

CONTENTS



001 과제 배경 및 목표



002 외래어 기반 G2P 개발의 필요성



003 개발 방향성 및 진행 방식



004 개발 언어, 도구 및 환경



004 개발 일정 및 역할 분담

1 과제의 배경 및 목표

과제의 배경

한국어 TTS 모델은 외래어와 한/영 혼용 텍스트의 정확한 발음 변환에 어려움이 있다. 외래어는 원래의 발음을 제대로 표현하지 못하고 혼용 텍스트의 경우 발음이 경우가 발생한다. 이로 인해 의미 전달에 혼동과 어려움이 있다. 따라서 혼동되는 이러한 문제를 해결하기 위해 정확한 발음으로 외래어와 한/영 혼용 텍스트를 처리할 수 있는 한국어 TTS 모델의 개발이 필요하고 이를 통해 자연스럽게 이해하기 쉬운 음성 변환 결과를 얻을 수 있다.



1 과제의 배경 및 목표

국제 음성 기호(International Phonetic Alphabet, IPA)란?

전 세계 언어의 다양한 소리를 표현할 수 있도록 국제음성학회에서 창안한 소리를 문자로 표기하는 기호 체계.

사전에 등록되지 않은 외래어는 IPA를 사용하여 발음을 표현.

IPA 체계를 사용해 변환한 발음이 일반적으로 쓰이는 발음과 차이가 클 경우, 별도의 보정이 필요.

자음(배에 의한 발동)

© 2019 IPA

	양순음	순치음	치음	치경음	후치경음	권설음	경구개음	연구개음	구개수음	인두음	성문음
폐색음	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
비음	m	ɱ		n		ɳ	ɲ	ɳ	ɴ		
전동음	β			ʀ					ʀ		
탄설음		ʋ		ɾ		ɽ					
마찰음	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
설측 마찰음				ɬ ɮ							
접근음		ʋ		ɹ		ɻ	j	ɰ			
설측 접근음				ɭ		ɮ	ʎ	ɱ			

한 칸의 오른쪽 기호는 유성음, 왼쪽 기호는 무성음이다. 음영 표시 영역은 조음이 불가능하다고 판단됨을 의미한다.

1 과제의 배경 및 목표

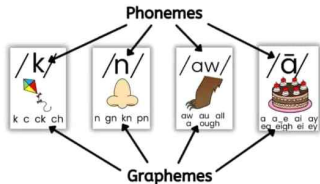
G2P(Grapheme-to-Phoneme)란?

입력으로 들어온 단어의 철자(표기)로부터 해당하는 발음을 매칭해주는 기술.

초기에는 수동으로 구축한 발음 사전이나 규칙을 기반으로 개발되었으나, 최근에는 지식베이스를 활용한 기계 학습 방식을 사용하여 더 높은 정확도를 보임.

장단기 메모리(LSTM) 방식을 주로 이용하고 인코더로 양방향 LSTM을 사용함.

음성 합성, 기계 번역, 음성 인식 등 다양한 작업에서 정확성과 품질을 향상시킬 수 있음.



1 과제의 배경 및 목표

과제의 목표

고유어나 한자어의 경우 한글이라는 문자로 표기되었기 때문에 문자열의 발음열 변환에 문제가 없다.

Ex) [국밥] -> [국뽕]

반면 외래어는 어원이 되는 원어를 구성하는 소리 체계가 한글의 체계와 다르기 때문에 우리가 원하는 정확한 발음으로 변환하기 힘든 경우가 많다.

Ex) 'b', 'p' -> 'ㅂ', 'ㅍ', 'ㅍ'

따라서 발음열 변환에 어려움이 있는 외래어에 대해 알맞은 변환이 되도록 모델을 학습시키는 것을 목표로 한다.

2 외래어 기반 G2P 개발의 필요성

외래어 기반 G2P 개발의 필요성

1. 사전에 미 등록된 외래어가 존재

기본적으로 우리말샘 사전과 표준국어대사전에 표기된 발음을 사용하지만 등록 되어 있지 않은 단어들은 별도의 규칙을 사용해 변환할 필요가 있음.

2. 일반적인 한국어 발음 규칙과의 차이

단순히 한국어 발음 규칙을 적용하여 외래어 단어에 적용하여 변환한다면 큰 성능 향상을 기대할 수는 없을 것임.

2 외래어 기반 G2P 개발의 필요성

외래어 기반 G2P 개발의 필요성

3. 표준 발음과 실생활 발음의 차이

사전에 등재된 발음을 사용하는 것을 원칙으로 하되, 실생활의 발음과 괴리가 너무 클 경우, 필요에 따라 수정이 필요.

3 개발 방향성 및 진행 방식

개발 방향성

TTS 기술은 많은 발전을 하였지만 정확한 외래어의 발음에서는 아직까지 부족한 부분이 많이 존재한다. 따라서 외래어 부분에서 기존의 모델과 성능적인 차이를 낼 수 있도록 개발은 진행한다. 단순히 발음 규칙 기반으로 변환하면 실제 발음과 일치하지 않는 경우가 생기고 앞 뒤 음운에 따라 발음이 달라지는 등 한계가 있기 때문에 딥러닝을 사용해서 성능을 개선한다. 또한 단어와 발음을 매핑한 데이터베이스인 기분석 사전을 이용하고 부족한 부분은 직접 데이터를 처리해주어서 체계적이고 정확성과 효율성을 높이는 방향으로 개발을 진행한다.

3 개발 방향성 및 진행 방식

개발 진행 방식

1. 데이터 수집 : 외래어에 대해 음운적 단위로 데이터를 수집.
2. 데이터 전 처리 : 수집한 데이터를 정제하고 구조화하여 모델 학습에 적합한 형태로 변환하고 예외적인 경우 직접적인 처리를 진행.
3. G2P 모델 구축 : 음운론적 단위로 변환하는 사용할 수 있도록 전 처리된 데이터를 사용하여 G2P 모델을 구축.
4. 모델 훈련 및 평가 : 구축한 G2P 모델을 학습시키고 성능 평가 및 수정.

4 개발 언어, 도구 및 환경

개발 언어, 도구 및 환경



5 개발 일정 및 역할 분담

개발 일정



May

G2p 모델 개발 관련 기술 공부, 착수 보고서 작성



June

외래어 데이터 수집 및 전 처리



July

G2P 모델 개발 및 데이터 학습, 중간 보고서 작성



August

테스트 및 디버깅, 기존 개발 모델과 성능 비교, 성능 발전 반향 모색



September

최종 보고서 작성 및 발표

5 개발 일정 및 역할 분담

역할 분담

이름	역할 분담
고상현	<ul style="list-style-type: none">• 학습용 데이터 전 처리• 시스템 테스트• 보고서 작성
권민규	<ul style="list-style-type: none">• 모델 학습 및 성능 평가• 시스템 테스트• 착수 발표 및 시연 준비
박건우	<ul style="list-style-type: none">• 모델 성능 비교 및 개선 방안 모색• 착수 발표 및 시연 준비• 학습용 데이터 개선
공 통	<ul style="list-style-type: none">• 사용되는 기술에 대한 공부 및 이해• 외래어 데이터 수집• 한국어 발음 규칙 이해

감사합니다

4조 RNG 아니조