pISSN 2005-8063
eISSN 2586-5854
2023. 12. 31.
Vol.15 No.4
pp. 1-10

말소리와 음성과학

Phonetics and Speech Sciences

한국음성학회지

https://doi.org/10.13064/KSSS.2023.15.4.001

Comparing the effects of letter-based and syllable-based speaking rates on the pronunciation assessment of Korean speakers of English*

Hyunsong Chung**

Department of English Education, Korea National University of Education, Cheongju, Korea

Abstract

This study investigated the relative effectiveness of letter-based versus syllable-based measures of speech rate and articulation rate in predicting the articulation score, prosody fluency, and rating sum using "English speech data of Koreans for education" from AI Hub. We extracted and analyzed 900 utterances from the training data, including three balanced age groups (13, 19, and 26 years old). The study built three models that best predicted the pronunciation assessment scores using linear mixed-effects regression and compared the predicted scores with the actual scores from the validation data (n=180). The correlation coefficients between them were also calculated. The findings revealed that syllable-based measures of speech and articulation rates were more effective than letter-based measures in all three pronunciation assessment categories. The correlation coefficients between the predicted and actual scores ranged from .65 to .68, indicating the models' good predictive power. However, it remains inconclusive whether speech rate or articulation rate is more effective.

Keywords: speech rate, articulation rate, pronunciation accuracy, prosody fluency

1. 서론

National Information Society Agency(2023) 사업의 일환으로 대규모 고품질의 음성 학습데이터 확보를 위해 '교육용 한국인의 영어 음성 데이터'가 AI Hub에 구축되었다(Han, 2023). 최근 이러한 학습자의 외국어 발화 또는 영어 발화 데이터 구축에 관한 연구는 국내외에서 많이 진행되었다. Rhee et al.(2003)은 초등학생, 고등학생, 일반 성인의 세 연령층과 7개 지역 및 남녀 성별에 따라 한국인 총 336명의 영어 발화를 녹음한 '한국인의 영어발음 음성코퍼스(Korean-spoken English corpus, K-SEC)'를 구축하였다. 이 코퍼스에는 독립어인 어휘 음성과, 문장 및 이야기인 연속 음성이 녹음되어 있어서 한국인 영어 학습자가 발화한 영어의 분절음 및 운율 연구에 큰 도움을 주고 있다. 하지만 이 코퍼스에는 표준 평가 준거에 의한 발음 평가 요소가 포

* This work was supported by the 2023 Sabbatical Leave Research Grant funded by Korea National University of Education.
** hchung@knue.ac.kr, Corresponding author
Received 30 October 2023; Revised 5 December 2023; Accepted 11 December 2023

함되지 않아 발음의 평가를 연구자에 따라 개별적으로 진행해야 하는 어려움이 있다.

Ishikawa(2014)는 ICNALE(The International Corpus Network of Asian Learners of English) 과제의 일환으로 중국, 홍콩, 인도네시아, 일본, 한국, 파키스탄, 필리핀, 싱가포르, 대만, 태국 등 10개 아시아 언어 배경의 영어 학습자 및 영어 원어민들 총 1,100명이 각각 60초씩 녹음한 독백 음성 파일 4,400개 및 참여자의 배경 자료, 발화 분석 자료 등을 제공하고 있다. 동일한 과제의 연속선상에서 Ishikawa(2019)는 중국, 홍콩, 인도네시아, 일본, 한국, 말레이시아, 파키스탄, 필리핀, 대만, 태국과 영어 원어민 등 총 425명 화자들의 45분 간의 인터뷰를 각각 녹화한 대화 영상 자료 4,250개 및 참여자의 배경 정보와 발화 분석 자료 등을 제공하고 있다. 두 코퍼스의 참여자는 영어 원어민을 제외하고 모두 Nation & Beglar(2007)의 '어휘 능력 검사'에 참여해 그 점수를 제출하고, 자신의 TOEFL이나 TOEIC, IELTS와 같은 공인 성적을 제출하였다. 이 코퍼스는 그 점수를 '유럽연합 공통 언어 표준 등급(Common European Framework of Reference for Languages, CEFR)'의 언어 능숙도 등급으로 변환하여 제공하고 있다. 이 데이터에는 아시아권의 다양한 언어 배경을 가진 영어 화자가 포함되어 있고, 표준 평가 준거에 의한 화자들의 언어 능숙도와 발음 평가 결과가 제시되어 있어서 영어 교육 영역에 활용하기가 쉬운 장점이 있다. 다만 독백 음성 파일은 음질이 상당히 떨어지고, 대화 영상 자료는 음성 파일이 따로 제공되지 않아서 연구를 위한 사전 처리 과정에 다소 어려움이 있다.

이에 반해 '교육용 한국인의 영어 음성 데이터'는 다양한 국적의 영어 화자 데이터나 다양한 지역의 한국인 영어 데이터를 제공하지는 않지만, 다양한 연령의 발음 데이터와 말하기 데이터를 대용량으로 제공하고 있고, 참여자의 언어 실력과 평가 점수를 제공하고 있다. 또, 훈련 데이터와 검증 데이터를 별도로 제공하고, 음성 파일의 음질이 분석 연구를 수행하기에 적합한 수준이어서 고무적이라고 할 수 있다.

이 데이터에서 주목한 것은 발화의 속도가 어떻게 측정되고 그것이 발음의 평가에 어떤 영향을 주는지였다. 일반적으로 음성학이나 응용언어학 연구에서는 발화의 속도를 측정할 때 휴지를 포함한 전체 발화에서 그 발화에 포함된 음절 수를 나누어 초당 또는 분당 음절 수를 계산한 음절 기반 발화 속도(speech rate), 휴지를 제외한 전체 발화에서 그 발화에 포함된 음절 수를 나누어 초당 또는 분당 음절 수를 계산한 음절 기반 조음 속도(articulation rate)(Choi & Jang, 2023; Chung, 2022; Kim & Jang, 2019; White & Mattys, 2007), 휴지 사이의 평균 음절 수 또는 단어 수를 계산한 평균 조음 길이(mean length of run, MLR)(Kang, 2010; Thomson, 2015; Yang, 2021) 등을 사용한다. 특이하게도 이 코퍼스에서는 이러한 일반적인 계산식을 사용하지 않고, 초당 철자 수를 사용하여 발화 속도와 조음 속도를 계산해 제공하고 있다. 이것은 Educational Testing Service(ETS)에서 개발한 발화 평가 프로그램인 SpeechRater^{SM}에 내장된 계산 방식에 포함된 속도 요소로, ETS 자체 연구에 의하면, 이 계산 방식의 철자 기반 속도가 음절을 기반으로 한 속도와 상관관계가 크고 이것을 바탕으로 예측한 점수와 실제 점수와의 상관계수가 .44로 비교적 적절한 성능을 제공하고 있는 것으로 기술되고 있다(Chen et al., 2018). 철자 수를 발화의 속도를 계산하는데 반영하면 데이터 처리의 시간과 부담을 줄일 수 있다는 장점이 크기 때문에 자동 평가를 신속하게 처리할 수 있다. 하지만, 영어의 철자가 발음을 제대로 반영하지 못하는 경우가 많기 때문에 과연 음성학에서 이용하는 음절 기반의 속도만큼 발화의 정확성 또는 유창성을 평가하는데 기여를 할 수 있는지 의문이 있을 수 있다. 아쉽게도 철자 기반 속도와 음절 기반 속도가 발음의 평가에 미치는 영향에 대해서 직접적으로 비교・분석한 연구는 찾기 힘들기 때문에 이에 대한 연구가 필요하다.

이러한 의문과 필요성을 바탕으로, 본 연구에서는 AI Hub에서 구축한 '교육용 한국인의 영어 음성 데이터'에서 제공하고 있는 철자 기반 발화 속도 및 조음 속도와 음성학 연구에서 주로 사용하고 있는 음절 기반 발화 속도 및 조음 속도가 코퍼스 내 훈련 데이터에 기반한 점수 예측 모델링과 모델에 의해 예측된 점수와 검증 데이터의 실제 점수와의 상관관계에 어떤 영향을 미치는지 살펴보고자 한다.

2. 연구 방법

2.1. 발화 자료

본 연구에서는 1장에서 언급한 것과 같이 AI Hub에 구축되어 있는 '교육용 한국인의 영어 음성 데이터'를 사용하였다(National Information Society Agency, 2023). 이 데이터는 영어 음성 인식, 통・번역, 교육용 AI 모델 등의 연구, 개발에 활용하고, 영어 교육을 위한 AI 기반 발음, 말하기 평가 시스템 개발을 위해 구축된 것으로 음성 발화는 wav 형식으로 되어 있고, 평가 점수와 발화에 대한 각종 정보가 담겨 있는 어노테이션(annotation) 파일은 json 형식으로 되어 있다. 이 데이터에는 발음 평가 데이터와 말하기 평가 데이터가 포함되어 있는데, 본 연구에서는 발음 평가 데이터를 활용하여 연구를 진행하였다. 발음 평가 데이터는 한국인 영어 학습자가 영어 문장을 낭독한 음성을 듣고, 발음 및 운율의 유창성을 평가한 점수와 틀린 발음의 유형 태깅 정보, 문장 단위의 철자 전사, 정답 음소, 발화 음소 라벨링 정보 및 발음 자질 정보가 어노테이션 파일에 기록되어 있다. 어노테이션 파일은 Python의 json 패키지를 이용해 파일에 포함된 다양한 자질을 추출한 후 csv 파일로 변환한 뒤 저장하여 활용하였다.

발음 평가 데이터는 다시 훈련 데이터(training data)와 검증 데이터(validation data)로 구분되어 있는데, 훈련 데이터에는 총 91,594개의 발화가 녹음되어 있고, 검증 데이터에는 총 11,450개의 발화가 녹음되어 있다. 본 연구에서는 그 중, 훈련 데이터에서 900개의 발화, 검증 데이터에서 180개의 발화를 추출하였다. 데이터를 추출할 때 음성 인식 결과물인 Speech-To-Text (STT) 결과물에 숫자가 포함되어 있을 경우, 실제 발음에 다양한 변이가 있을 수 있기 때문에 수가 포함된 발화는 제외하였다. 데이터는 13세, 19세, 26세의 데이터를 추출했는데, 13세는 녹음에 참여한 발화자 중 제일 어린 연령으로 중등학교를 막 시작

한 연령대를 대표하는 집단으로 가정하였다. 19세는 대학에 막 입학한 성인 화자를, 26세는 대학을 마친 성인 화자를 대표하는 연령으로 가정하였다. 모든 연령대의 데이터를 추출할 수 없었기 때문에 영어 학습량과 형태에 큰 변화가 있는 이 세 연령대를 추출하여 전 연령대를 대표할 수 있다는 가정 하에 분석을 진행하였다.

훈련 데이터에서는 연령별로 각각 300개의 발화를, 검증 데이터에서는 훈련 데이터의 20%에 해당하도록 연령별로 각각 60개의 발화를 추출하였다. 화자의 언어 실력은 상, 중, 하로 구분되어 있는데, 이것은 자신이 인지하는 언어 실력을 본인이 제출한 것이다. 13세는 훈련 데이터와 검증 데이터 모두 언어 실력이 '하'였고, 19세와 26세의 경우에는 훈련 데이터에서는 '상', '중', '하' 각각 100개씩, 검증 데이터에서는 각각 20개씩 추출하였다. 성별은 동수를 이루게 자료를 추출하였다. 추출된 자료의 구체적인 연령별, 성별, 수준별 분포는 표 1과 같다.

표 1. 추출된 발음 평가 데이터의 연령별, 성별, 수준별 분포
Table 1. Demographics of extracted pronunciation assessment data

연령	훈련 데이터		검증 데이터	
	여성	남성	여성	남성
13세	하 150	하 150	하 30	하 30
19세	상 50 중 50 하 50	상 50 중 50 하 50	상 10 중 10 하 10	상 10 중 10 하 10
26세	상 50 중 50 하 50	상 50 중 50 하 50	상 10 중 10 하 10	상 10 중 10 하 10

발음 평가 데이터에는 녹음 일자를 비롯해 발음 자질 등 총 40여 개의 어노테이션 항목이 있는데, 본 연구를 위해 그 중, 성별(gender), 연령(age), 언어 실력(ability), STT 결과물(sttText), 철자 기반 발화 속도(SpeechRate), 철자 기반 조음 속도(ArticulationRate), 초당 단어 수(wpsec), 발화 내 초당 단어 수(wpsecutt), 문장 단위 오류 유형 라벨링(tagging), 발음의 정확성 점수(ArticulationScore), 운율의 유창성 점수(ProsodyScore), 합산 점수(RatingSum)를 활용하였다. 평가는 두 명의 평가자가 각자 평가한 후 점수 차가 2점 이상일 경우 제3의 전문가가 최종적으로 점수를 조정하는 방식으로 이루어졌고(Han, 2023), 발음의 정확성 점수와 운율의 유창성 점수는 각각 1점부터 5점까지, 합산 점수는 두 점수를 합친 1점부터 10점까지 부여하였다.

어노테이션 파일에서 발화 속도와 조음 속도는 일반적으로 음성학 연구에서 많이 활용되는 초당 음절 수를 계산한 것이 아니고, 철자를 바탕으로 한 초당 글자 수를 계산한 것이기 때문에 초당 음절 수는 어노테이션 파일과 별도로 연구자가 직접 계산하였다.

개별 발화의 총 음절 수는 Python의 pandas, nltk와 cmudict 패키지를 활용하여 추출하였다. 발음 평가 데이터의 실제 발화 음성에 대한 스크립트인 STT 결과물 sttText를 위에 제시한 패키지에 포함된 발음 사전과 대조해 음절 수를 계산하였다. 휴지를 제외한 순수 음성 발화의 길이(phonation time)는 De Jong & Wempe (2008)의 Praat 스크립트를 활용해 추출하고, 어노테이션 파일에 포함된 PhonationTimeRatio(ptr)를 음성 발화의 길이에 대입해 휴지를 포함한 전체 발화 길이를 계산하였다. 예를 들면, 순수 음성 발화의 길이가 2.86초로 추출되고, 해당 음성 파일의 ptr이 0.94이면 순수 음성 발화의 길이를 ptr로 나누어 3.04초의 휴지를 포함한 전체 발화 길이를 도출할 수 있다. 해당 파일의 발화 스크립트를 기반으로 추출된 총 철자 수가 34개이고, 총 음절 수가 11개라고 가정할 때, 34개의 철자를 3.04초로 나누면 초당 11.18이라는 철자 기반 발화 속도(SpeechRate)가 계산되고, 2.86초로 나누면 초당 11.89라는 철자 기반 조음 속도(ArticulationRate)가 계산된다. 반면, 총 음절 수 11개를 3.04초로 나누면 초당 3.62라는 음절 기반 발화 속도(SpeechRate_s)를, 2.86초로 나누면 초당 3.85라는 음절 기반 조음 속도(ArticulationRate_s)를 구할 수 있다. 철자 기반 발화 속도와 조음 속도는 어노테이션 파일에 이미 포함되어 있기 때문에 본 연구에서는 그 정보를 그대로 사용하였고, 음절 기반 발화 속도와 음절 기반 조음 속도만 새롭게 계산하여 활용하였다.

문장 단위 오류 유형은 어노테이션 파일에 정조음(C), 음소 삭제(D), 음소 삽입(I), 음소 교체(S), 기타(O)로 분류되어 있는데, 개별 발화에 대해 전체 음소에 대한 각 유형의 개수를 구해 백분율로 변환하였다. 예를 들어 특정 발화에서 인식된 전체 음소 33개 중 정조음이 21개라면 63.64%로 계산하였다. 모델링에서는 각 비율을 '정조음비율(PctC)', '음소삭제비율(PctD)', '음소삽입비율(PctI)', '음소교체비율(PctS)', '기타비율(PctO)'로 각각 명명하였다.

2.2. 예측력 분석

어노테이션 파일에서 추출된 발화 속도 및 일부 운율 요소와 분절음 요소가 발음의 정확성 점수, 운율의 유창성 점수, 합산 점수에 어떤 예측력을 가지는지 분석하기 위해 R(R Core Team, 2023)에서 'lme4' 패키지(Bates et al., 2015)와 'lmerTest' 패키지(Kuznetsova et al., 2017)를 사용해 훈련 데이터에 대해 선형혼합효과분석(linear mixed effects regression)을 실행하였다. 평가 점수를 종속 변수로 두고 발화자(RecorderID)와 평가 유형(TestType)을 임의 효과(random effects)로 한 후, 어노테이션 파일에 있는 다양한 요소를 고정 효과(fixed effects)로 대입해 최적의 모델을 보여주는 요소가 무엇인지 분석하였다. 발화 속도의 경우 모델을 구성할 때 어노테이션 파일에 있는 철자 기반 발화 속도, 철자 기반 조음 속도, 초당 단어 수, 발화 내 초당 단어 수와 본 연구를 위해 새롭게 계산한 음절 기반 발화 속도, 음절 기반 조음 속도 중 각 하나씩만 선택하여 모델을 만들었다. 속도 요소를 모두 하나의 모델식에 반영할 경우 다중공선성(multicollinearity)으로 인해 독립 변수들 간의 강한 상관관계가 나타나기 때문이다. 이것은 '정조음비율'과 다른 오류 비율의 경우에도 마찬가지로 '정조음비율'이 높으면 오류 비율이 낮아지기 때문에 두 유형 중 한 유형만 모델에 반영하였다.

다양한 모델에서 도출된 AIC(akaike information criterion)를 비교하여 AIC가 가장 낮은 모델 3개를 선정해, 최적의 모델에

서 예측력이 큰 고정 효과가 무엇인지 분석하였다. AIC는 다양한 모델의 상대적 적합도를 비교해 최적의 모델을 선택하는 준거 공식의 하나이다. 관찰 데이터에 대한 모델의 적합도를 나타내는 로그 가능도(log-likelihood)가 클수록 그 값이 낮아지고, 모델의 매개변수(parameter)가 복잡할수록 불이익을 주는 구조이기 때문에 AIC가 낮을수록 모델의 적합도가 더 크다고 할 수 있다(Chakrabarti & Ghosh, 2011).

적합도가 큰 3개의 모델을 적용해 훈련 데이터를 대상으로 한 모델의 예측 평가 점수가 검증 데이터의 평가 점수와 얼마나 상관관계가 있는지 알아보기 위하여 상관관계 분석을 실시하였다. 상관관계 분석과 함께 예측 평가 점수와 실제 점수 간의 '평균 제곱 오차(mean sqaure error, MSE)', '평균 제곱근 오차(root mean square error, RMSE)', '평균 절대 오차(mean absolute error, MAE)'도 추출하였다.

3. 연구 결과

3.1. 발음의 정확성 점수 예측 모델

훈련 데이터에 있는 발음의 정확성에 대한 연령별, 성별, 언어 실력별 평균 점수는 표 2와 같다.

표 2. 훈련 데이터의 연령별, 성별, 언어 실력별 발음 정확성 평균 점수
Table 2. Mean pronunciation score of training data in terms of age, gender, and ability

연령	언어 실력	여성		남성		전체	
		평균	표준 편차	평균	표준 편차	평균	표준 편차
13세	하	3.60	.84	3.11	.84	3.36	.88
19세	상	3.91	.79	3.11	.59	3.90	.81
	중	4.23	.61	4.59	.39		
	하	3.88	.78	3.70	.80		
26세	상	4.51	.43	4.44	.49	3.91	.85
	중	3.82	.70	3.79	.68		
	하	3.67	.95	3.13	.78		

표 2를 보면 19세의 경우 남녀 모두 자신의 언어 실력이 '중'이라고 진술한 참여자의 발음 정확성 평균 점수가 가장 높고, 특히 남성의 경우에는 언어 실력이 '상'인 경우의 평균 점수가 가장 낮았다. 발음 평가의 외적 요인인 '연령', '성별', '언어 실력'은 본 연구의 주제가 아니기 때문에 더 깊이 분석하지는 않지만, 추후 언어 외적 요인이 발음 평가에 미치는 상호작용 등에 관한 추가 연구가 필요한 부분이다.

훈련 데이터의 어노테이션 파일에 포함된 다양한 요소를 고정 효과로 하여 발음의 정확성 점수에 대한 예측력 분석을 한 결과 표 3과 같이 다음 3개 모델의 AIC가 가장 낮아 가장 최적의 모델로 선정하였다.

표 3. 발음의 정확성 점수 예측 상위 3개 모델
Table 3. Three best models for articulation score

모델	모델식	AIC
M04	lmer(ArticulationScore ~ ArticulationRate_s + PctC + (1|RecorderID) + (1|TestType), data = data)	1,806
M03	lmer(ArticulationScore ~ SpeechRate_s + PctC + (1|RecorderID) + (1|TestType), data = data)	1,807
M01	lmer(ArticulationScore ~ SpeechRate + PctC + (1|RecorderID) + (1|TestType), data = data)	1,822

AIC, akaike information criterion.

각 모델에 대한 분석 결과는 표 4와 같다.

표 4. 발음의 정확성 점수 예측 상위 3개 모델의 분석 결과
Table 4. Results summary of three best models for articulation score

모델	고정 효과	추정치	표준 오차	*t*값	확률 (Pr(>|t|))
M04	잔차	2.244e+00	2.071e-01	10.83	8.62e-10***
	음절기반 조음속도	2.209e-01	3.227e-02	6.84	1.41e-11***
	정조음비율	8.080e-03	2.006e-03	4.02	6.15e-05***
M03	잔차	2.270e+00	2.061e-01	11.01	6.14e-10***
	음절기반 발화속도	2.172e-01	3.222e-02	6.74	2.80e-11***
	정조음비율	8.128e-03	2.007e-03	4.05	5.58e-05***
M01	잔차	2.186e+00	2.115e-01	10.33	7.21e-16***
	철자기반 발화속도	7.419e-02	1.308e-02	5.67	1.96e-08***
	정조음비율	8.760e-03	2.016e-03	4.34	1.57e-05***

*p<.05, **p<.01, ***p<.001.

발음의 정확성 점수에 대한 상위 3개 모델의 분석 결과를 살펴보면 'M04'의 경우 잔차(intercept)의 추정치(estimate)가 2.244로, 모든 예측 변수가 '0'일 때 발음 평가 점수를 '2.244'로 예측하고 있다. 모든 고정 효과 중 '음절기반조음속도'가 발음 평가 점수에 가장 크고 유의미한 긍정적 효과를 미치고 있다. 표 3을 보면 '음절기반조음속도'가 '1'씩 증가할 때마다, 발음 평가 점수는 '.2209'점 상승하는 것으로 예측하고 있다. 두 번째로 유의미한 효과를 보이는 것은 '정조음비율'로 '1%' 증가할 때마다, 발음 평가 점수가 '.0081'점 상승하는 것으로 예측하고 있다.

'M03'의 경우 모든 예측 변수가 '0'일 때 발음 평가 점수는 '2.27'로 예측되었고, '음절 기반 발화 속도'가 '1'씩 증가할 때마다 발음 평가 점수가 '.217'씩 상승하고, '정조음비율'이 '1%' 증가할 때마다, 점수가 '.0081' 상승하는 것으로 예측되었다. 'M01'은 모든 예측 변수가 '0'일 때 '발음 평가 점수'는 '2.186'으로 예측되었고, '철자기반발화속도'가 '1'씩 증가할 때마다 점수가 '.0742'씩 상승하고, '정조음비율'이 '1%' 증가할 때마다, 점수가 '.0088' 상승하는 것으로 예측되었다.

최적의 3개 모델을 검증 데이터에 적용했을 때 예측된 평가 점수와 실제 점수 간에 어느 정도의 일치성이 있는지 알아보기 위하여 상관관계 분석을 실시하였다. 상관관계 분석 결과는 표 5와 같다.

표 5. 발음 평가 예측 점수와 실제 점수와의 상관관계
Table 5. Correlations between predicted and actual pronunciation scores

모델	평균 제곱 오차	평균 제곱근 오차	평균 절대 오차	상관 계수	*p*값
M04	.456	.675	.534	.668	1.33e-24
M03	.453	.673	.533	.670	7.90e-25
M01	.460	.678	.543	.664	3.13e-24

표 5를 보면 'M03'이 발음 평가 예측 점수와 실제 점수와의 상관계수가 가장 크고(*r*=.67), AIC가 가장 낮아 최적의 모델로 선택된 'M04'는 상관계수가 .668로 미세하게 낮지만 큰 차이를 보인다고 할 수는 없다. 'M01'은 두 모델에 비해 상관계수가 .664로 다소 낮게 나왔다. 상관관계를 그림으로 나타내면 그림 1과 같다.

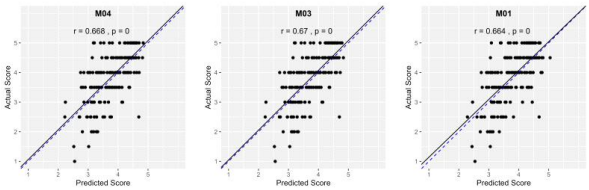

그림 1. 발음 평가 예측 모델의 상관관계 플롯
Figure 1. Correlation plots of articulation score models

그림 1에서 점선은 상관관계가 '1'일 때의 가상선을 나타내고, 실선은 실제로 드러난 예측 점수와 실제 점수의 산포도를 직선으로 나타낸 것이다.

3.2. 운율의 유창성 점수 예측 모델

훈련 데이터의 운율 유창성 점수의 연령별, 성별, 언어 실력별 평균 점수는 표 6과 같다.

표 6. 훈련 데이터의 연령별, 성별, 언어 실력별 운율 유창성 평균 점수
Table 6. Mean of prosody score of training data in terms of age, gender, and ability

연령	언어 실력	여성		남성		전체	
		평균	표준 편차	평균	표준 편차	평균	표준 편차
13세	하	3.54	.88	3.07	.97	3.31	.96
19세	상	3.89	.85	3.20	.81	3.89	.85
	중	4.23	.67	4.57	.39		
	하	3.84	.80	3.58	.80		
26세	상	4.54	.48	4.41	.52	3.87	.91
	중	3.71	.70	3.90	.66		
	하	3.62	1.00	3.04	1.03		

운율 유창성 평균 점수도 발음 정확성 평균 점수와 마찬가지로 19세에서 자신이 밝힌 언어 실력과 운율 유창성 평균 점수가 역전되는 상호작용을 관찰할 수 있다.

훈련 데이터의 어노테이션 파일에 포함된 다양한 요소를 고정 효과로 하여 운율의 유창성 점수에 대한 예측력 분석을 한 결과 표 7과 같이 다음 3개 모델의 AIC가 가장 낮아 가장 최적의 모델로 선정하였다.

표 7. 운율의 유창성 점수 예측 상위 3개 모델
Table 7. Three best models for prosody score

모델	모델식	AIC
M012	lmer(ProsodyScore ~ ArticulationRate_s + (1|RecorderID) + (1|TestType), data = data)	1,898
M011	lmer(ProsodyScore ~ SpeechRate_s + PctC + (1|RecorderID) + (1|TestType), data = data)	1,903
M04	lmer(ProsodyScore ~ ArticulationRate_s + PctC + (1|RecorderID) + (1|TestType), data = data)	1,905

AIC, akaike information criterion.

각 모델에 대한 분석 결과는 표 8과 같다.

표 8. 운율의 유창성 점수 예측 상위 3개 모델의 분석 결과
Table 8. Results summary of three best models for prosody score

모델	고정 효과	추정치	표준 오차	*t*값	확률 (Pr(>|t|))
M012	잔차	2.460	.167	14.75	3.16e-8***
	음절기반 조음속도	.320	.033	9.59	<2e-16***
M011	잔차	2.525	.165	15.32	2.59e-08***
	음절기반 발화속도	.309	.033	9.25	2e-16***
M04	잔차	2.082e+00	2.135e-01	9.75	1.65e-10***
	음절기반 발화속도	3.115e-01	3.409e-02	9.14	<2e-16***
	정조음비율	5.244e-03	2.119e-03	2.48	.0135*

*p<.05, **p<.01, ***p<.001.

'M012'의 경우 잔차의 추정치가 2.460으로, 모든 예측 변수가 '0'일 때 운율 평가 점수를 '2.460'으로 예측하고 있다. 이 모델에서는 '음절기반조음속도'만이 유일한 고정 효과로 사용되었고, 운율 평가 점수에 유의미한 긍정적 효과를 미치고 있다. '음절기반조음속도'가 '1'씩 증가할 때마다, 운율 평가 점수는 '.320'점 상승하는 것으로 예측하고 있다. 'M011'은 '음절기반발화속도'만이 유일한 고정 효과로 사용되었는데, 이 예측 변수가 '0'일 때 운율 평가 점수는 '2.525'로 예측되고, '음절기반발화속도'가 '1'씩 증가할 때마다 운율 평가 점수가 '.309'씩 상승한다. 'M04'는 모든 예측 변수가 '0'일 때 '발음 평가 점수'는 '2.082'로 예측되었고, '음절기반발화속도'가 '1'씩 증가할 때마다 점수가 '.312'씩 상승하고, '정조음비율'이 '1%' 증가할 때마다, 점수가 '.005'씩 상승하는 것으로 예측되었다.

운율의 유창성 점수에 대한 최적의 3개 모델을 검증 데이터에 적용했을 때 예측된 평가 점수와 실제 점수 간에 어느 정도의 일치성이 있는지 알아보기 위하여 상관관계 분석을 실시하였고, 그 결과는 표 9와 같다.

표 9. 운율 평가 예측 점수와 실제 점수와의 상관관계
Table 9. Correlations between predicted and actual prosody scores

모델	평균 제곱 오차	평균 제곱근 오차	평균 절대 오차	상관 계수	*p*값
M012	.496	.704	.569	.657	1.30e-23
M011	.493	.702	.568	.659	8.51e-24
M04	.503	.709	.572	.651	4.87e-23

'M011'이 운율 평가 예측 점수와 실제 점수와의 상관계수가 가장 크고(*r*=.659), AIC가 가장 낮아 최적의 모델로 선택된 'M012'는 상관계수가 .657로 'M011'보다 미세하게 낮지만 큰 차이를 보이지는 않는다. 'M04'는 두 모델에 비해 상관계수가 .651로 다소 낮지만 이 또한 크게 의미있는 차이라고 할 수는 없다.운율 평가 예측 모델의 상관관계를 그림으로 나타내면 그림 2와 같다.

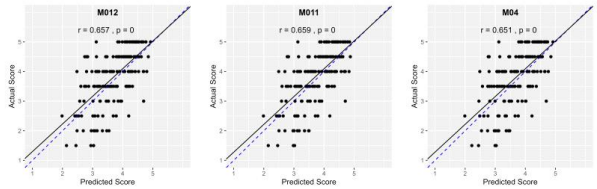

그림 2. 운율 평가 예측 모델의 상관관계 플롯
Figure 2. Correlation plots of prosody score models

그림 1과 마찬가지로 점선은 상관관계가 '1'일 때의 가상선을 나타내고, 실선은 실제로 드러난 예측 점수와 실제 점수의 산포도를 직선으로 나타낸 것이다.

3.3. 합산 점수 예측 모델

발음 평가 데이터의 어노테이션 파일에 있는 합산 점수는 개별 발화에 대한 총체적 평가가 아니고 단순히 발음 평가 점수와 운율 평가 점수를 합친 것이다. 하지만 여전히 개별 요소들이 합산 점수에 어떤 영향을 미치는지 분석하는 것이 필요하기 때문에 위의 절차와 동일하게 모델링을 진행하였다.

합산 점수의 연령별, 성별, 언어 실력별 평균 점수는 표 10과 같다.

표 10. 훈련 데이터의 연령별, 성별, 언어 실력별 합산 평균 점수
Table 10. Mean of prosody score of training data in terms of age, gender, and ability

연령	언어실력	여성		남성		전체	
		평균	표준편차	평균	표준편차	평균	표준편차
13세	하	7.15	1.67	6.18	1.76	6.66	1.78
19세	상	7.8	1.60	6.31	1.34	7.79	1.61
	중	8.46	1.23	9.16	.71		
	하	7.72	1.54	7.28	1.54		
26세	상	9.15	.86	8.85	.94	7.78	1.71
	중	7.53	1.35	7.69	1.28		
	하	7.29	1.92	6.17	1.71		

합산 점수의 평균은 발음 평가 점수와 운율 평가 점수를 합친 점수에 대한 평균이기 때문에 19세에서 '언어 실력'과 평가 점수의 상호작용을 여전히 관찰할 수 있다.

훈련 데이터의 어노테이션 파일에 포함된 다양한 요소를 고정 효과로 하여 합산 점수에 대한 예측력 분석을 한 결과 표 11과 같이 다음 3개 모델의 AIC가 가장 낮아 가장 최적의 모델로 선정하였다.

표 11. 합산 점수 예측 상위 3개 모델
Table 11. Three best models for rating sum

모델	모델식	AIC
M012	lmer(RatingSum ~ SpeechRate_s + PctI + PctD + PctS + (1|RecorderID) + (1|TestType), data = data)	3,001
M04	lmer(RatingSum ~ ArticulationRate_s + PctC + (1|RecorderID) + (1|TestType), data = data)	3,004
M011	lmer(RatingSum ~ SpeechRate_s + (1|RecorderID) + (1|TestType), data = data)	3,004

AIC, akaike information criterion.

구체적인 각 모델의 분석 결과는 표 12와 같다.

표 12. 합산 점수 예측 상위 3개 모델의 분석 결과
Table 12. Results summary of three best models for rating sum

모델	고정 효과	추정치	표준 오차	*t*값	확률 (Pr(>|t|))
M012	잔차	5.784	.332	17.4	2.53e-11***
	음절기반 발화속도	.511	.062	8.13	1.43e-15***
	음소삽입비율	−.033	.01	−3.17	.0016**
	음소삭제비율	.004	.013	.34	.731
	음소교체비율	−.01	.005	−1.97	.0486*
M04	잔차	4.452e+00	4.070e-01	10.97	5.62e-10***
	음절기반 조음속도	5.231e-01	5.299e-02	8.31	3.70e-16***
	정조음비율	1.206e-02	3.923e-03	3.07	.00218**
M011	잔차	5.423	.32	16.95	2.23e-07***
	음절기반 발화속도	.529	.062	8.56	<2e-16***

*p<.05, **p<.01, ***p<.001.

발음 평가 모델 및 운율 평가 모델과 달리 최적의 모델로 선택된 'M012'에서는 음소의 오류 요소들이 기여하고 있음을 알 수 있다. 이 모델에서 잔차의 추정치가 총 10점 가운데 5.784로, 모델에 사용된 모든 예측 변수가 '0'일 때 합산 점수를 '5.784'로 예측하고 있다. '음절기반발화속도'와 함께 '음소삽입비율', '음소교체비율'이 유의미하게 영향을 미치는 고정 효과로 분석되었고, '음소 삭제 비율'은 유의미한 영향이 없었다. 이 모델에서 '음절기반발화속도'가 '1'씩 증가할 때마다, 합산 점수는 '.511' 점 상승하는 것으로 예측하고 있다. 또 '음소삽입비율'과 '음소교체비율'이 각각 '1%' 증가할 때마다 합산 점수가 '−.033'점과 '−.01'점 하강하는 것으로 예측하였다.

'M04'는 '음절기반조음속도'와 '정조음비율' 모두 유의미한 영향을 끼치는 고정 효과로 사용되었는데, 두 예측 변수가 '0'일

때 합산 점수는 ‘4.452’로 예측되고, ‘음절기반조음속도’가 ‘1’씩 증가할 때마다 합산 점수가 ‘.523’씩 상승하고, ‘정조음비율’이 ‘1%’씩 증가할 때마다 점수가 ‘.0121’씩 상승하는 것으로 예측되었다.

‘M011’에서는 유일하게 사용된 고정 효과인 ‘음절기반발화속도’가 합산 점수 예측에 유의미한 영향을 미치는 것으로 나타났다. 이 변수가 ‘0’일 때 합산 점수는 ‘5.423’으로 예측되었고, ‘음절기반발화속도’가 ‘1’씩 증가할 때마다 점수가 ‘.529’씩 상승하는 것으로 예측되었다.

합산 점수에 대한 최적의 3개 모델을 검증 데이터에 적용했을 때 예측된 평가 점수와 실제 점수 간에 어느 정도의 일치성이 있는지 알아보기 위한 상관관계 분석 결과는 표 13과 같다.

표 13. 합산 예측 점수와 실제 점수와의 상관관계
Table 13. Correlations between predicted and actual rating sums

모델	평균 제곱 오차	평균 제곱근 오차	평균 절대 오차	상관 계수	*p*값
M012	1.739	1.319	1.040	.680	9.50e-26
M04	1.758	1.326	1.044	.675	2.95e-25
M011	1.725	1.313	1.039	.682	5.27e-26

AIC에서 예측한 순서와 달리 ‘M011’이 합산 예측 점수와 실제 점수와의 상관계수가 가장 크고(r=.682), AIC가 가장 낮아 최적의 모델로 선택된 ‘M012’는 상관계수가 .680으로 ‘M011’보다 미세하게 낮지만 큰 차이를 보인다고 보기는 어렵다. ‘M04’는 두 모델에 비해 상관계수가 .675로 다소 낮다. 합산 점수 예측 모델의 상관관계를 그림으로 나타내면 그림 3과 같다.

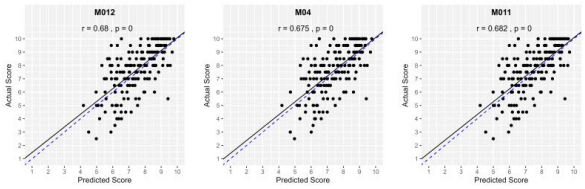

그림 3. 합산 점수 예측 모델의 상관관계 플롯
Figure 3. Correlation plots of rating sum models

이전 그림과 마찬가지로 점선은 상관관계가 ‘1’일 때의 가상선을 나타내고, 실선은 실제로 드러난 예측 점수와 실제 점수의 산포도를 직선으로 나타낸 것이다.

4. 논의 및 결론

AI Hub 발음 평가 데이터에서 제공하고 있는 철자 기반 발화 및 조음 속도와 본 연구에서 추출해 낸 음절 기반 발화 및 조음 속도 중 어떤 요소가 모델의 점수 예측에 더 큰 효과를 보이는지 살펴보았을 때, 철자 기반보다는 음절 기반의 발화 속도 및 음절 속도가 점수 예측 모델의 적합도와 예측성에 더 큰 효과를 보이는 것으로 나타났다.

발음의 정확성 점수 예측 모델에서는 음절 기반 조음 속도와 음절 기반 발화 속도가 모델식에 사용되었을 때, 철자 기반 발화 속도 및 조음 속도가 사용되었을 때보다 AIC가 더 낮아 더 적합도가 높은 것으로 분석되었다. 발음 평가 예측 점수와 실제 점수와의 상관관계 분석에서도 음절 기반 속도가 포함된 모델이 사용되었을 때 상관관계가 더 큰 것으로 나타났다. 음절 기반 발화 속도와 음절 기반 조음 속도 중 어느 요소가 더 큰 영향을 미치는지 살펴보면, AIC에서는 음절 기반 조음 속도가 포함된 모델이 더 낮아 더 적합하지만, 상관관계 분석에서는 반대로 음절 기반 발화 속도가 포함된 모델의 상관계수가 미세하게 높았다.

운율의 유창성 점수 예측 모델에서는 음절 기반 조음 속도만 고정 효과로 사용된 모델이 AIC가 가장 낮아 최적의 모델로 선택되었다. 두 번째로 적합한 모델은 음절 기반 발화 속도만을 고정 효과로 사용한 모델이었다. 상관관계 분석에서는 음절 기반 발화 속도가 사용된 모델의 상관계수가 음절 기반 조음 속도가 사용된 것보다 미세하게 높았다.

합산 점수 예측 모델에서는 음절 기반 발화 속도가 ‘음소삽입비율’, ‘음소삭제비율’, ‘음소교체비율’과 같은 음소 오류 비율과 함께 고정 효과로 사용된 모델의 적합도가 가장 높았다. 음절 기반 조음 속도가 ‘정조음비율’과 함께 사용된 모델이 그 뒤를 이었다. 음절 기반 발화 속도가 유일한 고정 효과로 사용된 모델은 세 번째로 적합한 모델로 분석되었다. 상관관계 분석에서는 적합도가 세 번째인 모델의 상관계수가 가장 높았고, 적합도가 가장 컸던 모델은 두 번째로 높았지만, 두 모델 모두 음절 기반 발화 속도가 포함된 모델이었다.

이러한 결과로 볼 때 모델의 적합성 분석과 상관관계 분석 모두에서 철자 기반 속도보다는 음절 기반 속도가 평가 점수를 예측하는데 더 적절한 역할을 하고 있다는 것을 알 수 있다. 철자를 기반으로 해서 발화 속도와 조음 속도를 산출하는 것이 음성처리와 분석의 신속성과 편의성 면에서는 긍정적인 기여를 하는 것이 틀림없지만, 한국어와 달리 영어는 철자와 발음 간의 연결이 불규칙하고 동일한 철자에 대한 발음의 변이가 다양해 발화의 특성을 온전히 담아낼 수 없는 한계가 존재할 수밖에 없다. 반면 음절 기반 속도는 발음 사전에 근거해서 실제로 발음된 발화의 속도 특성을 완벽하지는 않지만 비교적 근사치에 가깝게 계산할 수 있기 때문에 철자 기반의 속도 계산보다 생태학적 타당도(ecological validity)가 더 크다고 할 수 있다.

발화 속도와 조음 속도 중 어떤 것이 평가 점수의 예측에 더 큰 영향을 미치는지는 결론을 낼 수 없었다. 위에서 살펴본 것과 같이 적합성에서는 조음 속도가 포함된 모델이 더 낫다 하더라도 상관관계 분석에서는 미세하게 그 상황이 역전되는 경우가 있기 때문이다. 이것은 부분적으로 이 연구에 사용된 AI Hub 발음 평가 데이터의 특성 때문일 수도 있다. 파일을 분석했을 때 비교적 휴지가 적었고, 휴지 여부에 의해 좌우되는 발화 속도와 조음 속도의 차이가 크지 않은 발화가 많았다. 추후 발화 속도와 조음 속도의 차이가 큰 발음 평가 데이터로 분석하는 작

업이 필요할 것으로 판단된다.

발음의 정확성 점수 예측 모델에서 '정조음비율'이나 '음소 삽입비율'과 같은 분절음 요소가 더 큰 영향을 미치기보다는 발화 속도나 조음 속도와 같은 속도 요소가 더 큰 영향을 미치고 있음을 알 수 있었다. 평가자가 평가할 때 속도 요소에 민감하게 반응한다는 반증이 될 수 있을 것이다.

이 발음 평가 데이터의 어노테이션 파일에는 발화의 리듬 정보, 억양 정보 등은 담겨 있지 않아서 분석에 직접 반영할 수 없었다는 것은 이 연구의 큰 제한점이다. 발화의 속도뿐만 아니라 리듬 및 억양 요소 등을 모델에 반영해 평가 점수에 어떤 영향을 끼치는지 함께 볼 수 있다면 가장 이상적인 연구가 될 수 있지만, 리듬 및 억양 정보를 본 연구에 사용된 방대한 데이터에서 추출하기 위해서는 현재의 기술로 신속하게 자동화 또는 반자동화하는데 큰 한계가 있다. 우선 강제 정렬을 한 이후 분절음의 경계가 확정되어야 리듬을 계산할 수 있고 억양의 표기가 가능한데, 현재의 기술로는 강제 정렬의 정확성이 떨어져서 별도의 사후 조정이 수반될 수밖에 없다. 소규모의 평가 데이터를 활용한다면 이러한 작업이 가능하겠지만, 본 연구에 사용된 방대한 발화 자료에 이러한 작업을 하는 것은 비현실적이기 때문에 미래에 강제 정렬의 정확성이 높아져서 사후 조정이 필요 없는 시점에 추가 연구를 진행할 수 있을 것이다.

본 연구의 또 하나의 제한점은, 이 연구에 사용된 언어 실력 기준이 '유럽연합 공통 언어 표준 등급'과 같은 표준 평가 준거에 의한 것이 아니어서, 다른 대규모 발화 자료와의 비교가 쉽지 않다는 점이다. 추후 동일한 발화 자료를 이러한 표준 평가 준거에 맞게 개선하는 작업이 필요할 것으로 판단된다.

감사의 글

본 연구의 분석에 사용된 음성 데이터는 한국지능정보사회진흥원에서 '지능정보산업 인프라 조성' 사업의 일환으로 구축한 사업 결과물입니다. 음성 데이터의 세부 사항에 관해 개발자이신 한국외국어대학교 한승희 교수님의 도움을 받았습니다. 하지만, 논문에 오류가 있다면 전적으로 본 연구자의 책임임을 밝혀 둡니다. 발음 평가 데이터의 json 파일을 처리하는데 사용된 Python 스크립트는 2023년 7월 8일에 진행된 한국음성학회 산하 실험음성학연구회 하계워크숍에서 성신여자대학교 윤태진 교수님께서 제공해 주신 것을 수정하여 사용한 것입니다.

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Chakrabarti, A., & Ghosh, J. K. (2011). AIC, BIC and recent advances in model selection. In P. S. Bandyopadhyay, & M. R. Forster (Eds.), *Philosophy of statistics: Handbook of the philosophy of science* (pp. 583-605). North Holland: Elsevier.

Chen, L., Zechner, K., Yoon, S. Y., Evanini, K., Wang, X., Louskina, A., Tao, J., ... Gyawali, B. (2018). Automated scoring of nonnative speech using the SpeechRater^{SM} v. 5.0 engine. *ETS Research Report Series, 2018*(1), 1-31.

Choi, N., & Jang, T. Y. (2023). Evaluation of English speaking proficiency under fixed speech rate: Focusing on utterances produced by Korean child learners of English. *Phonetics and Speech Sciences, 15*(1), 47-54.

Chung, H. (2022). Relationships between rhythm and fluency indices and listeners' ratings of Korean speakers' English paragraph reading. *Phonetics and Speech Sciences, 14*(4), 25-33.

De Jong, N., & Wempe, T. (2008). Praat script syllable nuclei. [Computer script]. Retrieved from https://github.com/FieldDB/Praat-Scripts/blob/main/praat-script-syllable-nuclei-v2file.praat/

Han, S. H. (2023, June). Design and construction of Korean speech AI datasets for multilingual translation and language education. *Proceedings of the 2023 Spring Conference of the Korean Society of Speech Sciences* (pp. 16-27). Seoul, Korea.

Ishikawa, S. (2014). Design of the ICNALE spoken: A new database for multi-modal contrastive interlanguage analysis. In S. Ishikawa (Ed.), *Learner corpus studies in Asia and the world* (pp. 63-76). Kobe, Japan: Kobe University.

Ishikawa, S. (2019). The ICNALE spoken dialogue: A new dataset for the study of Asian learners' performance in L2 English interviews. *English Teaching, 74*(4), 153-177.

Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System 38*(2), 301-315.

Kim, M. S., & Jang, T. Y. (2019). Pauses and speech rates in assessing fluency of English speech. *The Korean Journal of Linguistics, 44*(3), 315-339.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1-26.

Nation, I., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher, 31*(7), 9-13.

National Information Society Agency. (2023). *English speech data of Koreans for education*. Retrieved from https://aihub.or.kr/

R Core Team. (2023). R: A language and environment for statistical computing (version 4.3.0) [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rhee, S. C., Lee, S. H., Kang, S. K., & Lee, Y. J. (2003). Design and construction of Korean-spoken English corpus (K-SEC). *Malsori, 46*, 159-174.

Thomson, R. I. (2015). Fluency. In M. Reed, & J. M. Levis (Eds.), *The handbook of English pronunciation* (pp. 209-226). Chichester, UK: John Wiley & Sons, Inc.

White, L., & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics, 35*(4), 501-522.

Yang, I. Y. (2021). Differential contribution of English suprasegmentals to L2 foreign-accentedness and speech comprehensibility: Implications for teaching EFL pronunciation, speaking, and listening. *Korean Journal of English Language and Linguistics, 21*, 818-836.

• **정현성 (Hyunsong Chung)** 교신저자
한국교원대학교 영어교육과 교수
충북 청주시 흥덕구 강내면 태성탑연로 250
Tel: 043-230-3554
Email: hchung@knue.ac.kr
관심분야: 실험음성학, 영어발음교육

철자 기반과 음절 기반 속도가 한국인 영어 학습자의 발음 평가에 미치는 영향 비교*

정 현 성

한국교원대학교 영어교육과

국문초록

본 연구에서는 AI Hub에 구축된 '교육용 한국인의 영어 음성 데이터'에 있는 발음 평가 데이터를 활용하여 철자 기반 발화 속도 및 조음 속도와 음절 기반 발화 속도 및 조음 속도 중 발음 정확성 및 운율 유창성, 합산 점수를 예측하는 모델에 어떤 요소가 더 유의미한 영향을 미치는지 분석하였다. 이를 위해 13세, 19세, 26세 연령별, 성별, 수준별로 이 코퍼스의 훈련 데이터에서 총 900개 발화를 추출하여 데이터에 포함된 다양한 요소를 활용해 평가 점수를 예측하는 선형효과분석을 실행하였다. 선형효과분석에서 최적의 세 개 모델을 통해 예측된 평가 점수를 검증 데이터에서 추출한 총 180개 발화의 평가 점수와 얼마나 상관관계가 있는지도 분석하였다. 분석 결과 발음의 정확성과 운율의 유창성, 합산 점수 예측 모델 모두 철자 기반 발화 속도와 조음 속도보다 음절 기반 발화 속도와 조음 속도가 평가 점수를 예측하는데 더 큰 영향을 주는 것으로 밝혀졌다. 모델에서 예측한 점수와 검정 데이터의 실제 점수와의 상관계수는 .65에서 .68 사이로 각 모델의 평가 점수 예측력이 나쁘지 않았다. 발화 속도와 조음 속도 간에 어떤 요소가 더 큰 영향을 미치는지는 본 연구를 통해 밝혀내지 못하였다.

핵심어: 발화 속도, 조음 속도, 발음의 정확성, 운율의 유창성

참고문헌

이석재, 이숙향, 강석근, 이용주 (2003). 한국인의 영어 음성 코퍼스 설계 및 구축. *말소리, 46*, 159-174.

정현성 (2022). 리듬 및 유창성 지수와 한국 화자의 영어 읽기 발화 청취 평가의 관련성. *말소리와 음성과학, 14*(4), 25-33.

한국지능정보사회진흥원 (2023). *교육용 한국인의 영어 음성 데이터*. Retrieved from https://www.aihub.or.kr/

한승희 (2023). 한국인 발화 다국어 AI Hub 데이터셋 설계 및 구축: 다국어 통번역 낭독체 데이터와 언어교육용 한국인 발화 외국어 음성 데이터. *한국음성학회 2023 봄학술대회 논문집* (pp. 16-27).

* 이 논문은 한국교원대학교 2023학년도 연구년교수 학술지원비 지원을 받아 수행한 연구의 결과임.