



Understanding L2 Speech Production: Implications for Teaching Speaking in EFL Classroom

On-Soon Lee

Sunchon National University, Korea

Jeongyeon Park

Dong-A University, Korea

As communicative competence has become a primary goal of English education in many Asian contexts, EFL curricula increasingly focus on students' speaking performance. Considering the demands of cultivating competent L2 speakers, this study investigates factors associated with L2 speaking performance in a Korean EFL college classroom setting. Fifty-one students enrolled in a basic English speaking course participated. They completed five tasks measuring two linguistic variables (i.e., learners' perception of segmental and suprasegmental features), a cognitive variable (i.e., short-term memory), and two language ability variables (i.e., listening comprehension ability and vocabulary size). They also completed a production task (i.e., picture narration). The results indicate that sensitivity to suprasegmental features (e.g., pause, stress, and intonation) and listening ability are associated with the quality of the learners' spontaneous speech production, while sensitivity to segmental information (e.g., minimal pairs), short-term memory, and vocabulary size are not. These findings suggest the importance of explicit instruction in suprasegmental features to improve L2 production as well as perception ability, and that integrating listening and speaking instruction in L2 curricula may be the most effective means of improving learners' speech.

Keywords: perceptual sensitivity, prosody, L2 speech, listening ability, EFL

Introduction

Interest in successful communication has increased in second language (L2) instructional and testing contexts in Asian settings in recent years, and findings from copious research demonstrate that pronunciation is crucial for L2 communicative success (see Liontas, 2018 for a review and discussion). A question for researchers and practitioners interested in English as a foreign language (EFL) oral proficiency is what an appropriate goal is for pronunciation in EFL contexts. Despite the emphasis on nativelike pronunciation in standardized tests (e.g., Test of English as a Foreign Language [TOEFL]; International English Language Testing System [IELTS]), it seems difficult to acquire nativelike English pronunciation in Asian educational settings. However, many researchers and educators have agreed that comprehensibility, a major construct in L2 pronunciation and speaking research, is congruent with the instructional goal of helping EFL learners achieve communicative success in real-world interactions (Derwing & Munro, 2009; Jenkins, 2000; Morley, 1994; Munro & Derwing, 2011). "Comprehensibility" is defined as listeners' perception that they understand the speech they hear, and is typically measured by the listeners' ratings of the ease of understanding (e.g., Munro & Derwing, 1999). If comprehensibility is

to be a realistic goal of teaching L2 pronunciation and speaking in EFL instructional settings, more research on how to apply or implement this goal in an actual classroom setting is necessary. One of the practical issues concerns whether nonnative-speaker teachers and practitioners should provide explicit instruction in L2 speaking and if so, what factors they must consider in their classroom implementation, as well as in the development of teaching methods and task design. To begin to address these needs, this study examines factors associated with L2 spontaneous speech production.

The assessment of L2 speech must also consider the multifaceted nature of comprehensibility (Kang, 2010; Saito et al., 2015, 2016). L2 phonetics research has analyzed a wide variety of linguistic dimensions of L2 speech. For example, recent studies have investigated the extent to which specific variables are related to L2 comprehensibility (Kang, 2010; Kang et al., 2010; Saito et al., 2015, 2016). Kang's (2010) study showed that L1-English speakers' comprehensibility judgments were related to learners' speech speed, but not to the suprasegmental factors of pausing or pitch.¹ Extending previous findings, Saito et al. (2015) found that raters' comprehensibility judgment of the speech of proficient English learners, but not of beginners, was related to the accuracy of their production of segments (i.e., consonants and vowels) and the segmental feature of prosody. Overall, the prior research findings on the effects of the production of segments and suprasegmentals on the comprehensibility of L2 speech are mixed. Furthermore, investigations of specific L1 populations such as EFL Korean college learners have demonstrated that proficiency in reading and writing does not necessarily entail the ability to perceive or produce sounds (e.g., in the Korean context, Kim & Nam, 2002; Yang, 2013). It is not uncommon for even fairly advanced students to mispronounce English words or fail to understand them when they hear them. This may be in part because they are unable to associate words with sounds accurately, despite being able to produce them in written form. Such inaccurate perception and production of speech including stress, pause, and intonation features can be expected to impact effective communication. Therefore, learning appropriate English pronunciation, including the use of both segments and suprasegmentals, is an important part of improving speaking and listening abilities.

One way to approach the question of how learners can practice suprasegmentals and segments in both speaking and listening is based on the research assumption that perception leads to production in L2 learning (e.g., Flege, 1995). Empirical research has demonstrated that better perception of language sounds including stresses and pauses enhances learners' L2 oral performance (e.g., Aliaga-García & Mora, 2009; Morley, 1991). It is assumed that practicing aurally perceiving suprasegmentals and segmentals as a classroom activity leads to more comprehensible speech production. Despite great interest in the relationship between L2 perception of sounds and learners' oral ability, it is still unclear how perceptual sensitivity to segmentals and suprasegmentals affects EFL learners' oral performance. The present study aims to fill this gap by examining how EFL learners' perception of segmental and suprasegmental information is linked to their oral performance.

Furthermore, based on previous studies' reports that general listening skills, cognitive ability (e.g., short term memory), and vocabulary knowledge have associations with learners' speech production (e.g., Astorga-Cabezas, 2015; Iwashita et al., 2008; Prebianca et al., 2014), the present study attempts to explore the extent to which these three variables affect EFL students' speech production. First, of the four language skills, listening is more closely related to speaking than to reading and writing (e.g., Astorga-Cabezas, 2015). Although speech is often linked with listening, there is no clear evidence of the direct impact of listening on L2 speech (Leather, 1999; McCandless & Winitz, 1986). Second, short-term memory is involved in the cognitive processes underlying the production of spontaneous speech, including accessing and retrieving items from the lexicon, as well as processes of combining phonetic and prosodic information (e.g., Levelt, 1989). Therefore, the phonological memory, as a key component of short-term memory capacity, is involved in all the processes of L2 spontaneous speech, so higher short-term memory capacity should facilitate more fluent and comprehensible speech. Third, the study measures vocabulary size via a receptive vocabulary knowledge test as an indirect indicator for general

¹ "Suprasegmental," as used in phonetics and phonology, is defined as "a vocal effect which extends over more than one sound segment in an utterance, such as a pitch, stress or juncture pattern" (Crystal, 2003, p. 446).

language proficiency (Meara, 1996). Although vocabulary knowledge is less frequently employed in relation to speaking skill, some studies support a relationship between speaking and vocabulary knowledge (e.g., De Jong et al., 2012; Koizumi & In'nami, 2013; Roberts, 2005), while others do not (Uchihara & Saito, 2019). This study seeks to clarify whether vocabulary size measures may be predictive of speaking skill by examining whether vocabulary knowledge affects L2 speech production.

The current study investigates the relationships among two linguistic variables (i.e., learners' perception of segmental and suprasegmental features), a cognitive variable (i.e., short-term memory), and two language ability variables (i.e., listening comprehension ability and vocabulary size), and how these variables predict performance in Korean EFL learners' spontaneous speech production in a quasi-experimental setting. The study hopes to provide information that will be useful in the evaluation and development of methods and materials for the teaching of speaking in EFL settings.

Background

Linguistic Variables: Segmentals and Suprasegmentals

Recent studies on L2 pronunciation for comprehensible speech have explored the linguistic dimensions involved in L2 speech production, analyzing several aspects of speech (e.g., segmentals, suprasegmentals, grammar, lexicon). For example, Derwing and Munro (2006) found that how easily listeners understand L2 speech was strongly related to segmental errors in an experimental setting. Despite such findings in support of the importance of segmental accuracy in L2 speech, few prior studies have attempted to examine the effect of learners' segmental perception on listeners' understanding of L2 speech production. Even fewer studies have examined the role of segmentals (consonants and vowels) in EFL learners' speaking performance at the college level in Asian settings, even though many curricula in these settings emphasize pronunciation teaching (e.g., Chang, 2008; Mehrpour & Makki, 2011).

As for suprasegmentals, Moyer (1999) pointed out that the appropriate use of intonation and pause is rarely focused on in educational settings, although many studies have argued that suprasegmentals influence EFL learners' speaking performance much more than segmentals do. In this regard, some researchers have investigated the role of suprasegmental features (e.g., the appropriate use of stress, intonation, and pauses) in how comprehensible L2 oral performance is assessed to be. Kang (2010) examined whether suprasegmental features (e.g., speech rate, pauses, stress, and pitch range) of L2 speech affected English native speakers' judgments of 5-minute speech samples of 11 international teaching assistants' presentations. The study found that speech rate alone accounted for 35% of the variance in the native listeners' judgments. This finding might be difficult to generalize in terms of the effects on L2 learners' oral performance, however, because the speech rate of nonnative speakers usually is slower than that of native speakers (e.g., Munro & Derwing, 2001). Extending Kang's 2010 findings, Kang et al. (2010) used computer-assisted acoustic analysis to further examine the effects of 29 suprasegmental features (p. 558) on native speakers' judgments of speech samples elicited in a testing setting (i.e., the iBT TOEFL). Fourteen of the 29 suprasegmental features accounted for 50% of the variance in judgments, indicating that suprasegmental features affect the assessment of L2 oral proficiency. Similar recent works have examined how variables such as pronunciation (segmental errors, word stress, intonation, and speech rate), lexical appropriateness, and grammatical accuracy affect L1 speakers' comprehensibility judgments of L2 speech (e.g., Akiyama & Saito, 2016; Isaacs & Trofimovich, 2012; Saito, 2017; Saito et al., 2016). For example, Saito et al. (2016) showed that most of the variables they tested (i.e., segmental accuracy, word stress, intonation, and speech rate) are related to L1 speakers' judgments of L2 speech in experimental settings.

Overall, these previous studies have focused on how linguistic variables affect L1 speakers' judgments of L2 speech, but have not found consistent effects across different types of speaking task and different settings (e.g., recorded vs. spontaneous speech; testing vs. experimental settings; Crowther et al., 2018).

Methodologically, the studies have relied on acoustic analysis (Kang, 2010; Kang et al., 2010) or raters' subjective evaluations of pronunciation and lexicogrammar (e.g., Kashiwagi & Snyder, 2008; Saito et al., 2016) of recorded speech. Hence, little is known about the relationship between EFL learners' perceptual sensitivity to linguistic variables and their speech production. Furthermore, research on how such variables affect L2 speakers' judgments of L2 speech has rarely been conducted in Korean classroom contexts. Therefore, the current study is conducted in a Korean EFL classroom setting and explores the question of whether suprasegmental perception activities that can be implemented in EFL instructional settings can lead to better L2 speech. To do so, it first assesses the learners' perceptual sensitivity to linguistic variables, and then examines the relationships between their perceptual measures and the comprehensibility of their spontaneous speech.

Language Ability and Cognitive Variables: Listening Comprehension, Vocabulary Size, and Short-term Memory

Research has shown various non-linguistic factors to be relevant to learners' speech production. Among the language skills related to oral production, listening comprehension plays an important role in foreign language learning (e.g., Rost, 2001). Most speech learning theories agree that listening and speaking skills are related, but there remains a practical debate around the question of how perception and production are linked.

Listening comprehension includes both bottom-up processing and top-down processing; it is generally agreed that both processes take place at the same time, and furthermore, that they interact as listeners understand spoken input. Top-down processing refers to listeners' use of background knowledge to understand a message, whereas bottom-up processing refers to using the incoming input (sounds and words) to understand the message. In top-down processing, for instance, listeners' knowledge of the appropriate use of suprasegmentals tells them what to expect next as they hear input (O'Malley et al., 1989, p. 421). In bottom-up processing, segments are decoded at the phonetic level, and then suprasegmental features, such as stress, pause, and intonation, provide cues to segmentation and meaning to interpret the spoken input. Likewise, suprasegmentals and segments are involved in both types of processes in listening comprehension. It is assumed that language learners cannot produce target language sounds accurately without first perceiving the sounds (e.g., Flege, 1995), and it has been reported that perceptual training on target sounds improves L2 learners' speech production as well as their perception (Borden et al., 1983; Bradlow et al., 1997; Wang et al., 2003). Such findings provide evidence that understanding how a language's sound system works is essential not only in listening comprehension but also in the development of L2 speaking ability (e.g., Astorga-Cabezas, 2015; Bozorgian, 2012).

Recent research has also shown that L2 learners' listening ability is positively related to their target language speaking ability, and that better listening comprehension enables EFL learners to perceive language input and facilitates their development of language skills (Bozorgian, 2012; Nation & Newton, 2009; Pavlenko, 2010; Vandergrift & Goh, 2012). An empirical study conducted by Astorga-Cabezas (2015) explored the efficacy of listening-based instruction for improving students' oral skills in English. The study found that not only pronunciation but also vocabulary, speech register, topicalization, and paralinguistic features were associated with listening skills (p. 53). A positive association between listening and speaking was also reported by Bozorgian (2012), who found that 701 IELTS test takers' listening scores were strongly correlated with their speaking scores ($r = .654$, $n = 1800$, $p < .000$; p. 661).

In addition to listening ability, appropriate vocabulary knowledge is essential to successful communication. Vocabulary knowledge is not only strongly related to language learners' reading and writing ability, but also influences their phonemic awareness, which is further related to their articulation, speech, and reading skills (e.g., Roberts, 2005). Although there is still a paucity of research on this topic, some studies have reported a relationship between speaking and vocabulary knowledge (e.g., De Jong et al., 2012; Koizumi & In'nami, 2013), implying that vocabulary knowledge is a potential predictor of EFL learners' speaking ability. Iwashita et al. (2008), for example, analyzed 200 speech samples produced by

L2 English speakers at different levels. The test takers participated in a five-task speaking test on the TOEFL iBT. The speech samples were holistically rated and objectively scored in terms of linguistic resources (e.g., grammatical accuracy, grammatical complexity, and vocabulary), phonology (e.g., pronunciation, intonation, and rhythm), and fluency (e.g., total pausing time, speech rate). Vocabulary was one of the features that most strongly impacted overall speech performance. The study examined vocabulary knowledge by counting proportions of low- and high-frequency vocabulary, and found that learners' oral performance was positively correlated with both tokens (i.e., number of words produced) and types (i.e., number of different words produced). Similarly, Koizumi and In'nami (2013) explored the relation between L2 speaking proficiency and vocabulary knowledge, and found substantial predictive effects of vocabulary size, vocabulary depth, and speech speed on learners' speaking proficiency. However, a contradictory result was also reported in one recent study: Uchihara and Saito (2019) found no correlation between L2 oral proficiency, in terms of comprehensibility, and vocabulary knowledge.

As for the cognitive variable, short-term memory has been shown to play a role in learners' speech production, and is associated with both storage and processing. Many researchers have agreed that short-term memory capacity is a good predictor of EFL learners' ability to perceive, store, recall, and reproduce phonological sequences (e.g., Baddeley, 2003; Baddeley & Hitch, 1974; Juffs & Harrington, 2011; Kormos, 2006; O'Brien et al., 2006; Nowbakht & Fazilatfar, 2019; Prebianca et al., 2014). Among the research on the relationship between short-term memory and second language acquisition, Prebianca et al. (2014) found that L2 short-term memory capacity is affected by L2 proficiency level. O'Brien et al. (2006) found that phonological working memory is correlated with the development of L2 oral fluency. Recent evidence also indicates that short-term memory is associated with listeners' ability to detect stress (e.g., Dupoux et al., 2001) and contrast between minimal pairs (e.g., Jacquemot et al., 2006). Although some research (e.g., Mizera, 2006) has found no relationship between short-term memory and L2 fluency, this study draws on widespread findings to hypothesize that learners' short-term memory will have a positive relationship with their oral ability. The existing literature focuses on short-term memory in the development of fluency and accuracy in L2 speech production; how the short-term memory is involved in perceiving suprasegmentals and segmentals, and the resulting influence on the comprehensibility of L2 speech production, is less explored.

Research Questions

Considering that improving learners' ability to produce comprehensible speech is one of the primary aims of English education, the research findings discussed above suggest pedagogical implications for Asian EFL classrooms. Understanding these linguistic and cognitive variables and the two language abilities related to L2 oral performance seems necessary to explain L2 speech production, and to provide insights into teaching comprehensible speech in EFL instructional settings. Furthermore, understanding of the relationships among factors may be insightful as a preliminary step so as to better interpret a predictor of L2 oral performance. Therefore, the present study addresses the following two research questions:

- 1) What relationships exist among five potential predictor variables for EFL learners' speech production: segmental perception, suprasegmental perception, listening comprehension, vocabulary size, and short-term memory?
- 2) Which of the five variables predicts performance in EFL learners' speech production?

Method

Participants

A total of 51 Korean EFL learners (14 female, 37 male; mean age = 24.5 years old) participated in the study. They were enrolled in a basic English-speaking course at a university in South Korea, and they are considered a homogenous basic level group because their TOEIC scores fell in the range of 500–550. Their majors were electronic engineering ($n = 40$) and business administration ($n = 11$).

Instruments

Six instruments were utilized to collect data: (a) a suprasegmental feature perception test, (b) a segmental feature perception test, (c) a listening comprehension test, (d) a vocabulary size test, (e) a short-term memory test, and (f) a timed picture-description test.

Suprasegmental feature perception test

In an in-class session before the experiment, the participants received instruction on marking the appropriate use of stress, pauses, and intonation while listening to recorded speech. In the suprasegmental feature perception test, the students listened to two 13-sentence passages, one 77 words long, and the other 79 words long. The participants were allowed to listen to each passage twice. They read along and marked the passages in a paper handout. Only sentences that were marked entirely correctly, including emphasis and pauses, received a score of 1. No partial credit was given. The highest possible score for each passage was 13 points. The two passages were taken from a TOEIC speaking textbook (*PAGODA TOEIC Speaking* 2016) used in Korean college classrooms for beginners or lower intermediate learners. The test took approximately 10 minutes.

Segmental feature perception test

The segmental feature perception test measured students' ability to distinguish similar sounds. Participants heard a sound and then selected the word including that sound from a minimal pair (e.g., *boat* vs. *vote*). The test included 12 items; the highest possible score was 12 points. It took about five minutes.

Listening comprehension test

The listening comprehension test was administered to measure students' general English listening comprehension ability, and to see whether their listening proficiency was associated with their speaking ability. The test included a total of 31 questions taken from a TOEIC practice test published by a major language school and publishing company in Korea (<http://www.ybm.co.kr/>). The students took about 25 minutes to complete the test.

Vocabulary size test

To measure students' receptive vocabulary knowledge, a Korean version of Nation and Beglar's (2007) 14000 level test was used (<https://www.victoria.ac.nz/lals/about/staff/paul-nation>). It included 140 multiple-choice questions with four choices, one of which corresponded to the target word embedded in a simple sentence. Students were allowed to skip unknown words but advised to refrain from wild guesses to improve the accuracy of the measurement. A learner's total score is multiplied by 100 to find their total

vocabulary size. For example, 35 out of 40 means the learner's vocabulary size is 3,500 word families. Thirty minutes were given to complete the test.

Short-term memory test

An online test (<http://faculty.washington.edu/chudler/neurok.html>) was employed to assess learners' short-term memory. They were presented with strings of letters of increasing length (2, 4, 6, 8, 10, 12 letters) for three seconds, and asked to immediately repeat them in the same order as presented. The test consisted of six items. One point was given for each correct answer; no partial points were given.

Timed picture-description test

Following previous L2 pronunciation studies (e.g., Derwing & Munro, 2009; Munro & Mann, 2005; Saito et al., 2016), learners' speech was elicited via a timed picture-description test. This task type helps to minimize the amount of conscious speech monitoring (e.g., Ellis, 2005), providing spontaneous speech data without frequent hesitations and dysfluencies. The students were given two pictures and asked to describe them in as much detail as possible. They were given 45 seconds to speak about each picture after 30 seconds of preparation time. All participants' productions were recorded, and they were assessed based on a seven-point Likert scale (0 = incomprehensible; 6 = highly comprehensible), following the evaluation of comprehensibility used by Kang (2010).

Two nonnative speakers of English, both of whom held doctorates in applied linguistics and had taught for at least three years at a college in South Korea, assessed the comprehensibility of each speech sample. The two raters reached 82% agreement. Discrepancies were resolved by averaging the scores.

Procedures

As Table 1 shows, all participants completed three separate sessions with intervals of at least one week between each session. In the first session, students took the listening comprehension test containing 31 listening comprehension questions. Before the second session, the suprasegmental instruction session was conducted in class to help the students understand the concepts of segmentals and suprasegmentals as well as to allow them to practice marking intonation, stress, and pausing. During the actual test, the students listened to the two recorded files and marked the sentences as they had during the practice session. Then the segmental perception test, in which the students distinguished between the 12 minimal pairs, followed. Next, they completed the timed-picture description test. In the third session, the students took both the vocabulary size test and the short-term memory test.

TABLE 1
Experimental Procedure: Order of the Six Tests

	Session 1	Session 2	Session 3
Tests	Listening comprehension test	Suprasegmental perception test Segmental perception test Picture-description test	Vocabulary size test Short-term memory test

Results

Relationships among Variables Involved in EFL Learners' Speech Production

The first research question examines the relationships that may exist among the five variables assessed by this study: two linguistic variables, segmental and suprasegmental perceptual sensitivity; two language ability variables, English listening comprehension ability and vocabulary size; and a cognitive variable, short-term memory. As Table 2 shows, the score on the picture-description test is a dependent measure, whereas all the others are independent measures as the potential predictor variables for EFL learners' speech production. The table presents the descriptive statistics for the participants' scores on the six tests. Mean scores on each test were calculated and used for further analysis.

TABLE 2
Descriptive Statistics for the Six Tests

Tests	<i>N</i>	<i>M</i>	<i>SD</i>
Listening comprehension test (Max = 100)	51	58	22.03
Vocabulary size test (Max = 14000)	51	6709.8	2806.08
Short-term memory test (Max = 12)	51	5.69	2.13
Suprasegmental perception test (Max = 13)	51	6.65	2.54
Segmental perception test (Max = 12)	51	10.02	1.46
Picture-description test (Max = 6)	51	3.35	0.76

TABLE 3
Correlations among Scores on the Six Tests

Tests	1	2	3	4	5	6
1	1					
2	.232	1				
3	.340*	.320*	1			
4	.355*	.304*	.186	1		
5	.016	.034	.094	.378**	1	
6	.376**	.258	.109	.423**	.229	1

Note. 1 = English listening proficiency scores; 2 = vocabulary size test scores; 3 = short-term memory test scores; 4 = suprasegmental perception test scores; 5 = segmental perception test scores; 6 = picture-description test scores
* $p < .05$; ** $p < .01$

Correlations between predictor variables were examined. The results appear in Table 3. First, learners' English listening proficiency was moderately related to their scores on the suprasegmental perception test ($r = .355$) and the short-term memory test ($r = .340$). These findings were expected, as listening comprehension proficiency as a general representation of learners' language skills is closely related to suprasegmental sensitivity (i.e., the ability to perceive such prosodic features as stress, pauses, and intonation). As the knowledge of suprasegmentals is involved in both bottom-up and top-down processing in listening, better listening comprehension scores can be explained by higher perceptual sensitivity to suprasegmentals, which provide a cue to segmentation and meaning in the interpretation of spoken input.

Further, as described in previous studies on the correlation between short-term memory and speech perception, short-term memory ability is one of the underlying factors influencing listening proficiency; this finding may imply that short-term memory indirectly influences listening comprehension ability. In line with the previous studies that have shown a crucial role of short-term memory in foreign language acquisition (e.g., Aliaga-García et al., 2011; Darcy et al., 2015), this study also found a relation, albeit weak, between EFL learners' listening proficiency and their short-term memory capacity. Greater short-term memory capacity might facilitate a learner's ability to perceive and store phonological sequences, leading to better comprehension, which means that learners with more short-term memory capacity would show better comprehension than learners with less short-term memory capacity. Short-term memory may help to encode linguistic information in the phonological loop as a function of storing sequential information, affecting the amount of aurally comprehended information learners can hold in memory at one time (e.g., Baddeley et al., 1998).

Second, vocabulary size was moderately related to short-term memory ($r = .320$) and to scores on the suprasegmental perception test ($r = .304$). As discussed, learners can hold different amounts of information in memory, which also leads to differences in vocabulary size (Papagno et al., 1991; Service, 1992), so different sizes of working memory in foreign language learning affect achievement in speech production. Vocabulary size also seems to be related to EFL learners' suprasegmental sensitivity, because letter-sound knowledge of words is associated with sound awareness (e.g., Roberts, 2005). Third, the scores on the suprasegmental perception test were also moderately related to the scores on the segmental perception test ($r = .378$). This finding was somewhat expected, because learners would be expected to learn both types of features while learning pronunciation and prosody in instructional contexts (e.g., Trofimovich & Gatbonton, 2006).

Fourth, in regard to spontaneous speaking ability, the participants' listening comprehension proficiency was moderately related to their scores on the picture-description test ($r = .376$), and their perceptual sensitivity to suprasegmentals was considerably related to the comprehensibility of their speech production in the picture-description test ($r = .423$). In the literature on the relationship between speaking and listening, better listening ability has been shown to be closely related to oral ability (Astorga-Cabezas, 2015; Bozorgian, 2012; Nation & Newton, 2009; Pavlenko, 2010). Hence, the moderate relationship between listening comprehension scores and the scores on the picture-description test supports the generalization that the deep processing of listening influences EFL learners' speaking ability. Interestingly, the observed relationship between perceptual sensitivity to suprasegmentals and higher scores on the picture-description test provides the empirical evidence supporting that perception and production are linked. However, the current study found no relationship between learners' perceptual sensitivity to segmentals and their speaking performance scores. Saito et al. (2015) found a relationship between accuracy in the production of segments and oral performance only for more proficient learners, which suggests that the finding of no relationship in the current study might be explained by the participants' low English proficiency; in other words, as beginners, they were not yet able to accurately perceive the English segments.

Predictors of EFL Learners' Oral Performance

To address the second research question, on the relationships between the potential predictor variables and the EFL learners' oral performance, a multiple regression analysis was conducted. Table 4 shows a summary of the analysis of variance for regression models for predicting the scores of learners' speech production in the picture-description test.

TABLE 4

Multiple Regression Analysis for Variables Predicting the Performance of the EFL Learners on the Picture-description Test

Model	Variable entered	R^2	Adjusted R^2	ΔR^2	ΔF	B	β	t	Sig.
1	Suprasegmental perception	.179	<i>N.A</i>	<i>N.A</i>	<i>N.A</i>	.13	.423	3.264	.002
2	Suprasegmental perception	.179	.162	.179	10.652	.1	.331	2.453	.018
	Listening proficiency	.237	.205	.058	3.664	.01	.258	1.914	.062

As Table 4 shows, among the five predictor variables, only the scores on the suprasegmental perception test and the English listening proficiency test were found to be significant predictors for the scores on oral performance. The two variables together accounted for 20.5% of the variance of the scores on the learners' speech production ($F(2, 48) = 7.447, p = .002, R^2 = .237, \text{Adjusted } R^2 = .205$). First, their suprasegmental perception scores only explained 17.9% of the variance of the scores on the picture-description test ($F(1,49) = 10.652, p = .002, R^2 = .179$). Next, the contribution of listening proficiency to their scores on the picture-description test was assessed after controlling for the scores on the suprasegmental perception test. As Table 4 shows, there was a significant F -value change ($p = .062$) when English listening proficiency was entered as an additional variable after the suprasegmental perception scores. This means that English listening proficiency was also a marginally significant predictor of the scores on the speaking test, explaining an additional 5.8% of the variance of the scores on the picture-description test. No other variables were found to be significant.

Standard beta coefficients were examined to determine whether the two variables that showed an association with the EFL learners' picture-description test scores played a positive or a negative role in the learners' oral performance. First, for the suprasegmental perception scores, positive beta weights indicate that higher suprasegmental perception scores correlated with better speech comprehensibility ($\beta = .423, t = 3.264, p = .002$). English listening proficiency also played a positive role, with higher scores on the English proficiency test also correlating with higher comprehensibility ($\beta = .258, t = 1.914, p = .062$). The current study's finding that suprasegmental perception is important in EFL learners' oral performance contributes to a better understanding of EFL learners' speech production. However, the finding of no significant relationship between segmental perception scores and the comprehensibility of EFL learners' oral performance was unexpected.

Discussion

The present study has explored the association between Korean EFL learners' speech production, segmental and suprasegmental sensitivity, short-term memory, listening comprehension ability, and vocabulary size. The results showed that EFL learners' perceptual sensitivity to suprasegmentals is the most predictive of their performance of spontaneous speaking. In contrast, learners' perceptual sensitivity to segments showed no predictive power regarding their oral performance. Among the learner variables, only listening comprehension proficiency was also found to be a predictor, albeit a weak one, of the EFL learners' oral performance. No other variables were found to be significant.

The current findings provide empirical evidence of the effect of suprasegmental features on the EFL learners' speech production. The finding of the crucial role of suprasegmental sensitivity in EFL learners' oral performance contributes to our understanding of the nature of EFL learners' speech proficiency in Korean college contexts. This study supports Celce-Murcia et al.'s (1996) emphasis on the importance of suprasegmentals over segmentals in successful communication. This specific result is consistent with findings from some previous L2 studies (e.g., Kang et al., 2010; Saito et al., 2016), although the finding

of no relationship between sensitivity to segmental information and oral performance is not. Because perceptual sensitivity of suprasegmentals appears to be a good predictor of EFL learners' speech, this study supports explicit instruction on suprasegmental features in order to foster successful communication ability, although it also acknowledges that the relative weight of instructional focus on segmentals versus suprasegmentals may need to be varied in different EFL classrooms.

Second, the positive relationship between EFL learners' listening comprehension proficiency and oral performance found in this study corroborates empirical evidence reported in some previous studies (e.g., Astorga-Cabezas, 2015), indicating that listening skill is a likely predictor of L2 speech. This finding suggests that the listening process is involved in the cognitive process of perceiving suprasegmentals by using cues to understand the meaning of spoken input. As seen from the correlation between scores on the suprasegmental perception test and listening comprehension ($r = .355$), learners are able to indirectly learn how to use suprasegmentals via the cognitive process of listening comprehension. They then may apply their knowledge of suprasegmentals to their oral performance, such as during a picture-description task. This implies that practicing listening that involves the perception of suprasegmentals may increase learners' sensitivity to phrasal and clausal boundaries, in turn improving their ability to appropriately use suprasegmentals in their L2 speech. This explanation supports the hypothesis that the better perception of suprasegmentals as well as segments during the listening process is linked to learners' oral ability so that better perception should indirectly lead to better production. Therefore, these findings generally support the relationship between perception and production.

This study found no relation between learners' short-term memory or vocabulary size and their scores on the oral performance, while several previous studies have shown these two variables' positive effects on L2 speech proficiency (e.g., Baddely, 2003; Iwashita et al., 2008; Koizumi & In'nami, 2013; O'Brien et al., 2006; Prebianca et al., 2014). One of the variables, short-term memory, seems likely to be related to spontaneous speech by L2 learners, in that such production needs a very high level of attention. Nonetheless, no significant relation was found. One possible explanation for this finding is that short-term memory is more involved in perception (a relation between short-term memory and listening comprehension was found; $r = .340$, Table 3) than in production. With respect to the learners' vocabulary size, the finding of no relationship to oral performance might have resulted from the raters' focus on the comprehensibility of the learners' pronunciation. Another possibility is that these beginning-level participants' vocabulary size was too small to be measured so that the effect of their vocabulary size cannot explain the learners' oral proficiency. The relation between EFL learners' vocabulary size and the quality of their production deserves closer attention in future research.

Pedagogical Implications

With the burgeoning interest in increasing EFL learners' speaking ability as part of communicative competence, the findings of the present study suggest some pedagogical implications for Korean college-level classroom settings, which may be applicable in other EFL contexts as well. First, in line with previous studies (e.g., Derwing & Munro, 2009; Kang et al., 2010; Saito et al., 2016), the present study suggests that instruction designed to increase students' sensitivity to suprasegmental features is crucial. In particular, instructions should specifically target EFL students' knowledge of appropriate pausing, word stress at the sentence level, and variation of intonation (Kang, 2010, p. 312). Such instruction would also raise learners' awareness, improving their ability to notice their pronunciation when involved in a variety of speaking activities (e.g., presentations, dialogues).

Although teachers of English may recognize the importance of these features for increasing overall oral fluency, many are unsure how to effectively integrate suprasegmental instruction into their classrooms (Levis & Grant, 2003). Harmer (2007) suggested that providing students plenty of opportunities to listen is obviously beneficial for their pronunciation, which includes not only their ability to produce individual sounds but also their ability to accurately employ pitch, intonation, and stress. Levis and Grant (2003) similarly asserted that instruction should focus on suprasegmental features rather than segmental ones, as the

former contribute to speaking ability more directly. Specifically, the researchers recommended storytelling, utilizing thought grouping or chunking (p. 16), as a useful way to practice suprasegmental features.

Another suggestion is the use of reading aloud as an instructional practice for teaching pronunciation in listening and speaking classes. While reading aloud is considered old-fashioned, this does not mean that it is no longer useful in language learning (e.g., Gibson, 2008). Some pronunciation books include short passages for reading aloud, and reading aloud activities could be introduced to students as part of their preparation for giving oral presentations and taking dictation. Such activities might encourage learners to focus on accurate articulation for better comprehensibility of L2 speech along with better pronunciation. Practicing reading aloud along with appropriate animation and expression in the classroom was also suggested by Nurani and Rosyada (2015) as a helpful means of drawing awareness to suprasegmental features, including intonation and stress patterns. Nevertheless, reading aloud alone is unlikely to guarantee better L2 speech over the long term, and more investigation into the best methods of teaching suprasegmentals in particular and pronunciation in general is needed.

Second, as listening ability turned out to be a significant predictor of oral performance in this study, combining listening and speaking skills as much as possible in the classroom should be highly recommended to L2 teachers. More importantly, increasing students' sensitivity in the context of speaking for meaningful communication is desirable (e.g., Murphy, 1991). Hence, this study suggests that L2 program designers, practitioners, and teachers should consider the integrative teaching of listening and speaking. Listening and speaking are naturally integrated in real-life communication, as well as linked by experimental research. There is no doubt that practicing perception and production at both segmental and suprasegmental levels can increase comprehensibility among interlocutors (e.g., between EFL learners and teachers or among EFL learners), leading to improved quality and quantity of communication in and outside the classroom. This also parallels the integrative approach of test settings. For example, the TOEFL iBT test includes an integrated speaking test, in which test takers are asked to speak on the basis of what they understand after listening to an academic lecture. This trend should not be neglected in instructional settings, yet many educators who are interested in the integrative approach to teaching English at the college level are hindered by practical concerns.

Conclusion

This study explored the extent to which linguistic variables and language skills could affect Korean EFL learners' speech production. To summarize the findings, EFL learners' suprasegmental sensitivity is a predictor of their speaking performance, and their listening comprehension ability also marginally predicts their speech production. However, the study found no relationship between speaking performance and sensitivity to segmental information, short-term memory, or vocabulary size.

The findings of the study lend support to the importance of emphasizing the appropriate use of suprasegmentals in the EFL classroom. Increasing their sensitivity to intonation, pause, and stress should help EFL learners improve their oral performance and the comprehensibility of their production. Curricula that include extended and focused exercises on such aspects of language would contribute to learners' improved communicative competence, which, in most Asian countries as well as elsewhere, corresponds to one of the main instructional goals of English education.

The Authors

On-Soon Lee (first author) is an assistant professor at Sunchon National University in Korea. Her current research interests include cognitive individual differences in L2 acquisition.

Email address: oslee@scnu.ac.kr

Jeongyeon Park (corresponding author) is an assistant professor at Dong-A University in Korea. Her research interests include language pedagogy and individual differences in language learning and teaching in EFL settings.

Email address: jpark93@dau.ac.kr

References

- Akiyama, Y., & Saito, K. (2016). Development of comprehensibility and its linguistic correlates: A longitudinal study of video-mediated telcollaboration. *The Modern Language Journal*, 100(3), 585-609.
- Aliaga-Garcia, C., & Mora, J. C. (2009). Assessing the effects of phonetic training on L2 sound perception and production. In M. A. Watkins, A. S. Rauber, & B. O. Baptista (Eds.), *Recent research in second language phonetics/phonology: Perception and production* (pp. 2-31). Newcastle, UK; Cambridge Scholars Publishing.
- Aliaga-García, C., Mora, J., & Cervino-Povedano, E. (2011). L2 speech learning in adulthood and phonological short-term memory. *Poznan Studies in Contemporary Linguistics*, 47(1), 1-14.
- Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, 36, 189-208.
- Baddeley, A., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, 8, 47-89.
- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105(1), 158-173.
- Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second language acquisition. *Journal of Memory and Language*, 44, 235-249.
- Borden, G., Gerber, A., & Milsrak, G. (1983). Production and perception of the /r/ -/l/ contrast in Korean adults learning English. *Language Learning*, 33(4), 499-526.
- Bozorgian, H. (2012). The relationship between listening and other language skills in international language testing system. *Theory and Practice in Language Studies*, 2(4), 657-663.
- Bradlow, A., Akahane-Yamada, R., Pisoni, D., & Yoh'ichi, T. (1997). Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Percept Psychophys*, 61(5), 977-985.
- Celce-Nurcia, M., Briton, D. M., & Goodwin, J. K. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge, UK: Cambridge University Press.
- Chang, H. J. (2008). Effectiveness of pronunciation training on suprasegmentals. *Korean Journal of Applied Linguistics*, 24(1), 85-108.
- Crowther, D., Trofimovich, P., Saito, K., & Isaacs, T. (2018). Linguistic dimensions of L2 accentedness and comprehensibility vary across speaking tasks. *Studies in Second Language Acquisition*, 40(2), 443-457.
- Crystal, D. (2003). *English as a global language*. UK: Cambridge University Press.
- Darcy, I., Park, H., & Yang, C. (2015). Individual differences in L2 acquisition of English phonology: The relation between cognitive abilities and phonological processing. *Learning and Individual Differences*, 40(1), 63-72.
- De Jong, N. H., Steinel, M. P., Florijn, A. F., Schoonen, R., & Hulstijn, J. H. (2012). Facets of speaking proficiency. *Studies in Second Language Acquisition*, 34(1), 5-34.
- Derwing, T., & Munro, M. (2009). Putting accents in its place: Rethinking obstacles to communication. *Language Teaching*, 42(4), 476-490.
- Derwing, T., Munro, M., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48(3), 393-410.

- Derwing, T., & Rossiter, M. (2003). The effects of pronunciation instruction on the accuracy, fluency, and complexity of L2 accented speech. *Applied Language Learning*, 13(1), 1-17.
- Derwing, T., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, 39, 379-397.
- Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2001). A robust method to study stress “deafness”. *Journal of Acoustical Society of America*, 110, 1606-1618.
- Flege, J. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Timonium, MD: York Press, 1995.
- Flege, J., Yeni-Komshitan, G., & Liu, S. (1999). Age constraints on second language acquisition. *Journal of Memory & Language*, 41, 78-104.
- Gan, Z. (2012). Understanding L2 speaking problems: Implications for ESL curriculum development in a teacher training institution in Hong Kong. *Australian Journal of Teacher Education*, 37(1), 43-59.
- Harmer, J. (2007). *How to teach English* (New ed.). Harlow: Pearson Education Limited.
- Isaacs, T., & Trofimovich, P. (2012). Deconstructing comprehensibility. *Studies in Second Language Acquisition*, 34, 475-505.
- Iwashita, N., Brown, A., McNamara, T., & O’Hagan, S. (2008). Assessed levels of second language speaking proficiency: How distinct? *Applied Linguistics*, 29(1), 24-49.
- Jacquemot, C., Dupoux, E., Decouche, O., & Bachoud-Lévi, A. (2006) Misperception in sentences but not in words: Speech perception and the phonological buffer. *Cognitive Neuropsychology*, 23(6), 949-971.
- Jenkins, J. (2000). *The phonology of English as an international language*. Oxford: Oxford University Press.
- Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System*, 38, 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.
- Kashiwagi, A., & Snyder, M. (2008). American and Japanese listener assessment of Japanese EFL speech: Pronunciation features affecting intelligibility. *The Journal of Asia TEFL*, 5(4), 27-47.
- Kim, M., & Hahn, P. (2016). *PAGODA TOEIC Speaking*. Seoul: Pagoda Books.
- Kim, M. H., & Nam, J. Y. (2002). Segmental vs. prosodic approach to listening comprehension in English. *Foreign Languages Education*, 9(2), 303-326.
- Koizumi, R., & In’nami, Y. (2013). Vocabulary knowledge and speaking proficiency among second language learners from novice to intermediate levels. *Journal of Language Teaching and Research*, 4(5), 900-914.
- Levis, J. M., & Grant, L. (2003). Integrating pronunciation into ESL/EFL classrooms. *TESOL Journal*, 12(2), 13-19.
- Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: the MIT Press.
- Leather, J. (1999). Second-language speech research: An introduction. *Journal of Research in Language Studies*, 49(1), 1-56
- Liontas, J. (2018). *The TESOL encyclopedia of English language teaching: Volume III*. NJ: Wiley Blackwell.
- McCandless, P., & Winitz, H. (1986). Test of pronunciation following one year of comprehension instruction in college German. *The Modern Language Journal*, 70(4), 355-362.
- Mehrpour, S., & Makki, M. (2011). Let’s see if learners’ level of proficiency in English affects their production of segmentals and suprasegmentals: A case of Iranian EFL learners. *The Journal of Asia TEFL*, 8(2), 119-143.
- Mizera, G. J. (2006). *Working memory and L2 oral fluency* (Unpublished doctoral dissertation). University of Pittsburgh, USA.

- Morley, J. (1991). The pronunciation component of teaching English to speakers of other languages. *TESOL Quarterly*, 25, 481-520.
- Morley, J. (1994). *Pronunciation pedagogy and theory: New views, new directions*. Washington, DE: Teachers of English to Speakers of Other Language (TESOL).
- 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, 451-468.
- Munro, M., & Mann, V. (2005). Age of immersion as a predictor of foreign accent. *Applied Psycholinguistics*, 26(3), 311-341.
- Munro, M. J., & Derwing, T. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System*, 34(4), 520-531.
- Munro, M. J., & Derwing, T. (2011). The foundation of accent and intelligibility in pronunciation research. *Language Teaching*, 44(3), 316-327.
- Murphy, J. (1991). Oral communication in TESOL: Integrating listening, speaking, and pronunciation. *TESOL Quarterly*, 25, 51-74.
- Nation, I. S. P., & Newton, J. (2009). *Teaching ESL/EFL listening and speaking*. New York, NY: Routledge.
- Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9-13.
- Nowbakht, M., & Fazilatfar, A. M. (2019). The effects of working memory, intelligence and personality on English learners' speaking ability. *The Journal of Asia TEFL*, 16(3), 817-832.
- Nurani, S., & Rosyada, A. (2015). Improving English pronunciation of adult ESL learners through reading aloud assessments. *Lingua Cultura*, 9(2), 107-112.
- O'Brien, I., Segalowitz, N., Collentine, J., & Freed, B. (2006). Phonological memory and lexical narrative, and grammatical skills in second language oral production by adult learners. *Applied Psycholinguistics*, 27, 377-402.
- Pavlenko, T. (2010). *Integration of listening and speaking in teaching English*. Ukraine: E-Teacher Summer Institute.
- Papagno, C., Valentine, T., & Baddeley, A. D. (1991). Phonological short-term memory and foreign language vocabulary learning. *Journal of Memory and Language*, 30, 331-347.
- Prebianca, C. V. V., Finardi, K. R., & Wessheimer, J. (2014). Working memory capacity across L2 speech proficiency levels. *Revista Brasileira de Linguística Aplicada*, 14(2), 441-462.
- Roberts, T. A. (2005). Articulation accuracy and vocabulary size contributions to phonemic awareness and word reading in English language learners. *Journal of Education Psychology*, 97(4), 601-616.
- Rost, M. (1991). *Listening in action*. UK: Prentice Hall International Ltd.
- Saito, K. (2017). Effects of sound, vocabulary, and grammar learning aptitude on adult second language speech attainment in foreign language classroom. *Language Learning*, 67(3), 665-693.
- Saito, K., Trofimovich, P., & Isaacs, T. (2015). 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., & Isaacs, T. (2016). Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholinguistics*, 37, 217-240.
- Service, E. (1992). Phonology, working memory, and foreign language learning. *Quarterly Journal of Experimental Psychology*, 45, 21-50.
- 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., & Gatbonton, E. (2006). Repetition and focus on form in processing L2 Spanish words: Implications for pronunciation instruction. *The Modern Language Journal*, 90(4), 519-535.
- Trofimovich, P., & Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism: Language and Cognition*, 15, 905-916.

- Uchihara, T., & Saito, K. (2019). Exploring the relationship between productive vocabulary knowledge and second language oral ability. *The Language Learning Journal*, 47(1), 64-75.
- Wang, Y., Jongman, A., & Sereno, J. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *The Journal of the Acoustical Society of America*, 113(2), 1033-1043.
- Yang, M. (2013). The effect of teaching prosody on L2 listening comprehension. *Foreign Languages Education*, 20(1), 27-44.

(Received May 15, 2020; Revised September 02, 2020; Accepted September 12, 2020)