Korean EFL Learner's Suprasegmental Features

Xin Hu[1] & Haiying Du[2]

[1] Shandong Vocational University of Foreign Affairs, China

[2] Professor, English Teaching Section, Shandong Vocational University of Foreign Affairs, China

Correspondence: Haiying Du, Professor, English Teaching Section, Shandong Vocational University of Foreign Affairs, China

Abstract

This study delves into some aspects of suprasegmental features such as syllable structure, stress, and rhythm and compares them between NS and NNS. It is investigated in spectrograms and sound waveforms that 1. On the aspect of syllable structure in English, the onset and the coda in English syllable structure are characterized to have a maximum of 3 and 4 consonant clusters, respectively. In contrast, Korean allows only 1 single consonant in onset and coda position. This cross-linguistic difference gives rise to the insertion of the neutral vowel /ɨ/ to break up the consonant clusters in English words, in which the inserted vowel forms an independent wave chunk. 2. Refer to stress in English, it is universally recognized as every single English word or sentence consist of its own stress. On the contrary, Korean lacks stress placed at the level of the word. It follows that Korean EFL learners tend to put an approximately equal prominence on every syllable in a word and to exhibit a tendency to put a strong prominence particularly on the first syllable of a word with more than 2 syllables, which is dubbed an 'initial prominence phenomenon'in this paper. 3. In relation to English rhythm, English is certainly a stress-timed rhythm, but Korean is a syllable-timed rhythm. The core differences between the stress-timed rhythm and the syllable-timed are on the form of 'foot', which is established when stressed and unstressed syllables occur in relatively regular alternating patterns in sentences, led to a phenomenon of that the number of feet depends on the timing of articulation within a whole sentence. "This paper finds significance in exhibiting suprasegmental features in visualization between NS and NNS, given that these features play a more important role than segmental ones. It can also serve a milestone for future researchers in the EFL phonetic filed."

1. Introduction

It is obvious that pronunciation is one of the important components of language skills. However, it has been neglected in EFL teaching and research, as well as treated as Cinderella, as dubbed in Kelly (1969). Since then, the pendulum of pronunciation teaching and research has begun to swing in the opposite direction with the emergence of the Reform Movement in phonetics in the 1890s. With the advent of this movement, there appeared an opposite approach in pronunciation teaching and research, which is referred to as the 'Analytic-Linguistic Approach' (Celce-Murcia et al., 1996). In the Analytic-Linguistic Approach, explicit influences of pronunciation pedagogy is enhanced. The clear information guidance can be demonstrated in different interactive speech software and websites (Lee, 2008). Moreover, Murphy (2003) also states that research in the different fields of L2 learning and teaching has shown that the use of explicit instruction can have useful effects in learning. Linguistic research on segmental features of phonemes gives birth to the development of Audiolingualism in the 1940s and 1950s. In the 1970s, we encountered the advent of 'Communicative Language Teaching' in the area of EFL teaching, which values 'fluency' over 'accuracy' in order to fulfill the primary purpose of language, that is, communication. The Communicative Approach, which emerged in the 1970s and is currently dominant in language teaching, is said to play a guiding role in today's pronunciation teaching and research. It follows that suprasegmental features of CLT in pronunciation receive more attention from both ESL and EFL researchers. More importantly, a more significant distinction is often detected between the two groups of learners in the area of suprasegmentals than segmentals. This has led me into research on suprasegmental features among Korean EFL learners. In this paper, I delve into the aspects of syllabification, stress, and rhythm, which are most deviant from those of English native speakers. To show such deviance, I have conducted an experiment on these aspects with one native speaker (NS) and non-native speaker (NNS). Based on

the performance of the NNS, I argue that such suprasegmental features play a significant role in enhancing communicative competence and accordingly suggest some implications ~~to~~ "for" pronunciation teaching to Korean EFL learners.

2. Literature Review

2.1 Research on Segmentals and Suprasegmentals

Prior to interpreting and stating the properties of suprasegmental features in English in the following chapters, I initially present the fundamental components of segmentals and suprasegementals and the shift of research from segmentals to suprasegmentals, which are considered as a prerequisite to the research in this thesis.

As a first approximation, segmentals are widely defined as "the basic inventory of distinctive sounds" which are combined to form a spoken language (Morley, 1991; Celce-Murcia et al., 1996; Florez, 1998). Being 'distinctive' here indicates that segments serve to distinguish words from one another in meaning and are thereby equivalent to phonemes in phonetics. Traditionally, segmental features have been taught more widely than suprasegmental features in the formal classroom. This is partly because they relate to letters in writing and easy to notice and work on (Gilakjani, 2012). Gilakjiani (2012) and Yates (2003), focusing on the "highly recognizable feature" of segmentals, further state that pronunciation is considered the way of producing sounds that are used to convey meaning when speakers pronounce consonants and vowels, which are naturally treated as segments. Accordingly, Derwing & Rossiter (2002) did research on adult EFL learners so as to verify if pronunciation problems are associated with segmental features. The results are quite convincing enough to indicate that these features are a main concern for the learners. This is partly due to the fact that they are so much concerned with segmental differences between their target and native language. Gradually, the focus on segmental features has been somewhat converted to suprasegmental features of pronunciation (Gilbert, 1993).

Anderson-Hsieh et al. (1992) compare the relative contributions made to intelligibility between segmental and suprasegmental features and find that the latter can earn higher scores in enhancing intelligibility. Accordingly, the attention of researchers has been shifted to the research on suprasegmental features. This shift has gone hand in hand with the new trend of language teaching which values 'communication' rather than 'accurate production' of sounds. As the two main components in pronunciation teaching and research, segmentals and suprasegmentals, have undergone fluctuations in the past several decades, so have the language teaching methods, ranging from Grammar-Translation (GT) in the 1800s to Communicative Language Teaching (CLT) in the 1970s. "In what follows, I will briefly introduce a historical review of pronunciation teaching, centering around 3 major language teaching methods such as Grammar-Translation, Audiolingualism, and Communicative Language Teaching. It can be roughly said that modern language teaching began with the Grammar-Translation Teaching method in the 1800s."The teaching method sees its basic tenet in teaching the grammar rules of the target language and translating it into the native language of learners based on these rules. It follows that it has focused on the segmental properties of words, grammar teaching, and written translation between the two languages. As the method neglects the oral communicative forms of language, its aim was simply to assist EFL learners to read and comprehend foreign language literature (Larsen-Freeman, 2003). "Affirmatively, Vienne (1994) also demonstrates that the grammar translation activities will increase EFL learners' awareness not only of their mother tongue and target language but of two language cultures. For the sake of accurate communication in terms of grammar and spelling, a segmental feature, the Grammar-Translation method still contributes to the underlying skills and exercises for EFL learners (Fish, 2003)." Despite of this, Florez (1998) states that pronunciation was almost irrelevant in this method and therefore was seldom taught in the past. Later on, to keep English pronunciation standardized, language teachers needed a universal device to train their students for pronunciation. To this need came the Reform Movement in the 1980s, based upon which researchers set up the International Phonetic Association (IPA) and established the International Phonetic Alphabet (IPA), which laid the foundation for development of English phonetics (Howatt, 1984).

Deeply influenced by the Reform Movement, Audiolingualism appeared in the 1940s and had been dominant until the 1950s. It held pride of place in language teaching on a par with structural linguistics (Fries, 1945).This teaching method highly valued 'accuracy' in pronunciation. It follows that most research and teaching was implemented with a main focus on segmental features such as individual phonemes in words (Morley, 1991). Accordingly, there ensued a host of pronunciation teaching skills ranging from articulatory explanations to imitation and pattern drills, which all emphasize accuracy in pronunciation. The approach attached great importance to the sentence pattern practice and paid attention to students' pronunciation habits. Meanwhile, there were auxiliary pronunciation tools that were used for the improvement of pronunciation of ESL learners, such as the phonetic symbols, pronunciation graphic tables, etc. Thus, the effectiveness of pronunciation was

improved by making an effective use of minimal word pair drills, which take two words that differ in only one sound in the same position. Bloomfield (1933) demonstrates that a minimal distinctive sound is used for both listening practice and guided oral production.

In the 1970s, we witness a great turning point in language teaching with the advent of Communicative Language Teaching (CLT). CLT places its primary goal on 'communication' rather than the 'accurate use' of individual phonemes and words. The method stresses the need to help students achieve a certain level of pronunciation skills for fluent communication, dubbed as a 'threshold level' in the literature above which communication is not threatened by misunderstanding caused by incorrect pronunciation (Celce-Murica, Brinton & Goodwin, 1996). Gradually, people recognized that the ultimate goal of pronunciation was not to reach the phonetic level of English native speakers but to attain basic pronunciation skills for communication. To attain these basic skills, CLT put its higher values in training suprasegmental features because these features have proved to play a more significant role than segmental ones in communication. With the in-depth of communicative language teaching (CLT) method, the shift of researchers' attention was made from 'accuracy' to 'fluency'.

From the previous reviews, we have noticed the transition that has been gradually brought to the forefront over the past decades. It has been transferred from experiential pronunciation teaching to scientific one with the former dubbed as 'Intuitive-Imitative' and the latter as 'Analytic- Linguistic', respectively, in Celce-Murica et al. (1996). Until now, in the process of training the learner's communicative skills on the basis of suprasegmental features, CLT has gradually clarified its goal and set its developmental direction for pronunciation teaching toward enhancing EFL learners' communicative competence and intelligibility, which will be discussed in detail shortly.

2.2 Suprasegmentals and Intelligibility

Pennington & Richards (1986) point out that the teaching of phonetics should include segmental and suprasegmental features, pronunciation habits, and intelligibility. "Intelligibility here is often categorized into 'functional' and 'comfortable' one in the literature (Moley, 1991)." Functional intelligibility here can be defined as an ability to make oneself relatively easily understood. "As for comfortable intelligibility, they point out that if most adults are unable to achieve native-like pronunciation, language teachers need to put the aim of pronunciation teaching at attaining comfortable intelligibility rather than accuracy." Celce-Murcia et al. (1996) also take intelligible pronunciation as a model and realistic goal of teaching pronunciation. "Moley (1991), along the same line, states that the goal of pronunciation should be changed from the acquisition of native-like pronunciation to the development of functional intelligibility, communicability, increased self-confidence, speech monitoring abilities, and speech modification strategies for use beyond the classroom. According to Abercrombie (1991), comfortable intelligibility is defined as pronunciation which can be understood by the audience with little or no consciousness. In this respect, functional and comfortable intelligibility are realized on the part of the speaker and the listener, respectively."

Two dichotomic terms of intelligibility can be said to be accuracy and fluency. The two concepts are best represented by Audio-Lingual Method (ALM) in the 1940s-1950s and CLT, respectively. Accuracy was highly valued in ALM, and thereby it can be achieved through intensive and explicit pronunciation teaching. On the other hand, fluency is one of the goals of CLT.

"According to them, EFL learners below this level will encounter problems in oral communication although they are excellent in grammar and vocabulary." In other words, the communicative target for EFL learners is universally recognized as intelligibility.

It is generally noted that the acquisition of pronunciation is usually associated with pronunciation models. There are, at large, 5 standard models of pronunciation such as Received Pronunciation (RP), General American (GA), Canadian, Australian, and Indian. The first two models have been recognized to be the most dominant ones of all. The pronunciation model that I adopt here in this paper is GA mentioned above, based upon which I analyze the aspects of one Korean EFL learner in comparison with one native speaker, focusing our main attention on intelligibility at which the threshold level is set. In fact, this course is complicated in that there is little agreement as to what phonological aspects threaten an EFL leaner's intelligibility and therefore there is relatively little information that guides the students to decide what aspects of superasegmental features can cause unintelligibility (Munro and Derwing, 1995). Accordingly, Kenworthy (1987) argues that intelligibility is the most sensible goal and sets intelligibility as being understood by a listener at a given time in a given situation, "is coherently connected with the point of view that all language learners are in two extreme broad ranges, and they need to be placed at a point where their pronunciation can be understood by the listener as a threshold level of pronunciation for non-native speakers of English (Hinofotis & Baily, 1980; Celce-Murcia et al., 1996). In

connection with this issue, Flores (1997) finds that three suprasegmental aspects such as syllable, stress, and rhythm--especially the proper length of stressed syllables in the specific time and the shortening of the unstressed syllables--are a big challenge in learning English pronunciation to meet intelligibility for EFL learners. Therefore, she suggests that EFL learners should start with training in the suprasegmental features of words, then phrases, and finally move on to sentences rather than segmental features." "Consequently, I support Kenworthy's argument so that I implement an experiment on those characteristics of suprasegmental features performed by NNS in comparison with NS. According to the performance of the NNS, I believe that it is appropriate for Korean EFL learners to overcome suprasegmental obstacles, mentioned above, such as syllabification, stress, and rhythm to meet the goal for intelligibility (Chela-Flores, 1997)."

3. Suprasegmental Performance in Comparison

In this chapter, I compare and contrast between a native speaker (NS) and a non-native speaker (NNS) in their performance of syllabification, stress, and rhythm. The NS is a female who comes from the US, who is considered to be authentic in the production of these suprasegmental features. The NNS is a 3rd year middle school student, a typical EFL learner who performs these features uniquely characteristic of Korean phonetics and phonology.

3.1 Vowel Insertion in Syllabification

3.1.1 Restriction on Onset and Coda in Korean

As illustrated in syllable structure (1), a nucleus is surrounded by a consonant cluster of onset and coda, which consists of one or more consonants across languages. Syllables in Korean, unlike in English, are characterized to strictly have 1 consonant in onset and coda position. (Note 1) This strict restriction on the number of onset and coda is seen to impose a huge burden on Korean EFL learners' syllabifying English words with more than 2 consonants in the positions. Let's consider the syllabification.

(1) a. Christmas /krɪsməs/ → /krɪsɨməs/

b. spring /sprɪŋ/ → /sɨpɨrɪŋ/

(2) a. ghost /gəʊst/ → /gəʊsɨtɨ/

b. risk /rɪsk/ → /rɪsɨkɨ/

In (1a), the word Christmas has a consonant cluster /kr/ in onset, which is not allowed in Korean syllable structure. Accordingly, the neutral vowel /ɨ/ is inserted to break this cluster. (Note 2) In (1b), the neutral vowel is likewise inserted after /s/ and /p/ in order to break the consonant cluster /spr/, virtually allowing only the single consonant /r/ to become the onset. In contrast, the 2 words in (2a,b), ghost and risk, contain consonant clusters, /st/ and /sk/, in coda, respectively. The principle of Korean syllable structure, mentioned above, holds here as well, resulting in the insertion of the neutral vowel.

It is noteworthy that Korean EFL learners largely insert this neutral vowel when a syllable ends in a stop or fricative, as illustrated below.

(3) a. hand /hænd/ → /hændɨ/

b. help /help/ → /helpɨ/

c. five /faɪv/ → /faɪvɨ/

As seen in (3a,b), hand and help end in stops, /d/ and /p/, respectively, in their syllabification. Now the neutral vowel in question is added to these consonants to become a nucleus, and it forms an independent syllable /dɨ/ in (3a) and /pɨ/ in (3b) along with /d/ and /p/ as its onset, respectively. A consequence arises here that the consonant clusters in coda, /nd/ and /lp/, are broken to satisfy the Korean syllable structure, which allows only a single consonant in onset and coda. Notably, the word five in (3c) ends in a single fricative /v/, satisfying the structure in question. Despite this, most Korean EFL learners are observed to insert the neutral vowel, consequently turning fricative coda /v/ into fricative onset /v/. I attribute this to the unique aspect of syllabification in Korean that fricatives are not realized, let alone released, in coda, which enforces them to become an onset along with /ɨ/ as its nucleus, resulting in two syllables.

We note that stops and fricatives in (3) above are rarely released when they occur in coda. Even when they are released, English native speakers never insert any vowel, something like /ɨ/, in order to form an extra syllable. This is due to the fact that English allows syllables to have consonant clusters consisting of more than two consonants. This cross-linguistic difference yields a difference in articulation, as illustrated below.
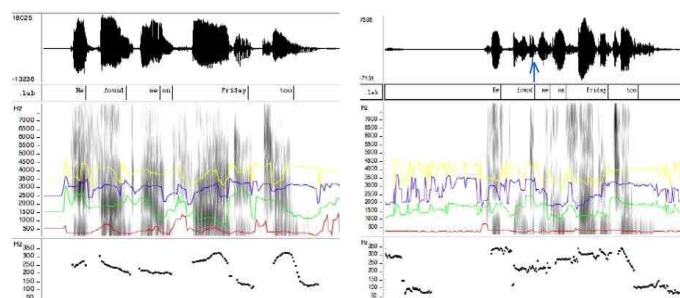
Figure 1. Syllabification of phoned "He phoned me on Friday, too."

Figure 1 above is a visual representation of the sentence He phoned me on Friday, too. A close look at the waveform of the word phoned /foʊnd/, which has a consonant cluster /nd/ in coda, reveals a difference in performance between an NS and an NNS. That is, the NNS is shown to have inserted the Korean neutral vowel /ɨ/ after /d/, pointed out by the arrow (↑), to form an independent syllable /dɨ/ along with the coda consonant /d/, resulting in two syllables, /foʊn•dɨ/, consequently breaking the consonant cluster /nd/ in coda. Thus, the vowel-insertion phenomena are often observed among Korean EFL leaners because Korean EFL learners are unconscious of vowel insertion when they produce and perceive English speech.

3.1.2 Linking

In Korean orthography, syllables in a word stand separate from and independent of each other, so that there exists a clear-cut boundary between syllables. This in part gives rise to a difference in performance of linking between an NS and an NNS. The principle of linking is that when a word that ends in a stop consonant is followed by a word that begins with a consonant, the stop consonant is usually not released. However, Korean EFL learners tend to insert the vowel /ɨ/ after the stop, as illustrated below.

(4) a: stop trying (p+t) /stop/ /traɪɪŋ/ → /stopɨ/ /traɪɪŋ/

b: big boy (g+b) /bɪg/ /bɔɪ/ → /bɪgɨ/ /bɔɪ/

c: bad judge (d+dʒ) /bæd/ /dʒʌdʒ/ → /bædɨ/ /dʒʌdʒ/

As shown above, Korean EFL learners often insert the vowel after the coda stops, virtually breaking the linkage between 2 consonants across 2 separate words. The visual sound waveform of the phrase late night shows this effect as below.
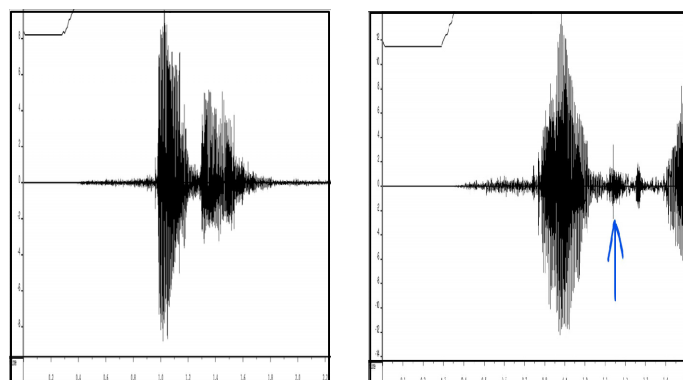

Figure 2. Linking of late night

The phrase in question consists of two words, late /leɪt/ and night /naɪt/, each of which contains one syllable. We note that the coda /t/ is followed by the onset /n/ in the phrase /leɪt•naɪt/. Now the NS is seen to have linked these 2 words (syllables) with the coda /t/ unreleased, while the NNS is seen to have inserted /ɨ/ after /t/ to form an independent syllable, as manifested by the arrow (↑), resulting in three syllables as in /leɪ•tɨ• naɪt/. (Note 3) Vowel insertion is frequently observed in the process of linking among Korean EFL learners and is discovered more noticeably in comparison with English native speakers, as exemplified in the above figure.

3.2 Stress and Word-Level Prominence

3.2.1 Equal Prominence

Stress plays a vital role in English because it is phonemic as illustrated below.

(5) a: reCORD /rɪˈkɔːd/ verb → REcord /ˈrekɔːd/ noun

b: rePORT /rɪˈpɔːt/ verb → REport /ˈrɪpɔːt/ noun

c: adDRSS /əˈdres/ verb → ADdress /ˈædres/ noun

d: exPORT /ɪkˈspɔːt/ verb → EXport /ˈekspɔːt/ noun

As seen in (5a,b,c,d), the position of stress on each pair of words changes their meaning. This phonemic feature is lexical in nature, so the stress position of every lexical item is predetermined. On the contrary, Korean lacks stress and never shows such effect; thus, stress shift does not cause the change of word meaning. It follows that Korean EFL learners have difficulties in placing stress on English words. Kang (2013) conducted an experiment of pronouncing word stress by Korean EFL learners in order to find their stress pattern of three words, different, polite, and interrupt. We note that stress falls on the first, second, and third syllable, respectively. He asked 29 Korean college students to record these words for the analysis of their stress pattern. His findings are as follows below, where E refers to 'evenly distributed over every syllable' and N/A to 'not applicable'.

Table 1. Stress Pattern (N=29)

Stress Position	DIfferent (1st)	poLITE (2nd)	interRUPT (3rd)
1st	24	11	10
2nd	0	5	0
3rd	0	N/A	5
E	5	13	14
Accuracy Rate (%):	82.7	17.2	17.2

We note that the average accuracy rate of 3 words (82.7%, 17.2%, 17.2%) is quite low, 39.03% as expected, What is noteworthy here is that the rates of the first syllable and the evenly distributed stress (E) pattern are very high, 44.82% and 48.27%, respectively.

Putting aside the first syllable pattern, I pay attention to the E pattern here and introduce the two different stress patterns of marvelous by NS and NNS.
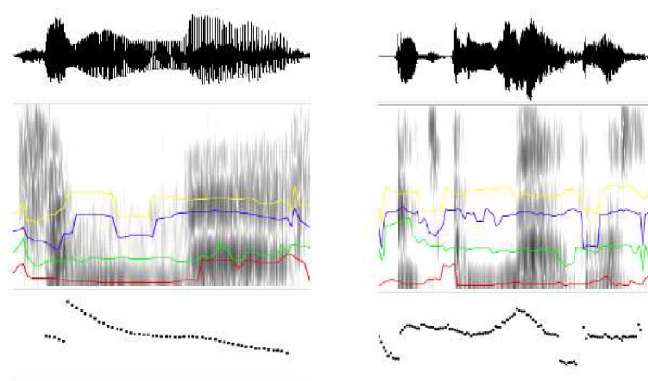

Figure 3. Stress Pattern of marvelous

Figure 3 above is a visual illustration of the word marvelous /ˈmɑːrvələs/ with the primary stress on the first syllable /mar/, which is more prominent than the neighboring unstressed syllables. It is shown that the waveform, spectrogram, and pitch contour, displayed in 3 layers, exhibit a different pattern between NS and NNS. First of all, the waveform of NS is rather continuous without interruption among 3 syllables. The spectrogram of NS manifests that he places stress on the first syllable as the first chunk appears to be thick. In contrast, NNS is seen to have inserted the neutral vowel /ɨ/ after the fricative /s/ and to display interrupted syllable boundaries as manifested by 4 distinct chunks of waveforms. The spectrogram performed by NNS shows 'evenly distributed

chunks' across three syllables. The visual representation of stress above exhibits the characteristics of Korean EFL learners' stress pattern, in which every syllable in a word appears to approximately receive equal prominence.

3.2.2 Initial Prominence Phenomenon

Lee & Rhee (2018) assert that Korean EFL learners are observed to apply Korean phonological rules to their speaking of English, in which the first syllable of (Korean) words is stressed. Kang (2013) confirms this fact by saying that Korean EFL learners tend to place a primary stress on the first syllable in English words, the placement of which he calls an 'initial prominence phenomenon'. A visualization of this stress pattern is shown below for expository purposes.
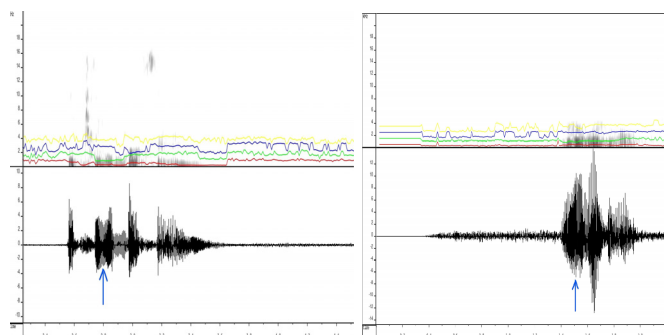
Figure 4. Stress Pattern of photography

As shown above, the second chunk of photography /fəˈtɑgrəfi/ is much thicker and darker in NS's waveform pattern, which represents that the stress is put on the second syllable of this word. In addition, the second chunk of photography in the NS's spectrogram forms a dark shade column, which exhibits the same effect as well. In contrast, the first chunk of the word for NNS indicated by the arrow (↑), is in a more darken shade than the other 3 chunks. Clearly, this constitutes an exemplary evidence for the 'initial prominence phenomenon', which is generally observed among Korean EFL learners.

3.3 Rhythm and Sentence-Level Prominence

Rhythm, in English, is comprised of stressed and unstressed syllables, the reduction of function words (Avery & Ehrlich, 1992), and the features of connected speech (Celce-Murica, 2010). "There are 2 major types of rhythm across languages: syllable-timed rhythm and stress-timed rhythm. In stress-timed rhythm, in particular, each stressed and unstressed syllable(s) that follow(s) together form a metrical unit, which is called 'foot', occurring in relatively regular alternating patterns. In general, a foot is primarily formed by combining stressed and unstressed syllables but not simply by concatenating words."

3.3.1 Content and Function Words

In English, content and function words often constitute a dichotomy in terms of pronunciation. As a first approximation, content words are words that have the most information in a sentence. Roach (2009) notes that in real connected speech, the effect of content and function words is a universal case of junction in which refers to the relationship between one weak and one strong sound in an utterance that alternates between them. Content words are generally categorized into syntactic attributes such as nouns, main verbs, adverbs, adjective, question words (who, where, what, when, why, how), demonstratives (this, that). In contrast, function words are considered as pronouns, prepositions, articles, 'to be' verbs, conjunctions (and, but), and auxiliary verbs (can, have, do, will).

The content words are usually easier to hear because they are given extra emphasis in an utterance. Words are emphasized by adding extra length to their stressed syllables as follows.

(6) a. Noun: This is my cat.

b. Main verb: What does it eat?

c. Adverb: Please come quickly.

d. Adjective: You did excellent work.

e. Question word: Why did you write the letter?

As illustrated above, the italicized words are taken to receive extra emphasis and length according to the dichotomic classification above.

3.3.2 Two Rhythms

Rhythm, in English, is comprised of stressed and unstressed syllables, the reduction of function words (Avery & Ehrlich, 1992), and the features of connected speech (Celce-Murica, 2010). As for connected speech, rhythm involves a variety of phonological aspects such as contractions, blends, reductions, linking words and phrases, assimilation, dissimilation, and epenthesis (Celce-Murcia, 2010). Rhythm is cross-linguistically divided largely into syllable-timed rhythm and stress-timed rhythm in terms of articulatory timing.

1) Syllable-Timed Rhythm

In syllable-timed rhythm, each syllable is given approximately equal time, so the overall length of a spoken sentence depends on the number of syllables in it, which is observed in Korean, Spanish, French, etc. In this rhythm, we do not witness the phenomenon of vowel reduction feature even if unstressed syllables exist. Additionally, the length of each syllable in question is equal (Schaetzel & Low, 2009). Accordingly, the length of an utterance in a syllable-timed language does not depend on the number of stressed syllables, but rather on the number of syllables. According to Avery & Ehrlich, (1992), EFL students who natively speak a syllable-timed language tend to assign equal weight to every syllable in English sentences, no matter whether the syllable is stressed or unstressed. This phenomenon might make their speech form a staccato-like rhythm that negatively influences the intelligibility of their English. A cross-linguistic comparison of 2 rhythms can be illustrated as below (Celce-Murcia et al., 2010).

(7) a. Los lib-ros es-tan en la me-sa. (Spanish) ˆ   ˋ ˆ   ˆ   ˆ ˋ ˆ

b. Les livres sont sur la ta-ble. (French) ˆ   ′   ˋ   ˋ   ˋ   ′

c. The books are on the table. (English)

As we note above, the sentences of Spanish and French, a language of syllable-timed rhythm, do not exhibit a primary stress (′) on words. By contrast, the primary stress (′) in English falls only on words that are classified as content words such as books and table.

Korean is typologically regarded as a language of syllable-timed rhythm, whereas English as a language of stress-timed rhythm. In syllable-timed rhythm, first of all, each syllable in a word is said to receive an approximately equal prominence, and its timing thus depends on the number of syllables in a sentence. Let us look at the paradigm of sentences below with the different number of words (syllables). (Note 4)

(8) a. Birds eat worms. (3 syllables)

b. The birds eat worms. (4 syllables)

c. The birds eat the worms. (5 syllables)

d. The birds will eat the worms. (6 syllables)

According to the basic tenet of syllable-timed rhythm, illustrated above, we expect that all the sentences will take different time for articulation with (8a) the shortest and (8d) the longest. This is due to Korean EFL learners' articulatory trait that they put an approximately equal prominence on every syllable, which in turn renders it to take equal time. It follows that the timing for a whole sentence is proportionate to the number of words (syllables) in it. The waveforms of (8a-d) are visualized as below for further discussion.
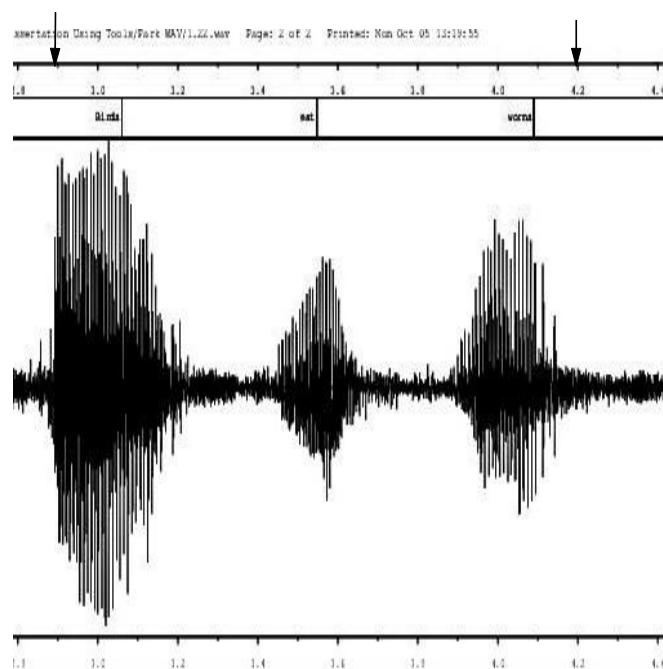
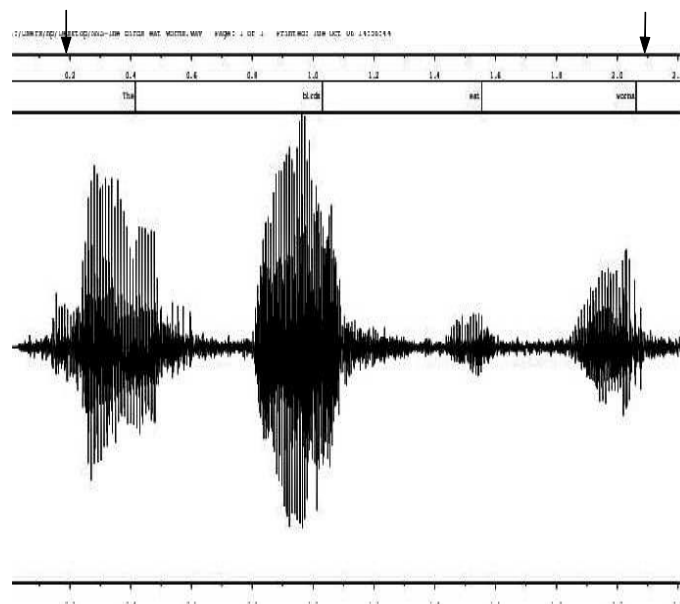Figure 5. The Rhythm Pattern of 3 Syllables by NNS "Birds eat worms."

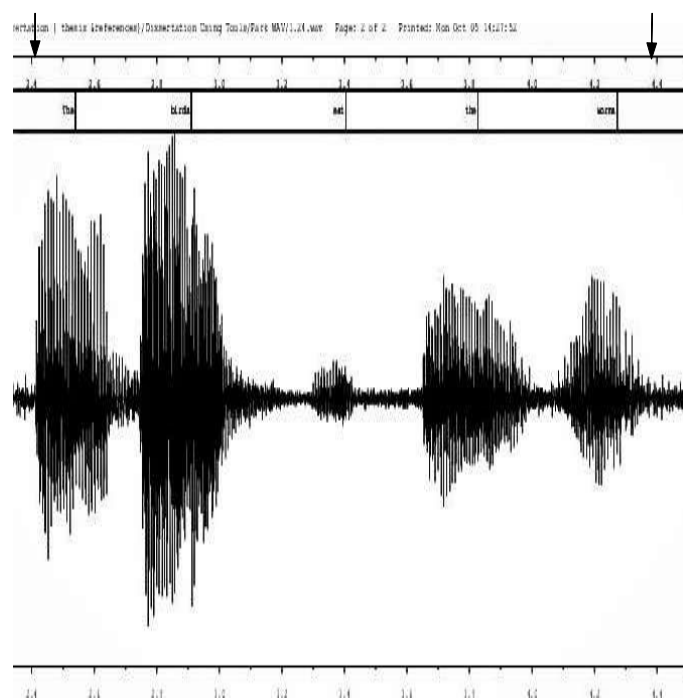Figure 6. The Rhythm Pattern of 4 Syllables by NNS "The birds eat worms."

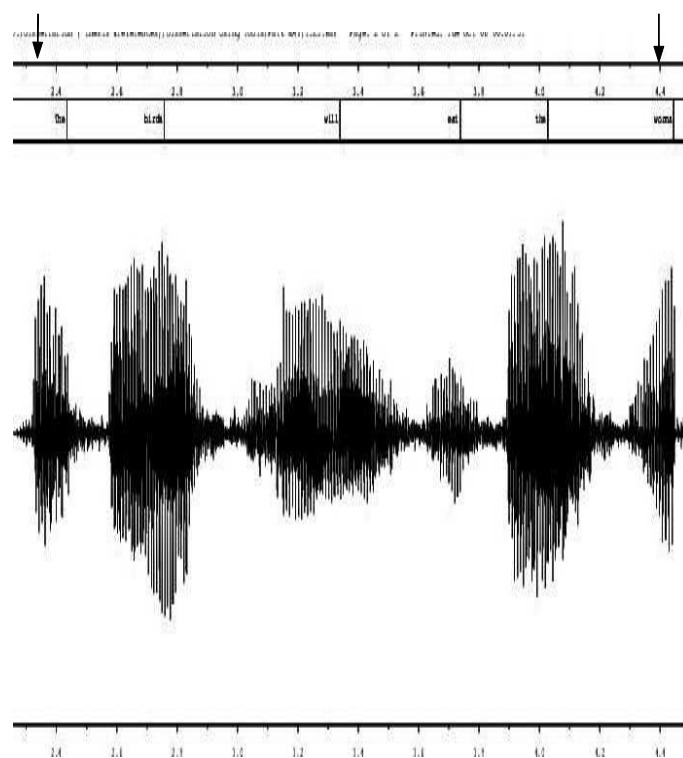Figure 7. The Rhythm Pattern of 5 Syllables by NNS "The birds eat the worms."

Figure 8. The Rhythm Pattern of 6 Syllables by NNS "The birds will eat the worms."

Figure 5 displays the waveform of sentence (8a) Birds eat worms. We obviously note three distinct chunks of waveforms, each one of which represents birds, eat, and worms, respectively. The sentence in (8b), The birds eat worms, consisting of 4 syllables (words), is displayed in visualization in Figure 6. It is obvious that this sentence is represented visually as four distinct chunks of waveforms. Interestingly, however, the third word (syllable), eat, is represented as the smallest waveform chunk despite it being a content word. The other 2 figures, Figure 7 and

8, exhibit the two sentences in (8c) and (8d), respectively. As displayed in these figures, the number of words (syllables) in both sentences is clearly represented in the form of wave chunks, 5 for (8c) and 6 for (8d), respectively. Even in these rhythm patterns, the content word eat is represented as the weakest in terms of prominence and the smallest in terms of waveform. Looking at the 4 rhythm patterns, displayed in Figure 5 through 8, we can conclusively say that the NNS has failed to differentiate between content words and function words, thereby placing an approximately equal prominence on each word (syllable). This in turn leads to different timing in each sentence, which will be illustrated shortly.

In Figure 5 through 8, I have indicated the starting point and the end point of articulation for the four sentences in (8) by the arrows (↓) so that we can calculate the timing of each sentence. The timing is thus calculated by subtracting the starting point from the end point. In Figure 5 for (8a), for instance, articulation starts at the point of 2.9 and ends at the point of 4.2, resulting in the timing of 1.3 seconds. Now we end up with 1.9, 2.0, and 2.1 for (8b), (8c), and (8d), respectively. It is obvious that the timing increases according to the number of words (syllables) of each sentence, consequently conforming to the basic tenet of syllable-timed rhythm.

2) Stressed-Timed Rhythm

The stress-timed rhythm is a rhythm in which stressed syllables occur at roughly equal intervals, regardless of the number of unstressed syllables occurring between the stressed ones (Clark & Yallop, 1995). Accordingly, the timing of stress-timed rhythm is determined not by the number of sentence segments or the number of syllables but by the number of prominent syllables. The pattern of stress-timed rhythm is illustrated in comparison with syllable-timed rhythm as follows (Prator & Robinett, 1985).

3.3.3 Stressed–Timed Rhythm and Foot

"When native speakers speak, they stress content words but reduce the sound of function words. Stressed and unstressed syllables usually occur in relatively regular alternating patterns in a phrase or sentence. A foot is formed as each stressed and unstressed syllable(s) that follow(s) it are grouped together to form a metrical unit. Rhythm in turn is created as feet are repeated in a phrase or sentence (Celce-Murcia et al., 2010).Especially in English, the stress-timed rhythm typically consists of regular and patterned feet of stressed and unstressed syllables." In this rhythmic pattern, the stressed syllable of a content word and the unstressed syllables of (a) function word(s) which follow(s) it are combined to be pronounced as a metrical unit, called 'foot' in English. Therefore, each foot is assumed to take an approximately equal time no matter how many syllables it contains. The rhythmic pattern of the sentences in (4) can be represented as (9), in which content words are in bold and the vertical line (｜) indicates a foot boundary.

(9) a. **Birds** ｜ **eat** ｜ **worms**. (3 feet / 3 syllables)

b. The **birds** ｜ **eat** ｜ **worms**. (3 feet / 4 syllables)

c. The **birds** ｜ **eat** the ｜ **worms**. (3 feet / 5 syllables)

d. The **birds** will ｜ **eat** the ｜ **worms**. (3 feet / 6 syllables)

We note above that each sentence in (9) contains a different number of syllables, but it unanimously has 3 feet regardless of the number of syllables in it. Now the syllables (words) in each section divided by boundaries are fused into a single unit called 'foot'. Sentence (9d), for example, has three feet with emphasis on the content words in bold, birds, eat, and worms, respectively. The other syllables (words), the, will, and the, are attached to the content words, birds and eat, consequently maintaining 3 feet likewise as in (9a-c). Now a stress-timed rhythm is established in (9a-d) by uttering 3 feet at regular intervals with each stressed syllable given more prominence than the other unstressed syllables. As each foot is produced at a regular interval, all the sentences in (9) are presumed to take generally the same amount of time because of the same number of feet regardless of a different number of syllables in each sentence. It follows that the timing of 4 sentences in (9) is roughly the same if they are articulated according to stress-timed rhythm. The rhythm patterns of the above 4 sentences are displayed below.
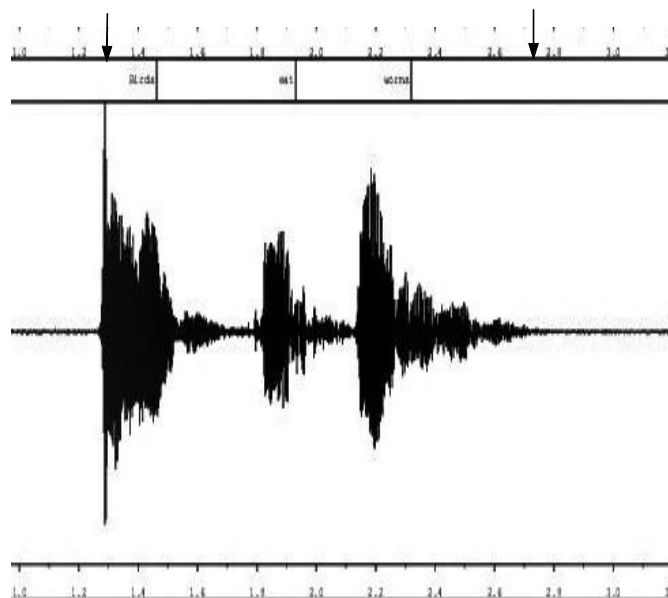
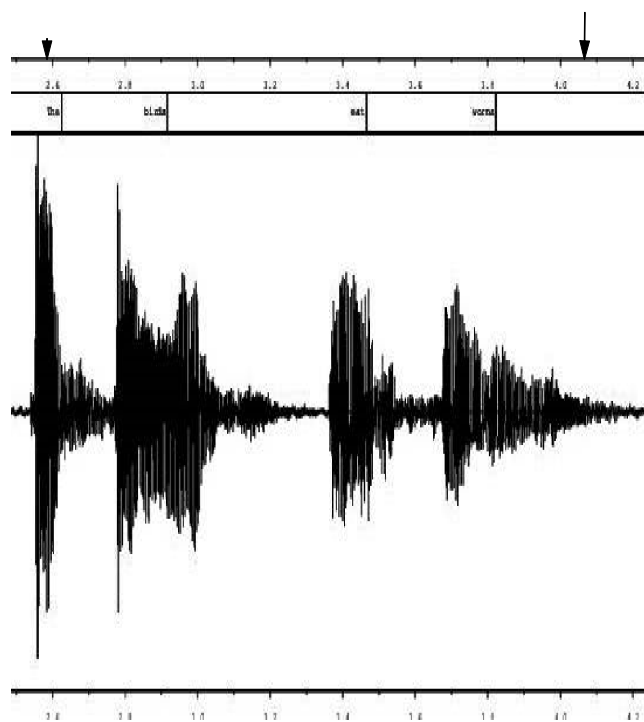Figure 9. The Rhythm Pattern of 3 Feet / 3 Syllables by NS "Birds eat worms."

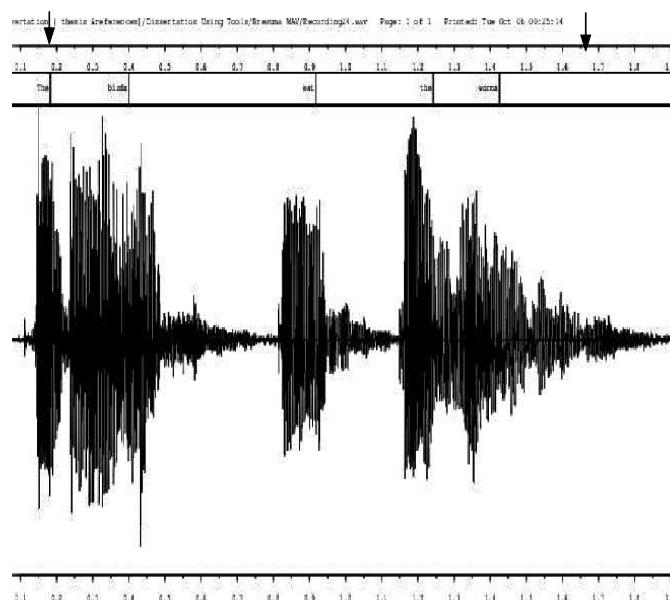Figure 10. The Rhythm Pattern of 3 Feet / 4 Syllables by NS "The birds eat worms."

Figure 11. The Rhythm Pattern of 3 Feet / 5 Syllables by NS "The birds eat the worms."
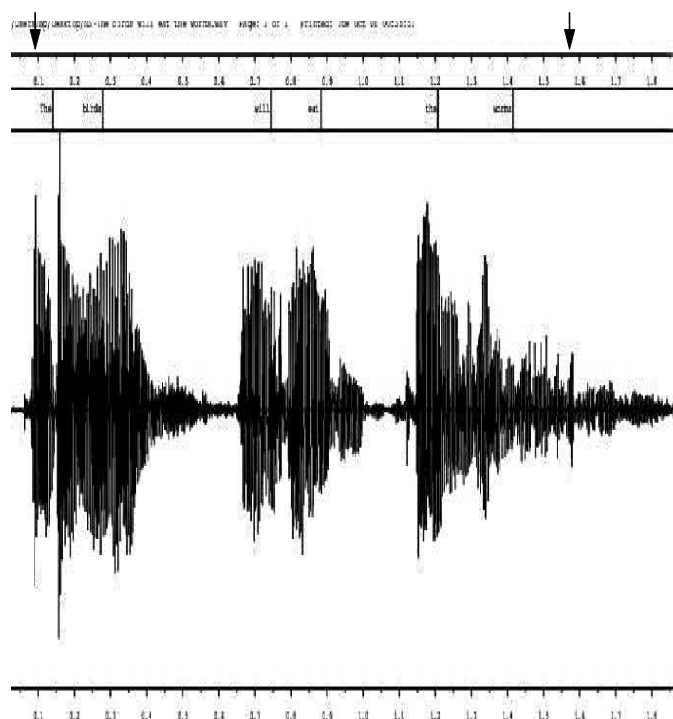

Figure 12. The Rhythm Pattern of 3 Feet / 6 Syllables by NS "The birds will eat the worms."

Figure 9 shows the waveform of sentence (9a) Birds eat worms. We clearly note three dark chunks of waveforms. Each of them represents birds, eat, and worms, respectively. The rhythm pattern of sentence (9b), The birds eat worms, with 4 syllables (words) is also illustrated and visualized in 3 distinct chunks in Figure 10. What concerns us here is the 1st chunk, the birds, which seemingly looks like 2 chunks. A sharp contrast arises here between the chunk of the birds by NNS in Figure 6 and that by NS in Figure 10. In the former, there exists a rather long interval between the and birds. This indicates that the NNS has articulated the two syllables (words) as two separate feet with an equal prominence on each of them. In the latter, however, we notice a short interval between the and birds, which enables us to say that the two are completely fused into a single metrical unit called foot.

The other 2 figures, Figure 13 and 14, exhibit the two sentences in (9c) and (9d), respectively. As displayed in

these figures, the number of words (syllables) in both sentences is clearly represented in the form of wave chunks, 3 for (9c) and 3 for (9d), respectively. Even in these rhythm patterns, the function words, the and will, are shown to be attenuated in their waveforms adjacent to the emphasized syllables. Looking at the 4 rhythmic patterns, illustrated in Figure 9 through 12, we can arguably say that the NS has created 3 metrical feet in each sentence in (9) by placing an approximately equal prominence on each metrical foot. This in turn leads to equal timing in each sentence according to the number of metrical feet, which will be illustrated shortly.

In Figure 9 through 12, I have indicated the starting point and the end point of articulation for the four sentences in (9) by the arrows (↓) so that we can calculate the timing of each sentence. The timing is thus calculated by subtracting the starting point from the end point. In Figure 11 for (9a), for instance, articulation starts at the point of 1.3 and ends at the point of 2.75, resulting in the timing of 1.45 seconds. Now we likewise end up with 1.45 seconds all for (9b), (9c), and (9d). It is obvious, therefore, that the equal timing for 3 metrical feet of each sentence is coherent to the basic tenet of stress-timed rhythm.

4. Conclusion

The findings of this study reveal that there exist huge differences in suprasegmental features between English and Korean in terms of syllable structure, stress, and rhythm. "First of all, Korean allows only one consonantal element in onset and coda position, which will pose a great difficulty for Korean EFL learners who articulate English syllables, in which a maximum of 3 and 4 consonants are allowed in onset and coda position, respectively. Accordingly, a process of resyllabification arises here." That is, Korean EFL learners are often seen to insert a neutral vowel /ɨ/ in order to break consonant clusters. This process is also shown to take place even with single stop and fricative consonants in coda position such as /p,b,t,d,k,g,s,z/, as visualized in Figure 2, in which the insertion of the vowel in the phrase late night, as a linking device, forms an independent wave chunk indicated by the arrow. According to stress, it represents two characteristics. One is that the lack of stress among Korean EFL Learners, naturally leads to placing an equal prominence on every syllable in a word, as displayed visually in Figure 3, in which 4 syllables of the word marvelous, articulated by NNS, are stressed approximately equally. The other characterization is what I call 'initial prominence phenomenon', in which they tend to put a primary stress on the initial syllable of multi-syllable words, as visualized in Figure 4, in which the first wave chunk of photography appears thicker and darker than the others. In correlation with rhythm, Korean is treated to belong to languages of syllable-timed rhythm, which is defined as a rhythm in which the timing is determined by the number of syllables (words) in a sentence. Given that, Korean exhibits an evenly distributed stress pattern among syllables in words, it is obvious that the articulation timing of a sentence is proportional to the number of syllables (words) contained in a sentence. This rhythmic pattern is exactly shown visually in Figures 5, 6, 7, and 8, which contain 3, 4, 5, and 6 syllables (words), respectively. Accordingly, the timing of these 4 sentences is differentiated and visually shown to be 1.3, 1.9, 2.0, and 2.1 seconds, respectively, consequently conforming to the basic tenet of syllable-timed rhythm. Ultimately, this paper attempts to exhibit different suprasegmental features in phonology between English and Korean, and it aims at applying the experimental analysis to enhance the level of intelligibility and intercultural communication for Korean EFL learners in their foreign language teaching and learning. "Given that," English pronunciation education in Korean EFL settings should swing its pendulum to the focus on the teaching of suprasegmental features. This paper finds its contribution in the English pedagogical approach to pronunciation teaching to Korean EFL learners by showing the visualized patterns of such features in comparison between a native and non-native speaker of English.

References

Abercrombie, D. (1991). Teaching pronunciation. In A. Brown (ed.), Teaching English Pronunciation. *A Book of Readings, 7*(5), 967-972. London: Routledge.

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgements of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning, 42*, 529-555. https://doi.org/10.1111/j.1467-1770.1992.tb01043.x

Avery, P., & Ehrlich, S. (1992). *Teaching American English Pronunciation*. Oxford: Oxford University Press.

Bloomfield, L. (1933). *Language*. New York: Henry Holt.

Celce-Murcia, M., & McIntosh, L. (1991). *Teaching English as a Second or Foreign Language*. Cambridge, MA: New Burry House Publishers.

Celce-Murcia, M., Briton, D. M., Goodwin, J. M., & Griner, B. (2010). *Teaching Pronunciation*: *A Course Book and Reference Guide* (2nd ed.). Cambridge: Cambridge University Press.

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching Pronunciation: A Reference for Teachers of English to Speakers of Other Language*. Cambridge: Cambridge University Press.

Chela-Flores, B. (1997). Rhythmic patterns as basic units in pronunciation teaching. *ONOMAZEIN, 2*, 111-134. https://doi.org/10.7764/onomazein.2.05

Clark, J., & Yallop, C. (1995). *An Introduction to Phonetics and Phonology*. Oxford: Blackwell.

Derwing, T. M., & Rossiter, M. J. (2002). ESL learners' perceptions of their pronunciation needs and strategies. *System, 30*, 155-166. https://doi.org/10.1016/S0346-251X(02)00012-X

Fant, Gunnar. (1970). *Acoustic Theory of Speech Production with Calculations Based on X-ray Studies of Russian Articulations*. Boston: Mouton. https://doi.org/10.1515/9783110873429

Florez, M. C. (1998). Improving adult ESL learners' pronunciation skills. *Phonetics, 25*(2), 169-186. https://doi.org/10.1006/jpho.1996.0040

Fish, S. E. (2003). *Is There a Text in This Class? - The Authority of Interpretative Communities* (12th ed.). Boston: Harvard University Press.

Fraser, H., & Perth, H. F. (1999). ESL pronunciation teaching: Could it be more effective? *Australian Language Matters, 7*(4), 7-8.

Fries, Charles C. (1945). *Teaching and Learning English as a Foreign Language*. Ann Arbor, MI: The University of Michigan Press.

Gilakjiani, Abbas P. (2012). The significance of pronunciation in English language teaching. *English Language Teaching, 5*(4), 96-107. https://doi.org/10.5539/elt.v5n4p96

Gilbert, J. B. (1993). *Clear Speech*. Cambridge: Cambridge University Press.

Hammer, J. (2007). *How to Teach English*. Harlow: Pearson Education.

Hinofotis, F. B., & Bailey, K. M. (1980). American undergraduates' reaction to the communication skills of foreign language teaching assistants. In J. C. Fisher, M. A. Clarke, & J. Schachter (Eds.), *On TESOL 80 Building Bridge: Research and Practice in Teaching English as a Second Language*. pp. 120-136. Washington, DC: TESOL.

Hoard, J. E. (1971). Aspiration, tenseness and syllabification in English. *Language, 47*, 133-140. https://doi.org/10.2307/412191

Howatt. A. P. R. (1984). *A History of Language Teaching*. Oxford: Oxford University Press.

Kang, Seung-Man. (2013). Stress, rhythm, and intelligibility in English. *The Jungang Journal of English Language and Literature, 55*(3), 53-76. https://doi.org/10.18853/jjell.2013.55.3.003

Kelly, L. G. (1969). *Centuries of Language Teaching*. Rowley, MA: Newbury House.

Kenworthy, J. (1987). *Teaching English Pronunciation*. London: Longman.

Larsen-Freeman, D. (2003). *Techniques and Principles in Language Teaching* (2nd ed.). Oxford: Oxford University Press.

Lee, J. K. (2008). An intonational assessment of English communicative competence of Korean college students with the curriculum. *Modern English Education, 9*(1), 256-282.

Lee, Jeong-Hwa., & Rhee, Seok-Chae. (2018). Acoustic analysis of Korean trisyllabic words produced by English and Korean speakers. *Phonetics Speech Science, 10*(2), 1-6. https://doi.org/10.13064/KSSS.2018.10.2.001

Lee, J. K. (2008). An intonational assessment of English communicative competence of Korean college students with the curriculum. *Modern English Education, 9*(1), 256-282.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility and intelligibility in the speech of second language learners. *Language Learning, 45*, 73-97. https://doi.org/10.1111/j.1467-1770.1995.tb00963.x

Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly, 25*(3), 481-520. https://doi.org/10.2307/3586981

Murphy, J. (2003). Pronunciation. In D. Nunan (ed.), *Practical English Language Teaching.* pp. 111-128. Boston: McGraw-Hill.

Pennington, M., & Richards, J. (1986). Pronunciation revisited. *TESOL Quarterly, 21*(2), 207-225. https://doi.org/10.2307/3586541

Prator, C., & Robinett, B. (1985). *Manual of American English Pronunciation* (4th ed.). New York: Holt, Reinehart, and Winston.

Pulgram, E. (1970). *Syllable, Word, Nexus, Cursus.* The Hague: Mouton. https://doi.org/10.1515/9783110815443

Roach, P. (2009). *English Phonetics and Phonology: A Practical Course.* (4th ed.). Cambridge: Cambridge University Press.

Schaetzel, K., & Low, E. (2009). CAELA network brief. *Teaching Pronunciation to Adult English Language Learners.* Washington, DC: Center for Applied Linguistics.

Selkirk, E. O. (1980). The role of prosodic categories in English word stress. *Linguistic Inquiry, 11*(3), 563-605.

Selkirk, E. O. (1981). Epenthesis and degenerate sllable in Cairene Arabic. *MIT Working Papers in Lingusitcs, 3*(2), 209-232.

Vienne, J. (1994). Towards a pedagogy of "translation in situation" perspectives. *Studies in Translatology, 2*(1), 51-59. https://doi.org/10.1080/0907676X.1994.9961222

Xu, L. H. (1991). Developing student confidence in speaking English. *Modern English Teacher, 3*(3), 74-89.

Yate, K. Y. (2003). *Teaching Linguistic Mimicry to Improve Second Language Pronunciation.* Ph. D. Dissertation, University of North Texas.

Notes

Note 1. Syllables in English are observed to maximally have 3 and 4 consonants in onset and coda, respectively, as in *spring /sprɪŋ/* and *contempts /kən'tempts/*.

Note 2. The neutral vowel /ɨ/, analogous to /ə/ in English, is uniquely found in Korean and constitutes a nucleus for any consonant excluded from onset as in *spring* /sprɪŋ/ → /sɨpɨrɪŋ/ (CCCVCC → CVCVCV).

Note 3. Note that the visual waveform pointed by the arrow (↑) in Figure 2 does not necessarily manifest the insertion of the neutral vowel /ɨ/. However, the sound recording of this phrase clearly shows the insertion of the vowel of our concern.

Note 4. In this paper, I use the terms *word* and *syllable* interchangeably because the distinction of them is blurred when words in a sentence are syllabified.

Copyrights