English Language & Literature Teaching, Vol.27, No.3 Fall 2021
https://doi.org/10.35828/etak.2021.27.3.27

Identifying Phonological Barriers in Korean-Accented English Using Artificial Intelligence Apps[*] †

Bohyon Chung
(Hanbat National University)

Chung, Bohyon (2021). Identifying phonological barriers in Korean-accented English using artificial intelligence apps. *English Language & Literature Teaching*, *27*(3), 27-45.

Foreseeing the future of English education, this study has incorporated artificial intelligence voice assistant mobile phone applications (AI apps) into English Language Teaching. The study explored: (1) to what extent the current AI apps can dictate Korean-accented English (KoE), and (2) the phonological elements that are commonly misrecognized by the current AI apps. To accomplish these goals, 108 Korean undergraduates, whose English proficiency levels were on average B2 (upper-intermediate level) according to the Common European Framework of Reference for languages (CEFR), participated in the study. Each participant read twenty-five sentences aloud, while AI apps produced transcripts of their dictation. Then the percentage of correctly recognized target words was calculated to determine speech intelligibility. The findings show that KoE attained an accuracy rate of 56.8 percent on average, ranging from 50.1 to 62.7 percent. The KoE barriers associated with the phonological features of less intelligible tokens were also investigated. By bringing attention to the features in comparison to the Lingua Franca Core targets (Jenkins, 2000), the findings suggest the necessity of further research into the feasibility of implementing AI apps within teaching practices, and expanding the boundaries of English as a lingua franca research to KoE, which has not been sufficiently studied.

[intelligibility/mobile applications/Korean-accented English/pronunciation teaching]]

---

[*] A portion of this paper was presented at *the 11th International Conference of English as a Lingua Franca at King's College London, 2018.*

† This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2017S1A5A8021743).

I. INTRODUCTION

Generation Z (Gen Z), defined by birth years from the mid-1990s to 2010, includes currently enrolled undergraduate students. It is acknowledged that Gen Zers matured during the internet era and possess a good command over smart and connective devices. Over 90% of Gen Zers in Korea revealed that using smart devices has reinforced online information seeking for everyday tasks such as acquiring directions, watching movies, reading books, using smart controls, ordering food delivery, and communicating with friends (Ministry of Science and ICT, National Information Society Agency, 2020).

Smart devices are the essential part of choice for most of the Gen Zers, but the use is also rising in automatic speech recognition (ASR) technology. According to a report (Petrock, 2020), it is estimated that 128 million people in the United States, which makes for 44.2 % of smart device users, use ASR-based virtual assistants every month, and this trend has continued growing. In the field of second language (L2) pronunciation training, there are several commercially available software packages featuring ASR technology: MyET from L Labs, Native Accent from Carnegie Speech, SpeechViewer by IBM, Tell Me More from Rosetta Stone, and others. In these programs, students listen to target words in a native English variety, rather than different types of English varieties in the world, and read out the given words. In response, corrective feedback on mispronounced aspects is provided, and learners can focus on correcting the problematic pronunciations.

These ASR-based programs appear to help language learners, as Hincks (2002) argued that the programs allow users automatically identify unique problems (either prosodic or segmental features) on an individual basis (e.g., MyET, SpeechViewer and Tell Me More show spectrograms and speech waves; Native Accent provides visualizations of articulatory help). In addition, it is revealed that ASR pronunciation programs substantially reduce the level of students' anxiety, as learners can be excessively conscious of their own accents (Choi, 2012). Simultaneously, these programs bear some disadvantages. Kim and Jung (2018) argued that the current programs were regarded as providing unbalanced language models (either general American or Received Pronunciation English), and only offer access to a limited range of content for learners to practice, while requiring high maintenance fees for upgrades and support.

Alternatively, ASR technology implemented in artificial intelligence mobile applications (AI apps), such as Google Assistance and Apple Siri, seems to be a viable option for L2 pronunciation training. The advantages offered by AI apps are offset the disadvantages of ASR-based pronunciation training programs. First, it is possible to correspond at the sentence level, not merely repeating pre-selected minimal pairs pronunciation training software requires. Furthermore, AI apps take advantage of Gen Zers existing familiarity with their platforms at no additional cost (Sung, Chang, & Liu, 2016). Sung, Chang, and Liu

(2016) made a useful comparison between the effects of learning tools and students' performances. The results revealed that mobile devices, including mobile phones, personal computers, e-books, and PDAs, have a stronger influence on student performance than conventional teaching methods, such as lecture, group discussion, and experiment. In addition, the mobile devices presented unique characteristics such as tailored settings and real-time access to information as well as instant responses and feedback. The reviewed studies suggest that using mobile devices in English pronunciation teaching would benefit Korean learners who extensively use mobile devices daily. Nonetheless, only when instructional approaches combine with AI apps will the Gen Zers be able to maximize their learning experiences. The present study, therefore, utilizes AI apps to measure the level of intelligibility and to identify phonological barriers of Korean-accented English (KoE) communication.

II. LITERATURE REVIEW

In early L2 pronunciation studies, goals for pronunciation instruction mainly focused on stipulating native language (L1) features. These teaching methods include classical oral approaches, such as the audio-lingual method and situational language teaching. However, the goals of L2 pronunciation practice have been gradually replaced with the notion of intelligibility, being clear enough to be understood by speakers with different language backgrounds (Levis, 2018). A number of papers have built a case that the speech of L2 speakers should have pronunciation qualities that are intelligible enough to foster mutual understanding and, therefore, communication (Derwing & Munro, 2015; Jenkins, 2000; Matsuura, 2012; Smith & Nelson, 2006; Walker, 2010). Among others, Jenkins (2000, 2007) proposed Lingua Franca Core (LFC) features, which establish the minimum phonological requirements for intelligible communication from the perspective of non-native speech. The study (Jenkins, 2000) emphasized that (1) all the consonants except dental fricatives and dark /l/ , (2) consonant clusters in initial and medial word positions, (3) vowel length difference, and (4) tonic stress placement should be properly articulated. Over recent decades, the LFC list (Jenkins, 2000) was replicated for a wide range of speakers from various language backgrounds to provide more context-specific evidence: in European (Wach, 2011), in the Middle Eastern (Rahimi & Ruzrokh, 2016; Walker, 2010), in Southeast Asian (Deterding & Mohamad, 2016), and in an English as a Lingua Franca (ELF) context, namely, a various country-of-origin group (Deterding, 2013). These studies provide useful information for teaching English pronunciation in accommodation with a specific language context.

Nevertheless, it remains challenging to identify which LFC features should be taught to

Korean-speaking learners of English to achieve an intelligible level of English pronunciation. Despite some KSLs appearing to be tolerant when listening to the speech of shared-L1 speakers (Chung, 2018), KoE was once characterized by English language teaching practitioners as "weird, harsh, and quarrel-like" in the previous study (Jenkins, 2007). While increasing number of KoE studies has been carried out, previous studies have sought the phonological characteristics of KoE in the view of native English speakers, chiefly focusing on an inability of KSLs to articulate English sounds. It is likely that KSLs exhibit conflation of sounds between voiceless labiodental fricative /f/ and other stop sounds (stopping fricative) (Lee, 2001); assimilations occurring between /z/ and /ʤ/ (Major & Kim, 1999) and between liquids /r/ and /l/ (Chung, 2011; Kim, 2015; Shin, 2015); and excessive vowel epenthesis in KoE appeared due to the consonant-vowel constraint in Korean (Ha, 2014; Kim & Kochetov, 2011; Lee, 2001).

Simultaneously, it should be observed that human raters generally assessed pronunciation in the previous studies. In the study by Smith and Nelson (2006), raters from different lingual backgrounds evaluated the speech in various English varieties. Furthermore, transcriptions written by human raters from English native or non-native countries served as a data source for the analysis (Chung, 2011; Deterding, 2013; Kim, 2018; Major & Kim, 1999). In the study on KoE intelligibility by Chung (2013), the researcher transcribed 28 speakers' interactions, and Kim (2018) invited two native speakers and two non-native speakers to rate suprasegmental features of KoE. Additionally, in Major and Kim (1999)'s study, five raters assessed palatalization in KoE. These studies were careful to ensure interrater reliability; however, the potential for the reduction of internal consistency remains plausible when individual differences of human raters are considered.

The existing literature suggests that KoE features can be misheard by listeners; however, the objective assessment measures of general speaking abilities and which factors should be prioritized have been investigated less despite their expected usefulness for teaching L2 pronunciation. With the purpose of the LFC features to ensure the intelligibility between human interlocuters, it is worth to investigate whether the same core features would be beneficial to the area of human-AI apps interactions, especially when considering the proficiency in using internet-connected mobile devices among Gen Zers. In the current study, the overall intelligibility level of KoE was measured by means of utilizing AI apps. Based on the transcription generated by AI apps, the phonological barriers of KoE speech were examined to narrow down the priorities in L2 pronunciation training for KSLs. Specifically, the following questions were addressed in the study:

1. To what extent can the current AI apps dictate KoE?
2. Which phonological elements are misrecognized by contemporary AI apps?

III. METHODS

1. Participants

Korean-speaking learners of English (KSLs) enrolled in general English courses at a university located in Daejeon, South Korea, were invited to participate in the study, over a two-semester period in 2018. Overall, four groups of 27 individuals, totaling 108 students participated in the study. Students whose primary language is Korean, but who have been exposed to general American English throughout the school year, ranged in age from 18 to 28 years (M = 20.09). More than half (66.7%) of the participants in the study were younger than 20 years of age, and none of them identified as disabled. Female students comprised 37% (N = 40), whereas the other 63% (N = 68) were male students. According to the mean score of the Oxford Placement Test (OPT), all participants were at English level B2 of CEFR (M = 122.55), corresponding to the upper-intermediate level of English language proficiency. The English language homogeneity between four groups of students was confirmed by a one-way ANOVA test without a statistically significant difference [F (3, 96) = 116.450, p = 0.217]. More than 82% of the students reported that they do not have any study abroad experiences. Other 17.6% of participants studied in eight regions: Australia (5.3%), Canada (10.5%), China (15.8%), the EU (15.8%), Japan (5.3%), Malta (5.3%), the Philippines (26.3%), and the US (15.8%).

2. Stimuli

Each group of participants was presented with an audio script to read aloud. There were four sets of scripts labeled with numbers, *set 1, 2, 3,* and *4.* Each set contained twenty-five items, totaling 100 sentences, each containing one target word. The audio scripts were adapted from the listening session of OPT (Allan, 2004) with permission from the author. A sample item illustrating the target word in bold, follows:

*I gather you've been having trouble with your **hearing**.*

Different sets of audio scripts were assigned to each of the four participating groups. For example, group 1 was given *set 1*, group 2 was given the *set 2*, and so on. The homogeneity among different sets was secured. No statistical differences were found in regard to 1) the readability of each test item in the value of Flesch Reading Ease (FRE) [M = 82.67, SD = .316, F(3, 96) = 31.554, p = .457], 2) the CEFR levels assigned to target words [M = 2.82, SD = 1.52, F(3,96) = 2.428, p = .066], and 3) the number of words contained in a sentence

[M = 10.2, SD = 3.090, F(3, 96) = .111, p = .953], as shown in Table 1. Hence, test material variables should not have influenced the results. Cronbach's alpha coefficient revealed the test instrument with 100 items to be highly reliable (α = 0.891).

TABLE 1
Description of Test Instruments

Set	Items Included	FRE	Target Words Level	Sentence Length (words)
1	#1-25	82.83 (Easy)	A2	9.92
2	#25-50	86.83 (Easy)	A2	10.80
3	#51-75	81.79 (Easy)	B1	10.60
4	#76-100	79.23 (Fairly easy)	A2	10.92

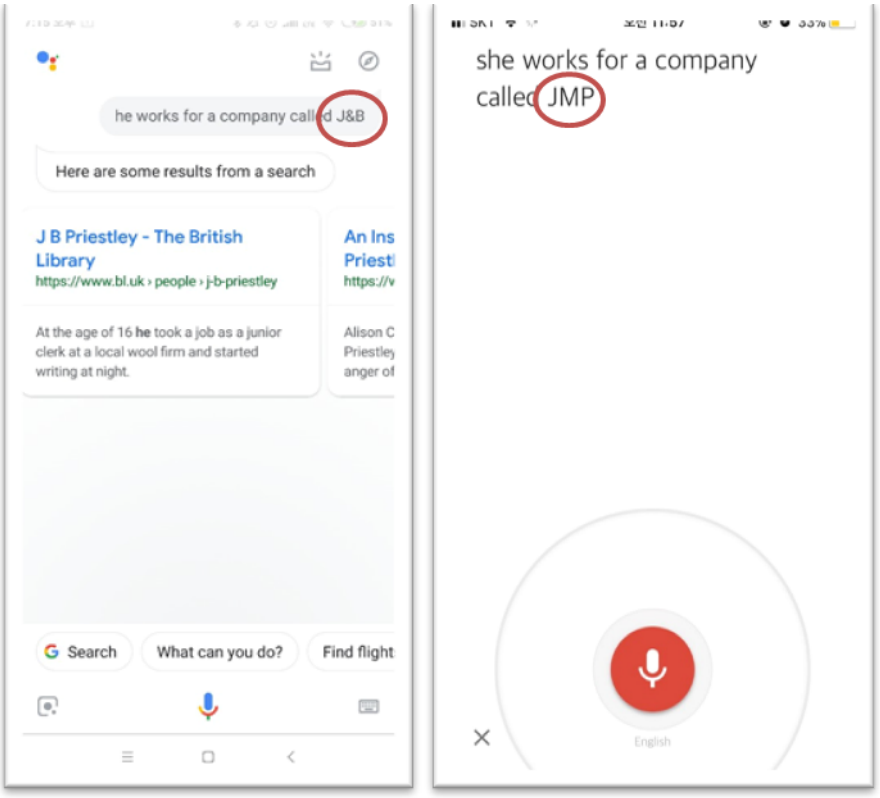
Note. Different sets of the audio scripts were given to four different groups, each comprised of 27 participants.

3. Data Collection

Data collection was conducted after regular class hours in week 7 of both semesters and took approximately 30 minutes to gather. Those who signed institutional review board consent forms were asked to read aloud consecutive sentences from one of the four sets of the audio scripts, each with twenty-five sentences. Immediately after each sentence was read aloud, they were required to take screenshots of transcription, showing how well the AI apps understood each sentence (Figure 1), and to submit them online. During the process, the participants could repeat the sentence within the time allotted but were unapprised of the target words in each sentence. Participants were familiarized with the AI apps by audibly reading voice commands and capturing the transcribed screenshots prior to the actual data collection procedure.

Reliable AI apps should be used to test the intelligibility level of KoE speech. The type of AI apps used in the current study was Google Assistant. Even if Apple Siri was built into the mobile phones of some participants, all students were asked to use Google Assistant either on the app or via the Google Chrome Internet browser. The justification for choosing Google Assistant was because in a recent study (Munster & Thompson, 2019), it was found that Google Assistant recognized voice commands with a 93% accuracy rate, while Apple Siri recognized the same commands at only 83%. Moreover, the annual performance of Google Assistant showed a consistent improvement, but with slight differences (7% vs. 5%). Aside from high performance, familiarity with the Google Android system was also taken into consideration. Google continues to be the more popular among Koreans: seven-in-ten mobile users in Korea indicated that they use Google mobile platforms (DMC Report, 2020). In addition, Google Assistant does not require a dedicated device while navigating the Google Chrome browser. It is, therefore, available on all mobile phones, regardless of mobile phone types used by participants.

FIGURE 1
Screenshots of a Test Item (No. 80) with the Target Word Emphasis



4. Data Analysis

The analysis methodology adopted in the present study was both, quantitative and qualitative: 1) Once data (screenshots of audio transcriptions) were collected, the target words embedded in each sentence were coded as either 0 or 1, representing incorrect and correct answers. Then, the recognition accuracy rate for each target word, was calculated. 2) The qualitative analysis process involved three steps. First, the most commonly misrecognized items (in other words, variables/ divergent forms) for each target word were identified on the basis of the transcriptions produced by AI apps. Subsequently, researchers repeated reading the target words along with their misrecognized forms to tag key factors.

In this process, intelligibility factors gathered from Chung and Bong (2019) served as the core reference for categorization. The study examined how KSLs likely to perceive KoE, and three other varieties of English, distinguished from other native-speaker oriented studies.

Additionally, the study identified intelligibility factors to incorporate binary features that would increase or decrease intelligibility. Those intelligibility factors were adopted in this study and the list with their brief descriptions are presented in Table 2. Among the nine factors in the list below, four are consonant-feature factors (F1.1-F2.2, voicing and liquids), three are vowel-feature factors (F3.1-F4.1, vowel quality and quantity), and the remaining two are related to suprasegmental features. These factors provided a baseline to integrate the target words into groups and to further compare them qualitatively and quantitatively.

TABLE 2
Intelligibility Factors to KSLs (Adapted from Chung & Bong, 2019)

Categorization	Factor	Description	Effect
Voicing	[F1.1] [-voice]	Voiceless consonants to be either in onset, medial, or coda position	+
	[F1.2] [+voice]	Voiced consonants regardless of word positions	-
Liquids	[F2.1] Lateral /l/	Lateral /l/ regardless of word positions	+
	[F2.2] Rhotic /r/	Rhotic /r/ regardless of word positions	-
Vowel Quality	[F3.1] Front, central vowels	/i, ɪ, e, ɛ, æ, ə, ɜ/	+
	[F3.2] Back vowels	/u, ʊ, o, ɔ, ɑ/	-
Vowel Quantity	[F4.1] Vowel length	Diphthongs, epenthetic vowel, and coda voicing effect	+/-
Suprasegmental features	[F5.1] Stress	Stress in polysyllabic words	-
	[F5.2] Re-syllabification	Syllable assimilation, consonant elision	-

Note. + facilitating intelligibility, - impeding intelligibility

IV RESULTS AND DISCUSSION

1. KoE Intelligibility on AI Apps

The first research question was: To what extent can the current AI apps dictate KoE? The findings reveal a 56.8% success rate on average. In other words, when KSLs speak in English to AI apps, they are capable of conveying their intended words within contexts with a 56.8% success rate on average. Comparing by participant groups, the range of recognition success rates, or intelligibility scores, is from 50.1% to 62.7%, with no statistically significant differences between four participant groups [F (3, 96) = 1.687, p = 0.175]. In addition, the recognition success rate is not affected by the difficulty of target words. A Pearson correlation test was conducted between the CEFR level of target words and their mean rates. The results

show a negative relationship with no significant differences (r = -.407, n = 100, p = .259), suggesting that the difficulty level of target words is less likely to have influenced the intelligibility scores.

A couple of studies provide a useful reference point for understanding the intelligibility criteria. In a study by Smith and Nelson (2006), the number of listeners with 60% or more correctly answered rates for a cloze test was examined in the analysis. In another clinical investigation (Monsen, 1981), less than 59% of speech intelligibility was reported to result in listeners who experience difficulty in understanding speech. Therefore, the cut-off for the intelligibility of continuous speech is 60%. Thus, KSLs whose intelligibility rate was 56.8% should be considered a low intermediate level. It is plausible to expect that KSLs would experience some challenges in making AI apps understand KoE.

In regard to categorizing each target word into nine intelligible factors adapted from a recent study (Chung & Bong, 2019), six items out of one-hundred were excluded from the intelligibility factors classification, including “40p” in item 12, “why” in item 21, “pictures” in item 26, “notice” in item 29, “Penny’s” in item 39, and “free” in item 84. For these items, there were few or no misrecognized words generated by AI apps, because they were relatively intelligible, and were therefore considered not representative. Thus, a total of ninety-four items were categorized according to the intelligibility factors. The entire list of target words and their most frequently misrecognized pairing, according to intelligibility factors, is presented in the Appendix.

In seeking to prioritize factors affecting intelligibility, mean of intelligibility rates were compared by nine factors. The findings revealed that intelligibility rates significantly differed statistically between nine factors. The most and least intelligible factors revealed a statistically meaningful difference. Target words involving rhotic /r/ sounds were significantly less intelligible than those affected by the re-syllabification feature [55.2 ±34, F (8, 85) = 2.220, p = 0.034]. Other factors presented slightly different intelligibility rates but did not present significant differences. Hence, the intelligibility factor, according to the intelligibility scores, can be ranked as follows, less intelligible first: rhotic /r/ < front, central vowels ≒ stress ≒ [+Voice] ≒ vowel length ≒ back vowels ≒ lateral /l/ ≒ [-Voice] < re-syllabification.

TABLE 3
Means and Standard Deviations of Intelligibility Scores per Factor

Factor	Counts (%)	*M % (SD)*	Intelligibility Rank*
[F1.1] [-voice]	15 (16)	64.2(23.785)	2
[F1.2] [+voice]	15 (16)	50.5(25.250)	6
[F2.1] Lateral /l/	7 (7.4)	64.1(24.990)	3
[F2.2] Rhotic /r/	7 (7.4)	41.4(23.394)	9
[F3.1] Front, central vowels	15 (16)	44.1(18.271)	8
[F3.2] Back vowels	8 (8.5)	62.0(26.651)	4
[F4.1] Vowel length	14 (14.9)	54.9(22.923)	5
[F5.1] Stress	6 (6.4)	46.3(14.962)	7
[F5.2] Re-syllabification	7 (7.4)	75.4(15.904)	1

Note. N of participants = 27; *1 represents the most intelligible factor and 9 least intelligible.

Although a relatively less intelligible level was observed on average, a couple of additional findings are noteworthy. On the one hand, higher scores of KoE intelligibility were achieved in four factors, namely F1.1 voiceless consonant, F2.1 Lateral /l/, F3.2 back vowels, and F5.2 re-syllabification. They are above the threshold of intelligibility (> 60%), suggesting English spoken by KSLs was intelligible. On the other hand, intelligibility factors could be explained by the sources as emerging from the interplay of L1 phonological influences. The target words involved voiceless consonants in [F1.1] are likely to be recognized more successfully, which can be attributed to the absence of voiced consonants in the Korean language (Chang, 2007; Choi, Kim & Cho, 2016; Kang, 2008; Shin, 2015). In [F2.1], target words including lateral /l/ improved intelligibility scores as Korean liquids have two allophonic realization /l/ and /ɾ/ (Chung, 2011; Kim, 2015; Shin, 2015).

2. Classification of KoE Barriers

The second research question was posted to explore the phonological barriers that are misrecognized by the current up-to-date AI apps. Among the nine intelligibility factors classified, five factors present less than 60% of the intelligibility rates, suggesting those factors are likely to contribute to instances of misrecognition by Korean learners. Accordingly, three target words with the lowest mean rates of intelligibility are presented for each of the less intelligible factors to provide information regarding distinct characteristics arising from each factor. On the basis of the phonological features impeding intelligibility, KoE barriers are illustrated in Table 4.

In [F1.2] [+voiced] factor, target words involving [ð], [z], and [d] have the lowest intelligibility rates. The greatest difficulty occurred when a token featured /th/ in intervocalic position, which was falsely recognized as [st] on most AI apps. This finding merits significant attention as interdental sounds have often claimed to be substituted by allophones

of /s/ produced by Korean learners (Major & Kim, 1999), suggesting a greater possible barrier to KoE intelligibility. Voiced alveolar obstruent sounds [d, z] followed intervocalic /th/ positioned tokens as second and third frequently misrecognized words. For those consonants, AI apps misunderstood those sounds with voiceless counterparts (i.e., [t, s]), respectively.

TABLE 4
Three Most Misrecognized Target Words by AI Apps per Less Intelligible Factors

Target Word	Misrecognized word (%)*	Context	*M % (SD)*
[F1.2] [+voice] related KoE Barrier: involving [ð], [z], and [d]			
loa**th**ed	lo**s**t (17.4)	I really _____ in the late	15 (.362)
joy**s**	Joy**ce** (21.7)	wonder what _____ they	15 (.362)
Dennis	**t**ennis (46.7)	played _____ very	41 (.501)
[F2.2] Rhotic /r/ related KoE Barrier: Rhotic /r/ in all word positions			
rowed	**l**oad (12.5)	top oarsman _____ to his	11 (320)
ho**rr**id	ho**l**iday (50)	a _____ day, isn't	26 (.447)
ra**r**ely	rea**ll**y (57.9)	were _____ dangerous in	30 (.465)
[F3.1] Front, central vowels related KoE Barrier: /æ-ɪ; æ-ɜ; ɪ-ə/ contrasts			
Andean	**I**ndian (60.9)	got an _____ buffalo.	11 (.320)
b**a**th day	b**ir**thday (71.4)	Richard's _____ on	19 (.396)
un**e**xceptional	un**a**cceptable (26.3)	behavior was _____.	26 (.447)
[F4.1] Vowel length related KoE Barrier: longer vowel sounds, inserting epenthetic vowel			
sh**o**d	sh**o**t (36.4)	be _____ immediately.	19 (.396)
Finni**sh**	fini**shed** (45)	if she's _____?	26 (.447)
old	**a**ll	They are _____ things	26 (.447)
[F5.1] Stress related KoE Barrier: word stress			
personNEL	PERsonal (81.4)	for the _____ side	22 (.424)
TWO days'	ToDAY's (21.4)	through _____ money	41 (.501)
PIrated	paRAde (12.5)	depend on _____decisions	41 (.501)

Note. N = 27; * Most frequently misrecognized words are shown with percentage of total incorrect words

With respect to [F2.2] rhotic /r/ factor, AI apps are likely to fail to recognize the speech, arising from the words with /r/ in word-initial, but also in word medial positions. Liquid contrasts have been considered as an important pair in English pronunciation teaching in Korea because the Korean language utilized only the /l/ phoneme. Thus, the common conflation between English liquid sounds by KSLs is to be expected. The findings confirmed /r/ and /l/ conflation by KSLs in all three cases, where /l/ was in place of /r/ both in initial and medial word positions, as in rowed (load), horrid (holi-) and rarely (really). This finding was congruent with previous studies that identified liquid sound assimilation by KSLs (Chung, 2011; Kim, 2015). Noticeably, in these two consonant related features, namely [F1.2] [+voiced] and [F2.2] rhotic /r/, the word position was not a strong determinant. The findings suggest that consonants in word medial and final positions can still create environments reducing intelligibility to Korean students, not limited to the word-initial position, as suggested in the core LFC features (Jenkins, 2000, 2007).

On account of [F3.1] front, central vowels factor, two front vowels with tongue height differences /æ, ɪ /, caused the source of the problem, as in Andean, bath, and unexceptional. The contrast between target words and their misrecognized forms indicate that /æ/ is likely to be replaced with /ɪ/ or /ɜ/, and substitute /ɪ/ for /ə/. These vowel pairs are likely to be considered as infrequent phonemes in that they are distinct in acoustic properties, creating less confusion. That is, there are not many minimal pairs including /æ-ɪ; æ-ɜ; ɪ-ə/ contrasts. By comparison, it was illustrated that the English vowel /æ/ produced by KSLs tends to be perceived as /eɪ, ɛ/ and /ɪ/ as /i/ (Bent & Kewley-Port, 2010; Heo & Park, 2012). As such, more front vowel quality has been suggested for the speech produced by KSLs. Confusion between these rather uncommon pairs /æ-ɪ; æ-ɜ; ɪ-ə/ by Korean learners, however, suggests that front and central vowels, in general, ought to be given greater attention. Since /æ, ɪ, ɜ -/ sounds are considered "new" vowels to Korean learners, owing to the fact that no lax vowels exist in L1 (Kim, 2010; Lee & Cho, 2020, Shin, 2015), front and central vowels bear tremendous potential to hamper intelligibility of KoE.

In the same vein, [F4.1] vowel length factor presented typical L1 phonological influences. Korean does not have tense-lax distinction (Kim, 2010, Shin, 2015) and does not allow consonant clusters in onsets or codas resulting in vowels with longer duration (Heo & Park, 2012), or adding an epenthetic vowel [ɨ] (Ha, 2014; Kim & Kochetov, 2011; Lee, 2001). In that sense, given target words, "Finnish," and "old," were misrecognized as "finished," and "all," with the second and third lowest intelligibility rates. Regarding a target word "shod," the misrecognized pair was "shot." This pair contrasted by vowel length, but this case should consider more features than the phonological ones. Owing to the pair varying in word difficulty levels (CEFR level, shod: C1 vs. shot: A1), AI apps might have selectively identified with the easier word, instead of the one relatively more difficult.

Lastly, as for F5.1 stress factor, the lowest intelligibility rates were found in the target words "personnel, two days', and pirated." At the cost of unclear stress assignment, each target word was falsely recognized as "personal, today's, and parade." Korean is commonly referred to as a tonal language with no lexical stress (Jeon, 2015), and is not a stress-timed, nor a syllable-timed language. In this circumstance, it is highly likely that KSLs may find it difficult to grasp which syllable is stressed and unstressed, causing KSLs to speak less intelligible English.

V. CONCLUSION

The primary goal of this study was to measure the intelligibility of KoE using AI apps. The results indicated that 56.8% of KoE is intelligible for recognition by AI apps, suggesting that Korean learners (KSLs) may find it a struggle to have AI apps understood in KoE. Based on the transcription generated by AI apps, KoE barriers were outlined with reference to the factors affecting intelligibility to KSLs (Chung & Bong, 2019). Target words involving rhotic /r/, front and central vowels, stress, voiced consonants, and vowel length reduce intelligibility scores (< 60%), largely considered less intelligible. At the other extreme are target words including voiceless consonants, lateral /l/, back vowels, and re-syllabification phenomena, which appeared intelligible (> 60% of intelligibility rates). The limited scope of analysis in the current study, however, prevented from testing the other higher levels of understanding: comprehensibility and interpretability (Smith & Nelson, 2006). Alternatively, future studies delving further into the KoE intelligibility in a discourse level of interaction rather than in a sentence-level investigation would encourage meaningful discussions.

VI. PEDAGOGICAL IMPLICATION: MOST USEFUL FEATURES TO TEACH FOR KSLS

According to identified KoE barriers, the extent of LFC features targeting KSLs has been explored. Table 5 compares the original scope of LFC features (Jenkins, 2000) and the Korean-context specific LFC.

First, in regard to consonant inventory, it may not be adequate to eliminate [θ] and [ð] from the consonant inventory targeting KSLs, contrary to the original LFC features. A target word categorized in the consonant voicing category, loathed, was likely to be recognized as lost by AI apps, and merely 15% of the KSLs correctly performed. It was voiced dental fricative [θ] in the medial position replaced with [s] that caused the problem. Voiced consonants, [z] and [d], also caused KoE problems, as in joys and Dennis. Because English fricatives and affricates are more widely distributed than Korean consonants, KSLs will benefit significantly with practices focusing on the sounds. In addition, the KoE barriers indicate that more emphasis on rhotic /r/ regardless of word positions, should be included in pronunciation training in the Korean context. Consonant clusters in the initial and medial positions did not result in misrecognition in the current study, nor did connected speech processes, elision, or assimilation across syllable boundaries.

Next, in terms of vowels, confusion between front and central vowels are likely to be a contributing factor in degrading KoE intelligibility. Some of the test items involving the front or central vowels, such as bath, ill, mass, mini, and set, were misrecognized, potentially due

to more front vowel quality produced by KSLs. In particular, /æ-ɪ; æ-ɜ; ɪ-ə/ contrasts ought to be reflected in pronunciation training. Furthermore, vowel length variation, owing to coda consonant voicing, vowel epenthesis, and lack of lax vowels in L1 vowel inventory, may be worth noting for KSLs and should retain priority importance in teaching pronunciation.

Finally, although it was originally implied that word stress is less important than stress in LFC (Jenkins, 2000), the findings showed that misplaced stress could reduce intelligibility scores. There were cases when assigning stress to wrong syllables resulted in misrecognition by AI Apps, such as personnel, two days, and pirated. Thus, where to apply stress, even at the word level, should not escape the attention of KSLs.

TABLE 5
ELF Users and KSLs Pronunciation Targets

	ELF users target (Jenkins, 2000)	KSLs target
Consonantal inventory	All sounds except [θ], [ð], and [ɬ]	All sounds involving [ð], [z], and [d]
	Rhotic /r/ only	Rhotic /r/ in all word positions
	Intervocalic [t] only	-
	Close approximations to core consonant sounds generally permissible	Voicing realization of fricatives and affricates
Consonant clusters	Word initial and medial	-
	L1 rules of elision	-
Phonetic requirements	Aspiration after /p, t, k/	-
	Consonant effect on preceding vowel length	Consonant effect on preceding vowel length
Vowel sounds	Vowel length	longer vowel sounds, inserting epenthetic vowel
	nurse/ɜ:/ vowel	/æ-ɪ; æ-ɜ; ɪ-ə/ contrasts
Stress	Nuclear stress	Word stress

Note. ELF: English as lingua franca; KSLs: Korean-speaking learners of English

The novel approach in the current study, which used AI applications to assess intelligibility level, was one of the main accomplishments. The findings provide evidence that overall intelligibility can be measured by AI apps and how the speech is realized, including the fluency of the speech, can be captured by transcriptions of AI apps. In contrast, the current commercial pronunciation training programs featuring voice recognition technology provide feedback or scores primarily through comparison of sound patterns at the syllabic level, only encouraging practice in a matter of making sounds similar to native speakers. Transcriptions produced by AI apps, however, provide useful information, which is helpful for locating the sounds and words that are less intelligible in their communication. Likewise, the use of AI apps gives researchers and English learners the capacity for an objective understanding of how KoE speech in general may sound to other people.

Therefore, for researchers and teaching practitioners, it is possible to develop a list of which phonemic confusions can directly hinder the intelligibility level by target students.

The previous studies may cast doubt on internal validity as human evaluators are involved in data analysis. This internal validity issue may be addressed by utilizing AI apps, by providing a great alternative to human raters, offering an objective assessment. Moreover, using the AI apps for speech intelligibility measurements offer Gen Z learners of English the advantage of a platform and learning method with which they are already familiar. Predicting how their speech may have been heard by other people by using their mobile devices is helpful, both inside and outside of the classroom. This would allow them to work towards a solution to enhance their speech intelligibility, thereby increasing English learners' autonomous learning abilities in pronunciation training. Furthermore, if AI development ultimately yields technology that can understand the locutionary meaning of context, students will no longer need to obsess over minor issues in their pronunciation, as well as being concerned by teachers' oral pronunciation feedback. Rather, they will be able to effectively understand the main features affecting their intelligibility by checking their pronunciation through AI apps. Thus, AI apps are more likely to be utilized in the future as self-directed learning tools for language learners.

ACKOWLEDGEMENT

I am grateful to the reviewers for their insights and to Dave Allan for permission to reuse the listening section of OPT1.

REFERENCES

Allan, D. (2004). *Oxford placement test 1*. Oxford: Oxford University Press.

Bent, T., & Kewley-Port, D. (2010). Across-talker effects on non-native listeners' vowel perception in noise. *Journal of Acoustical Society of America, 128*(5), 3142-3151.

Chang, C. B. (2007). The status of voicing and aspiration as cues to Korean laryngeal contrast. *US Berkeley Phonology Lab Annual Report, 3*, 330-344.

Choi, J., Kim, S., & Cho, T. (2016). Phonetic encoding of coda voicing contrast under different focus conditions in L1 vs. L2 English. *Frontiers in Psychology, 7*, 1-17. https://doi.org/10.3389/fpsyg.2016.00624

Choi, S. (2012). Effects of a Web-Based English Learning Program for Korean Adult Learners. *Studies in English Language & Literature, 38*(4), 353-377.

Chung, B. (2018). A study on the linguistic prejudice against Korean-accented English. *English Language & Literature Teaching, 24*(2), 1-20.

Chung, B., & Bong, H. K. M. (2019). A study on the relation between intelligibility and attitudes. *English Teaching, 74(*2), 103-123. https://doi.org/10.15858/engtea.74.2.201906.103

Chung, C-W. (2011). The phonological behavior of liquids in English: Similarities and differences. *The Research Institute of Humanistics, 6*, 297-329.

Derwing, T. M., & Munro, M. J. (2015) *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.42

Deterding, D. (2013). *Misunderstanding in English as a lingua franca: An analysis of ELF interactions in South-East Asia.* Boston, MA: Mouton de Gruyter.

Deterding, D., & Mohamad, N. R. (2016). The role of vowel quality in ELF misunderstandings. *Journal of English as a Lingua Franca, 5*(2), 291-307. https://doi.org/10.1515/jelf-2016-0021

DMC Report. (2020, May 27). The future and current state of mobile Internet in 2020: (2) Contents and service [PDF file]. Retrieved from https://www.dmcreport.co.kr/report/trendBrief/freeView?reportcode=DMCTBF20200069&drtopdeth=RPT_TYPE_1

Ha, H-J. (2014). Counteracting Korean English learners' epenthesis of the vowel [ɨ]. *Concordia Working Papers in Applied Linguistics, 5,* 229-239.

Heo, Y., & Park H. (2012). American learners' identification of Korean vowel categories. *Harvard Studies in Korean Linguistics, 14*, 77-84.

Hincks, R. (2002). Speech recognition for language teaching and evaluating: A study of existing commercial products. Proceedings of the 7th International Conference on Spoken Language Processing, Denver, Colorado, USA, 733-736.

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford, England: Oxford University Press.

Jenkins, J. (2007). *English as a lingua franca: Attitude and identity*. Oxford, England: Oxford University Press.

Jeon, H-S. (2015). Prosody. In L. Brown & J. Yeon (Eds.), *The handbook of Korean linguistics* (pp. 41-58). Malden, MA: Wiley-Blackwell.

Kang, Y. (2008). Tensification of voiced stops in English loanwords in Korean. *Harvard Studies in Korean Linguistics, 12*, 179-192.

Kim, J-E. (2010). Perception and production of English front vowels by Korean speakers. *Phonetics and Speech Sciences, 2*, 51-58.

Kim, J-E. (2018). An analysis of English pronunciation for high-level proficiency adult learners. *Phonetics and Speech Sciences, 10*(2), 39-44. https://doi.org/10.13064/KSSS.2018.10.2.039

Kim, J-Y. (2015). Second language acquisition: Phonology. In L. Brown, & J. Yoon (Eds).

The Handbook of Korean Linguistics (pp. 373-388). Oxford: Wiley Blackwell. https://doi.org/10.1002/9781118371008

Kim, S., & Jung, H. (2018). A study on the utilization of speech recognition technology in foreign language learning applications: focusing on English and French speech. *Journal of Digital Contents Society, 19*(4), 621-230. https://doi.org/10.9728/dcs.2018.19.4.621

Kim, K., & Kochetov, A. (2011). Phonology and phonetics of epenthetic vowels in loanwords: Experimental evidence from Korean. *Lingua, 121,* 511-532. https://doi.org/10.1016/j.lingua.2010.10.012

Lee, J-A. (2001). Korean speakers. In M. Swan & B. Smith (Eds.) *Learner English: A teacher's guide to interference and other problems* (pp. 325-342). Cambridge, England: Cambridge University Press. https://doi.org/10.1017/CBO9780511667121.023

Lee, S., & Cho, M-H. (2020). The impact of L2-learning experience and target dialect on predicting English vowel identification using Korean vowel categories. *Journal of Phonetics, 82,* 1-20. https://doi.org/10.1016/j.wocn.2020.100983

Levis, J. (2018). *Intelligibility, oral communication, and the teaching of pronunciation.* Cambridge, England: Cambridge University Press. https://doi.org/10.1017/9781108241564.003

Major, R. C., & Kim, E. (1999). The similarity differential rate hypothesis. *Language Learning, 49*(1). 151-180. https://doi.org/10.1111/j.1467-1770.1996.tb01244.x

Matsuura, H. (2012). Japanese learners' evaluative reactions and affective responses toward English accents. *The Shogaku Ronshu, 80*(4), 21-32.

Ministry of Science and ICT, National Information Society Agency. (2020, December). *The survey on smartphone overdependence* (Publication No. NIA VII-RSE-A-200117). Retrieved September 7, 2021, from the World Wide Web: http://www.index.go.kr/potal/main/EachDtlPageDetail.do?idx_cd=3061

Monsen, R. B. (1981). A usable test for the speech intelligibility of deaf talkers. *American Annals of the Deaf, 126*(7), 845-852.

Munster, G., & Thompson, W. (2019). *Annual digital assistant IQ test.* Retrieved September 7, 2021, from the World Wide Web: https://loupventures.com/annual-digital-assistant-iq-test/

Petrock, V. (2020). *Voice assistant and smart speaker users 2020.* Retrieved September 7, 2021, from the World Wide Web: https://www.emarketer.com/content/voice-assistant-and-smart-speaker-users-2020

Rahimi, M., & Ruzrokh, S. (2016). The impact of teaching lingua franca core on English as a foreign language learners' intelligibility and attitudes towards English pronunciation. *Asian Englishes, 18*(2), 141-156. https://doi.org/10.1080/13488678.2

016.1173466

Shin, J. (2015). Vowels and consonants. In L. Brown & J. Yeon (Eds.), *The handbook of Korean linguistics* (pp. 3-21). Malden, MA: Wiley-Blackwell. https://doi.org/10.1002/9781118371008.ch1

Smith, L. E., & Nelson, C. L. (2006). World Englishes and issues of intelligibility. In B. B. Kachru, Y. Kachru, & C. L. Nelson (Eds.), *The handbook of World Englishes* (pp. 428-445). Malden, MA: Blackwell. https://doi.org/10.1002/9780470757598.ch24

Sung, Y-T., Chang, K-E., & Liu, T-C. (2016). The effects of integrating mobile devices with teaching and learning on students' learning performance: A meta-analysis. *Computers & Education, 94*, 252-275. https://doi.org/10.1016/j.compedu.2015 11.008

Wach, A. (2011) Native-speaker and English as a lingua franca pronunciation norms: English majors' views. *Studies in Second Language Learning and Teaching, 1*(2). 247-266. https://doi.org/10.14746/ssllt.2011.1.2.5

Walker, R. (2010). *Teaching the pronunciation of English as a Lingua Franca.* Oxford, England: Oxford University Press. https://doi.org/10.1080/17501229.2011.607918

APPENDIX

The List of Target Words and Their Most Frequently Misrecognized Tokens by AI Apps

Factor	Item	Target Word	Misrecognized	Factor	Item	Target Word	Misrecognized
[F1.1]	6	**F**riesian	**d**epression	[F3.1]	40	m**i**ni	m**a**ny
[-voice]	81	**H**A	**e**ighty	Front,	47	m**e**nial	m**e**nu
	96	pre**c**edent	pre**s**ident	Central	58	m**a**ss of	m**e**ss of
	55	fi**s**cal	phy**s**ical	Vowels	51	ch**air**ed	ch**ar**ged
	17	some **k**inds	some**t**imes		57	s**e**t	s**ai**d
	59	conte**xt**	conta**ct**		72	**i**ll	w**e**ll
	20	**St**eve	**b**eef		97	**ter**rain	**tr**ain
	4	presen**t**	presen**ce**		83	w**o**men	w**o**man
	99	fa**c**e	Fai**th**		35	l**a**sh	l**i**st
	7	**f**ounded	**b**ounded		45	s**ea**ling	s**e**lling
	10	**sh**e's	**h**e's		93	He **a**nd	He **i**n
	37	text**s**	tex**t**		46	p**e**rsecution	pr**o**secution
	1	**h**earing	**f**eeling	[F3.2]	91	r**o**d	t**oa**d
	11	**ch**ess	**C**os	Back	87	sc**a**rred	sc**a**red
	56	tigh**t**	Ty**pe**	Vowels	31	C**o**nsul	C**ou**ncil
[F1.2]	9	loa**th**ed	los**t**		68	l**au**nch	l**un**ch
[+voice]	15	joy**s**	Joy**ce**		71	**ar**med	b**e**yond
	74	**D**ennis	**tennis**		27	w**ou**ldn't	wh**a**t
	79	**n**ewer	**y**our		36	t**a**lk	t**oo**k
	5	**Eu**ston	**Hou**ston		86	**au**tistic	**ar**tistic
	41	**b**uyable	**v**iable	[F4.1]	44	sh**od**	sh**o**t
	65	ca**b**in up	Ke**v**in up	Vowel	69	Finnish	finish**ed**
	67	**b**ison	**v**ison	Length	73	**o**ld	**a**ll

	64	**d**ry	**t**ry		75	cru**de**	cru**dités**
	53	ri**d**ing	wri**t**ing		50	slo**t**	sl**ow**
	3	**gr**ows	**cl**osed		54	hear**ten**ed	heart **a**nd
	80	J**&**B	J**M**B		23	glass	glass**es**
	2	compo**s**ition	compe**t**ition		89	**et**ernal	**int**ernal
	49	**v**iew	**y**ou		98	to da**te**	tod**ay**
	28	**bl**ended	**spl**endid		16	c**a**n	c**a**n't
[F2.1]	18	tearfu**l**	(quie)t here fo**r**		8	my **ow**n	my **o**n
Lateral	76	barre**l**	bette**r**		38	hatch	hatch**ed**
/l/	94	loya**l**	lawye**r**		24	injur**ed**	injur**y**
	77	arriv**al**	arrive		25	c**a**n see	c**a**n't see
	34	**gl**ass	**gr**ass	[F5.1]	92	person**nel**	**per**sonal
	62	**cl**ouds	**cr**oss	Stress	42	**two** days'	**to**day's
	63	be**ll**y	be**rr**y		100	**pi**rated	**pa**rade
[F2.2]	88	**r**owed	**l**oad		48	**co**mmittee	**co**medy
Rhotic	14	ho**rr**id day	ho**l**iday		60	de**fia**nce	**de**fense
/r/	33	ra**re**ly	rea**ll**y		52	**d**elicate	**d**elete
	70	**tri**als	**t**ails	[F5.2]	66	stock**s**	stock
	22	**r**eading	**l**eading	Re-	95	free	frea**ky**
	43	I **r**eckon	**l**ike	syllabifi-	13	clock**s**	clock
	19	**re**gain	**l**eague	cation	32	**wa**shing	(tracking)ing
[F3.1]	90	**A**ndean	**I**ndian		78	inven**tory**	inven**tor**
Front,	85	b**a**th day	b**ir**thday		30	collect**ed**	collect
Central	61	un**e**xceptional	un**a**cceptable		82	hair-drier**s**	hair-dryer

Note. Target words are sorted by mean rates of intelligibility in descending order (from less to more intelligible tokens); the main source of misrecognition by AI apps is highlighted in bold.

Examples in: English

Applicable Languages: English

Applicable Levels: Tertiary

Bohyon Chung (Lecturer)
Nomad College
Hanbat National University
125, Dongseodae-ro, Yuseong-gu, Daejeon, South Korea
Email: bchung@hanbat.ac.kr

Received: July 15, 2021
Reviewed: August 15, 2021
Revised version received: September 4, 2021