

종합설계 프로젝트 제안서

팀 코드	A6	제출일	2026년 3월 17일
팀 이름	O(1)		
과제 명	"ECHO" English Correction & Hearing Optimization		
팀 구성원	김형준, 김지용, 백지영, 이찬영, 신영환, 최현지		
과제 개요	기존 발음 학습 앱의 사용자 불만은 두 가지로 수렴된다. ① 발음을 제대로 인식하지 못하는 것, ② 그리고 왜 틀렸는지 알려주지 않는 것이다. 전자는 MDD 시스템의 한계에서 비롯된다. L2-ARCTIC 데이터셋(6명 화자, 29,087 음소)에 Wav2Vec2-Base/Large, HuBERT-Large 세 모델로 베이스라인 실험을 수행해본 결과는 다음과 같았다: ① MDD F1은 최대 0.6078에 그쳤고, FRR(5.38~7.48%)과 FAR(36.85~42.77%) 사이에 trade-off가 존재했다. ② 한국어 화자(YKWK)에서는 MDD F1이 0.4974~0.5166, FAR은 최대 54.52%로 더 두드러졌다. 이는 한국인 특유의 발음 습관으로 인한 /r/·/l/ 혼동, /θ/→/s/ 대치, 자음군 모음 삽입 등 L1 간섭 오류 패턴을 범용 모델이 충분히 포착하지 못했기 때문이다. 결국 이러한 탐지 수준으로 인해, 교정 피드백의 부재는 당연한 결과였다. 또한 잘못된 쉼도잉 반복은 오히려 발음 오류를 악화시키고 있었다.		
해결 방안	ECHO는 한국인 대상 영어 발음 교정 서비스를 제공하며, 위 두 문제를 다음과 같이 해결한다: ① 탐지 성능 개선: FRR 감소 우선. HuBERT 기반 음소 인식기에 Canonical Sequence Injection을 적용해 정발음 인식 오류율(PER C)을 낮추고, 데이터 불균형(정발음 25,710 vs 오발음 4,318)은 Focal CTC Loss로 보완한다. 한국어 화자 성능 저하에는 AI-Hub 한국인 영어 발화 데이터셋을 추가 학습에 활용한다. ② 피드백 개선: 탐지된 오류 음소와 유형(대치·삭제·삽입)을 LLM에 전달해 한국어 음소 체계 기반 교정 가이드를 생성한다. [아학편]을 참고한 프롬프트 엔지니어링으로 IPA 없이 즉시 따라 할 수 있는 안내를 목표로 하며, '구강 단면 이미지 시각화'와 '최소대립쌍 훈련 콘텐츠'도 제공한다. ③ 범용성 개선: 학습 목적(여행·학술·비즈니스 등)별 스크립트 선택 또는 직접 입력 기능을 구현한다.		
달성 목표	L2-ARCTIC 기준 성능 평가와 소프트웨어 파이프라인 정상 동작을 기준으로 하며, 성공 기준은 다음과 같다. ① FRR을 4%대로 감소시키되 MDD F1(현재 0.6078) 하락 없이 유지. ② LLM 피드백의 음성학적 타당성을 자체 벤치마크로 평가. ③ 단일 발화(~5초) 기준 end-to-end 추론을 CPU 환경에서 실시간 응답 속도로		

	달성. 소프트웨어 테스트는 모듈 단위 테스트, 통합 테스트, API 안정성 검증을 수행한다.
기타	역할분담 1. 음소 탐지 모델 학습 및 실험 - 신영환 2. 백엔드 API 서버와 오디오 처리 - 김지용 3. 프론트엔드 인터페이스와 피드백 시각화 - 이찬영, 최현지 4. QA와 데이터 처리 - 김형준, 백지영 기타 기능 1. 사용자 리텐션을 위해 제한 시간 내 정확한 발음을 요구하는 속도 챌린지 2. 손실 회피 심리를 활용한 출석 환급 이벤트 도입

참고문헌

- [1] Kissling (2018). Pronunciation Instruction Can Improve L2 Learners' Bottom-Up Processing for Listening. *The Modern Language Journal*, 102, 653-675.
- [2] Scott et al. (2019). The motor system's contribution to speech perception. *Psychonomic Bulletin & Review*, 26, 1004-1013.
- [3] 체계적 문헌고찰 (2025). A Systematic Review of Research on the use of Shadowing for Second Language Pronunciation Teaching. *Journal of Second Language Pronunciation*.