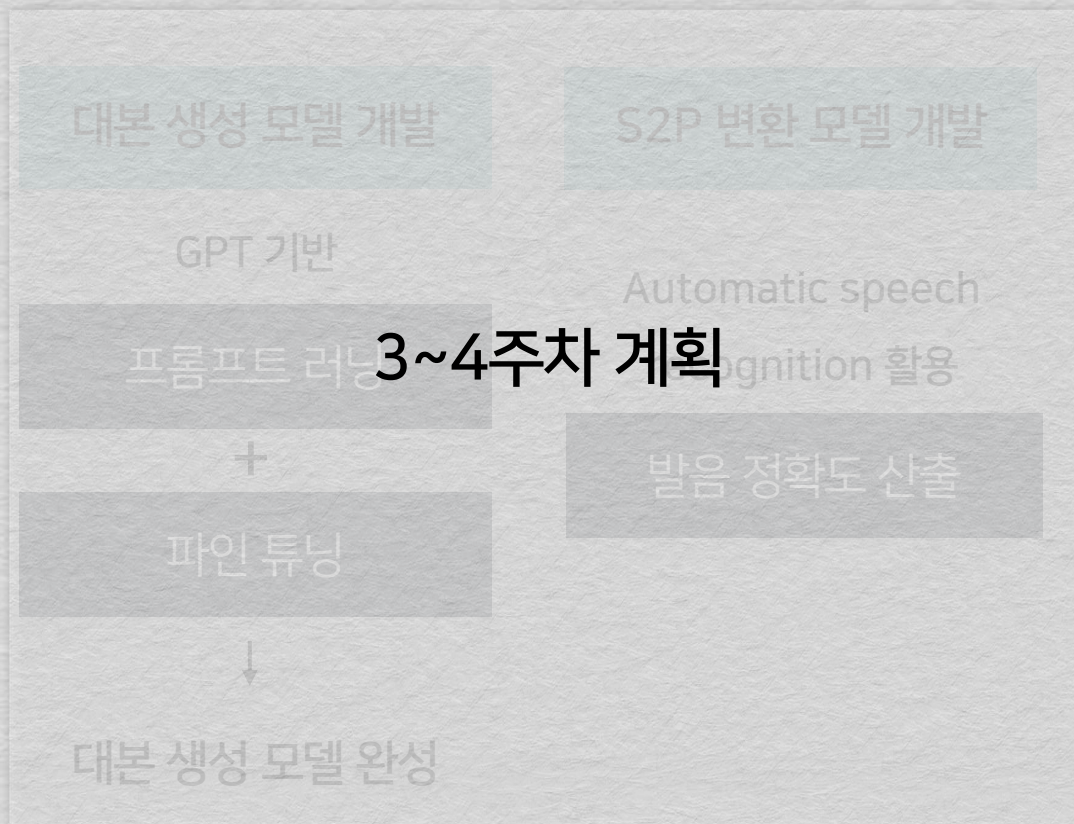


지표 계산 공식 결정

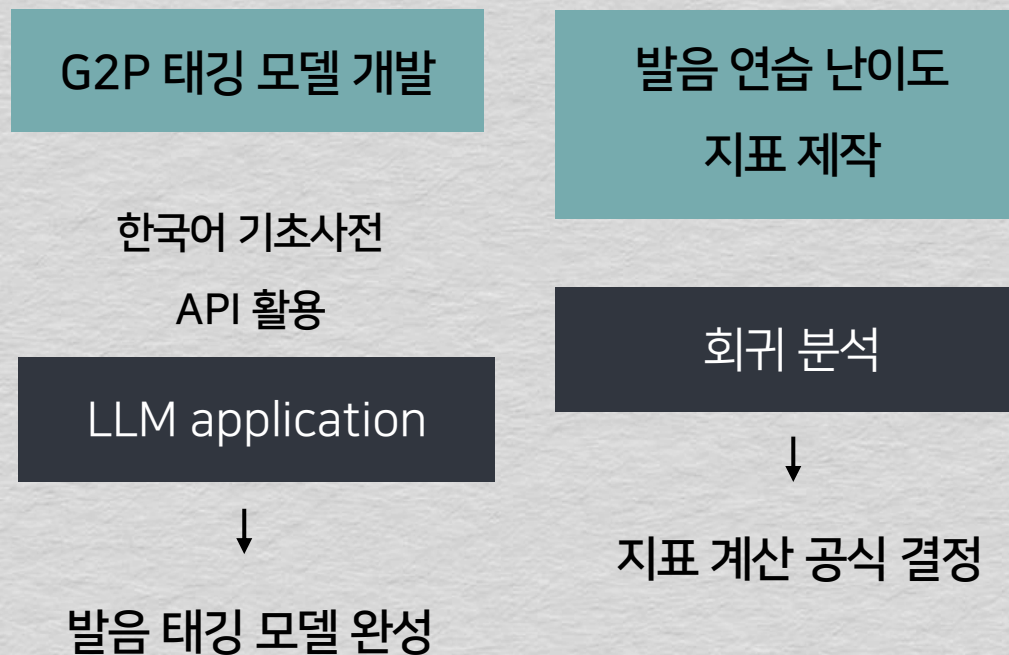
1~2주차 계획



# 1. 2주차 피드백 분석 및 구현 흐름

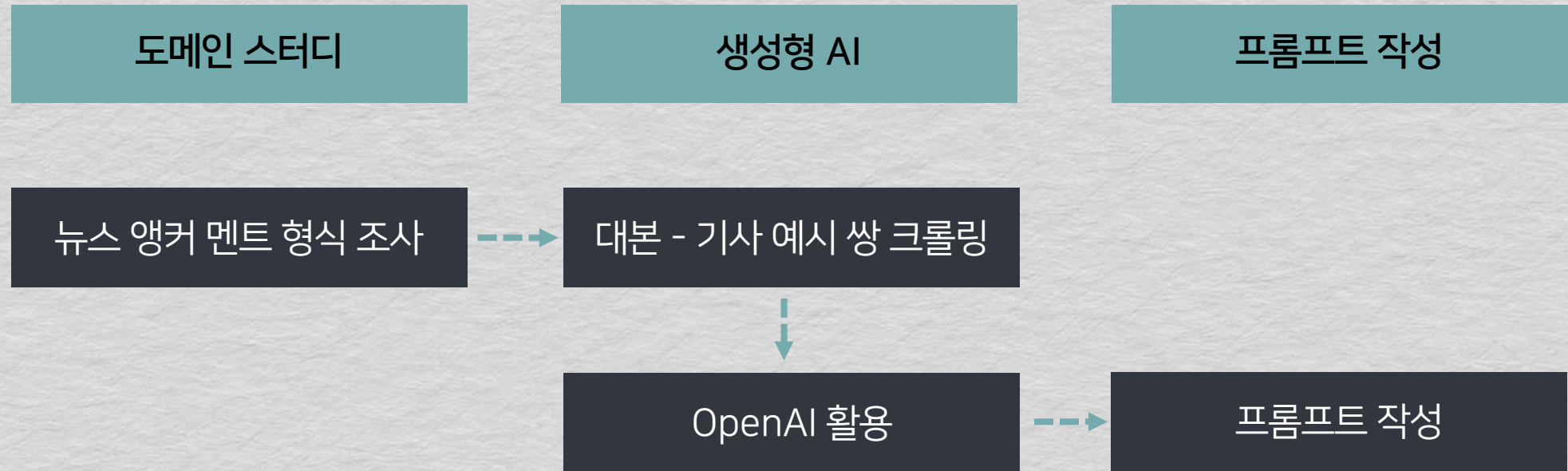


## 3~4주차 계획



## 1. 2주차 피드백

### 뉴스 대본 생성 흐름 리뷰





# 1. 2주차 피드백

## 뉴스 대본 생성 흐름 리뷰

	앵커 멘트	기사
1	임차인 보호를 위한 전세보증금 반환 보증 제도, 그런데 막상 전세 사기에서 이 보증이 무용지물이 되는 경우가 많다고 합니다. ...	전세사기 피해자가 아무 잘못을 하지 않았음에도 일방적으로 보증을 취소할 수 있도록 하는 주택도시보증공사(HUG)의 약관 조항에 대해 ...
2	우주항공청이 나사(NASA)와 공동 개발한 태양 망원경, 코덱스가 미국 플로리다 케네디 우주센터에서 우리 시간 어제 오전 11시 29분 발사됐습니다. ...	한국천문연구원이 미국 항공우주국(NASA)과 공동으로 개발한 태양 관측 망원경인 '코로나그래프(코덱스·CODEX)'가 ...
...	...	...



Fine tuning 과  
few-shot learning 진행

58개의 대본 - 기사 예시 쌍 크롤링



## 1. 2주차 피드백

### 뉴스 대본 생성 흐름 리뷰



너는 예비 아나운서들을 위해 뉴스 기사를 [앵커 멘트]형식으로 바꿔주는 생성형 AI야.

예시로 같은 주제 대해 같은 내용을 다루고 있는 앵커 멘트와 기사 쌍 58개가 있어.

이것들을 참고해서 아래의 조건에 맞게 작성해줘.

- 조건1. 이 기사는 뉴스에서 보도될 거야. 이 리포터 보도 앞에 나오는 [앵커 멘트]를 작성해줘.  
이 앵커 멘트는 흥미를 유발하거나, 요약하거나, 향후 소개된 리포트 내용을 이해하는데 도움이 되는 내용이어야 해.
- 조건2. 원본 기사에서 가장 중요하다고 생각하는 사항을 소개해줘. 만약 비중이 비슷한 여러 사항이 있을 경우 두세 정보를 모두 소개해줘.

...(중략)...

도메인 스터디를 통해 프롬프트 작성

## 1. 2주차 피드백

### 뉴스 대본 생성 흐름 리뷰

[실제 앵커 멘트] 도박 빚이 있던 40대가 일면식도 없는 운전자를 흥기로 살해하고, 시신을 유기한 사건에서 범행 후 피해자의 지갑에서 훔친 돈으로 로또를 구입한 사실이 드러났습니다.

경찰은 어젯밤 40대 남성 A씨에 대해 구속영장을 신청했으며, 로또를 사는 등 쓰고 남은 5만 7천 원을 압수했습니다.

[생성된 앵커 멘트] 주차장에서 처음 본 남성의 금품을 빼앗고 살해유기한 40대 남성이 검거됐습니다. 증거인멸을 위해 피해자 차량에 불까지 질렀는데, 빼앗은 돈 12만 원으로 일부 복권을 산 것으로 드러났습니다.

기사 예시 쌍과 작성한 프롬프트를 통해 뉴스 대본 생성

## 1. 2주차 피드백

### 뉴스 대본 생성 흐름 리뷰 | 파인 튜닝의 문제점

58개의 기사-대본 쌍으로 Fine-tuning 할 경우,  
데이터가 적어 **과적합의 가능성**을 배제할 수 없음

추후 서비스가 적용되면 대량의 대본 생성을 해야 하기 때문에  
**비용 문제**가 발생할 수 있음



파인 튜닝 없이, 프롬프트만으로 대본 생성을하기로 결정!



## 1. 2주차 피드백

### 뉴스 대본 생성 흐름 리뷰 | 파인 튜닝의 문제점

58개의 기사-대본 쌍으로 Fine-tuning 할 경우,  
데이터가 적어 **과적합의 가능성**을 배제할 수 없음



추후 사례  
대신 기존의 Few-shot learning의 프롬프트를  
**CoT(Chain of Thought)**로 개선하고자 함



파인 튜닝 없이, 프롬프트만으로 대본 생성을하기로 결정!

# 1. 2주차 피드백

## CoT 프롬프트



### 예시 기사에 대한 CoT

#### 1. 기사의 핵심 키워드 파악하기

... 가장 중요한 키워드는 '지방의회 의원들', '딥페이크', '경찰'이야.

#### 2. 전문 용어 설명하기

... 이 기사에서는 '딥페이크'가 전문 용어이고, 합성 가짜 음란물이라는 의미를 가져.

#### 3. 기사의 핵심을 파악하기

... 예를 들어 이 기사에서는, 지방의회 의원을 상대로 한 딥페이크 피해 사례가 전국적으로 확산하고 있다는 점과 경찰이 대응에 나섰다라는 점이야. ...

예시 기사-대본쌍을 주고, 5단계에 따라 기사에서 대본 형식으로 변환하는 과정과 예시를 설명

# 1. 2주차 피드백

## CoT 프롬프트



### 예시 기사에 대한 CoT

#### 4. 객관성과 중립성 유지하기

... 예를 들어, "날이 갈수록 확산하고 있습니다" 같은 문장에서 '날이 갈수록'이라는 표현은 감정적인 요소를 포함할 수 있어. 이를 '전국적으로 확산하고 있습니다'와 같은 객관적인 표현으로 수정해야 해.

#### 5. '하십시오'의 격식체로 작성하기

... 예를 들어, "경찰이 수사에 착수했다."를 "경찰이 수사에 착수했습니다"로 변경해서 정중하고 격식을 갖춘 표현으로 만들면 돼.

예시 기사-대본쌍을 주고, 5단계에 따라 기사에서 대본 형식으로 변환하는 과정과 예시를 설명



## 1. 2주차 피드백

### 뉴스 대본 생성 | case 비교

Zero shot	뉴스 앵커 대본 작성을 위한 기본 조건 6가지와 타깃 기사만 제공
One shot	뉴스 앵커 대본 작성을 위한 기본 조건 6가지와 타깃 기사 + 대본-기사 예시 쌍 1개 제공
One shot + CoT	뉴스 앵커 대본 작성을 위한 기본 조건 6가지와 타깃 기사 + <b>추론 과정이 포함된</b> 대본-기사 예시 쌍 1개 제공

3가지 경우에 대해서 결과물을 비교하고자 함

## 1. 2주차 피드백

### 뉴스 대본 생성 | case 비교 ② 경제 분야

[원본 기사 제목] 美금리 내려 4.5~4.75%...파월 "사퇴 요구받아도 안할것"

**[실제 앵커 멘트]** 미국 대선 직후 열린 통화정책 회의에서 지난 9월에 이어 또 금리 인하가 단행됐습니다. 파월 연준 의장은 트럼프가 사퇴를 요구해도 물러나지 않겠다며 트럼프 당선이 금리 정책에 영향을 미치지 않을 거라고 밝혔습니다.

앵커 대본의 특징을 기준으로 판단했을 때  
Zero-shot prompting, One-shot prompting의 결과보다  
One-shot CoT prompting의 결과가 실제 앵커 멘트와 더 유사함



## 1. 2주차 피드백

### 뉴스 대본 생성 | case 비교 ② 경제 분야

[원본 기사 제목] 美금리 내려 4.5~4.75%...파월 "사퇴 요구받아도 안할것"

One-shot CoT 프롬프트를 사회 분야 기사로 작성하여

[실제 앵커 멘트] 미국 대선 직후 열린 통화정책 회의에서 지난 9월에 이어 또 금리 인하가

수치 관련 내용 처리에 부족한 부분이 있었다고 판단

단행됐습니다. 파월 인준 의장은 트럼프가 사퇴를 요구해도 물러나지 않겠다고 트럼프 당선이 금리

정책에 영향을 미치지 않을 거라고 밝혔습니다.



CoT 프롬프트에 IT\_과학, 문화, 경제 분야 기사-대본 쌍을

예시로 추가하여 Few-shot learning 진행

Zero-shot prompting, One-shot prompting의 결과보다

One-shot CoT prompting의 결과가 실제 앵커 멘트와 더 유사함



## 1. 2주차 피드백

### 뉴스 대본 생성 | case 비교 ② 경제 분야

[원본 기사 제목] 美금리 내려 4.5~4.75%...파월 "사퇴 요구받아도 안할것"

[Few-shot CoT prompting] "미국 중앙은행인 연방준비제도가 기준금리를 또다시  
인하했습니다. 이번 인하는 지난 9월에 이어 두 번째로, 한국과 미국의 금리 격차가 1.5%로  
줄어들었습니다. 한편, 트럼프 당선인이 연준 구성원에 대한 인사를 바꿀 가능성이 제기되면서,  
통화정책에 대한 관심이 집중되고 있습니다."

One-shot CoT prompting의 결과에 비해  
수치 관련 내용이 간결해졌음을 확인

# 1. 2주차 피드백

## S2P 모델 개발

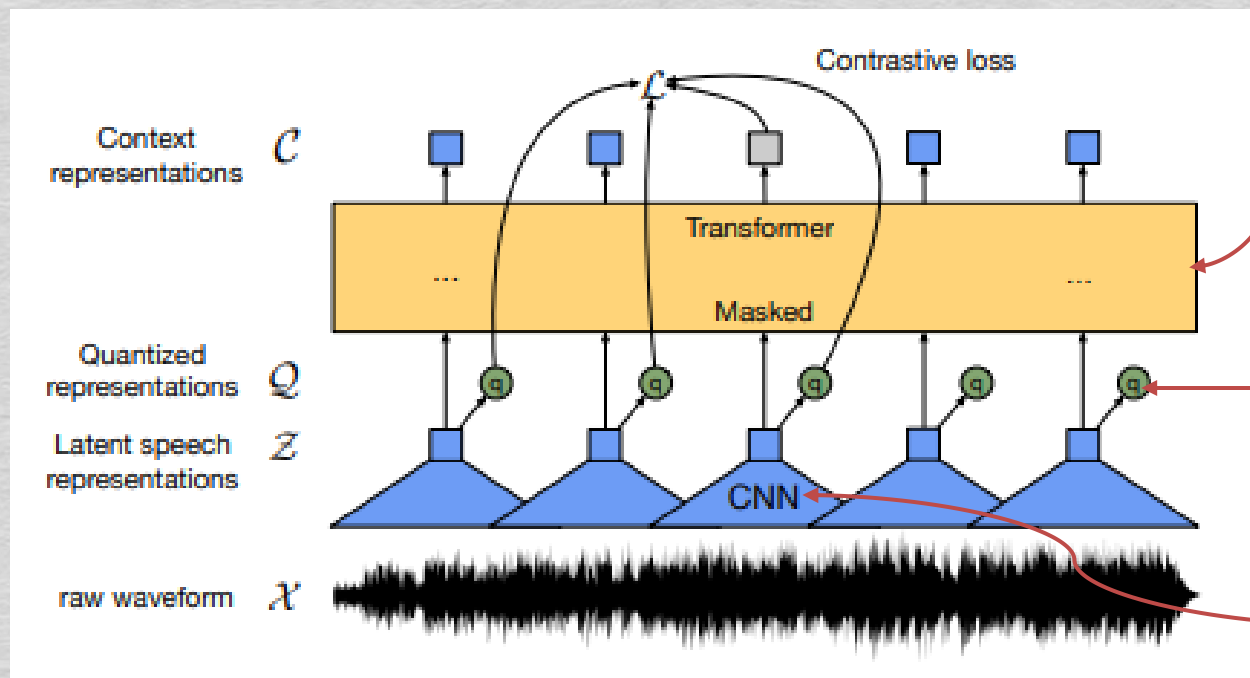
X : 문장을 읽은 음성 데이터 (16kHz)	문장의 음성적 표기 (발음열)
[0.0, -3.0517578125e-05, -6.103515625e-05, -3.051758125e-05, ... -6.103515625e-05]	이는 지난 시비월 십사일 개봉한 ...
[0.0, -3.0517578125e-05, 0.0, 0.0, ... -3.0517578125e-05]	반면교사로 삼아야 ...
...	...

주어진 음성을 음소열로 변환하기 위해 데이터셋을 준비 후

Wav2Vec 2.0 기반 모델을 파인튜닝 해 S2P task를 학습시키고자 함

# 1. 2주차 피드백

## S2P 모델 개발



[Wav2Vec 2.0 핵심 아키텍처]



Speech의 전체 Context가  
반영된 벡터 시퀀스를 만들어내는

**Transformer**



연속적인 특성을 갖는 벡터를  
이산적인 특성을 가진 벡터로  
매핑하는 **양자화 모듈**



음성 데이터를 잠재 벡터 시퀀스로  
변환하는 **Feature Encoder**



# 1. 2주차 피드백

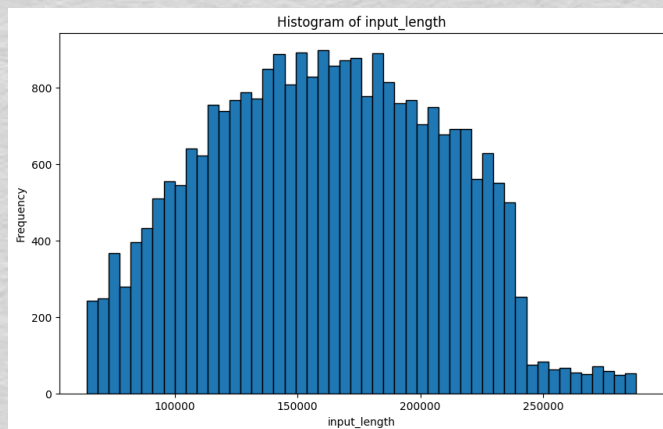
## S2P 모델 개발

X : 문장을 읽은 음성 데이터 (16kHz)	Y : 문장의 음성적 표기 (음소열)
[0.0, -3.0517578125e-05, -6.103515625e-05, -3.051758125e-05, ... -6.103515625e-05]	ㅇ   ㄴ — ㄴ ㅈ   ㄴ ㅈ ㄴ ...
[0.0, -3.0517578125e-05, 0.0, 0.0, ... -3.0517578125e-05]	ㅂ ㅈ ㄴ ㅁ ㅈ ㄴ ㄱ ㅍ ㅈ ㅈ ㄴ ㄴ ...
...	...

발음열을 총합 40개의 자음과 모음으로 분리해 얻은 음소열을  
Target sequence로 설정

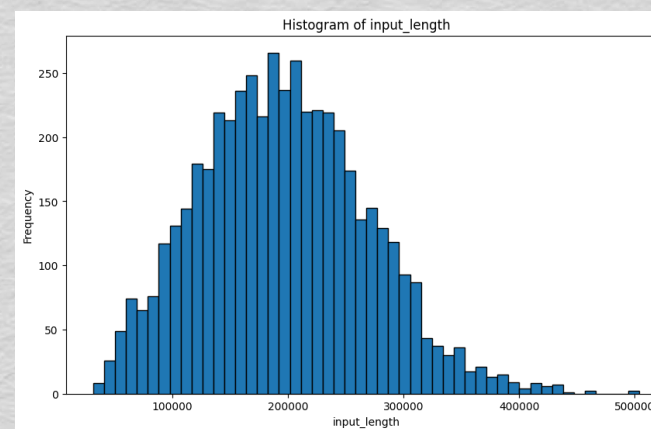
# 1. 2주차 피드백

## S2P 모델 개발 | ③ 데이터셋 재구축



훈련 셋 : 샘플 28217개

시간 길이 : 4초 이상 18초 미만



검증 셋 : 샘플 4980개

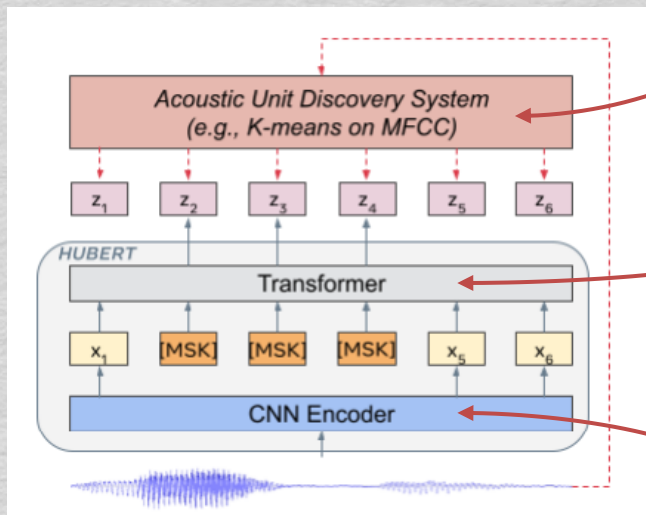
샘플 수와 음성 녹음 시간을 달리하여

훈련과 검증 데이터셋을 재구성



# 1. 2주차 피드백

## S2P 모델 개발 | 파인튜닝 재시도



[HuBERT 핵심 아키텍처]



Wav2Vec 2.0의 양자화 모듈을  
대체하는 **K-means 앙상블**



Speech의 전체 Context가  
반영된 벡터 시퀀스를 만들어내는  
**Transformer**



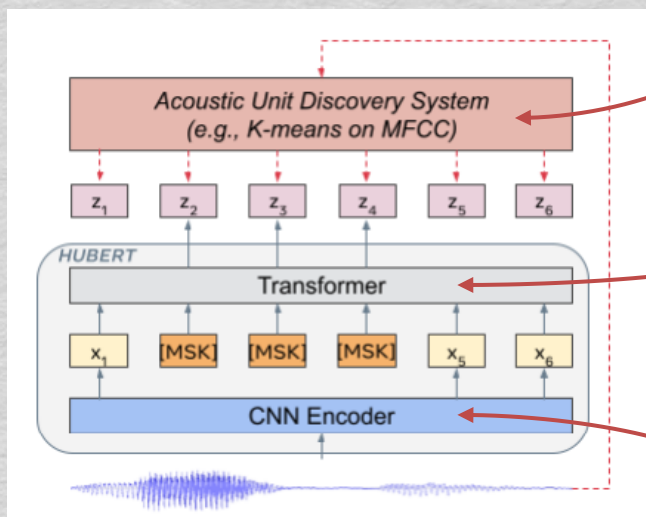
음성 데이터를 잠재 벡터 시퀀스로  
변환하는 **Feature Encoder**

부족한 훈련 데이터셋의 수를 극복하기 위해 **한국어로 4000시간**  
**음성 사전학습이 된 버전의 모델**이 존재하는 **HuBERT**로 모델 변경



# 1. 2주차 피드백

## S2P 모델 개발 | 파인튜닝 재시도



[HuBERT 핵심 아키텍처]



Wav2Vec 2.0의 양자화 모듈을  
대체하는 **K-means 앙상블**



Speech의 전체 Context가  
반영된 벡터 시퀀스를 만들어내는  
**Transformer**



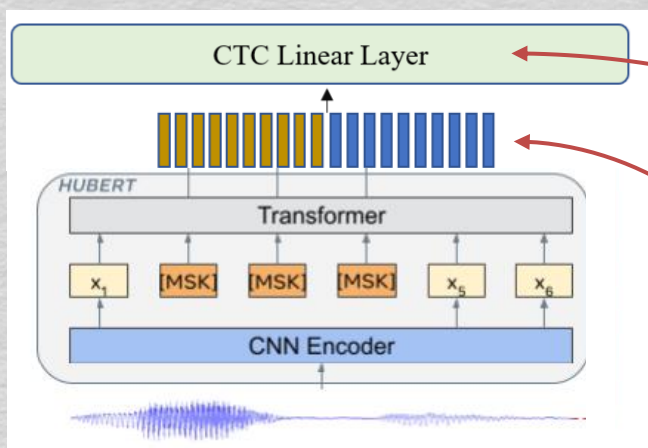
음성 데이터를 잠재 벡터 시퀀스로  
변환하는 **Feature Encoder**

이산화 방식과 사전학습의 Loss 함수를 제외할 때  
이는 Wav2Vec 2.0과 동일함

# 1. 2주차 피드백

## S2P 모델 개발 | 파인튜닝 재시도

CTC Loss를 최소화하는 방향으로 파인튜닝 중  
파라미터 최적화가 진행됨



[HuBERT for CTC]

출력 벡터를 토큰 집합의 크기와  
같은 수의 차원을 갖는 확률 벡터로  
선형 변환하는 레이어

Transformer의 출력 벡터 시퀀스

사전학습 된 HuBERT의 가중치를 초기치로 삼고

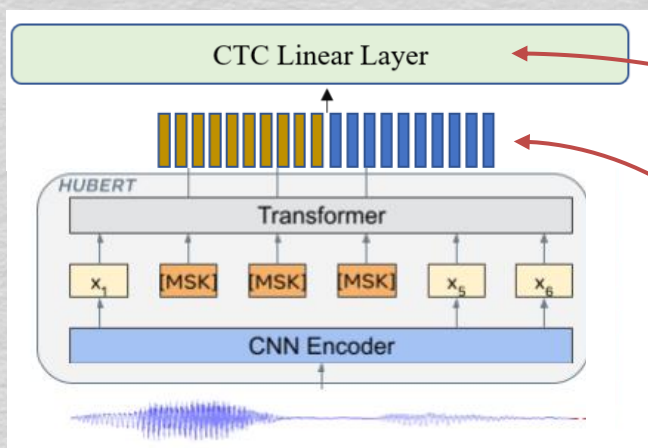
아키텍처 맨 위에 선형 레이어를 추가해 S2P task를 위한 파인튜닝을 진행



# 1. 2주차 피드백

## S2P 모델 개발 | 파인튜닝 재시도

CTC Loss를 최소화하는 방향으로 파인튜닝 중  
파라미터 최적화가 진행됨



[HuBERT for CTC]

출력 벡터를 토큰 집합의 크기와  
같은 수의 차원을 갖는 확률 벡터로  
선형 변환하는 레이어

Transformer의 출력 벡터 시퀀스

선형 레이어 뿐만 아니라 **트랜스포머 최상 층 레이어의 파라미터들까지** 업데이트 시도



# 1. 2주차 피드백

## S2P 모델 개발 | 파인튜닝 재시도

음성 데이터를 입력 받은 모델이  
내놓은 예측 발음열

PER	Prediction
0.014925	이저네는 특빠란 공무 여귀를 신청하는 경우가 아니라면 구니니나 경차
0.030675	여귀를 만들기 원서는 새로 사지를 찌거나 저거도 육개월 이내에 찌
0.052174	이저네는 위장 변장으로 오   바들 수 익끼 때무네 특꺼운 뿔테 양경은
0.014085	강까넨던 여권 사진 규정이 와놔됐찌만 여저니 주이가 필요암니다
0.037037	해외로 나갈때 업써서는 안되는 걸 바로 여귀니쵸
0.125	또 뿔테 양경을 차공애도 무방아다
0.061538	먼저 양쪽 귀가 반드시 보여야 한다는 규정이 삭제됐습니
0.051282	제복과 굼보게 차공도 허용됐습다
0.054545	그런데 외교부가 어제부터 일부규정을 와놔엿습다
0.060345	시보년 마네 부자핀 호프집 3 주인 사린 사건 버미네게 검차리 무기징년

주어진 문장에 KoG2P를 적용해 만든 정답 발음열  
(문법적으로 100% 정확한 건 아님)

Reference
이저네는 특빠란 공무 여귀를 신청하는 경우가 아니라면 구니니나 경차리더라도 여권 사지는 사보글 익꼬 찌거야만 햐습니다
여귀를 만들기 위해서는 새로 사지를 찌거나 저거도 육개월 이내에 찌근 사지를 제추래야 하는데 지금까지는 이 여권 사진 규정이 상당히
보닌 화기니 어려울 정도에 여권 사지는 해외에서 익꼭 심사알 때 부리이글 바들 수 익끼 때무닙니다
이저네는 위장 변장으로 오   바들 수 익끼 때무네 두꺼운 뿔테 양경은 가급쩍 지양아도록 돼 이썸습니다
강까넨던 여권 사진 규정이 와놔됐찌만 여저니 주이가 필요암니다
해외로 나갈 때 업써서는 안 되는 걸 바로 여귀니쵸
또 뿔테양경을 차공애도 무방암니다
먼저 양쪽 귀가 반드시 보여야 한다는 규정이 삭제됐습니다
제복과 굼보게 차공도 허용됐습니다
그런데 외교부가 어제부터 일부 규정을 와놔엿습니다
시보년 마네 불짜핀 호프집 여주인 사린사건 버미네게 검차리 무기징년글 내려달라고 재팘부에 요청엿습니다

Test Sample size : 4627

음성이 학습에 사용되지 않은 새로운 발화자인

아나운서 교육생 11명의 음성 데이터를 사용하여 S2P 성능 평가를 다시 진행

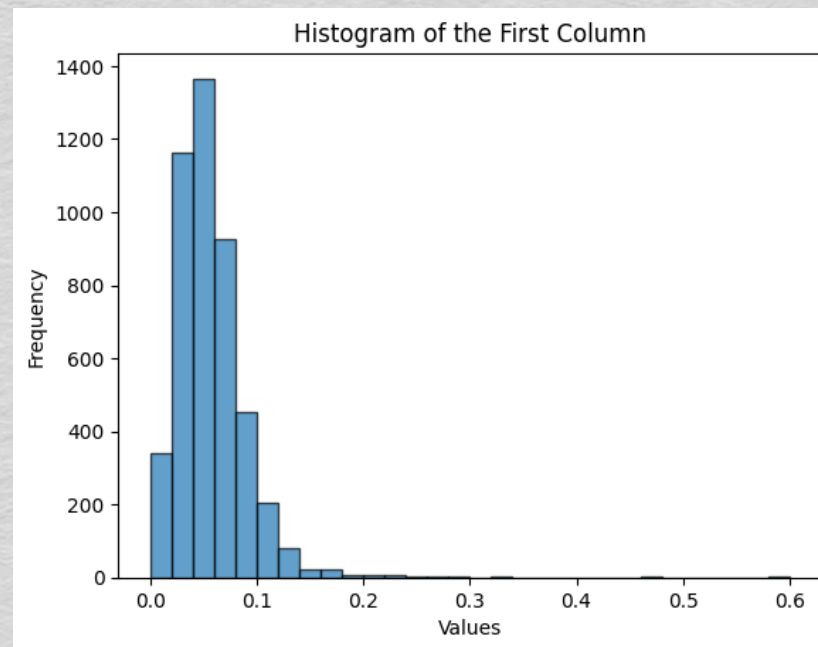
# 1. 2주차 피드백

## S2P 모델 개발 | 파인튜닝 재시도

PER	Prediction
0.014925	이저네는 특빠란 공무 여귀늘 신청아는 경우가 아니라면 구니니나 경차i
0.030675	여귀늘 만들기 원서는 새로 사지를 찌거나 저거도 육ㄱ개월 이내에 찌근
0.052174	이저네는 위장 변장으로 오   ㅁ바들 수 익끼 때무네 특꺼운 뽉ㅌ 양경은
0.014085	강까넨뎌 여권 사진 규정이 와놔땡찌만 여저니 주이가 필요암니다
0.037037	해외로 나갈때 업써서는 안되는 건 바로 여귀니쵸
0.125	또 뽉태 양경을 차공애도 무방아다
0.061538	먼저 양쪽 귀가 반드시 보여야 한다는 규정이 삭제됐습ㄴ
0.051282	제복과 굼보게 차공도 허용됐습다
0.054545	그런데 외교부가 어제부터 일부규정을 와놔얏습다
0.060345	시보년 마네 부자핀 호프찹 ㅈ 주인 사린 사건 버미네게 검차리 무기징녀

Avg PER : 0.05

준수한 예측 성능을 확인



PER distribution

02

음성 임베딩



## 2. 음성 임베딩

S2P를 이용해 산출해낸 PER 이외에도 발음 정확도 점수로  
사용할 만한 지표가 있는지 확인



아나운서와 아마추어의 음성 임베딩 벡터를 비교할 수는 없을까?

## 2. 음성 임베딩

S2P를 이용해 산출해낸 PER 이외에도 발음 정확도 점수로  
사용할 만한 지표가 있는지 확인



두 임베딩 벡터 시퀀스 간의 거리를 계산해보자!





## 2. 음성 임베딩

### 스피치 임베딩 | 음성 녹음



아나운서 음성 데이터는 있으나,  
동일한 대본을 읽는 일반인 음성 데이터가 없으므로 직접 녹음하기로 결정






## 2. 음성 임베딩

### 스피치 임베딩 | 음성 녹음



**모든 녹음 항목** 죽음의 녹음 시작...

🔍 검색 

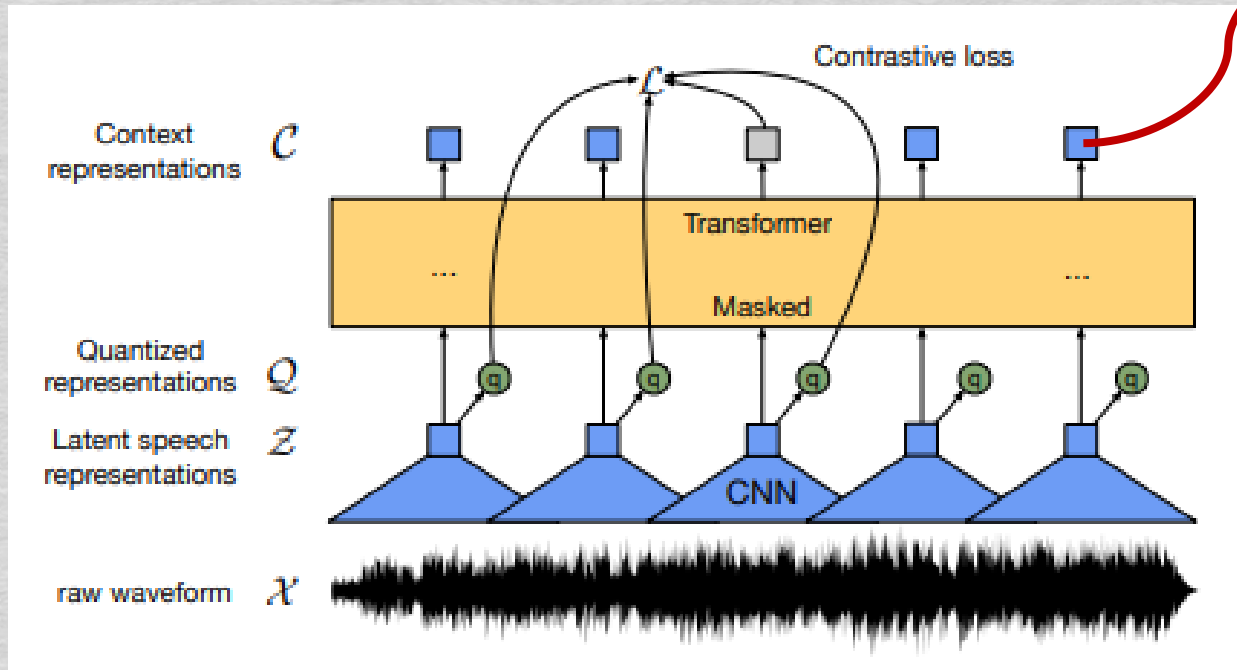
 " "라고 말하십시오. 

SPK001YTNSO487M005	0:11
월요일	
SPK001YTNSO487M004	0:15
월요일	
SPK001YTNSO487M003	0:17
월요일	

선택한 음성 파일 500개에 대해 각자 100개씩 녹음

## 2. 음성 임베딩

Wav2vec 2.0 transformer output



Transformer 최종 출력값



입력값(음성데이터)에 대한  
임베딩 벡터 시퀀스



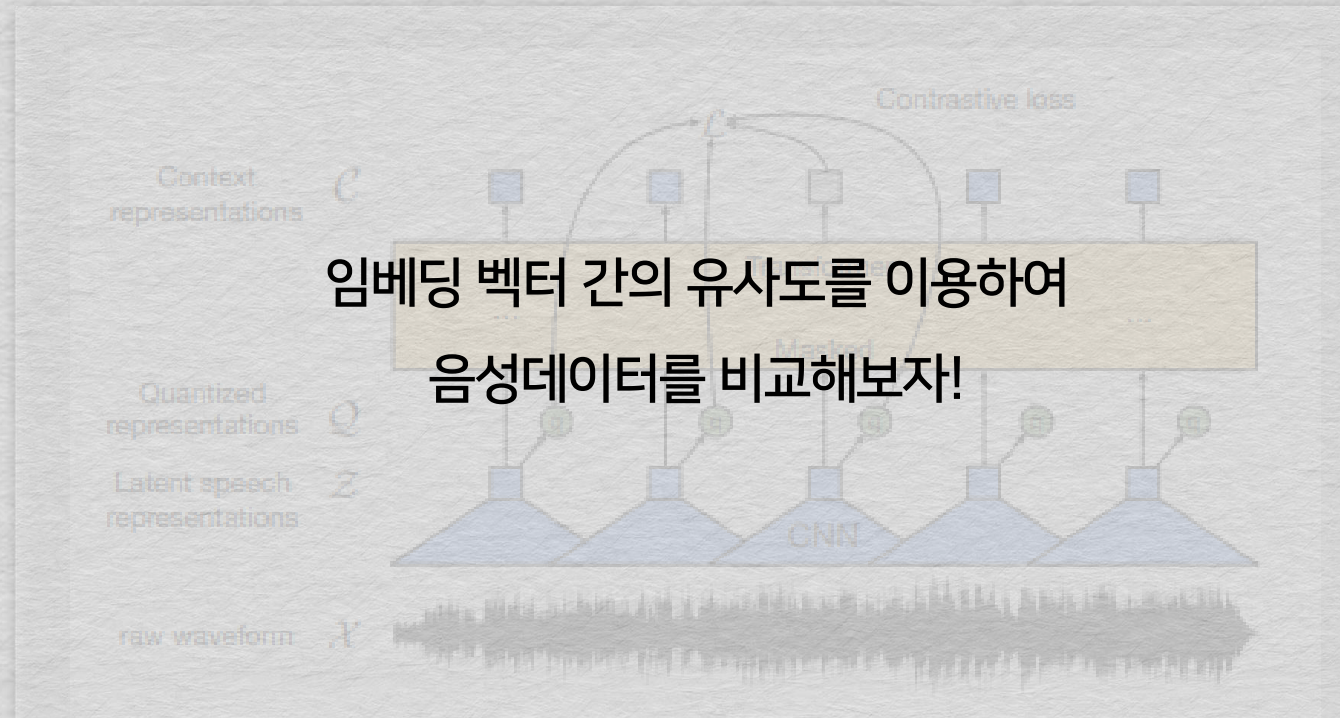
음성의 음향적 특성을  
반영하고 있음





## 2. 음성 임베딩

Wav2vec 2.0 transformer output



Transformer 최종 출력값



입력값(음성데이터)에 대한  
임베딩 벡터 시퀀스



음성의 음향적 특성을  
반영하고 있음



## 2. 음성 임베딩

DTW

일반적으로 시계열과 같은 시퀀스 데이터의 유사도를 계산할 때는  
유클리디안 거리, DTW 등의 지표를 활용



비교하고자 하는 데이터의 길이가 서로 다르다면..?





## 2. 음성 임베딩

### DTW

DTW(Dynamic Time Warping)

시계열 형태의 시퀀스 데이터 간의 유사성을 측정하는 알고리즘

문장의 길이나 발화 속도가 달라도, 이것을 고려하여 유사성을 측정할 수  
있기 때문에 발화 데이터 비교에 유용하게 쓰일 수 있을 것으로 판단

## 2. 음성 임베딩

### DTW | DTW거리 계산

File_name	Embedding
SPK056YTNEC188M001.m4a	[[-7.9821e-02, -1.5583e-01, 3.7928e-01, ..., 3.4168e-04, 1.0458e-01, 3.4776e-01], [-5.6692e-02, -1.9426e-01, 3.6262e-01, ..., -4.8365e-02, 9.5932e-02, 3.8288e-01], ..., [-5.6692e-02, -1.9426e-01, 3.6262e-01, ..., -4.8365e-02, 9.5932e-02, 3.8288e-01]]
SPK056YTNEC188M002.m4a	...
...	...
SPK003YTNSO243F006.m4a	[[-4.2790e-01, -4.4997e-01, 4.5860e-01, ..., -1.2737e-01, 2.1961e-01, -5.1946e-01], ..., [-4.7955e-01, -2.7390e-01, 7.1928e-02, ..., -7.8745e-02, 2.2913e-01, -3.2672e-01]]

File_name	Embedding
SPK056YTNEC188M001.wav	[[-0.0434, -0.1400, 0.5398, ..., 0.0285, 0.0699, 0.3954], [0.0423, -0.1907, 0.4688, ..., -0.0421, 0.0671, 0.3927], ..., [0.0081, -0.3055, 0.3829, ..., -0.0869, 0.1008, 0.3407]]
SPK056YTNEC188M002.wav	...
...	...
SPK003YTNSO243F006.wav	[[-0.3806, -0.3076, 0.4051, ..., -0.0778, 0.2498, -0.3241], [-0.3906, -0.3173, 0.4713, ..., -0.0443, 0.2455, -0.4257], ..., [-0.5052, -0.2719, -0.0030, ..., 0.0012, 0.2361, -0.3149]]

직접 녹음한 발화데이터에 대한  
임베딩 벡터 시퀀스



## 2. 음성 임베딩

### DTW | DTW거리 계산

File_name	Embedding
SPK056YTNEC188M001.m4a	[[ -7.9821e-02, -1.5583e-01, 3.7928e-01, ..., -3.4168e-04, 1.0458e-01, 3.4776e-01], [-5.6692e- 02, -1.9426e-01, 3.6262e-01, ..., -4.8365e-02, 9.5932e-02, 3.8288e-01], ..., [-5.6692e-02, 1.9426e-01, 3.6262e-01, ..., -4.8365e-02, 2.88e-01]]
SPK056YTNEC188M002.m4a	...
...	...
SPK003YTNSO243F006.m4a	[[ -4.2790e-01, -4.4997e-01, 4.5860e-01, ..., 1.2737e-01, 2.1961e-01, -5.1946e-01], ..., [-4.7955e-01, -2.7390e-01, 7.1928e-02, ..., -7.8745e-02, 2.2913e-01, -3.2672e-01]]

전직 아나운서가 대본을 읽은  
발화 데이터에 대한  
... 임베딩 벡터 시퀀스 ...

File_name	Embedding
SPK056YTNEC188M001.wav	[[ 0.0434, -0.1400, 0.5398, ..., 0.0285, 0.0699, 0.3951], [ 0.0423, -0.1907, 0.4688, ..., -0.0421, 0.0671, 0.3927],... [ 0.0081, -0.3055, 0.3829, ..., -0.0869, 0.1008, 0.3407]]
SPK056YTNEC188M002.wav	...
...	...
SPK003YTNSO243F006.wav	[[ -0.3506, -0.3076, 0.4051, ..., -0.0778, 0.2498, -0.3241], [-0.3906, -0.3173, 0.4713, ..., -0.0443, 0.2455, -0.4257],... [-0.5052, -0.2719, -0.0030, ..., -0.0012, 0.2361, -0.3149]]

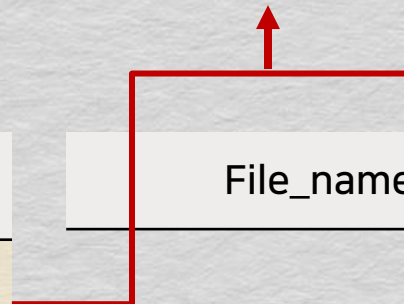
## 2. 음성 임베딩

### DTW | DTW거리 계산

두 임베딩 시퀀스 간의 DTW 거리 계산!

File_name	Embedding
SPK056YTNEC188M001.m4a	[[-7.9821e-02, -1.5583e-01, 3.7928e-01, ..., 3.4168e-04, 1.0458e-01, 3.4776e-01], [-5.6692e-02, -1.9426e-01, 3.6262e-01, ..., -4.8365e-02, 9.5932e-02, 3.8288e-01], ..., [-5.6692e-02, -1.9426e-01, 3.6262e-01, ..., -4.8365e-02, 9.5932e-02, 3.8288e-01]]
SPK056YTNEC188M002.m4a	...
...	...
SPK003YTNSO243F006.m4a	[[-4.2790e-01, -4.4997e-01, 4.5860e-01, ..., -1.2737e-01, 2.1961e-01, -5.1946e-01], ..., [-4.7955e-01, -2.7390e-01, 7.1928e-02, ..., -7.8745e-02, 2.2913e-01, -3.2672e-01]]

File_name	Embedding
SPK056YTNEC188M001.wav	[[ 0.0434, -0.1400, 0.5398, ..., 0.0285, 0.0699, 0.3951], [ 0.0423, -0.1907, 0.4688, ..., -0.0421, 0.0671, 0.3927], ..., [ 0.0081, -0.3055, 0.3829, ..., -0.0869, 0.1008, 0.3407]]
SPK056YTNEC188M002.wav	...
...	...
SPK003YTNSO243F006.wav	[[-0.3506, -0.3076, 0.4051, ..., -0.0778, 0.2498, -0.3241], [-0.3906, -0.3173, 0.4713, ..., -0.0443, 0.2455, -0.4257], ..., [-0.5052, -0.2719, -0.0030, ..., -0.0012, 0.2361, -0.3149]]





## 2. 음성 임베딩

### DTW | DTW거리 계산 결과

산출된 DTW 거리를

발화자의 발음 정확도 점수로 삼기로 결정!

File_name	Text	DTW_distance
SPK056YTNEC188M001	지난해 상장폐지 직전까지 갔다가 회생 기회를 얻은 경남제약이 다시 상장폐지 위기에 ...	38.576573
SPK056YTNEC188M002	레모나로 유명한 경남제약은 감사의견 한정을 받아 상장폐지 사유가 발생하자 김주선 대...	46.494312
...	...	...
SPK003YTNSO243F005	이에 대해 춘천소년원 측은 십대의 경우 대장암 발병이 흔치 않은 데다 이 군이 당시...	40.794254
SPK003YTNSO243F006	이와 관련해 전문의들은 청소년 대장암 발병을 의심하기 어려운 것은 사실이지만 이 군...	50.403858

# 03

## GZP



### 3. G2P

#### G2P

G2P (Grapheme to Phoneme)

텍스트 형태의 문자를 음성(음소)으로 변환하는 기술

이 프로젝트에서 G2P를 사용하는 이유



발음 주석 태깅을 위해



S2P 변환에 쓰일 더 정확한 발음열을 만들기 위해

S2P 변환에 쓰일 target sequence가 발음열

### 3. G2P

#### G2P

G2P (Grapheme to Phoneme)

텍스트 형태의 문자를 음성(음소)으로 변환하는 기술

이 프로젝트에서 G2P를 사용하는 이유



G2P를 처음부터 만들기는 어려우므로  
이미 만들어진 KoG2P를 활용하기로 함!

S2P 변환에 쓰일 target sequence가 발음열



### 3. G2P

KoG2P

KoG2P

한국어의 문자열로부터 발음열을 생성하는 파이썬 기반 G2P 패키지

Input text: 윤아가 예감을 먹는다



Output text: 유나가 예가들 멍는다

### 3. G2P

#### KoG2P의 문제점

##### KoG2P 결과의 문제점

1. 장음, 단음을 구별하지 못함
2. 첫 음절이 아닌 음절 초성의 'ㅎ'이 사라짐  
EX) '보호' -> [보오]
3. 'ㄴ' 받침을 'ㄹ'이나 'ㅇ'으로 변환  
EX) '권고' -> [굉고]
4. 띄어쓰기가 있는 경우 음운 변동이 적용되지 않음
5. 문장 부호가 사라짐
6. 숫자를 처리하지 못함
7. 영어를 처리하지 못함





### 3. G2P

#### KoG2P의 문제점

##### KoG2P 결과의 문제점

1. 장음, 단음을 구별하지 못함
2. 첫 음절이 아닌 음절 초성의 'ㅎ'이 사라짐  
EX) '보호' -> [부호]
3. 'ㄴ' 받침을 'ㄹ'이나 'ㅇ'으로 변환  
EX) '권고' -> [경고]
4. 띄어쓰기가 있는 경우 음운 변동이 적용되지 않음
5. 문장 부호가 사라짐
6. 숫자를 처리하지 못함
7. 영어를 처리하지 못함

**KoG2P를 직접 개선해보기로!**

두...



### 3. G2P

#### KoG2P 개선 (3) 규칙 기반

##### 1. 'ㄴ' 받침 변환 현상

EX) 권고[권:고] -> [꺠고], 망원경[망:원경] -> [망꺠경]

##### 2. 한자어에서의 된소리 변환 반영 X

EX) 효율적으로[효율찌그로] -> [효율꺠그로]

##### 3. 용언의 활용형 속 '저, 처, 찌'는 [저, 처, 찌]로 변환 X

EX) 가져[가저] -> [가꺠], 다쳐[다처] -> [다꺠]

##### 4. 음절의 끝소리 규칙이 부분적으로 반영 X

EX) 꽃 한 송이[꼬 탄 송이] -> [꼰 한 송이], 밭 아래[바 다래] -> [반 아래]

...

KoG2P가 잘못 변환했던 발음들을 바탕으로 비슷한 오류 유형을 정리



### 3. G2P

#### KoG2P 개선 (3) 규칙 기반

##### 1. 숫자와 영어 처리

EX) 'IMF는 2023년에 세계 경제를 전망한다고 발표했습니다.'

기존 KoG2P: [는 녀네 세계 경제를 점망안다고 발표앤쌔니다.]

개선 결과: [아이엠에프는 이쳐니십쌔년에 세계 경제를 전망한다고 발표했습니다.]

##### 2. 첫 음절이 아닌 음절에서 초성이 'ㅎ'일 때 'ㅎ'이 사라지는 현상을 복원

EX) 전투화: [전투와] -> [전투화]

KoG2P가 처리하지 못한 단어들을 규칙 기반으로 처리

### 3. G2P

#### KoG2P 개선 (3) 규칙 기반

3. 종성 'ㄴ'이 초성 'ㄱ, ㅂ'을 만나면 'ㅇ, ㅁ'으로 바뀌는 현상을 복원

EX) '신발': [심발] -> [신발], '권고': [꺨고] -> [권고]

4. 음절의 끝소리 규칙 후 비음화가 잘못 적용된 현상을 원래 규정대로 복원

EX) '뱃머리': [뱌머리] -> [뱌머리]

5. 한자어에서 'ㄹ' 뒤 'ㄷ, ㅅ, ㅈ'는 된소리로 발음

EX) 효율적: [효율적] -> [효율적]

KoG2P가 처리하지 못한 단어들을 규칙 기반으로 처리



### 3. G2P

#### KoG2P 개선 (3) 규칙 기반

##### 규칙 기반 개선의 한계점

1. 용언의 활용형 '저, 쩌, 처'는 '저, 찌, 처'로 발음 반영 안 됨
2. 띄어쓰기가 되어 있는 문장에서 실질형태소가 올 때의 연음과 형식형태소가 올 때의 연음을 구분하지 못함
3. 장단음 구분 못함
4. 함수를 하나씩 실행할 땐 제대로 변환이 됐지만 한꺼번에 실행했을 때 부분적으로 변환이 안 되는 문제 발생



### 3. G2P

#### KoG2P 개선 (3) 규칙 기반

##### [원본 문장]

2020년부터 2년간 투자자 6천여 명에게 천4백억 원을 가로챈 사기 일당이 붙잡혔습니다.  
비상장주식에 투자하면 500%까지 수익을 얻을 수 있다고 속인 걸로 드러났습니다.



##### [기존 KoG2P 실행 결과]

년부터 년간 투자자 천여 명에게 천배억 위를 가로챈 사기 일당이 붙잡혔습니다 비상장주시게  
투자아면 까지 수이글 어들 수 인따고 소긴 걸로 드러났습니다

아라비아 숫자로 적힌 부분이 사라짐, 'ㅎ' 사라짐 등의 문제점을 확인



### 3. G2P

#### KoG2P 개선 (3) 규칙 기반

##### [원본 문장]

2020년부터 2년간 투자자 6천여 명에게 천4백억 원을 가로챈 사기 일당이 붙잡혔습니다.

비상장주식에 투자하면 500%까지 수익을 얻을 수 있다고 속인 걸로 드러났습니다.



##### [개선한 KoG2P 실행 결과]

이치니심년부터 이년간 투자자 육천여 명에게 천사배격 워늘 가로챈 사기 일당이 붙잡혔습니다.

비상장주시게 투자하면 오백퍼센트까지 수이글 어들 수 인따고 소빈 걸로 드러났습니다.

숫자가 한국어 발음으로 변환되었고, 사라진 'ㅎ' 이 복원되었음을 확인

# 04

## 발음 난이도 지표 제작



## 4. 발음 난이도 지표 제작

### 설명변수 지정 | 발음 난이도 정의

올바른 발음 방법을 모르는 경우

ex) 장단음, ㄴ첨가

발음 난이도 = 길이 관련 변수 + 물리적 난이도 + 지식적 난이도

→ 제대로 발음하기 힘든 경우 ex) 간장공장공장장...



선행 연구와 국립국어원의 표준 발음법, EDA 결과를 고려해 설명변수를 선정

## 4. 발음 난이도 지표 제작

### 설명변수 지정 | 물리적 어려움 ①

#### 음운 반복 지수

반복되는 음운 비율 x 반복 평균 횟수

↘ 두 번 이상 나온 음운의 비율

Sentence = '간장공장공장장'

음운 반복 횟수 : {'ㄱ': 3, 'ㄴ': 1, 'ㅈ': 4, 'ㅇ': 6, 'ㅏ': 5, 'ㅑ': 2})

반복되는 음운 비율 =  $(3+4+6+5+2) / (3+1+4+6+5+2)$

반복 평균 횟수 :  $(3+4+6+5+2) / 5$

**Score : 3.81**



## 4. 발음 난이도 지표 제작

### 설명변수 지정 | 물리적 어려움 ②

#### 유사 음운 지수

유사한 음운이 반복되는 횟수

'ㄱ': ['ㄱ', 'ㄲ'] 처럼 모든 음운에 대해 유사 음운을 매핑함

'ㄱ': ['ㄱ', 'ㄲ'], 'ㄴ': ['ㄴ', 'ㄷ'] 처럼 유사 음운들을 매핑한 후

유사 음운 등장 횟수를 count함

Sentence = '도토리가 문을 도로록, 드르록, 두루록 열었는가? 드로록, 두루록, 두르록 열었는가.'

음운 반복 횟수: {'ㄷ': 7, 'ㅌ': 1, 'ㄹ': 16, 'ㄱ': 9, 'ㅁ': 1, 'ㄴ': 5, 'ㅇ': 5, 'ㅅ': 2, 'ㄴ': 7, 'ㄷ': 8, 'ㄴ': 9, 'ㅣ': 1, 'ㅏ': 3, 'ㅑ': 2, 'ㅓ': 2}

Score : 6.01



## 4. 발음 난이도 지표 제작

설명변수 지정: 물리적 어려움 ②

반복 음운 지수와 유사 음운 지수를 동시에 고려한 이유

유사 음운 지수

유사한 음운이 반복되는 횟수

Sentence = '간장공장공장장은 강공장공장장이고 콩공장공장장은 콩공장공장장이다.'

유사 음운 지수 : 0.09, 반복 음운 지수 : 9.67

Sentence = '도토리가 문을 도로록, 드르륵, 두루룩 열었는가? 드로록, 두루룩, 두르륵 열었는가.'

유사 음운 지수 : 0.69, 반복 음운 지수 : 6.01

Sentence = '도토리가 문을 도로록, 드르륵, 두루룩 열었는가? 드로록, 두루룩, 두르륵 열었는가.'

평균 : 0.22

음운 반복 횟수: ['다': 7, '니': 1, '로': 16, '느': 9, '고': 1, '나': 5, '이': 5, '사': 2, '노': 7, '누': 6,

'ㅡ': 9, 'ㅣ': 1, 'ㅏ': 3, 'ㅑ': 2, 'ㅓ': 2}

언뜻 보면 비슷한 유형인 것 같은 문장도

Score : 6.01

지수의 대소가 상반되게 도출됨



## 4. 발음 난이도 지표 제작

### 설명변수 지정

$$\text{발음 난이도} = \text{길이 관련 변수} + \text{물리적 난이도} + \text{지식적 난이도}$$

길이 관련 변수	물리적 난이도	지식적 난이도
문장 길이 변수	음운 반복 지수	비음화 난이도 변수 1
어절 길이 변수	유사 음운 지수	비음화 난이도 변수 2
받침의 난이도 변수	단모음 난이도 변수	유음화 난이도 변수
	이중모음 난이도 변수	경음화 난이도 변수
	유기음화 난이도 변수	

최종적으로 각 문장 별 발음 난이도 지표 리스트 산출

## 4. 발음 난이도 지표 제작

### 설명변수 지정

Sentence = '강원도소방본부에 따르면 이날 새벽 6시 44분쯤 홍천군 서석면 서울양양고속도로 서울 방향 서석터널 입구에서 4중 추돌 사고가 났습니다.'



[17, 3.41, 0, 6.44, 0.58, 0,  
0, 0, 3, 2, 0, 0]

Sentence = '트럼프가 경제 외적인 문제에도 관세 카드를 수시로 꺼내겠다 의지를 드러내면서 한국 역시 방위비 분담금 문제 등에서 관세가 협상 카드로 동원될 가능성에 대비해야 한단 지적도 나옵니다.'



[26, 3.0, 0, 7.65, 0.43, 3,  
0, 0, 4, 1, 0, 0]

각 문장마다 12개의 난이도 지표에 대한 output 도출



## 4. 발음 난이도 지표 제작

### 통합 데이터셋 ②

DTW Distance	sen_length	chars_per_words	count_double_final_consonants	repeated	similarity	...
38.5765727112995	180	3.15	2	38.22	0.47	...
46.49431231257330	87	3.2	0	22.19	0.5	...
64.81745801901350	84	3.52	4	23.9	0.48	...

반응변수가 DTW인 데이터셋

## 4. 발음 난이도 지표 제작

### 통합 데이터셋

상훈-1, 동길-2, 윤아-3, 수진-4, 민서-5

DTW Distance	sen_length	chars_per_words	count_double_final_consonants	...	Group	...
38.5765727112995	180	3.15	2	...	1	...
46.49431231257330	87	3.2	0	...	1	...
64.81745801901350	84	3.52	4	...	2	...

이때, 발화자 구분을 위해 group 열 추가



## 4. 발음 난이도 지표 제작

### Linear Mixed Model

#### LMM(Linear Mixed Model)

고정 효과(fixed effects)와 무작위 효과(random effects)가 모두 포함된 모델



$$Y = X\beta + Z\gamma + \epsilon$$

fixed effects   random effects

발음 정확도 점수 데이터는 소수의 개인으로부터 반복 측정되어 얻어진 데이터이므로  
데이터 내 자기상관을 고려해야 함

## 4. 발음 난이도 지표 제작

### Linear Mixed Model

#### LMM(Linear Mixed Model)

고정 효과(fixed effects)와 무작위 효과(random effects)가 모두 포함된 모델



$$Y = X\beta + Z\gamma + \epsilon$$

fixed effects   random effects

Random effect를 모델에 포함하는 것이 correlation 구조를 반영하는 핵심!



## 4. 발음 난이도 지표 제작

### Linear Mixed Model

통합 데이터셋 ②의 Group 변수를 Random effect로,  
앞서 사용한 12개의 설명변수를 Fixed effect로 설정한 뒤  
DTW 거리를 종속변수로 하여 LMM에 적합

## 4. 발음 난이도 지표 제작

Linear Mixed Model(log transformed) | 모델 적합 결과 ① 계수 추정치, R-squared

Fixed Effects	Estimated coefficients
Intercept	1.915435
Sen_length	0.038245
Chars_per_words	0.119442
Count_double_final_consonants	0.015098
Repeated	0.080393
Similarity	0.289449
Count_vowel_sequence	-0.003865

Fixed Effects	Estimated coefficients
Count_je_rye	0.022773
count_h_after_bdjg	0.039775
gyeong_eum	0.008353
bi_eum_1	0.026137
bi_eum_2	-0.006532
yu_eum	0.023435

R-squared : 0.7960



## 4. 발음 난이도 지표 제작

Linear Mixed Model(log transformed) | 모델 적합 결과 ② p-value

Fixed Effects	Estimated coefficients
<b>Intercept</b>	<b>1.915435</b>
<b>Sen_length</b>	<b>0.038245</b>
<b>Chars_per_words</b>	<b>0.119442</b>
Count_double_final_consonants	0.015098
<b>Repeated</b>	<b>0.080393</b>
<b>Similarity</b>	<b>0.289449</b>
Count_vowel_sequence	-0.003865

Fixed Effects	Estimated coefficients
Count_je_rye	0.022773
<b>count_h_after_bdjg</b>	<b>0.039775</b>
gyeong_eum	0.008353
<b>bi_eum_1</b>	<b>0.026137</b>
bi_eum_2	-0.006532
yu_eum	0.023435

P-value가 0.05 미만은 독립변수는 위 6개인 것을 확인

## 4. 발음 난이도 지표 제작

Linear Mixed Model(log transformed) | 최종 회귀식

모델 설명력과 잔차 plot 결과를 바탕으로 모델이 데이터의 구조를  
잘 설명하고 있다는 결론을 내림





## 4. 발음 난이도 지표 제작

Linear Mixed Model(log transformed) | 최종 회귀식

$$y = 1.915435 + 0.038245x_1 + 0.119442x_2 + 0.080393x_3 + \\ 0.289449x_4 + 0.0399775x_5 + 0.026137x_6$$

앞서 선택한 6개의 독립변수로 최종 회귀식 완성!

# 05

## 결론 및 제언



## 5. 결론 및 제언

### 성과 요약



규칙 기반 G2P(자소-음소 변환) 모델 구현



발화 음성 데이터를 발음 나는 대로 텍스트로 변환하는 S2P 구현



발음 난이도 지표 생성

## 5. 결론 및 제언

### 한계점

규칙 기반 G2P(자소-음소 변환) 모델에서 발생하는  
불규칙한 발음 오류를 완벽히 개선할 수는 없었음

S2P를 하기 전 KoG2P만을 이용하여 발음열 라벨을 만들었기 때문에  
문법적으로 정확한 출력을 만들어낼 수 없었음

발음 난이도 지표를 생성할 때  
형태소 분석이 필요한 문법 변동을 포함할 수 없었음



- 자연어처리팀 주제분석 -

감사합니다