

Phonetic Features Influencing Native English Speakers and Korean EFL Teachers' Comprehensibility and Accentedness Judgments

Hyunji Kim

(Korea National University of Education)

Hyunsong Chung*

(Korea National University of Education)

Kim, Hyunji, & Chung, Hyunsong (2023). Phonetic features influencing native English speakers and Korean EFL teachers' comprehensibility and accentedness judgments. *Studies in Foreign Language Education*, 37(4), 83-107.

The purpose of this study was to investigate and compare the phonetic features that influence the judgments of comprehensibility and accentedness made by native English speakers (NESs) and Korean English teachers (KETs). Further, it discusses implications for pronunciation instruction and assessment in Korean EFL context. Adapted from Yang's (2021) research design, 10 NESs and 10 KETs each rated comprehensibility and accentedness for 30 speech samples, and the researcher analyzed which phonetic features of the speech samples influenced their judgments. The results of a multiple regression analysis revealed that the predictors of NESs' comprehensibility judgments were the number of syllable structure errors and VarcoV, which is one of rhythm metrics, while the predictors of their accentedness judgments were the number of syllable structure errors, the number of segmental errors, and articulation rate. KETs were influenced by articulation rate, the number of syllable structure errors, and VarcoV when they made judgments on comprehensibility and accentedness. Based on the results, the guidelines for English pronunciation instruction and assessments conducted in Korean EFL contexts are suggested.

I. INTRODUCTION

The Communicative Language Teaching (CLT) approach prioritizes fluency over accuracy and sometimes minimizes emphasis on some pronunciation aspects as they are not seen as major obstacles to successful communication. In this context, many researches focused on the relationship between comprehensibility and accentedness. Generally, it is known that L2 pronunciation instruction should help L2 learners achieve maximum comprehensibility (e.g., Derwing & Munro, 1997; Munro & Derwing, 1995a, 1995b, 1999), and linguistic errors leading to accentedness do not necessarily impede comprehensibility. Meanwhile, some

* First author: Kim, Hyunji; Corresponding author: Chung, Hyunsong

prior research tried to identify the linguistic features that contribute to comprehensibility and accentedness judgments (Derwing et al., 1998; Flege et al., 1995; Kang, 2010; Munro & Derwing, 2001; Trofimovich & Issacs, 2012; Yang, 2021). Uncovering which linguistic features are more closely linked to comprehensibility than to accentedness could offer valuable guidance for pronunciation instruction.

So far, a significant portion of research on L2 comprehensibility and accentedness has relied on native English speakers as evaluators. There have been only a few studies that explored and compared judgments of L2 speech by both native and nonnative English speakers (Bent & Bradlow, 2003; Crowther et al., 2016; Foote & Trofimovich, 2018; Saito et al., 2017). Within this context, the primary objective of this study is to uncover and compare which phonetic features impact the judgments of comprehensibility and accentedness by NESs and KETs. The aim was to shed light on defining optimal goals for English pronunciation instruction and to draw implications for speaking assessments, particularly those conducted by KETs.

II. LITERATURE REVIEW

1. Assessing L2 Pronunciation

In the realm of second language pronunciation instruction, two main guiding principles are commonly recognized: the “Native Principle” and the “Intelligibility Principle.” The Native Principle posits that the ultimate objective of teaching and learning second language (L2) pronunciation is to attain a pronunciation that closely mirrors that of native speakers (Levis, 2005). However, it has been observed by many experts in the field that most L2 speakers often retain a foreign accent, particularly if they begin learning the L2 after reaching puberty (Flege et al., 1995). Additionally, achieving a pronunciation very similar to that of native speakers is rare and may only be feasible when the learner’s native language shares significant linguistic similarities with the target language (Saito et al., 2020). Given these circumstances, an increasing number of researchers have introduced the concept of the Intelligibility Principle as a crucial criterion for evaluating L2 pronunciation (Crowther, 2020). Derwing and Munro (2015) contended that the focus in assessing and instructing L2 pronunciation should be on effective communication rather than striving for native-like pronunciation. Similarly, Levis (2018) underscored the importance of setting attainable goals in L2 learning, specifically emphasizing comprehensibility, intelligibility, and effective communication.

2. Comprehensibility and Accentedness

The present study centers its attention on comprehensibility and accentedness, both of

which are grounded in how listeners intuitively perceive L2 speech. Accentedness pertains to how listeners perceive the degree to which L2 speakers can reproduce the sound patterns of native speakers (Derwing & Munro, 2009). Conversely, comprehensibility deals with the extent to which a listener encounters difficulty when attempting to grasp L2 speech. Besides these two measures, intelligibility is also a widely used scale when evaluating L2 pronunciation. Intelligibility is the degree to which the utterance the speaker intended to deliver is actually understood. This study excluded intelligibility and only targeted comprehensibility and accentedness for two reasons. According to Levis (2006), intelligibility refers to how raters subjectively perceive their level of understanding or the ease of L2 speech in a broad sense. Also, typically measured through orthographic transcription (e.g. Derwing & Munro, 2009), intelligibility was not assessed in this study due to methodological inappropriateness of dictation, where raters evaluated speech samples reading the same paragraph.

Numerous efforts have been made to investigate the connection between comprehensibility and accentedness (Munro & Derwing, 1995a, 1995b, 1999). Munro and Derwing ascertained that even heavily accented L2 speech can still be highly understandable (Munro & Derwing, 1995a). These results were replicated in their subsequent investigations, which delved into how comprehensibility and accentedness impact the time it takes to process sentences (Munro & Derwing, 1995b), as well as in other studies involving non-native speakers from diverse language backgrounds (Derwing & Munro, 1997). To summarize, comprehensibility and accentedness are interrelated yet separate concepts.

3. Linguistic Features Influencing Comprehensibility and Accentedness

Several studies have concentrated on identifying linguistic factors that underlie listeners' immediate and instinctive assessments of comprehensibility and accentedness (Derwing et al., 1998; Flege et al., 1995; Kang, 2010; Munro & Derwing, 2001; Yang, 2021). Munro and Derwing (2006) found that segmental differences with high functional importance, such as those involving English sounds like /ɪ/ and /l/ rather than /s/ and /θ/, affect both comprehensibility and accentedness evaluations, while those with low functional importance have only a minimal impact on comprehensibility. Kang et al. (2010) demonstrated that suprasegmental features can account for a significant portion of the variability in ratings for comprehensibility and oral proficiency.

Recent efforts have sought to determine the relative importance of various suprasegmental features on comprehensibility and accentedness. Kang (2010) examined 11 suprasegmental variables to gauge their impact on comprehensibility and accentedness, and revealed that comprehensibility were primarily linked to speaking rate while accentedness was most accurately predicted by measures of pitch range and word stress. Yang (2021) affirmed that the key predictors of comprehensibility were the number of pauses, sentence

stress appropriateness, and articulation rate, whereas accentedness ratings correlated most strongly with speech rate and prominence frequency through a multiple regression analysis.

In a broader context, Issacs and Trofimovich (2012) and Trofimovich and Issacs (2012) unveiled that judgments of comprehensibility were best anticipated by type frequency, word stress, and grammatical accuracy. Conversely, rhythm and word stress emerged as the foremost predictors of accentedness ratings. Saito et al. (2016) explored the factors influencing comprehensibility and accentedness concerning the oral proficiency of non-native speakers and found out that learners at different proficiency levels exhibited distinct determinants of these attributes.

4. Listener Factors

Judgments of comprehensibility and accentedness can also be influenced by factors related to the listeners or raters themselves. Some research has identified factors that impact how listeners judge L2 comprehensibility and accentedness. Kennedy and Trofimovich (2010) demonstrated that listeners tended to find L2 speech more understandable when they were familiar with accented speech. Additionally, certain studies have revealed that listeners' linguistic or teaching backgrounds can also impact their assessments of comprehensibility and accentedness (Saito et al., 2016; Saito et al., 2017). Some research explored the distinctions and commonalities between judgments of L2 comprehensibility made by native and L2 listeners (Foote & Trofimovich, 2018; Saito et al., 2020).

To the best of the author's knowledge, there has been limited research that made comparison between evaluations of L2 speech made by both native and nonnative English speakers (Bent & Bradlow, 2003; Crowthe et al., 2016; Foote & Trofimovich, 2018; Gordon-Salant et al., 2019; Saito et al., 2017). Additionally, there is a scarcity of studies that have investigated the differences in factors influencing judgments of comprehensibility and accentedness between NESs and nonnative speakers of Korean origin. For example, Park (2022) explored how L2 English utterances by Korean speakers are assessed by both native English speakers and native Koreans, concurrently examining linguistic factors that impact how listeners assess comprehensibility.

To address this gap in the realm of research on L2 pronunciation, this study aims to compare the phonetic features that influence judgments of L2 comprehensibility and accentedness by both NESs and Koreans. Specifically, this research intends to focus on Koreans who teach English in Korean EFL context. This choice is driven by the fact that a significant proportion of L2 learners in Korea receive instruction in pronunciation and speaking skills from KETs. To do so, the researchers set the following research questions:

- (1) Which phonetic features influence NESs' and Korean EFL teachers' judgments of comprehensibility and accentedness respectively?

- (2) How are the phonetic features that influence NESs' and Korean EFL teachers' judgments of comprehensibility and accentedness different?

III. METHODOLOGY

1. Participants

1) Speech Samples

30 speech samples were chosen from the Speech Accent Archive (George Mason University, 2023), which is a collection of speech accents representing various language backgrounds. Only recordings from female speakers were utilized to avoid potential complications arising from gender difference. These 30 speech samples consist of Korean English learners with differing levels of English proficiency and ages ranging from 18 to 57.

All speakers were given the same English paragraph consisting of 5 sentences. The paragraph employed commonly used English vocabulary while incorporating a range of English sounds and sound combinations. The provided elicitation paragraph (George Mason University, 2023) reads as follows:

Please call Stella. Ask her to bring these things with her from the store: Six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the kids. She can scoop these things into three red bags, and we will go meet her Wednesday at the train station.

2) Raters

The raters were 10 native speakers of North American English (2 males and 8 females) and 10 Korean EFL teachers (3 males and 7 females). All NESs except one was born in the United States of America and their length of residence in Korea varied from none to 10 years. Among 10 KETs, 4 of them are teaching in Korean high schools while the remaining 6 are teaching English in middle schools in Korea. Their teaching experiences varied from 3 to 16 years, all of them having plenty of experience in English speaking assessment. The background information about NES and KET raters is in Table 1 and Table 2. The raters rated 30 speaking samples on a 7-point Likert scale of comprehensibility and accentedness simultaneously. All actions involving the participants in this study adhered to the ethical standards established by the institutional committee. Approval for the study was obtained from the Institutional Review Board (IRB) of Korea National University of Education under reference [KNUE-202305-SB-0059-01].

Table 1. Background Information About the NES Raters

Gender	Age	Nationality	LOR in Korea
Male	50s	USA	10 years
Male	20s	USA	1 year
Male	30s	USA	none
Male	40s	USA	none
Male	20s	South Korea	2 years
Male	30s	USA	none
Female	30s	USA	none
Female	50s	USA	7 years
Male	40s	USA	5 years
Male	20s	USA	none

Table 2. Background Information About the KET Raters

Gender	Age	School	Teaching Experience
Female	30s	High school	8 years
Female	20s	High school	3 years
Male	30s	High school	6 years
Female	20s	High school	5 years
Female	30s	Middle school	7 years
Male	40s	Middle school	15 years
Female	30s	Middle school	7 years
Female	40s	Middle school	12 years
Female	50s	Middle school	16 years
Male	30s	Middle school	10 years

2. Procedures

The current study followed a three-step process. Initially, the researcher examined a range of phonetic features in the 30 speech samples using the voice analysis software, Praat version 6.2.17 (Boersma & Weenink, 2022). Figure 1 provides an illustration of the speech sample analysis conducted with Praat. Subsequently, the raters assessed L2 comprehensibility and accentedness for each of the 30 speech samples using a seven-point Likert scale. To prevent the sequence of speech samples from influencing the participants' judgments, the speech samples were presented in a randomized order utilizing Praat's 'PermuteBalancedNoDoublets' function. Lastly, the researcher analyzed and compared the phonetic characteristics that influenced judgments of L2 comprehensibility and accentedness between the NES listeners and the KET listeners.

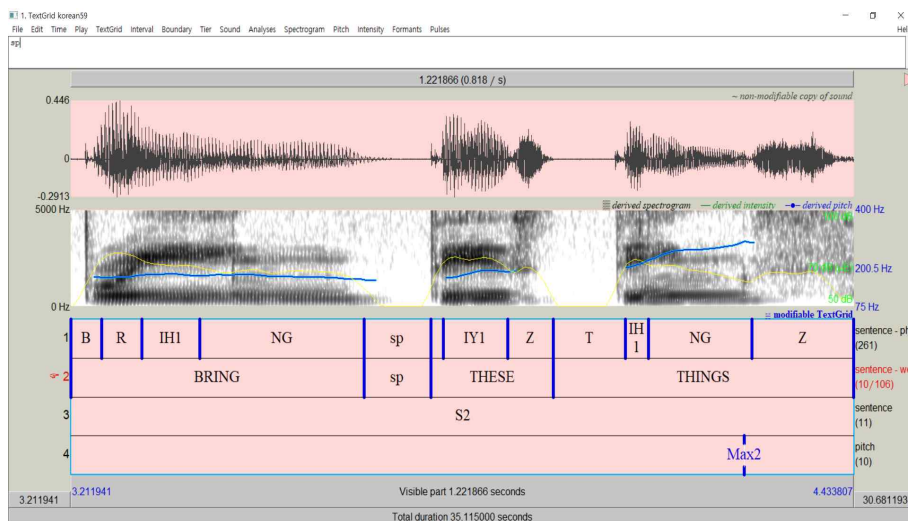


Figure 1. An Example of Speech Sample Analysis Using Praat

3. Data Analysis

1) Independent Variables

In this study, two segmental features and eight suprasegmental features were examined and analyzed as independent variables.

(1) Segmental Errors

The analysis of the speech samples involved the number of segmental errors and syllable structure errors, in accordance with previous studies (Issacs & Trofimovich, 2012; Saito et al., 2015; Trofimovich & Issacs, 2012). Segmental errors encompassed instances of phonemic substitutions, such as pronouncing “snake” as “snack,” while phonetic substitutions (e.g., [t] vs. [r]) were not classified as segmental errors. For syllable structure errors, the examination included errors such as vowel and consonant epenthesis (insertion) and elision (deletion), like adding an epenthetic schwa at the end of “ask” or omitting /z/ at the end of “kids.”

(2) Pitch Range

Earlier studies have explored the impact of alterations in voice pitch on speech comprehensibility or accentedness (Issacs & Trofimovich, 2012; Kang, 2010; Trofimovich & Issacs, 2012; Yang, 2021). The underlying concept posits that a broader pitch range entails

more paratones, which indicate shifts in discourse at the topic level. This is believed to contribute to enhanced comprehensibility (Kang et al., 2010).

In this study, the pitch range was computed by determining the average difference between the highest and lowest F0 values within each sentence, following the approach outlined by Yang (2021). Given that the elicitation paragraph comprised five sentences, the differences in pitch for these five sentences were averaged to yield a single pitch range value for each speech sample. These pitch range measurements were conducted using Praat (version 6.2.17), with a voicing threshold set at 0.45.

(3) Pause

Two types of pause measurements were integrated into the experiment. Firstly, the number of pauses was calculated from the speech samples. Yang (2021) found that comprehensibility had the strongest correlation with the number of pauses. Similarly, Trofimovich and Issacs (2012) identified that pause frequency exhibited a modest yet noteworthy connection with comprehensibility, but not with accentedness. The second pause measurement involved determining the duration of pauses. Previous research highlighted a significant correlation between pause duration and accentedness (Kang, 2010; Trofimovich & Baker, 2006).

In this study, a pause was defined as any interruption in the flow of speech lasting longer than 100 milliseconds within a tonality, as defined by Yang (2021). Any pause with a duration less than 100 milliseconds was regarded as an inherent articulation pause and was excluded from the analysis (Kang, 2010; Trofimovich & Baker, 2006; Yang, 2021). Pauses occurring at tonal boundaries were also excluded from the analysis since they typically sound natural. Given the absence of filled pauses like ‘*hmm*’ or ‘*uuh*’ in any of the speech samples, this research solely considered silent pauses.

(4) Speech Rate

Speech rate was assessed in two distinct manners, following the methodology employed by Yang (2021). Firstly, the mean length of run (MLR) was calculated, representing the average number of syllables uttered between pauses lasting longer than 100 milliseconds. Kang (2010) and Issacs & Trofimovich (2012) observed that MLR was a strong indicator of comprehensibility, while Yang (2021) determined that it had a notable impact on listeners’ assessments of accentedness. Secondly, the articulation rate was computed as the average number of syllables articulated per minute, excluding pause time. In the findings of Kang (2010) and Yang (2021), the articulation rate emerged as a significant predictor for both comprehensibility and accentedness.

(5) Rhythm

According to Abercrombie (1967) and Pike (1945), English exhibits a stressed-timed rhythmic pattern, characterized by stressed syllables occurring at relatively regular time intervals. To maintain this consistent interval between stressed syllables, unstressed syllables are often pronounced quickly, reduced, or occasionally omitted. In order to illustrate the stress-timed nature of English rhythm, prior studies such as Issacs and Trofimovich (2012) and Saito et al. (2015) utilized a measure known as the vowel reduction ratio. This metric calculates the ratio of correctly reduced syllables to the total count of required vowel reduction instances in polysyllabic words and function words.

However, in the present study, a rhythm index called VarcoV was computed. According to Kim (2021), both VarcoV and nPVI-V serve as effective indicators for distinguishing between groups with pronounced foreign accents and those with milder foreign language accents. VarcoV represents the standard deviation of vocalic interval duration expressed as a fraction of mean vocalic duration (White & Mattys, 2007). On the other hand, Low et al. (2000) found that “nPVI-V is compiled by calculating the difference in duration between each pair of successive measurements, taking the absolute value of the difference, and dividing it by the mean duration of the pair. The differences are then summed and divided by the number of differences, and the result is multiplied by 100 (p. 383).” In this study, nPVI-V was not included in the analysis. This decision was made because the regression model for comprehensibility and accentedness achieved a higher R-squared score when VarcoV was included in the model as compared to when nPVI-V was one of the independent variables used to assess rhythm.

(6) Tonic Stress

In Yang (2021), the term ‘sentence stress’ was employed to signify the standard sentence-level stress, excluding stress used for contrastive purposes. Many linguists, such as Bresnan (1971) and Schmerling (1976), have used the term ‘sentence stress’ to refer to this type of sentence-level prominence. However, in this study, the term ‘tonic stress’ was adopted, aligning with British scholars like Halliday (1967), Wells (2006), and Brazil (1997).

Tonic stress measures encompassed ‘space,’ which represents the proportion of prominent words relative to the total number of words (Vanderplank, 1993), as well as the ‘tonic stress appropriacy ratio.’ Prominent words were identified as stressed words that typically exhibited characteristics such as extended duration, higher pitch (lower pitch in the case of phrases with a rising boundary tone), and greater amplitude compared to other words within the same sentence. Both auditory and instrumental analyses using Praat were employed to identify stressed words. When calculating ‘space,’ the precision of tonic stress placement was not taken into consideration. In Yang (2021), it was found that ‘space’

significantly influenced listeners' judgments of accentedness.

Drawing inspiration from Mo (2010) and Yang (2021), the 'tonic stress appropriacy ratio' was calculated by dividing the number of appropriately placed tonic stresses by the expected number of tonic stresses. The expected placement of tonic stress was determined by marking prominence and boundaries based on auditory perception. Due to variations in boundaries within each speech sample, the expected placement of tonic stress was supposed to differ for each individual speaker as in speech samples #1 and #2 in the appendix. The expected tonic stress placement was determined by the researchers based on Celce-Murcia et al. (2010) and Ladefoged and Johnson (2015). As per Yang (2021), 'tonic stress appropriacy' significantly impacted judgments of L2 comprehensibility.

2) Dependent Variables

In the rating phase of the experiment, comprehensibility and accentedness were assessed using a 7-point Likert scale, following the methodology of Derwing & Munro (1997), Munro & Derwing (1999), and Yang (2021). In this study, the raters were instructed to assign a higher score for comprehensibility when the speech delivered by a Korean English learner was easily understood. For accentedness, the raters were directed to assign a lower score if the speech exhibited a strong accent. Consequently, a speech that received a high score in accentedness could be considered to have minimal accent and resemble that of a native speaker's speech.

All raters were given an explanation of the concepts of comprehensibility and accentedness in advance, and were provided with practice in evaluating two to three sample speeches before the actual experiment. The rating experiment was conducted using Praat and conducted online through the Zoom platform. The researcher conducted individual online sessions with each rater for the experiment. Figure 2 displays a screenshot of the online rating experiment.

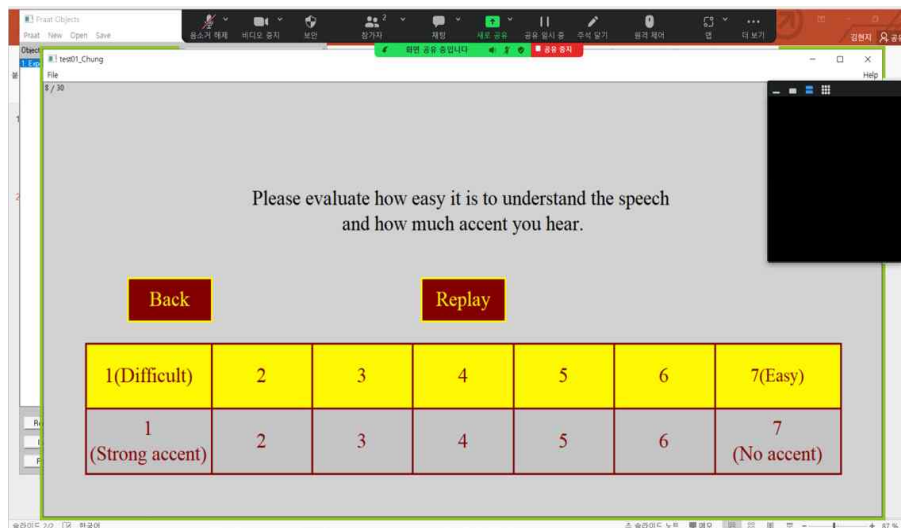


Figure 2. A Screenshot of the Online Rating Experiment

3) Statistical Treatment

The data analysis was performed using the Statistical Package for Social Science (SPSS 22). Initially, two intraclass correlation coefficient analyses were carried out to assess the internal consistency of the ratings provided by both NESs and KETs for comprehensibility and accentedness, respectively. Subsequently, in order to identify the phonetic features that influence the judgments of comprehensibility and accentedness by both NESs and KETs, 4 stepwise multiple regression analyses were conducted. Each analysis includes 10 independent variables and single dependent variable, which comprises the averaged scores for comprehensibility and accentedness, of the NES raters and the KET raters each.

IV. FINDINGS

1. Inter-Rater Reliability

The results of the analysis of the NES and KET raters' scoring are presented in Table 3.

Table 3. Inter-Rater Reliability: Intraclass Correlation Coefficients

	Intraclass Correlation	95% Confidence Interval	
		Lower Bound	Upper Bound
NESs_Comprehensibility	.948	.915	.972
NESs_Accentedness	.961	.936	.979
KETs_Comprehensibility	.911	.854	.952
KETs_Accentedness	.923	.873	.958

The intraclass correlation coefficients among the NES raters were found to be .948 for comprehensibility and .961 for accentedness. This signifies a substantial level of consensus among the NES raters in their evaluations of both comprehensibility and accentedness. Similarly, the KET raters exhibited a strong consensus in their assessments. The intraclass correlation coefficients for the KET raters were .911 for comprehensibility and .923 for accentedness, confirming their consistent judgments regarding both aspects. Given the high reliability of evaluation scores within their respective groups, the scores for each item were averaged to generate a final set of scores.

2. Phonetic Measurements

The descriptive statistics of measurements of 10 phonetic features analyzed from 30 speech samples are shown in Table 4.

Table 4. Descriptive Statistics of Measurements of 10 Phonetic Features

	<i>N</i>	<i>M</i>	<i>SD</i>	Min.	Max.
N of Segmental Errors	30	2.467	2.360	0.00	8.00
N of Syllable Structure Errors	30	4.767	3.803	0.00	13.00
Pitch Range	30	255.266	82.140	145.67	426.61
N of Pauses	30	5.633	4.881	0.00	19.00
Duration of Pauses	30	1.53	1.62	0.00	5.94
Mean Length of Run	30	6.300	2.091	2.64	12.00
Articulation Rate	30	520.253	100.448	335.55	759.49
VarcoV	30	50.635	5.486	38.30	61.37
Space	30	0.287	0.116	0.16	0.78
Tonic Stress Appropriacy Ratio	30	0.341	0.174	0.00	0.65

The number of segmental errors observed ranged from 0 to 8, with an average of 2.467, while the number of syllable structure errors varied from 0 to 13, with a mean of 4.767. In the case of pitch range in this study, there was greater variability among speech samples compared to Yang (2021). The results for mean length of run were consistent with those found in Yang (2021), but this was not the case for articulation rate. The articulation rate demonstrated greater variation among speech samples compared to Yang (2021). In terms of the rhythm metric VarcoV, it exhibited variations across speech samples, ranging from 38.30 to 61.37, with an average of 50.635. These findings closely resembled the results observed in Choe (2019), where the mean VarcoV among Korean English learners was 52.29. Regarding the proportion of prominent words to the total number of words (space), this varied from 16% to 78% across speech samples. However, considering the mean space variable was 29%, it can be inferred that only a few speech samples showed notably high space values. Regarding the tonic stress appropriacy ratio, the average was approximately 34%, ranging from 0% to 65%. This indicates that none of the speaker participants assigned tonic stress perfectly, and some participants used tonic stress entirely inaccurately.

3. Rating Scores

Table 5 describes the descriptive statistics of rating scores by rater participants.

Table 5. Descriptive Statistics of Rating Scores by Rater Participants

	<i>N</i>	<i>M</i>	<i>SD</i>	Min.	Max.
NESs_Comprehensibility	30	4.893	1.084	2.8	6.7
NESs_Accentedness	30	3.833	1.351	1.6	6.2
KETs_Comprehensibility	30	4.957	.903	3.2	6.4
KETs_Accentedness	30	4.350	1.196	1.9	6.1

Both NESs and KETs assigned higher ratings for comprehensibility compared to accentedness, similar to the findings in Yang (2021). Figure 3 and Figure 4 depict scatterplots illustrating averaged comprehensibility and accentedness scores assigned by NESs and KETs for thirty speech samples. It is evident from both graphs in figure 3 and figure 4 that for the majority of the speech samples, the comprehensibility score exceeded the accentedness score, and the score disparity between NESs' comprehensibility and accentedness is greater than that observed between KETs' comprehensibility and accentedness.

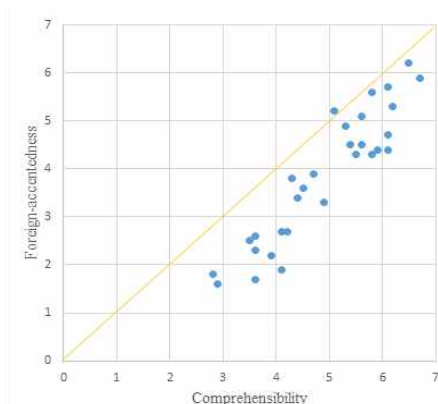


Figure 3. Scores Rated by NESs

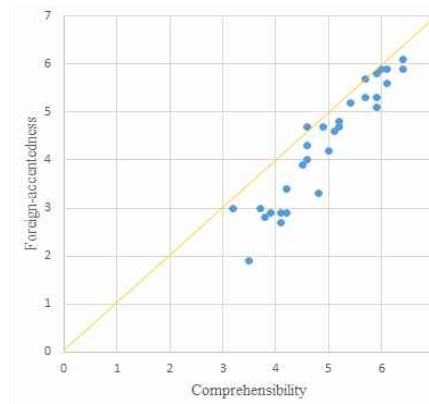


Figure 4. Scores Rated by KETs

In the comparison of the mean comprehensibility and accentedness rating scores between two groups, no statistically significant distinction emerged between the judgments by NESs and KETs. Nevertheless, when simply comparing the average values, it became apparent that KETs allocated higher scores than NESs in both types of evaluations. Specifically, it was noted that KETs displayed a more lenient stance regarding accentedness in contrast to NESs (Average: 3.83 (NESs); 4.35 (KETs)).

4. Phonetic Features Influencing NESs' Comprehensibility and Accentedness Judgments

To determine the phonetic features influencing the evaluations of comprehensibility and accentedness by NES raters, the researcher performed correlation analyses involving each respective dependent variable and the 10 independent variables. The findings from these analyses can be found in Table 6.

Table 6. Correlations of Phonetic Features With NESs' Comprehensibility and Accentedness Judgments

Comprehensibility	<i>r</i>	Accentedness	<i>r</i>
N of Syllable Structure Errors	-.781**	N of Syllable Structure Errors	-.731**
Articulation Rate	.593**	N of Segmental Errors	-.589**
VarcoV	.551**	Articulation Rate	.583**
N of Segmental Errors	-.509**	VarcoV	.535**
Duration of Pauses	-.459**	Space	-.473**
Space	-.463**	N of Pauses	-.398*
N of Pauses	-.454**	Tonic Stress Appropriacy Ratio	-.398*
Tonic Stress Appropriacy Ratio	-.412*	Duration of Pauses	-.387*
Mean Length of Run	.280	Mean Length of Run	.243
Pitch Range	-.004	Pitch Range	-.046

Note. *r* indicates Pearson correlations; $p < .01^{**}$, $p < .05^*$

Based on the outcomes of the correlation analyses, the most prominent factor associated with both rating constructs was the presence of syllable structure errors (comprehensibility: $r = -.781$, $p < .01$; accentedness: $r = -.731$, $p < .01$). Neither MLR nor pitch range exhibited any significant correlations with the judgments provided by NES raters. The remaining 7 variables had moderate correlations with both types of assessments conducted by NESs. An interesting observation is that the number of segmental errors had a more pronounced impact on NESs' judgments of accentedness compared to their assessments of comprehensibility. Conversely, the tonic stress appropriacy ratio and duration of pauses displayed stronger correlations with comprehensibility scores than with accentedness, aligning with the findings in Yang (2021).

Table 7 and Table 8 provide the conclusive overview of the stepwise multiple regression models, illustrating the impact of phonetic variables on the comprehensibility and accentedness judgments made by NESs ($VIF < 1.36$, $p < .05$).

Table 7. Final Regression Model of NESs_Comprehensibility

Factors	B	β	<i>t</i> -value	sig. <i>p</i>	Step Entered	R^2 Change	VIF
N of Syllable Structure Errors	-.196	-.677	-6.334	.000	1	.610	1.099
VarcoV	.070	.347	3.249	.003	2	.110	1.099

Final model $R^2 = .720$, $F(2, 28) = 34.653$, $p < .000$, Adjusted $R^2 = .699$, Durbin Watson = 1.232

Table 8. Final Regression Model of NESs_Accentedness

Factors	B	β	<i>t</i> -value	sig. <i>p</i>	Step Entered	R^2 Change	VIF
N of Syllable Structure Errors	-.179	-.494	-4.110	.000	1	.534	1.351
N of Segmental Errors	-.208	-.357	-3.228	.003	2	.137	1.143
Articulation Rate	.004	.257	2.177	.039	3	.051	1.300
Final model $R^2 = .721$, $F(3, 27) = 22.441$, $p < .000$, Adjusted $R^2 = .689$, Durbin Watson = 1.969							

In the case of comprehensibility judgments made by NES raters, the optimal model featured two variables: the number of syllable structure errors and VarcoV ($R^2 = .720$, $F(2, 28) = 34.653$, $p < .000$). As for accentedness, the most effective model consisted of three variables: the number of syllable structure errors, the number of segmental errors, and articulation rate ($F(3, 27) = 22.441$, $p < .000$). All variables associated with segmental errors exhibited strong negative correlations with NES judgments of accentedness (number of syllable structure errors: $\beta = -.494$; number of segmental errors: $\beta = -.357$). This suggests that speech samples featuring a higher incidence of errors in segmental sounds were perceived as more accented by NESs. Additionally, articulation rate, as one of the predictors, was also found to influence accentedness ($\beta = -.257$). In essence, faster speech from Korean English learners was perceived as more native-like by NESs.

5. Phonetic Features Influencing KETs' Comprehensibility and Accentedness Judgments

Correlation analyses to examine the relationship between KETs' ratings for comprehensibility and accentedness and the 10 independent variables were also performed. The findings from these analyses are detailed in Table 9.

Table 9. Correlations of Phonetic Features With KETs' Comprehensibility and Accentedness Judgments

Comprehensibility		Accentedness	
	<i>r</i>		<i>r</i>
Articulation Rate	.717**	Articulation Rate	.776**
N of Syllable Structure Errors	-.689**	N of Syllable Structure Errors	-.756**
Duration of Pauses	-.581**	VarcoV	.613**
VarcoV	.577**	Space	-.562**
N of Pauses	-.561**	N of Segmental Errors	-.527**
Tonic Stress Appropriacy Ratio	-.520**	Duration of Pauses	-.519**
Space	-.508**	N of Pauses	-.517**
N of Segmental Errors	-.474**	Tonic Stress Appropriacy Ratio	-.472**
Mean Length of Run	.353*	Mean Length of Run	.371*
Pitch Range	.072	Pitch Range	.038
Note. <i>r</i> indicates Pearson correlations; $p < .01$ **, $p < .05$ *			

In terms of comprehensibility scores by KETs, articulation rate displayed the strongest

correlation ($r = .717, p < .01$). In contrast, pitch range did not demonstrate a significant association with comprehensibility. The rest of the variables revealed moderate correlations with comprehensibility judgments. Similarly, akin to their assessment of comprehensibility, KETs' ratings for accentedness showed the strongest correlation with articulation rate ($r = .776, p < .01$). Notably, a strong correlation was also identified between accentedness judgments and the number of syllable structure errors ($r = -.756, p < .01$). VarcoV, space, and the number of segmental errors were observed as moderate contributors that had a more pronounced impact on accentedness assessments compared to comprehensibility. Duration of pauses, the number of pauses, and tonic stress appropriacy ratio also demonstrated moderate correlations with accentedness evaluations, albeit with a lesser degree of influence compared to comprehensibility. MLR exhibited a significant positive correlation with accentedness. Similar to comprehensibility, there was no discernible relationship between pitch range and accentedness.

The model summaries for stepwise multiple regression analyses involving phonetic variables in relation to KETs' judgments of comprehensibility and accentedness can be found in Table 10 and Table 11. In both models, there were no signs of collinearity, with VIF values consistently below 1.50. A significance threshold of $p < .05$ was applied to ascertain the significance of the tests.

Table 10. Final Regression Model of KETs_Comprehensibility

Factors	B	β	<i>t</i> -value	sig. <i>p</i>	Step Entered	R^2 Change	VIF
Articulation Rate	.004	.397	3.183	.004	1	.515	1.499
N of Syllable Structure Errors	-.103	-.426	-3.682	.001	2	.162	1.289
VarcoV	.044	.264	2.276	.031	3	.054	1.293
Final model $R^2 = .730, F(3, 27) = 23.459, p < .000$, Adjusted $R^2 = .699$, Durbin Watson = 1.830							

Table 11. Final Regression Model of KETs_Accentedness

Factors	B	β	<i>t</i> -value	sig. <i>p</i>	Step Entered	R^2 Change	VIF
Articulation Rate	.005	.429	4.748	.000	1	.602	1.499
N of Syllable Structure Errors	-.152	-.476	-5.677	.000	2	.200	1.289
Varco V	.060	.270	3.216	.003	3	.056	1.293
Final model $R^2 = .858, F(3, 27) = 52.493, p < .000$, Adjusted $R^2 = .842$, Durbin Watson = 2.015							

Among the 10 variables considered, it was evident that articulation rate, the number of syllable structure errors, and VarcoV had a significant impact on the comprehensibility judgments made by KETs (articulation rate: $\beta = .397$; the number of syllable structure errors: $\beta = -.426$; VarcoV: $\beta = .264$). Similar to NES raters, KET raters were influenced by

both the number of syllable structure errors and VarcoV when determining comprehensibility. What is particularly noteworthy is that articulation rate exerted a more substantial influence than these two variables. This outcome suggests that KETs found it easier to comprehend speech samples characterized by a rapid speaking pace, fewer errors in syllable structure, and appropriate rhythm. This regression model accounted for approximately 73% of the variability in comprehensibility scores provided by KETs ($R^2 = .730$) and was statistically significant ($F(3, 27) = 23.459, p < .000$).

The predictors for accentedness judgments by KETs mirrored those for comprehensibility, albeit with slightly different coefficients (articulation rate: $\beta = .429$; the number of syllable structure errors: $\beta = -.476$; VarcoV: $\beta = .270$). These three variables were able to predict accentedness judgments by KETs more accurately than comprehensibility. The final model explained approximately 85% of the variability in accentedness scores ($R^2 = .858$).

V. DISCUSSION

1. Comprehensibility and Accentedness Judgments by NESs

To reveal the features influencing NESs judgments of comprehensibility and accentedness, both correlation analyses and stepwise multiple regression analyses were conducted. The outcomes of the correlation analyses highlighted that the number of syllable structure errors displayed the strongest correlation with both NESs' evaluations of comprehensibility and accentedness. This suggests that for L2 learners, correctly articulating English words without omitting or inserting any segments can contribute significantly to rendering their speech more comprehensible and native-like.

It is worth highlighting that the variables related to tonic stress displayed an unexpected negative correlation with NESs' judgments of comprehensibility and accentedness, contrary to the findings in Yang (2021). In this study, the researcher followed Yang's (2021) methodology, which involved individually defining tonal characteristics for each speaker and assessing prominence within those tonal boundaries. However, the sentences in speech samples were notably longer and structurally more complex compared to those in Yang (2021). As the speaker participants inserted more pauses into their speech, the length of tonality decreased and this allowed for a more precise assignment of tonic stress. It is generally known that an excessive number of pauses within tonality can disrupt the comprehension of L2 speech (Hahn & Dickerson, 1999). It seems that when assessing L2 speech, the frequency of pauses takes precedence over the appropriateness of tonic stress. Consequently, the tonic stress appropriacy ratio exhibited a negative correlation with the frequency of pauses.

In the outcomes of the stepwise multiple regression analyses, the number of syllable structure errors and VarcoV emerged as significant factors influencing NESs' judgments of

comprehensibility. When it comes to accentedness ratings, the number of syllable structure errors, the number of segmental errors, and articulation rate played a significant role. It is crucial to highlight that in this study, syllable structure errors were identified as the primary factor impacting NESs' judgments of both L2 comprehensibility and accentedness, unlike other related studies (Isaacs & Trofimovich, 2012; Saito et al., 2017). While the accurate pronunciation of segmentals significantly affected accentedness, it did not carry as much weight in comprehensibility judgments among native speakers. Considering that the ultimate goal of English teaching is to enhance learners' communication skills, L2 pronunciation instruction and learning should prioritize reducing segmental errors related to insertion or omission, rather than solely focusing on correct segmental sounds. It is noteworthy that the extent to which L2 learners maintain English rhythm can have a considerable influence on NESs' perceptions of L2 speech. This finding of this study aligns with previous research (e.g., Kim, 2021; Polyanskya et al., 2017) which demonstrate that rhythmic L2 speech is more likely to be perceived as native-like by NESs. In this study, it was additionally revealed that rhythmic L2 speech not only influences the judgments of L2 accentedness by NESs but also has a significant impact on comprehensibility judgments. This aligns with the findings of Kim and Chung (2016), who similarly identified a more stress-based rhythm in the speech of highly proficient speakers, resulting in elevated rhythmic indices including VarcoV. These insights suggest a clear directive for English instructors to prioritize the enhancement of speech rhythm in their L2 learners when imparting English instruction.

2. Comprehensibility and Accentedness Judgments by KETs: How are They Different From Those by NESs?

First of all, there were no statistically significant differences between the comprehensibility judgments provided by NESs and KETs, nor between the accentedness ratings given by the two groups. This implies that KETs' judgments on L2 speech closely paralleled the judgments made by NESs. It can be concluded that KETs, who have received academic training and sufficient instruction, are well-qualified as English assessors, comparable to NESs, as in other relevant studies (e.g., Kang & Ahn, 2012). Nevertheless, it is also worth noting that the average ratings provided by KETs were consistently higher than those given by NESs for both evaluation criteria. This disparity was particularly noticeable in terms of accentedness judgments. These differences may be attributed to the shared linguistic background between KETs and the speaker participants, as well as their familiarity with English spoken with a Korean accent. Similarly, in Bent and Bradlow (2003), non-native listeners showed a tendency to rate speakers from their own native language community as being as intelligible as native speakers.

Second, based on the outcomes of the correlation analyses, all phonetic variables except

pitch range exhibited correlations with KETs' judgments of comprehensibility and accentedness. Among these 9 phonetic factors, articulation rate displayed the most robust correlation with their judgments of comprehensibility. This is in contrast with the findings that identified the number of syllable structure errors as the most influential factor influencing NESs' judgments. A similar pattern emerges from the findings of stepwise multiple regression analyses conducted between KETs' assessments of comprehensibility and accentedness and the 10 phonetic features. These results indicated that articulation rate, the number of syllable structure errors, and VarcoV were the most effective predictors for both kinds of judgments. This suggests that, KETs place the highest importance on the speech speed when evaluating L2 speech, while speech rate does not significantly affect NESs' judgments of comprehensibility. The similar result was shown in Park (2022). In Park (2022), it was revealed that the nonnative listeners paid relatively more attention to speech rate than native listeners. The significance of speaking rate as the most potent predictor in KETs' assessments of L2 speech could potentially impact their overall approach to pronunciation teaching and evaluation. However, since articulation rate does not significantly affect NESs' judgments of comprehensibility, KETs do not necessarily need to excessively emphasize fast speech unless the goal for their L2 learners is to achieve native-like pronunciation. Instead, it is more desirable for L2 learners to practice speaking English accurately, without omitting or adding segmental sounds in words, while also focusing on achieving a rhythmic speech pattern.

VI. CONCLUSION

The findings of the study offer several valuable pedagogical insights for English language classrooms, particularly in the context of Korea. Firstly, the results emphasize again that the ultimate goal of L2 pronunciation education should prioritize enhancing the comprehensibility of L2 speech rather than native-like pronunciation. To make this goal achievable, it is essential for English instructors to clearly distinguish comprehensibility from accentedness, recognizing that they are fundamentally distinct constructs. Secondly, the outcomes of the study offer valuable insights into where pronunciation education should place its emphasis specifically to help L2 students effectively achieve their linguistic objectives. While it is widely acknowledged that the primary aim of L2 pronunciation instruction is to enhance learners' capacity to communicate effectively, educators have often lacked specific guidance on which pronunciation elements should receive priority (Derwing & Munro, 2009). The results of the study suggests that it is necessary to provide L2 learners with ample guidance and feedback on pronunciation accuracy, particularly addressing issues related to the unnecessary insertion of segments. The analysis of the evaluation experiment revealed that Korean English learners inserting unnecessary vowel after a consonant, as in pronouncing 'three red bags' as /θuri rɛdə bægzə/, hindered

comprehensibility for NES raters. This is a pronunciation error commonly made by Korean English learners due to the influence of their native language, which does not allow consonant clusters. Therefore, addressing this directly in pronunciation education can assist learners in achieving their learning goals.

The present study also provides significant insights into the English speaking assessments carried out by KETs. The absence of a notable difference in the rating scores assigned by NESs and KETs, coupled with the strong internal consistency observed among KETs' evaluations, reinforces the argument that KETs can serve as assessors with a level of reliability comparable to NESs. Therefore, assessments of speaking proficiency conducted by KETs should not be undervalued solely because English is not their native language. However, it is also worth acknowledging that the majority of KETs are also EFL learners who are naturally acquainted with a Korean accent. Consequently, there is a possibility that they may exhibit leniency or make assessments that are not entirely accurate. For instance, this study revealed that KETs are influenced by speech rate when evaluating comprehensibility. It implies that KETs may prioritize speech rate in their teaching and assessment practices. Nonetheless, since speech rate does not significantly affect NESs' judgments of comprehensibility, increasing the speed of L2 speech may not be a desirable objective in L2 speaking education. Therefore, it is advisable for KETs to undergo specialized training to enhance the professionalism and reliability of their speaking assessments.

This study is subject to several limitations that should be considered. Firstly, the adoption of scripted speech materials in this study was deliberate, aiming to regulate variations in vocabulary and grammar. However, a future study adopting natural speech may yield different results. For example, Kim and Chung (2011) discovered that Korean L2 speakers exhibit more English rhythmic characteristics when reading scripted speech materials compared to when generating spontaneous speech. Future research employing unscripted speech samples from spontaneous conversations could offer broader insights. Additionally, the phonetic features analyzed in the speech samples were limited to specific phonetic features. Specifically, the analysis only considered two variables for segmental factors: the number of segmental errors and the number of syllable structure errors. Future studies should aim to comprehensively explore how linguistic features, encompassing not only phonological aspects but also various linguistic dimensions such as grammar, vocabulary, and discourse, influence the L2 speech judgments by both NESs and KETs. Also, in this study, the ROL of the ten NES raters varied widely from 0 to 10 years. This broad range could have influenced the assessments made by NES raters. Therefore, in future studies, it would be beneficial to control for a more accurate judgment of comprehensibility and accentedness by NES raters. Lastly, this research solely compared the assessments of comprehensibility and accentedness made by KETs to those made by NESs to analyze the assessments of KETs. Given the increasing prevalence of English as a lingua franca (ELF) and the recognition of English as an international language (EIL), it is essential

to acknowledge the diversity of English users worldwide. Consequently, it would be valuable to conduct a comprehensive comparative analysis of KETs' speaking assessments with evaluations conducted by individuals from diverse linguistic backgrounds.

REFERENCES

- Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh University Press.
- Bent, T., & Bradlow, A. (2003). The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America*, 114, 1600-1610.
- Boersma, P., & Weenink, D. (2022). Praat: Doing phonetics by computer (version 6. 2. 17) [Computer program]. Retrieved from <https://www.fon.hum.uva.nl/praat/>
- Brazil, D. (1997). *The communicative value of intonation in English*. Cambridge University Press.
- Bresnan, J. W. (1971). Sentence stress and syntactic transformations. *Language*, 47(2), 257-281.
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching pronunciation: A course book and reference guide* (2nd edition). Cambridge University Press.
- Choe, W. K. (2019). The realization of English rhythm by Busan Korean speakers. *Phonetics and Speech Sciences*, 11(4), 81-87.
- Crowther, D., Trofimovich, P., & Issacs, T. (2016). Linguistic dimensions of second language accent and comprehensibility: Nonnative listeners' perspectives. *Journal of Second Language Pronunciation*, 2, 160-182.
- Crowther, D. (2020). *Global dimensions of second language pronunciation: A meta-analysis of accentedness, comprehensibility, and intelligibility research*. Paper presented at the 2020 Conference of the American Association for Applied Linguistics (AAAL), Denver, Col., United States.
- Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition*, 19, 1-16.
- Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching*, 42(4), 476-490.
- Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research* (Vol. 42). John Benjamins Publishing Company.
- Derwing, T. M., Munro, M. J., & Weibe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48(3), 393-410.
- Flege, J. E., M. J. Munro, & I. R. MacKay. (1995). Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America*, 97(5), 3125-3134.

- Foote, J. A., & Trofimovich, P. (2018). Is it because of my language background? A study of language background influence on comprehensibility judgments. *Canadian Modern Language Review*, 74(2), 253-278.
- George Mason University (2023, October 22). *The speech accent archive*. <https://accent.gmu.edu>
- Gordon-Salant, S., Yeni-Komshian, G. H., Bieber, R. E., Jara Ureta, D. A., Freund, M. S., & Fitzgibbons, P. J. (2019). Effects of listener age and native language experience on recognition of accented and unaccented English words. *Journal of Speech, Language, and Hearing Research*, 62(4S), 1131-1143.
- Hahn, L. D., & Dickerson, W. B. (1999). *Speech craft: Discourse pronunciation for advanced learners*. The University of Michigan Press.
- Halliday, M. A., K. (1967). Notes on transitivity and theme in English: Part 2. *Journal of Linguistic*, 3(2), 199-244.
- Issacs, T., & Trofimovich, P. (2012). Deconstructing comprehensibility: Identifying the linguistic influences on listeners' L2 comprehensibility ratings. *Studies in Second Language Acquisition*, 34, 475-505.
- Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System*, 38(2), 301-315
- Kang, O., Rubin, D., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *The Modern Language Journal*, 94(4), 554-566.
- Kang, S-h., & Ahn, H. (2012). A comparative study on criteria and tasks in Korean English speaking assessment by native and non-native raters. *Language Research*, 48(2), 241-262.
- Kennedy, S., & Trofimovich, P. (2010). Language awareness and second language pronunciation: A classroom study. *Language Awareness*, 19, 171-185.
- Kim, K., & Chung, H. (2011). A study on the rhythm of Korean learners' of English in read aloud and spontaneous speech. *Studies in Foreign Language Education*, 25(1), 25-56.
- Kim, H., & Chung, H. (2016). An analysis of the rhythm and reduced vowels of Korean adult learners of English. *Studies in Foreign Language Education*, 30(3), 117-140.
- Kim, S. (2021). Foreign accentedness rating and rhythm in L2 English. *The Mirae Journal of English Language and Literature*, 26(4), 101-120.
- Ladefoged, P., & Johnson, K. (2015). *A course in phonetics* (7th edition). Cengage.
- Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, 39, 369-377.
- Levis, J. M. (2006). Pronunciation and the assessment of spoken language. In R. Hughes (Ed.), *Spoken English, TESOL and applied linguistics: Challenges for theory and practice*, (pp. 245-270). Palgrave Macmillan.
- Levis, J. M. (2018). *Intelligibility, oral communication, ad the teaching of pronunciation*.

Cambridge University Press.

- Low, E. L., Grabe, E., & Nolan, F. (2000). Quantitative characterizations of speech rhythm: Syllable-timing in Singapore English. *Language and Speech, 43*(4), 377-401.
- Mo, Y. (2010). *Prosody production and perception with conversational speech* (Order No. 3455774) [Doctoral dissertation, University of Illinois at Urbana-Champaign]. ProQuest Dissertations and Theses Global.
- Munro, M. J., & Derwing, T. M. (1995a). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45*(1), 73-97.
- Munro, M. J., & Derwing, T. M. (1995b). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech, 38*(3), 289-306.
- Munro, M. J., & Derwing, T. M. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech the role of speaking rate. *Studies in Second Language Acquisition, 23*(4), 451-468.
- Park, M. S. (2022). Comprehensibility of Korean college students' English utterances and its linguistic factors assessed by native and nonnative listeners. *The Journal of Learner-Centered Curriculum and Instruction, 22*(8), 223-235.
- Pike, K. L. (1945). *The intonation of American English*. University of Michigan Press.
- Polyanskaya, L., Ordin, M., & Busa, M. G. (2017). Relative salience of speech rhythm and speech rate on perceived foreign accent in a second language. *Language and Speech, 60*(3), 333-355.
- Saito, K., Macmillan, K., Tran, M., Suzukida, Y., Sun, H., Magne, V., & Murakami, A. (2020). Developing, analyzing and sharing multivariate datasets: Individual differences in L2 learning revisited. *Annual Review of Applied Linguistics, 40*, 9-25.
- Saito, K., Trofimovich, P., & Issacs, T. (2015). Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholinguistic, 37*(2), 217-240.
- Saito, K., Trofimovich, P., & Issacs, T. (2017). Using listener judgments to investigate linguistic influences on L2 comprehensibility and accentedness: A validation and generalization study. *Applied Linguistics, 38*(4), 439-462.
- Saito, K., Trofimovich, P., Issacs, T., & Webb, S. (2016). Re-examining phonological and lexical correlates of second language comprehensibility: The role of rater experience. In T. Issacs & P. Trofimovich (Eds.), *Interfaces in second language pronunciation assessment: Interdisciplinary perspectives*, (pp. 141-156). Multilingual Matters.
- Schmerling, S. F. (1976). *Aspects of English sentence stress*. University of Texas Press.
- Trofimovich, P., & Baker, W. (2006). Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition, 28*(1), 1-30.

- Trofimovich, P., & Issacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism: Language and Cognition*, 15(4), 905-916
- Vanderplank, R. (1993). 'Pacing' and 'spacing' as predictors of difficulty in speaking and understanding English. *ELF Journal*, 4(2), 117-125
- Wells, J. C. (2006). *English intonation: An introduction*. Cambridge University Press.
- White, L. & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, 35, 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.

APPENDIX

Examples of Tonic Stress Analysis

(The passage was excerpted from George Mason University (2023). Boundaries of tonality are separated by slash. The expected tonic stress placement was determined by the researchers based on Celce-Murcia et al. (2010) and Ladefoged and Johnson (2015). One underline indicates that tonic stress must be placed exactly in that position, not elsewhere. Two underlines indicate that within the same tonality, only one of the two positions must receive tonic stress. If both receive tonic stress, it was marked as incorrect. Actual places of tonic stress are bolded.)

Speech sample #1

Please call **Stella**. / Ask her to bring these **things** with her from the store: / Six spoons of fresh snow peas, / five thick slabs of blue **cheese**, / and maybe a **snack** for her brother **Bob**. / We also need a small **plastic** snake and a big **toy** frog for the kids. / She can scoop these **things** into three **red** bags, / and we will **go** meet her **Wednesday** at the train station.

Speech sample #2

Please **call** Stella. / Ask her to bring these things with her **from** the store: / Six spoons of fresh snow peas, / five / **thick** slabs of / blue **cheese**, / and maybe a **snack** for her brother Bob. / We **also** need a small plastic snake / **and** a / **big** / toy **frog** for the kids. / She can scoop these **things** / into **three** red bags, / and we will go meet **her** / **Wednesday** / at the train station.

〈Korean Abstract〉

김현지, 정현성. (2023). 원어민과 한국인 EFL 교사가 이해도 및 외국어 말투를 평가하는 데 영향을 미치는 음성적 특성. *외국어교육연구*, 37(4), 83-107.

본 연구의 목적은 원어민과(NES)와 한국인 EFL 교사(KET)가 이해도와 외국어 말투를 평가하는 데 영향을 미치는 음성적 특성을 조사하는 것이다. Yang (2021)의 연구 디자인을 참고하여, 10명의 원어민과 10명의 한국인 EFL 교사가 각각 30개의 한국인 영어 발화에 대한 이해도와 외국어말투 평가를 진행하였으며, 연구자는 이들의 판단에 영향을 미치는 발화의 언어적 특성을 분석하였다. 다중 회귀 분석 결과, 원어민의 이해도 판단에 영향을 미치는 주요 요인은 음절 구조 오류 개수와 VarcoV였고, 외국어말투 판단에 영향을 미치는 주요 원인으로는 음절 구조 오류 개수, 분절음 오류 개수 및 발음 속도였음이 드러났다. 한편 한국인 EFL 교사의 경우에는 발음 속도, 음절 구조 오류 개수 및 VarcoV가 이해도 및 외국어 말투를 판단하는 데 가장 큰 영향을 미치는 것으로 밝혀졌다. 이러한 결과를 기반으로, 본 연구는 한국 영어교육 환경에서의 바람직한 영어 발음 교육 및 평가를 위한 지침을 제안한다.

Key words: L2 pronunciation, comprehensibility, accentedness / L2 발음 교육 및 평가, 이해도, 외국어 말투

Examples in: English

Applicable Languages: English

Applicable Levels: University

Kim, Hyunji

Master's Student

Dept. of English Education, Korea National University of Education

250 Taeseongtabyeon-ro, Gangnae-myeon, Heungdeok-gu, Cheongju-si, Chungbuk, Korea 28173

TEL: (043)-230-3554

E-MAIL: sally5242@naver.com

Chung, Hyunsong

Professor

Dept. of English Education, Korea National University of Education

250 Taeseongtabyeon-ro, Gangnae-myeon, Heungdeok-gu, Cheongju-si, Chungbuk, Korea 28173

TEL: (043)-230-3554

E-MAIL: hchung@knue.ac.kr

Received in October 09, 2023

Revised version received in November 04, 2023

Revised version accepted in November 06, 2023